

Open Science ved Aarhus Universitet
Data management baseret på FAIR principperne

Version 2.0

Udarbejdet af

Birte Christensen-Dalsgaard

Versionskontrol

Versionsnr.	Dato	Forfatter	Ændringer	Status
1.0	2021-11-18	BCD		Udkast til behandling i forskningsudvalget
1.1	2022-01-03	BCD	Strategiske målepunkter, organisation - efter første møde med Anne Marie Pahuus og Hans Erik Bøtke	Internt arbejdsblad
1.2	2022-01-14	BCD	De strategiske målepunkter er gjort mere mundrette, resume er opdateret samt afsnit om organisering opdateret efter behandling på forskningsudvalgsmøde samt efter møder med og input fra Anne Marie Pahuus og Hans Erik Bøtke. Konsekvensrettelser er gennemført.	Nyt udkast til prodekaner
1.3	2022-02-03	BCD	Efter diskussion med Niels Mejlgaard og Brian Vinter	Udstræk af rapporten med fokus på Data management
2.0	2022-06-29	BCD	Efter input fra Anne-Mette Hvas er koordinationsaktivitet initieret af OS-forum og med et kommissorium. Efter input fra Niels Mejlgaard er de specifikke mål udeladt og skal aftales. Konsekvensrettelser gennemført i dokumentet.	Udkast til OS-forum

Indhold

1	Baggrund.....	1
2	Strategiske mål	4
3	Aktiviteter.....	6
4	Omkostninger	7
4.1	Central organisation.....	7
4.2	Fakultetsforankring.....	8
5	Anbefalinger.....	9
6	Bilag 1: Udtræk fra den nationale strategi	10
6.1	Det forventes at AU adresserer nedenstående områder:.....	10
6.2	Det forventes, at AU kan agere i forhold til en national (DeiC) og international agenda....	11
7	Bilag 2: Den nationale følgegruppes fokusområder.....	12
8	Bilag 3: Kommissorium for Data management koordinationsgruppen	13

1 Baggrund

Som beskrevet i "Open Science ved AU, Organisation" betyder digitaliseringen af såvel dataindsamling som de omkringliggende processer nye muligheder for forskningssamarbejder og er basis for Open Science agendaen, som arbejder mod åbenhed, stringens og transparens i alle faser af forskningsprocessen.

EU understøtter Open Science agendaen og ser adgang til forskningsresultater som en strategisk komponent i forhold til en øget konkurrenceevne for industrien i EU. Dette vil berøre forskere ved AU, idet ansøgninger til EU's forskellige programmer skal vedlægges en data management plan, hvor der redegøres for planer om publicering/deling af såvel artikler som andre komponenter af forskningens output i form af data, metoder og computerkoder i henhold til FAIR principperne (Findable, Accessible, Interoperable, Reusable). Elementerne i FAIR er illustreret nedenfor i figur 1 kopieret fra Australien¹.



Figur 1 Mulige elementer i at gøre data FAIR.

At data er FAIR betyder, at de kan findes og potentielt genbruges af andre, universitetsforskere såvel som udviklere og forskere i virksomheder. Det betyder ikke, at data nødvendigvis er frit tilgængelige, men at de er "As open as possible, as closed as necessary" og at deres eksistens er delt. Data management er processen, som ikke kun sikrer, at data samt relevant procesinformation indsamles struktureret og løbende, så de kan publiceres, men som også er med til at sikre forskningsintegritet² gennem transparens i hele forskningsprocessen.

Datadreven forskning, hvor forskningsresultater ikke alene bygger på egne data, men også på andres data, er ved at vinde frem. Astronomer har i mange år haft succes med at samarbejde om det såkaldte "virtuelle

¹ Fra <https://www.andis.org.au/working-with-data/fairdata/training>

² I forskriften: "Håndtering og opbevaring af forskningsdata skal inden for alle fagområder gennemføres i henhold til gældende regler, så der sikres transparent og troværdig forskning"

observatorie”, hvor data fra mange fysiske observatorier samles. Det samme gælder inden for sprogvidenskab, hvor forskere har baseret forskning på CLARIN VO³, som giver adgang til mere end 1,2 mio. tekster på forskellige sprog. Inden for partikelfysik har man på CERN⁴ samlet mere end to petabytes af åbne data, som sammen med en politik om at dele koder og workflows har været med til at booste området. Endelig kan nævnes, at de såkaldte metastudier også er ved at vinde frem inden for sundhed ligesom hele indsatsen omkring personlig medicin bygger på, at data er tilgængelige for analyser.

Med et datadrevet samfund bliver algoritmer og data ”guldet” og mange argumenterer derfor for, at universiteterne og forskerne skal beholde retten til brugen af deres data og ikke gentage fortidens synder med at forære al forskning i form af artikler til tidsskrifter. En række tidsskrifter kræver i forbindelse med review, og ofte også i forbindelse med publicering, enten aflevering af eller adgang til en sikker og valid version af data og computerkode. Adgangen kan ofte ske ved brug af en såkaldt persistent identifikator, PID, i stedet for, at forskeren ”forærer” tidsskriftet sine data. At få en PID kræver, at data (og kode) sikres mod sletning og ændringer; dette kan ske via et data publicerings repository. Der eksisterer allerede flere både generelle (f.eks. Zenodo) og fagspecifikke repositories, ligesom der lægges op til at skabe et dansk repository i DeIC regi. Pointen er, at der hos samtlige forskere skal skabes en forståelse for vigtigheden af og mulighederne for at bevare og styre brugen af egne data.

Det skal endvidere bemærkes, at det datadrevne samfund også vil stille andre krav til nyuddannede kandidaters kompetencer, herunder deres viden om, og erfaring med, korrekt håndtering af data og arbejde med data.

Ministeriets netop udgivne strategi, ”National Strategi for data management baseret på FAIR principperne”⁵, er baseret på samme observation. Strategien er en del af udmøntningen af [EU’s direktiv om åbne data](#) og den danske implementering af dette i PSI-loven. Den skal dermed bidrage til at opfylde forskningens og samfundets forventning til stigende tilgængelighed af offentligt finansieret forskning. Strategien peger på initiativer, som sikrer, at dansk forskning forbliver en integreret del af det internationale data-økosystem ved at dele data og integrere med bl.a. European Open Science Cloud (EOSC). Den Nationale Strategi opererer med en række principper og med deres implementering med dertil knyttede aktører. Strategien peger på nødvendigheden af såvel en teknisk infrastruktur som en supportinfrastruktur, som kan assistere forskerne i alle faser af forskningsprocessen. I bilag 1 er udtrukket alle punkter, som kræver handling fra universitetets side – enten i form af ting, som bør leveres eller aktiviteter, universitet bør forholde sig til. Samtidig har styrelsen nedsat en følgegruppe⁶, hvis kommissorium identificerer 13 punkter, hvor universiteterne forventes i samarbejde at levere resultater. De 13 punkter er indsat i bilag 2.

I forhold til forskningen vil et øget fokus på håndtering af data i hele livscyklus sikre:

³ Se <https://vo.clarin.eu/?1>

⁴ Se: <https://home.cern/news/press-release/knowledge-sharing/cern-announces-new-open-data-policy-support-open-science>

⁵ <https://www.deic.dk/data-management/nationalt-samarbejde/ny-national-strategi-fair-data-management>. AUs høringssvar blev behandlet på ledelsesmødet 27/1 2021

⁶ <https://ufm.dk/aktuelt/nyheder/2021/uddannelses-og-forskningsstyrelsen-nedsaetter-en-folgegruppe-der-skal-fa-ny-national-strategi-for-forskningsdata-til-at-blive-til-virkelighed>

- **Transparens og troværdighed** via åben adgang til forskningsdata og publikationer
- **Ansvarlighed og reproducerbarhed** af forskningsresultater

Afslutningsvis skal nævnes Rigsarkivets ”Bekendtgørelse om anmeldelse af digitale forskningsdata skabt af statslige myndigheder”⁷. Der er netop gennemført en række tests med anmeldelse og aflevering fra forskergrupper. Disse test har vist, at arbejdet med anmeldelse og aflevering simplificeres ved at have haft en god data management praksis gennem projektførelsen, at det kan være endog meget tidskrævende at gennemføre en aflevering og endelig, at universiteterne kan blive pålagt⁸ selv at stå for bevaring af data i (internationale) formater, som Rigsarkivet ikke kan håndtere. For nogle forskere vil dette være en fordel, da de driver et open acces data repository, men der kan komme udfordringer med finansieringen ved projektporteføljens ophør, for andre vil det være problematisk, hvis forskerne ikke havde planlagt en bevaring ud over de fem år foreskrevet i Code of Conduct. Under alle omstændigheder rejser det en interessant problemstilling omkring finansiering af egenudviklede data repositories, som bevares efter en bevaringsbestemmelse fra Rigsarkivet.

Der er derfor rigtig god grund til at implementere AUs egen digitaliseringsstrategi, hvori der står: ”Aarhus Universitet vil støtte de faglige miljøer med teknologi og viden til at sikre overholdelse af gældende krav – herunder eksterne krav ifbm. myndighedsbetjening – til performance, dataadgang, datadeling og databeskyttelse inden for forskning (Research Data management og Open Access).” (cit. digitaliseringsstrategien).

I det tilhørende roadmap⁹ er dette konkretiseret ved: *Den digitale understøttelse af forskningsområdet skal ifølge digitaliseringsstrategien ske således, at forskerne har frihed og understøttet adgang til at vælge it-løsninger i alle forskningens faser inden for rammer, der skal sikre, at universitetet kan opretholde tilstrækkelig beskyttelse og sikkerhed i håndteringen af data (Data management).* I forhold til dette roadmap udestår den sidste af tre aktiviteter: *Udarbejdelse og implementering af data management 5 års plan.* Dette notat giver en plan for implementeringen og skitserer, hvordan den kan implementeres.

Notatet har primært fokus på data, men rigtig mange af argumenterne vil også dække over eksempelvis metoder og koder. Det opstiller en række strategiske mål samt giver forslag til en organisering, som kan sikre gennemførelsen af disse. Det er skrevet under antagelse af, at beslutningen om at etablere af en relevant teknisk infrastruktur gennemføres. Denne infrastruktur vil sikre, at AU lever op til ”Den danske kodeks for integritet i forskning”¹⁰, hvor der står: *Institutionerne er ansvarlige for at have sikre dataopbevaringsfaciliteter, der er i overensstemmelse med gældende krav om fortrolighed samt andre relevante bestemmelser og retningslinjer, f.eks. vedr. behandlingen af personlige data.*

⁷ <https://www.retsinformation.dk/eli/ta/2020/514>

⁸ Bevaringsbestemmelse indeholder en passus om ”indtil Rigsarkivet kan håndtere...” og ”hvis man ønsker at slette, skal Rigsarkivet adviseres mindst 6 mdr. før)

⁹ https://medarbejdere.au.dk/fileadmin/www.medarbejdere.au.dk/Strategi_og_ledelse/samlet_roadmap_for_digitalisering_2018-2020_v5.pdf

¹⁰ <https://ufm.dk/en/publications/2014/the-danish-code-of-conduct-for-research-integrity>

2 Strategiske mål

Nedenstående strategiske mål arbejder mod både at arbejde med forskernes praksis i forhold til data og sikre, at AU leverer den nødvendige assistance.

Alle forskerne på AU:

1. Forholder sig til FAIR-principperne for data såvel som for andre af forskningens output som koder og metoder
2. Integrerer data management i forskningsprocesserne og sikrer derigennem transparens og integritet af forskningens resultater
3. Bidrager til god praksis og klare standarder for håndtering af såvel data som metadata gennem hele forskningens livscyklus, herunder bl.a. dataindsamling, kuratering og opbevaring både under og efter projekters gennemførelse, inklusive valg af licenser og brug af persistente identifikatorer

Aarhus Universitet understøtter disse mål ved at sikre, 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

I forhold til A og B gælder, at AU stiller en basis videns- og teknisk infrastruktur til rådighed. Avanceret brug vil skulle finansieres via projekter.

Det skal tilstræbes, at AU følger de nationale og internationale udviklingstendenser og standardiseringer på Open Science området generelt og for data management specielt.

For at de strategiske mål 1-3 ovenfor opfyldes, kræves såvel etablering af den tekniske infrastruktur som at der arbejdes med information, kompetenceudvikling og en løbende dialog med forskerne om muligheder og perspektiver i at dele og publicere eksempelvis data. Målet er ikke, at alle forskere publicerer deres data og/eller metadata, men at alle har forholdt sig oplyst og kritisk til beslutninger i forbindelse med publicering. Ud over et fokus på, hvad der kan deles med andre, skal alle forskere også se god data management som en integreret del af forskningsprocessen – både af hensyn til transparens i forskningen og for at sikre, at nødvendig information i forbindelse med publicering af eksempelvis data skabes løbende. Endvidere bør interesserede forskergrupper deltage i fagligt orienteret nationale og internationale projekter, som har til formål at definere gode praksisser og identificere relevante standarder for metadata, ontologier, identifikatorer osv. - på tværs af institutioner. Dette spiller direkte sammen med EU's strategi i forhold til Open Science Cloud og ministeriets strategi, som begge har fokus på og arbejder for, at data bliver så FAIR som muligt.

En nødvendig forudsætning for en succesfuld gennemførelse af de strategiske punkter er at Open Science agendaen bliver ledelsesforankret. Ledelsen skal sikre, at der frigøres de nødvendige ressourcer til at sikre en gennemførelse af de strategiske punkter A-D. Det betyder, at der ud over en sikring af den tekniske infrastruktur og etablering af en relevant support infrastruktur, også tages et ledelsesansvar både i forhold til meritering, og i forhold til at bidrage til og understøtte at forskerne arbejder med at definere praksisser og at skabe miljøer, hvor der kan eksperimenteres med forskellige aspekter af Open Science.

Det skal besluttes, hvordan Open Science forum vil sikre en monitorering af, hvordan udrulningen af den nationale strategi måles på AU.

3 Aktiviteter

På fakulteter skal organiseres et eller flere Open Science centre; om de kaldes Open Science eller noget andet er mindre væsentligt, det vigtigste er, at forskerne har et veldefineret sted at henvende sig lokalt parallelt med det centrale tilbud.

AU-IT fortsætter arbejdet med at etablere og udbrede ERDA/SIF i den projektorganisation, som er etableret.

Der nedsættes en data management koordinationsgruppe, som skal sikre vidensdeling om og koordination (hvor relevant) af de decentrale aktiviteter i forhold til punkterne A og B i de strategiske mål:

- 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.)

Koordinationsgruppens formål er derfor at sikre, at der på et operationelt plan etableres fælles løsninger hvor dette er relevant og hensigtsmæssigt og koordinerer tiltag således, at de forskellige centre kan trække på hinandens ekspertise, deles om supportopgaver og genbruge hinandens løsninger.

Koordinationsgruppen vil arbejde sammen med et centralt forankret **Single Point of Contact, SPOC**, som sikrer, at forskere kan henvende sig et sted med spørgsmål vedrørende alle dele af datahåndtering.

Kommissoriet for data management koordinationsgruppen er vist i bilag 3.

4 Omkostninger

Det er klart, at der er en sammenhæng mellem AUs ambitionsniveau og omkostningerne. Nedenfor er givet et bud på tre ambitionsniveauer for såvel den centrale som den decentrale del af AUs indsats på forskningsdata management området – for såvel den centrale del som den fakultetsorienterede del.

Baseret på erfaringerne i forbindelse med BCDs interviewrunde i sommeren 2020, beskæftiger AU i dag 10-15 personer, som helt eller delvist understøtter data management og flere, som aktivt arbejder med forskellige aspekter af Open Science. Flere af disse udfører specialiserede opgaver og vil ikke umiddelbart kunne indgå i en fælles pool af ressourcer. Imidlertid vil alle kunne bidrage til opbygning af en bedre forståelse for, hvad god data management-praksis går ud på og bidrage til, at det etableres. Desuden bør der indledes forhandlinger med KB omkring brug af 6-10 FTE til data management opgaven.

4.1 Central organisation

Lavt: Der ansættes en Open Science koordinator, som har det overordnede ansvar, og en leder af SPOCen, som sikrer at relevante processer implementeres. Der laves en aftale med AUL/KB om omprioritering af deres ressourcer, så de kan være drivende også for den fakultetsmæssige forankring og med Forsknings-IT om at varetage driften af den tekniske infrastruktur. Open Science-koordinatoren er AUs kontakt til nationale og internationale organisationer og arbejder tæt sammen med lederen af SPOCen. SPOCen sikrer, at henvendelser håndteres og fungerer som Front Office i forhold til DeiC. OS koordinatoren sikrer at relevante emner som for eksempel undervisning og informationskampagner kommer på Data management koordinationsgruppens møder. Kommunikation sker med jævne mellemrum via universitets nyhedsbrev og gennem fakulteternes kommunikationsafdelinger. AU-IT understøtter arbejdet med at sikre, at basissystemerne fungerer.

OS-koordinatoren er formand for Data management koordinationsgruppen og deltager i OS-forum og har det overordnede ansvar for, at relevante problematikker sættes på dagsordenen.

Middel: Dette er en udvidelse af det lave niveau, hvor der nu etableres en OS-funktion med 2-3 FTE heraf mindst 1 teknisk orienteret fuldtidsperson plus Open Science-koordinatoren. Kontakten til omverdenen sker via OS-funktionen. Funktionen vil have ressourcer til at arbejde med definition af lokale/DeiC-tjenester samt til minimal deltagelse i eksempelvis arbejde i internationale fora som RDA og vil have ressourcer til begrænset deltagelse i forbindelse med EOSC. Arbejde med eksempelvis undervisning koordineres og der arbejdes mod identifikation og evt. udvikling af relevante undervisningsmoduler. Medlemmer af OS-funktionen deltager i undervisning og generelt i kompetenceudvikling. På dette ambitionsniveau vil det give mening, at AUL/KB spiller en mere aktiv rolle omkring publiceringsmuligheder og i indgåelse af aftaler med datatidsskrifter.

Højt: Dette er en udvidelse af ”middel”, således at der etableres et OS-kontor med 4-6 FTE hvoraf mindst 2 er fuldtidspersoner (en teknisk og en data steward/kommunikativ profil) ud over Open Science-koordinatoren. Man kan forestille sig, at de resterende 1-3 FTE kommer ved at nøglepersoner fra fakulteterne

eller IT midlertidigt knyttes helt eller delvist til OS-kontoret. Kontakten til omverdenen sker via OS-kontoret. OS-kontoret vil også have ressourcer til at arbejde med definition af lokale/DeiC-tjenester samt til aktiv deltagelse i eksempelvis arbejde i internationale fora som RDA og vil have ressourcer til deltagelse i forbindelse med EOSC. Kan deltage aktivt i kompetenceudvikling. Medlemmer af OS-kontoret deltager i undervisning og generelt i kompetenceudvikling. Medlemmer af OS-kontoret kan evt. ”udlånes” til fakulteter i forbindelse med projekter.

4.2 Fakultetsforankring

Fakulteterne opretter en veldefineret indgang til deres Open Science support; det kan være et nyoprettet supportcenter eller det kan være et eksisterende, som allerede varetager en supportrolle over for forskerne. Som allerede nævnt, vil der være behov for at trække på en række forskellige kompetencer. Som ved bemanningen af den centrale funktion opereres med forskellige ambitionsniveauer:

Lavt: Opgaverne i OS-supportcentret håndteres primært af AUL/KB gennem en omprioritering af deres ressourcer – i tæt dialog med de data managers, som allerede i dag er aktive. Ved alle fakulteter er der i dag allerede enten dedikerede data managers eller forskere, som arbejder med data management. Denne/disse vil blive involveret som specialister både i forhold til undervisning og support. En specialist samt en AUL/KB-medarbejder fra hvert fakultet deltager i det koordinerende udvalg; AUL/KB forventes at være drivende. Kommunikation til medarbejderne udarbejdes i høj grad centralt, men med de decentrale kommunikationsafdelinger som afsendere.

I dette scenarie vil AUL/KB fremadrettet være forskerens naturlige sparringspartner i forhold til data management, herunder specielt hvordan man gør data FAIR.

Middel: Fakultetet har hver 2-4 FTE, som arbejder med forskellige aspekter, herunder at identificere manglende tjenester, understøtte relevant publicering, deltage i nationale og internationale projekter etc. Kommunikation til medarbejderne sker decentralt, men med en klar central koordinering. AUL/KB deltager aktivt som generalister og har en vigtig rolle i forbindelse med eksempelvis data management-planer og assistance ved publicering. AUL/KBs deltagelse sker ved omprioritering af deres ressourcer.

Fakultetet vil her optræde som forskerens naturlige sparringspartner i forhold til data management. På visse fakulteter vil denne funktion kunne bygges sammen med rådgivning omkring IT, herunder HPC.

Højt: Fakulteterne har etableret SPOC'en som et Open Science-kontor med repræsentation af alle relevante kompetencer (IT-support, GDPR, sikkerhed, strukturer (ontologier, persistente identifikatorer, etc.) og med ressourcer til at sikre national og international forankring gennem deltagelse i relevante projekter. Kommunikation til medarbejderne sker primært decentralt, men dog med en klar central koordinering.

Fakultetet signalerer hermed at Open Science er vigtigt, ved at omprioritere og samle alle relevante kompetencer, så de nemt kan assistere i alle aspekter af datahåndtering. På visse fakulteter vil Open Science-kontoret kunne bygges sammen med rådgivning omkring computing, herunder HPC, og vejledning om andre aspekter af Open Science som eksempelvis Citizen Science.

5 anbefalinger

Implementeringen vil være en iterativ proces og det er vigtigt løbende at evaluere, om de igangsatte tiltag fungerer og hvilke, der savnes. Arbejdet med data management generelt er tæt koblet til implementeringen af en teknisk platform, som understøtter opbevaring og publicering af data. Et fælles lagersystem forventes at være operationelt Q3 2022.

Det anbefales, at:

- Der etableres som minimum det lave ambitionsniveau for den centrale organisation. Dette vil give mulighed for at begynde en vidensdeling og få etableret et antal kurser og/eller workshops og påbegynde ”train the trainer” workshops
- OS centre ved fakulteterne etableres på middelniveau mhp. at håndtere data management og evt. andre opgaver som HPC. Ved hvert fakultet identificeres en ansvarlig for området, som får afsat tid til at deltage i fælles aktiviteter på tværs af AU
- Der udarbejdes et roadmap med fokus på aktiviteter i 2022-2023
- Der igangsættes relevante informationsaktiviteter – på både universitets- og fakultets/institutniveau

Det anbefales at:

- AU i samarbejde med andre universiteter spiller en aktiv rolle i forhold til infrastrukturaktiviteter i regi af DeiC og internationalt.

Det anbefales videre, at:

- AU etablerer en enhed under AU IT, forsknings-IT, som står for drift og udvikling af lager og tilhørende computerfaciliteter og som i samarbejde med andre universiteter kan spille en aktiv rolle som DeiC-løsningsudbyder i forhold til lager og services forbundet med Data management samt HPC. Enheden vil kunne finansieres 50% via DeiC
- Forskere støttes og opfordres til at deltage i nationale og internationale projekter, som har til formål at etablere best practice for implementering af Open Science
- AU deltager i dialogen omkring meritering og omkring udarbejdelse af en politik for bevaringsværdige data
- Det sikres, at Ph.D. introduceres til god data management praksis og får viden om FAIR data som led i deres uddannelse

6 Bilag 1: Udtræk fra den nationale strategi

Nedenfor er de aktiviteter, hvor AU som forskningsinstitution har en rolle/forpligtelse, trukket ud fra strategi-dokumentet. Der er ikke sket nogen form for redigering af de tekster, som er medtaget.

6.1 Det forventes at AU adresserer nedenstående områder:

indsatsområde	Forventninger til universitet
FAIR data management er en praksisændring inden for mange fagområder	<ul style="list-style-type: none">• Styrkelse af kommunikation og dialog om FAIR data management.• Starte en proces med fokus på at gøre flere outputs i forskningen meritgivende og at sikre en anerkendelse af data management med henblik på FAIR. Dette skal gøres på nationalt niveau, med inddragelse af internationale initiativer. Meriteringen bør tage højde for forskelligheder i forskningsområder.
Understøtte faggrupper i at definere egen implementering af FAIR-principper	Forskningsinstitutionerne skal sikre, at der udarbejdes beskrivelser af god data management-praksis inden for forskellige fagområder, gerne med afsæt i international praksis. Disse skal som minimum indeholde beskrivelse af, hvordan metadata kan indsamles som en integreret del af arbejdsprocessen. Dette kan indebære: <ul style="list-style-type: none">• Udarbejdelse af kortsigtede mål for tilgængeliggørelse og deling af data fx med udgangspunkt i metoder omkring modenhedsmodeller¹¹.• Hjælpe faggrupper med at identificere, udvikle og/eller publicere standarder, metoder, værktøjer og infrastruktur til at gøre deres data FAIR som har udspring i, eller som i hvert fald spiller sammen med internationale løsninger.• Arbejde med praksis for udvalg af data og med metoder til kvalitetssikring af data og metadata.• Udarbejde principper for korrekt (gen)brug af data.
Understøtte forskningens krav til dokumentation gennem hele forskningsprocessen	<ul style="list-style-type: none">• Der skal inden for fagområderne udarbejdes 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.
Etablere vidensinfrastruktur som leverer nødvendig hjælp og vidensressourcer til data management og FAIR	<ul style="list-style-type: none">• Forskningsinstitutionerne støtter opbygning af lokal data stewardship support til understøttelsen af institutionernes forskere.• Relevant efteruddannelse i FAIR, data management og data stewardship skal udvikles og tilbydes til forskere og data stewards på relevante faglige niveauer.• Viden om data management og FAIR-principperne skal stilles til rådighed på relevante trin af forskernes uddannelse, fx. på PhD-skolerne.

¹¹ Fx. CMMI Institutes *Data management Maturity (DMM)*, (<https://cmmiinstitute.com/data-management-maturity>) eller RDAs FAIR Data Maturity Model: <https://www.rd-alliance.org/group/fair-data-maturity-model-wg/outcomes/fair-data-maturity-model-specification-and-guidelines-0>

6.2 Det forventes, at AU kan agere i forhold til en national (DeiC) og international agenda

indsatsområde	Forventninger til universiteterne – og DeiC
Understøtte bevaring af forskningsdata på både kort og lang sigt	<ul style="list-style-type: none"> • Samarbejde, også med internationale aktører, om udvikling af metoder og infrastruktur til langtidsbevaring af forskningsdata, som er tilpasset forskningens behov for diversitet i objekt- og filtyper, såvel som skalering i forhold til datamængder. • Samarbejde om at definere kriterier for datas værdi ift. genanvendelighed og langtidsopbevaring.
Understøtte forskningens krav til dokumentation gennem hele forskningsprocessen	<ul style="list-style-type: none"> • Der bør på nationalt plan stimuleres og faciliteres en diskussion om håndtering af risikoen ved FAIR data. Risikofaktorer er eksempelvis fortrolighed, ophavsret, anden immaterialretlig regulering, dual use, GDPR samt nationale sikkerhedsinteresser. • Institutionelle retningslinjer for brugen af datalicenser skal udarbejdes og udbredes, eventuelt i nationalt samarbejde, primært baseret på internationale standarder.
Etablere teknisk infrastruktur, der understøtter data management og tilgængeliggørelse af FAIR data (DeiC og universiteterne):	<p>Specificer og byg danske dele af den tekniske infrastruktur:</p> <ul style="list-style-type: none"> • Etabler et generelt tilgængeligt nationalt trusted data repository hvor forskere kan gemme og eksponere metadata og data samt tilknytte persistente identifikatorer (PID) og licens. • Specificer og tilbyd komponenter nationalt som efterspørges af forskerne, fx. PID-systemer, ontologi-database og andet der kan lette arbejdet med at gøre data FAIR. • Udbyg den nationale tjeneste for data management-planer med fagtilpassede skabeloner samt tilpasninger der kan gøre data management-planer anvendelige i forbindelse med anmeldelse til Rigsarkivet. • Sørg for at anbefalede tjenester og infrastrukturer bliver certificeret efter internationale standarder i nødvendigt omfang, og at governance og politikker sikrer korrekt administration af adgangen til data, også efter at forskere har forladt institutionen.
Etablere vidensinfrastruktur som leverer nødvendig hjælp og vidensressourcer til data management og FAIR	<ul style="list-style-type: none"> • Der etableres koordinering og samarbejde på tværs af forskningsinstitutionerne om tilvejebringelsen af data stewardship service, hvor dette er gavnligt ud fra eksempelvis et ressourcemæssigt hensyn, fx. til større tværgående projekter. Større nationale satsninger koordineres under DeiC.

7 Bilag 2: Den nationale følgegruppens fokusområder.

Følgegruppen skal udarbejde en handlingsplan, der adresserer indsatser, der kræver en særlig national koordinering, og skal særligt adressere følgende forslag til konkrete indsatser ved at koordinere arbejdet med:

1. at få faggrupper på tværs af forskningsinstitutionerne til at definere egen implementering af FAIR-principper.
2. at udvikle metoder og infrastruktur til langtidsbevaring af forskningsdata, som er tilpasset forskningens behov for diversitet i objekt- og filtyper, såvel som skalering i forhold til datamængder.
3. at definere kriterier for datas værdi ift. genanvendelighed og langtidsopbevaring.
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.
5. at udarbejde og udbrede institutionelle retningslinjer for brugen af datalicenser primært baseret på internationale standarder.
6. at udbygge den nationale tjeneste for data management-planer med fagtilpassede skabeloner samt tilpasninger der kan gøre data management-planer anvendelige i forbindelse med anmeldelse til Rigsarkivet.
7. at anbefalede tjenester og infrastrukturer bliver certificerede efter internationale standarder i nødvendigt omfang, og at governance og politikker sikrer korrekt administration af adgangen til data, også efter at forskere har forladt institutionen.
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.
9. udvikling af relevant efteruddannelse i FAIR, data management og data stewardship, der skal tilbydes til forskere og data stewards på relevante faglige niveauer.
10. at viden om data management og FAIR-principperne stilles til rådighed på relevante trin af forskernes uddannelse, f.eks. på PhD-skolerne.
11. at der etableres en liste over udgifter til FAIR data management-relaterede udgifter (software, lagerplads, årsværk), som løbende opdateres. Listen er vejledende, og kan anvendes af fonde og universiteter, som udgangspunkt for udarbejdelse og vurdering af budgetter. Listen udarbejdes af et udvalg med repræsentanter fra Danske Universiteter og forskningsfondene.
12. at der på tværs af institutionerne samarbejdes om sikkerhed med henblik på at sikre danske forskningsdata mod uønsket deling med 3. part, som ikke selv lever op til de grundlæggende værdier for åbne data.
13. at problemstillinger vedr. GDPR-regler beskrives med det formål at drøfte, hvorledes deling af forskningsdata på tværs af institutioner og landegrænser bliver så smidig og agil som muligt. Arbejdet kan afvente Datatilsynets offentliggørelse af nye vejledninger med fortolkning af GDPR-reglerne inden for forskningsområdet.

8 Bilag 3: Kommissorium for Data management koordinationsgruppen

Kommissorium for Data management koordinationsgruppen Version 2.1

1. Formål

Open Science (OS)-Forum ønsker at styrke samarbejdet mellem de fakultets- og institutforankrede data management initiativer, som skal sikre, at forskere ved AU kan opfylde nationale og internationale krav/forventninger i forhold til god data management praksis, herunder at publicere **F**indable, **A**ccessible, **I**nteroperable and **R**eusable (FAIR) data.

Data management koordinationsgruppens formål er at sikre vidensdeling om og koordination (hvor relevant) af de decentrale aktiviteter i forhold til punkterne A og B i de af ledelsen vedtagne strategiske mål:

- 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.)

Koordinationsgruppens formål er derfor at sikre, at der på et operationelt plan etableres fælles løsninger, hvor dette er relevant og hensigtsmæssigt og koordinerer tiltag således, at de forskellige centre kan trække på hinandens ekspertise, deles om supportopgaver og genbruge hinandens løsninger.

Det er en forudsætning, at medlemmerne af koordinationsgruppen er orienterede om tiltag på de øvrige fakulteter og sikrer, at medarbejderne i de decentrale supportenheder har et overblik over, hvor forskellige typer ekspertise er forankret, således at de kan sikre, at forskerhenvendelser vedrørende håndtering af data bliver adresseret på rette sted i forhold til spørgsmål om aftaler, rettigheder, proces, analyse, publicering og arkivering, samt at sikre, at data og de tilhørende metadata gøres FAIR (FAIRifisering).

2. Opgaver

Data management koordinationsgruppen skal sikre koordination og fælles tilgang til en række konkrete initiativer, herunder:

- brug af ERDA og SIF
- fonde og EU's krav i forhold til FAIR

- Data Management Plan (DMP)-skabeloner og best practice
- vejledning om værktøjer til løbende dokumentation af forskningsprojekter (f.eks. e-labbøger)
- alle elementer i FAIR, herunder Persistent Identifiers, maskinlæsbare metadata og licenser
- publiceringsplatforme – fordele og ulemper
- Informationskampagner (gerne en besøgsrunde til institutter omkring FAIR og ERDA/SIF)
- Undervisningstilbud

Det er ikke et mål, at alle skal kende til alt, men at man kender begreber og ved, hvor man kan få hjælp.

Data management koordinationsgruppen skal arbejde med tiltag, som skal sikre, at hele supportinfrastrukturen kan fungere som et samlet hele. Dette betyder bl.a.:

- Fælles løsning på styring af henvendelser, enten ved fælles system eller ved brug af veldokumenterede metoder til at udveksle informationer (Application Programming Interfaces (APIer))
- Koordination/fælles metoder til at dokumentere såvel indsamling som den videre bearbejdning af data – hvor relevant
- Koordination af/vidensdeling om, hvordan data og metadata publiceres (herunder også brug af licenser og standarder for data såvel som metadata)

Data management koordinationsgruppen kan også komme med input i forhold til relevante nye initiativer.

3. Sammensætning

Data management koordinationsgruppen har følgende medlemmer:

- Open Science koordinator
- En teknisk/support person fra hvert fakultet
- Medarbejder fra Det Kongelige Bibliotek
- Systemansvarlig eller projektleder for ERDA/SIF

Systemkoordination i forhold til ERDA/SIF samt forbindelsen til DeiC er forankret i den centrale Open Science support funktion og er derfor ikke nævnt eksplicit.

Hertil kommer, at der løbende kan indkaldes andre kompetencer så som:

- jura (GDPR)
- IT-sikkerhed
- FAIR

4. Arbejdsform

Data management koordinationsgruppen nedsættes for en periode på 3 år. Efter 3 år evalueres behovet for en fortsættelse af koordinationsgruppen samt evt. justering af kommissoriet.

Mødefrekvens vil afhænge af udestående opgaver. Det forventes, at der vil være månedlige møder i efteråret 2022 i forbindelse med udrulningen af ERDA/SIF og i forbindelse med informationskampagner omkring såvel FAIR som mulighederne i ERDA/SIF.

OS-koordinatoren er ansvarlig for, at der årligt udarbejdes en rapport til OS Forum.

5. Sekretariatsbetjening

Data management koordinationsgruppen sekretariatsbetjenes af den centrale Open Science Support funktion, som er organisatorisk forankret i AU Forskning.