

Dato: 2020-03-09

Slutafrapportering på DM Forum aktivitet

Aktivitet	Making AnaEE DK data FAIR	Aktivetsperiode	1/6/19 – 31/12/19
Bevillingshaver	Københavns Universitet, Svend Christensen (med projektleder Katrine Vendelboe)	Samlet bevilling	287.800 DKK
Resume af aktiviteten	<p><u>Aktivitetens formål</u></p> <p>Det primære formål med aktiviteten har været at facilitere en læringsproces for alle parter involveret:</p> <ol style="list-style-type: none"> 1. De deltagende databehandlere/forskere (herfra "databehandlere") fra hhv. KU, AAU og DTU havde brug for en forståelse af hvordan man i praksis implementerer FAIR data på deres universiteter, for at komme skridtet videre fra at diskutere de overordnede formål med FAIR. Dette behov udsprang delvist af et krav om implementering af FAIR data fra bevillingsgiver til AnaEE Denmark. 2. DeIC havde behov for at få en dybere forståelse af de strukturer der arbejder for og imod implementering af FAIR data i hhv. større internationale infrastrukturer samt på de danske universiteter. Formålet på længere sigt er bedre at kunne supportere universiteterne og disses medarbejdere i implementeringen af FAIR data på alle niveauer. <p><u>Proces</u></p> <p>Inden projektstart blev der afholdt to møder mellem DeIC og KU, for at afdække behov og definere projektets indhold og formål.</p> <p>Venligst se process overblik sidst i rapporten.</p>		
Leverancer	<p>AnaEE DK Landing page: https://erda.ku.dk/vgrid/AnaEE_DK/</p> <p>DOI til de 3 datasæt, som indgik i projektet. Landing page vil fortsat fungere som landing page for kommende datasæt fra AnaEE DK partnere og er linket til via AnaEE DK hjemmesiden: https://anaee.dk/access/.</p>		
Evt. andet output	<p>Projektet er blevet formidlet på AnaEE workshop i November med deltagere fra alle europæiske AnaEE lande med det formål at vise et eksempel på hvordan implementeringen af FAIR data nationalt kan gribes an.</p>		
Læring	<p><u>Forståelse af data</u></p> <p>Det viste sig at være svært for databehandlerne at sætte sig ind i de nye terminologier og koble eksempelvis metadata terminologi med den meget specifikke forskningsterminologi. Gruppen brugte meget tid på at forklare og forstå hinandens specialist områder – særligt imellem DeIC konsulenter og databehandlerne. Fremadrettet vil gruppen anbefale at tænke i at uddanne specialister, som på en eller anden måde kan specialisere sig i mere specifikke (om end evt. brede) forskningsområder og kommunikere direkte med de databehandlere, der skal FAIRificere deres data.</p> <p><u>Forståelse af FAIR</u></p> <p>Det krævede en del diskussion i gruppen at komme til enighed om forståelsen af FAIR. Denne del af processen har et langt stykke af vejen været en succes. Til gengæld er det gået op for gruppen i processen af universiteterne ikke nødvendigvis har opnået samme forståelse af FAIR samt vigtigheden af samme.</p> <p>Der har været mange tekniske forhindringer i udmøntning af DOI undervejs i processen samt opdagelsen af de begrænsninger, der følger efter i den videre udvikling af de publicerede datasæt. KU som tilbyder brugen af ERDA til dets ansatte og deres samarbejdspartnere har indsat flere</p>		

sikkerhedsforanstaltninger grundet hensyn til GDPR og videnskabsetik, som desværre begrænser FAIRheden af de datasæt, der publiceres. Under mødet med repræsentanter fra ERDA gik det op for projektgruppen at begrænsningerne ikke alle er grundet i manglende udvikling af systemet, men i højere grad grundet disse sikkerhedsforanstaltninger som indtil videre har haft højere prioritet på universitetet.

Processen

Der var lavet en *forventningsafstemning* mellem DeiC og projektleder forud for projektstart. Denne har dog ikke formået at forventningsafstemme helt nok i forhold til hvad udgangspunktet for deltagerne var med hensyn til forhåndsviden om data, metadata, FAIR osv. Det er nok noget man skal være opmærksomme på fremadrettet at få afklaret.

Det viste sig at være sundt for projektets fremdrift at have *Zoom møder hver anden uge*, selvom det ind imellem kunne virke som meget. I det individuelle arbejde med metadata og søgen efter standarder imellem møderne var der meget ofte spørgsmål, der kunne blive besvaret ret hurtigt af vores DeiC konsulenter fordi vi mødtes ofte. Aktivitetsniveauet i september-november, midt i projektet, var meget højt og alle deltog med fuldt engagement i næsten alle møder og arbejdede effektivt i mindre grupper imellem møderne.

Det var yderligere en stor hjælp for deltagerne at vi ændrede formatet for vores sidste møde fra at være opsamling og evaluering til at invitere ERDA repræsentanter til at besvare en masse af de spørgsmål, der var opstået undervejs. Det fungerede som en god afrunder på projektet og gav anledning til tanker om videre samarbejde.

Fremtiden

AnaEE DK skal arbejde videre med at FAIRificere data, både nationalt og i forbindelse med AnaEE internationalt. Erfaringerne med dette projekt er blevet delt med internationale partnere, som skal i gang med lignende processer og AnaEE DK kan derfor fungere som foregangsland for andre.

Der er blevet skabt et solidt link imellem AnaEE DK deltagerne og DeiC og blevet lavet uformelle aftaler om videre samarbejder, hvor man evt. kan komme de næste skridt i processen.

Databehandlerne i projektet kan for sin vis anses som udlært i de basale metoder til at FAIRificere data og har mulighed for at hjælpe til med at videndele i egne organisationer.

Opmærksomhedspunkter

- Universiteterne har brug for støtte til forståelsen af FAIR og de krav der stilles til infrastruktur.
- Der er ufordringer i forhold til at finde en balance mellem videnskabsetik, GDPR, generel sikkerhed og FAIR, som alle parter bør være opmærksomme på.
- Det kan anbefales at være åben og realistisk i forhold til at formulere at det, at gøre data FAIR inden for et specifikt forskningsområde kræver en del opstartsressourcer i form af tid til diskussioner, men når først de basale aftaler og standarder er på plads, er det ikke nødvendigvis en fortsat stor opgave.
- Det er svært at lave en kulturændring hos forskerne før infrastrukturen er på plads til at facilitere ændringerne.

Deltagere

Databehandlere/forskere

Aarhus Universitet: Johannes (Jeroen) Pullens

Danmarks Tekniske Universitet: Teis Nørgaard Mikkelsen

Københavns Universitet: Klaus Steenberg Larsen (IGN) og Simon Fiil Svane (PLEN)

Konsulenter

DeiC: Rene Belsø, Anders Sparre Conrad, Diba Terese Markus

Rigsarkivet: Anne Sofie Fink Kjeldgaard

Projektleder

Københavns Universitet: Katrine Vendelboe

Opgave/møde/fase	Dato/deadline	Indhold/beskrivelse
1. workshop	01-07-2019	Heldagsworkshop for DeiC og alle projektdeltagere: Introduktion til FAIR konceptet Præsentation af faciliteter og potentielle datasæt til projektet Beslutning om projekt tidsplan og deadlines Beslutning om budget Beslutning om hvilke datasæt, der skulle bruges i projektet
Microsoft Teams	03-07-2019	Værktøj til intern kommunikation i projektgruppen
ICOS, FAO, WMO og andre metadata standarder tilgængelige for alle projektdeltagere	01-08-2019	Det var besluttet at gruppen skulle bruge allerede definerede standarder for metadata i så vid udstrækning som muligt.
IPR sikret	15-08-2019	AU og KU havde behov for at afklare IPR med universiteternes jurister for at sikre de kunne dele de nødvendige data.
Universiteternes infrastrukturer afklares	15-08-2019	Undersøg muligheder for at udmønte DOI gennem universiteternes egne datarepositories samt hvilke muligheder universiteterne ellers kunne tilbyde. Det viste sig AU ikke kunne tilbyde DOI hvorfor AUs deltager i projektperioden benyttede sig af KUs ERDA. KUs system ERDA havde andre begrænsninger i systemet, som gjorde muligheder for reel FAIRificering svær.
Skype møde 1	06-09-2019	Møde i projektgruppen: <ul style="list-style-type: none"> • Identificering af variabler for de enkelte platforme • Identificering af relevante metadata standarder
Skype møde 2	20-09-2019	Møde i projektgruppen: <ul style="list-style-type: none"> • Beslutning om variabler • Beslutning om standarder
Metadata skabeloner udfyldt	01-10-2019	Første udkast til metadata for de enkelte datasæt udfyldt
Skype møde 3	07-10-2019	Møde i projektgruppen: <ul style="list-style-type: none"> • Videndeling om diverse opståede problemer og forståelses problematikker i beskrivelsen af metadata.
Skype møde 4	21-10-2019	Møde i projektgruppen:

		<ul style="list-style-type: none"> • Fortsat problemløsning i forhold til metadata • Videndeling i forhold til brug af ERDA og overvejelser om brug af andre systemer
Quality Assurance	01-11-2019	Beskrivelse af quality assurance for hver platform på plads.
Skype møde 5	05-11-2019	<p>Møde I projektgruppen:</p> <ul style="list-style-type: none"> • Fortsat diskussion om metadata beskrivelser af problemer med brugen af ERDA • Enighed om at invitere ERDA udviklere fra KU til final meeting.
Final workshop	12-11-2019	<p>Sidste møde I projektgruppen:</p> <ul style="list-style-type: none"> • Udover projektgruppen deltog 3 repræsentanter fra KUs fællesadministration, som på forskellige niveauer arbejder med udvikling af ERDA. Projektgruppen kunne give feedback og stille spørgsmål direkte til ERDA repræsentanter, som kunne komme med løsningsforslag og fortælle om de videre planer for udvikling af systemet. • Vi forsøgte at udmønte en DOI for et endeligt datasæt, men havde tekniske problemer, som blev løst et par dage senere.
Data for alle platforme publiceret med DOI og landing page	10-01-2020	<p>Landing page: https://erda.ku.dk/vgrid/AnaEE_DK/</p> <p>Forsinkelsen skyldtes at KU PLENs datasæt gerne skal opdateres årligt og det gav mest mening at publicere et endeligt datasæt ved årets afslutning for at undgå dobbeltarbejde.</p>