

Dato: 2020-01-30

## Slutrapportering på DM Forum aktivitet

Aktivitet	<b>FAIRifyHUM - FAIRify humanistiske forskningsdata</b>	Aktivitetsperiode	1-3-2019- 31-12-2019
Bevillingshaver	Lene Offersgaard, KU	Samlet bevilling	Kr 342.400
Resume af aktiviteten	<p>Aktiviteten tog udgangspunkt i anbefalinger fra GoFAIR-initiativet<sup>1</sup> og har fokuseret sig på humanistisk forskning.</p> <p>Aktiviteten har inddraget grupper på AAU, AU/KB og KU der har arbejdet med at gøre nogle konkrete forskningsdata (mere) FAIR. Dette er sket ved at give information og support vedr. FAIR.</p> <p>Arbejdet har været relateret til DIGHUMLAB og de to ESFRI'er CLARIN og DARIAH.</p> <p>Målet var gennem samarbejde med konkrete forskergrupper at afdække hvordan de gør deres forskningsdata mere FAIR og at tilbyde konkret hjælp og vejledning til hvordan der tildes PID'er<sup>2</sup> til data, hvordan data metadateres, tildes licenser, arkiveres i datarepositorier mv. Hensigten med aktiviteten var at give en bedre baggrund for at DK kan blive en tydeligere GoFAIR-aktør fra 2020.</p> <p>Værdien er opstået på de medvirkende institutioner og forskergrupperne beskriver i en kort case hvordan de håndterer FAIR i den virkelige verden, herunder de udfordringer de har mødt.</p>		
Leverancer	<p><b>Leverance 1: Liste af udvalgte forskergrupper</b></p> <p>Der har været kontakt til en række forskergrupper i DK med nuværende relationer til DARIAH og CLARIN, og grupper uden relationer. På den baggrund er der udvalgt et antal grupper, der har arbejdet konkret med at FAIRificere deres data. Vedrørende de inddragede forskergrupper er der af rapporteret på DM Forum med overblik, samt skriftlig short-liste. <a href="#">URL</a></p> <p><b>Leverance 2: Casebeskrivelser med opsummering</b></p> <p>Case-beskrivelse fra hver gruppe om hvad FAIR-udfordringerne er og om de konkrete resultater, der er opnået i et resume fra hver forskergruppe. Der er også en opsummering af hvilke udfordringer der viste sig for forskerne. Opsummeringen giver et indtryk af hvilke udfordringer, der er i arbejdet med at gøre konkrete forskningsdata mere FAIR. <a href="#">URL</a></p>		
Evt. andet output	<p>Der er nu forskningsdata tilgængelige på følgende PID'er:</p> <p><a href="http://hdl.handle.net/20.500.12115/41">http://hdl.handle.net/20.500.12115/41</a></p> <p><a href="http://hdl.handle.net/20.500.12115/40">http://hdl.handle.net/20.500.12115/40</a></p>		
Læring	<p>Der opsummeres følgende fra arbejdet:</p> <p>Der er i en række tilfælde etiske udfordringer ved at gøre humanistiske forskningsdata FAIR, nogle data vil pga. GDPR og deres følsomhed ikke kunne deles. I en del situationer vil end ikke metadata kunne deles, da de også kan indeholde sensitiv information.</p> <p>I de seks af de syv cases har vilkår for deling af data været en udfordring. Der er brug for vejledning om håndtering af copyright, forståelse af gældende kontrakter og licensforhold for data. Der kan også være brug for midler til at frikøbe copyright-belagte data.</p>		

<sup>1</sup> <https://www.go-fair.org/countries/guidelines-involvement-go-fair-initiative-country-level/>

<sup>2</sup> PID: Persistent Identifier, fx DOI.

I fremtiden bør man være mere opmærksom på at afklare rettigheder til data og muligheder for at dele data allerede ved projekternes start.

Der efterspørges guidelines og beskrivelser af best-practice vedrørende deling af tekstdata, men det vil også kunne være relevant for andre typer af data. Det er også påpeget at der er brug for vejledning om forståelse af licensforhold og tildeling af licenser til datasæt.

Det generelle princip om at data *gøres findable* med at de tilknyttes en PID<sup>3</sup> har kun i begrænset omfang været muligt inden for projektets løbetid. I to cases er data blevet gjort FAIR ved at dele dem via CLARIN's dataarkiv i Danmark - clarin.dk. I tre tilfælde afventer deling af data stadig forhandling med rettighedshaverne. I et tilfælde ligger data åbent tilgængelige på url'er men disse er pt ikke omsat til PID'er.

At gøre et datasæt *findable* og *accessible* bl.a. med en PID og høstbare metadata, vil ikke altid være den bedste måde for forskerne at få adgang til forskningsmaterialet. Søgemuligheder i datasættene vil i en del tilfælde være et godt redskab for forskerne for at få et indtryk af materialet, og dette aspekt er det også vigtigt at bemærke i alle bestræbelserne på at FAIRificere forskningsdata.

Metadatering og dokumentation af data tager tid, og det er nødvendigt at anerkende at opgaven kræver resurser.

Endnu et standpunkt er, at hvis man ikke kan dele sine forskningsdatasæt, der er baseret på udvalgte dele af åbne datasæt, så tilskyndes forskere til hver især at opbygge deres egne uddrag af tilgængelige data og dermed udfordres sammenligneligheden af resultater og på den måde kan troværdigheden af forskningsresultaterne måske anfægtes.

Det er derfor vigtigt at det i fremtiden bliver lettere at indgå aftaler med datasættens rettighedshavere om tilladelser til videredistribution af forskningsdatasæt.

#### Deltagere

Mette Kia Krabbe, AU  
Marianne Ping Huang, AU  
Birte Christensen-Dalsgaard, AU  
  
Anette Lykke Hindhede, AAU  
Andreas Lindenskov Tamborg, AAU  
Anders Kristian Munk, AAU  
Anders Koed Madsen, AAU  
Benjamin Brink Allsopp, AAU  
Bent Egberg Mikkelsen, AAU  
Thomas Ryberg, AAU  
Karsten Kryger Hansen, AAU  
  
Tanya Karoli, KU  
Krista Stinne Greve Rasmussen, KU  
Frans van der Sluis, KU  
Haakon Lund, KU  
Dorthe Dunker, KU  
Dorte Haltrup Hansen, KU  
Lene Offersgaard, KU

<sup>3</sup> PID: Persistent Identifier, fx DOI.