

Afdækning A

Bevillingsgiveres krav

Introduktion

Denne afdækning skal besvare spørgsmålet: Hvilke krav stiller bevillingsgivere (EU, forskningsråd mv.) pt. til research data management? Hvad er det, forskningsprojekterne skal leve op til?

Afdækningen har form af en oversigt over krav fra væsentlige indenlandske og udenlandske bevillingsgivere med relevans for dansk forskning. Listen over bevillingsgivere er ikke fuldstændig.

Afdækningen skal informere Styregruppens arbejde med planlægnings-dimensionen "politikker for data management".

EU Horizon 2020

Som følge af H2020 skal hver bevillingshaver sikre OA til alle publikationer relateret til projektets forskningsresultater senest seks måneder efter offentliggørelsen (12 måneder for artikler i samfundsvidenskab og humaniora). Publikationerne skal som minimum gratis kunne læses online, downloades og printes, og så vidt muligt også kopieres, distribueres, søges, linkes til i fuldtekst, crawles mhp. indeksering og data mines. Kommisionen accepterer både selvarkivering (grøn OA), hvor den frie adgang til publikationen kan udskydes (embargoperiode/eksklusivperiode), og OA publisering, hvor der er umiddelbar adgang til publikationen.

Som en nyskabelse gennemfører Kommisionen indenfor rammerne af H2020 endvidere en pilot på OA til offentligt finansierede forskningsdata, **Open Research Data Pilot**¹.

Forskere i projekter, der deltager i piloten bliver bedt om at stille de underliggende data, der er nødvendige for at validere resultaterne, der ligger til grund for deres publikationer, og andre videnskabelige oplysninger til rådighed for andre forskere, industrier og borgere. OA til forskningsdata fortolkes som retten til at tilgå og genbruge data ihht. de konkrete betingelser fastsat i den pågældende bevillingsaftale².

For 2014-2015 vil emneområder, der deltager i piloten få tilskud på 3 mia €.

Emneområder omfattet af piloten

- Fremtidige og fremspirende teknologier
- Forskningsinfrastruktur - specifikt e-infrastrukturer

¹ http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf og http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

² Model Grant Agreement, Version 1.0 11 December 2013, artikel 29.3
http://ec.europa.eu/research/participants/data/ref/h2020/mga/gga/h2020-mga-gga-multi_en.pdf

- Lederskab inden for støtte-og industrielle teknologier - informations-og kommunikationsteknologier
- Samfundsmaessig udfordring: Sikker, ren og effektiv energi - specifikt intelligente byer og lokalsamfund
- Samfundsmaessig udfordring: Klima, miljø, ressourceeffektivitet og råvarer, eksklusive råvarer
- Samfundsmaessig udfordring: Europa i en foranderlig verden - inkluderende, innovative og reflekterende samfund
- Videnskab med og for samfundet

Pilotens krav

1. Detaljerede data management planer (DMP), der dækker individuelle datasæt, skal i en første version leveres som en tidlig, særskilt leverance indenfor de første seks måneder af projektets levetid. DMPer skal overholde den af kommissionen anviste skabelon³.
2. Data fra projekter omfattet af piloten skal deponeres i et repositorie. Kommissionen peger ikke på en specifik infrastruktur, men nævner specielt Open Access Infrastructure for Research in Europe (OpenAIRE)⁴.
3. 3. part skal så vidt muligt have mulighed for at tilgå, data mine, udnytte og formidle de deponerede data, dvs data skal følges af en licens, som muliggør genanvendelse. Kommissionen nævner her specifikt Creative Commons licensering som en mulighed.
4. Data skal suppleres med information om værktøjer og instrumenter, der er brugt i projektet og som er nødvendige for at validere resultaterne, fx software, kode, algoritmer, analyseprotokoller mv. Så vidt muligt skal disse værktøjer stilles til rådighed også.

Det anbefales, at supplere med et persistent, ikke-proprietært, åbent og interoperabelt ID for forskere og data, således at forskere er unikt identificerbare og data unikt henførbar. Kommissionen nævner her ORCID og DataCite som bæredygtige muligheder.

Type af data omfattet af piloten

1. Data, inklusiv tilhørende metadata, der er nødvendige for at validere de resultater, der ligger til grund for en publikation. Disse data skal tilgængeliggøres så hurtigt som muligt.
2. Anden data, inklusiv tilhørende metadata, beskrevet i DMP. Tilgængeliggøres på det tidspunkt, der er anført i DMP.

Opt in

Forskere indenfor andre områder kan vælge frivilligt at deltage i piloten. Sådanne forskningsprojekter monitoreres og modtager hjælp på lige fod med projekter omfattet af piloten.

Opt out

Et forskningsprojekt kan på et vilkårligt tidspunkt helt eller delvist fravælge piloten af forskellige grunde, fx:

- Hensyn til kommercial eller industriel udnyttelse (IPR).
- Behov for konfidentialitet af sikkerhedsmæssige årsager.
- Krav om overholdelse af persondatabeskyttelsesregler.
- Hvis projektets formål undermineres af at gøre data frit tilgængeligt.
- Hvis projektet ikke genererer/indsamler forskningsdata.
- Andre legitime grunde til ikke at deltage i piloten.

Grunde til opt out skal være beskrevet i DMP.

³ Annex 1: Data Management Plan (DMP) template
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

⁴ <https://www.openaire.eu/>

Data Management Plan krav ifm. piloten

- DMP skal som minimum opdateres midtvejs i projektperioden og ved projektslut, men derudover ved tilføjelse af nye datasæt osv. andre væsentlige ændringer i projektet⁵.
- DMP skal indeholde:
 - Datasæt reference og navn (ID for det datasæt der skal produceres)
 - Beskrivelse af datasæt, herunder dets oprindelse, natur og skala, og hvem det vil have værdi for, reference til publikation, information om evt. sammenlignelig data og mulighederne for data integration og genbrug.
 - Eventuelle standarder og metadata.
 - Beskrivelse af data deling, herunder procedurer for adgang, tekniske mekanismer for formidling og nødvendig software for genbrug, typen af adgang (open eller begrænset til definerede grupper), ID på repositoriet og type (institutionelt/disciplinspecifikt mv.)
 - Begrundelse i tilfælde af, at data ikke kan deles.
 - Beskrivelse af arkiverings- og langtidsbevaringsprocedurer (inklusiv storage og back up), herunder hvor lang tid data skal bevares, estimerede volumen og forventede omkostninger forbundet hermed, og hvordan disse tænkes dækket.

Godtgørelse og støtte

Ekstraomkostninger som følge af deltagelse i piloten godtgøres gennem tilsvarende forhøjet bevilling. Deltagende projekter vil have adgang til specifik teknisk hjælp bl.a. realiseret via projekter under the Research Infrastructures Work Programme 2014-15 (call 3 e-Infrastructures).

Krav til ansøgere udenfor piloten

Projekter under piloten, skal overholde ovennævnte krav, men for alle projektansøgninger under "Research and Innovation actions" og "Innovation actions" gælder det, at de indeholder en sektion om data management, og hvor det er relevant, skal ansøgere give en kort, generel beskrivelse af deres data management politik; hvilke typer af data projektet genererer/indsamler, hvilke data standarder der vil blive brugt, og hvordan data udnyttes og/eller deles/gøres tilgængelig for verifikation og genbrug (alternativt en forklaring, hvis dette ikke kan lade sig gøre), samt hvordan data kurateres og bevares.

Danske, offentlige forskningsråd og fonde

For nuværende består forskningsrådsystemet af følgende råd og fonde:

- Det Frie Forskningsråd (DFF)
- Det Strategiske Forskningsråd (DSF)
- Rådet for Teknologi og Innovation
- Højteknologifonden
- Danmarks Grundforskningsfond

Regeringen indgik oktober 2013 en bred politisk aftale med alle Folketingets partier om at de konkurrenceudsatte midler til strategisk forskning, teknologiudvikling og innovation i DSF, Højteknologifonden og Rådet for Teknologi og Innovation samles i en ny fond – Danmarks Innovationsfond.

Det er forventningen, at den nye fond er etableret senest pr. 1. april 2014. Indtil lovens ikrafttræden fortsætter råddene og fonden deres aktiviteter⁶.

I det følgende gennemgåes kun krav vedr. behandling, opbevaring og formidling af forskningsdata stillet af hhv. DFF og DSF, sådan som de er redegjort for af Dansk Data Arkiv (DDA) marts 2014.

I opslag fra DFF efterår 2013 og forår 2014 formuleres krav til forskningsdatamanagement som følger:

⁵ Annex 2: Additional guidance for Data Management Plans

http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

⁶ <http://fivu.dk/forskning-og-innovation/rad-og-udvalg/det-strategiske-forskningsrad/dsf-filer/orientering-om-overgangsperioden-frem-til-estableringen-af-danmarks-innovationsfond.pdf>

"1.4 Tilgængelighed af forskningsdata og publikationer

Dansk Data Arkiv

Hvis dit projekt omfatter indsamling/køb af kvantitative eller kvalitative data inden for forskningsområderne samfundsvidenskab, sundhedsvidenskab og historie, skal du i tilfælde af bevilling aflevere det indsamlede materiale i dokumenteret stand til Dansk Data Arkiv. Arkivet indhenter, bevarer og gør forskningsdata tilgængelige for andre brugere. Du kan læse mere på arkivets hjemmeside: <http://www.sa.dk/dda/>.⁷

I brevet til ansøgere til DFF Sundhed og Sygdom, der informerer om, at ansøgeren vil modtage bevilling fra rådet gentages kravet (citat, p. 1):

"Endvidere skal relevant datamateriale afleveres til Dansk Data Arkiv (DDA) i dokumenteret stand. Oplysninger om Dansk Data Arkiv findes på www.sa.dk/dda/."

I brevet til ansøgere til DFF Samfund og Erhverv, der informerer om, at ansøgeren vil modtage bevilling fra rådet gentages kravet også (citat, p. 3):

"Det påhviler bevillingshaver at sørge for, at det indsamlede datamateriale i dokumenteret stand afleveres til Dansk Data Arkiv (DDA), DDA Guide vedlægges. Forskningsrådet har bemyndiget DDA til at fastsætte en frist for afleveringen af datamaterialet og har orienteret DDA om din bevilling."

Der en aftale om arkivering af data indsamlet med støtte fra DFF Kultur og Kommunikation. Men i praksis er der aldrig arkiveret data med begrundelse i krav fra DFF Kultur og Kommunikation.

I Rapportskema, der udfyldes til aflæggelse af faglig midtvejs- og slutrapport for bevillinger findes under punkt '7. Publikationer afført af det finansierede projekt' felt til afkrydsning af, om data er afleveret til arkivering i DDA.

Aftaler omkring arkivering af data har været drøftet med en komitéen for DSF Individ, Sygdom og Samfund, men er ikke truffet endelig aftale.

DDA (primært DDA Sundhed) rådgiver forskere om procedure for håndtering af personhenførbar data i forskningsprocessen ud fra de definitioner og procedure der er beskrevet af Datatilsynet:

"Forskningsdata der indeholder personoplysninger skal slettes /destrueres eller anonymiseres, når de ikke længere skal benyttes til det angivne formål. Der er imidlertid en væsentlig modifikation til kravet om, at senere behandling ikke må være i strid med det oprindelige formål, idet det tilføjes, at senere behandling af oplysninger, der alene sker i historisk, statistik eller videnskabeligt øjemed, ikke anses for uforenelige med de formål, hvortil oplysningerne er indsamlet (jf. § 5, stk. 2). Endelig er der yderligere en bestemmelse, som muliggør en fortsat opbevaring af oplysninger efter endt brug, idet oplysningerne kan overføres til arkiv efter reglerne i arkivloven (jf. § 14). Hvis oplysningerne overføres til arkivering i Statens Arkiver, hvilket i praksis vil sige Dansk Data Arkiv, anses lovens krav om sletning som opfyldt."⁸

Loven rummer også mulighed for at forskningsdata med personhenførbar oplysninger, der er arkiveret i SA/DDA kan udleveres til et forskningsprojekt med et samfundsrelevant formål, der er anmeldt og godkendt af Datatilsynet.

DDA samarbejder med Datatilsynet om at regler og procedure for behandling af personfølsomme og personhenførbar data overholdes og har viden og rutine til at behandle datamaterialer med følsomme og personhenførbare data efter Datatilsynets anvisninger.

⁷ http://fivu.dk/forskning-og-innovation/tilskud-til-forskning-og-innovation/opslag/2013/filer-til-opslag/dk_dff_opslag_e2013_f2014-ydun.pdf p. 8.

⁸ Metode & Data, nr. 94: *Persondataloven - regler og praksis for god databehandlingsskik - Forskning med personoplysninger* (<http://samfund.dda.dk/ddakatalog/MogD/md94.pdf>), specialkonsulent Camilla Daasnes, Datatilsynet

Danske fonde

I perioden 2009-11 beviliggede danske, private fonde 5,6 mia. kr. til dansk forskning, mens Det Frie Forskningsråd og Det Strategiske forskningråd til sammenligning gav 6,7 mia. i samme periode. Fondene er altså af væsentlig betydning for landets forskningsfinanciering. Over 90% af alle beviliggede midler til offentlig forskning fra danske, private fonde kommer fra disse syv – nævnt i faldende rækkefølge: Novo Nordisk Fonden, Lundbeckfonden, VELUX Fondene, Carlsbergfondet, TrygFonden, Den Obelske Familiefond og Realdania⁹.

Generelt lægger mange af fondende stor vægt på formidling af resultater fra de aktiviteter, de støtter. Denne formidling finder både sted fra de støttede aktiviteter og fra fondene selv. Formidling er her imidlertid et bredt begreb, der i nogle tilfælde primært dækker forskningsbaseret undervisning mhp. uddannelsen af nye kandidater og forskere indenfor de respektive fondes felt.

Novo Nordisk Fonden

Fonden stiller i al væsentlighed kun krav om at gældende lovgivning vedr. håndtering af sensitive data.¹⁰

Lundbeck fonden

Af fondens information om generelle vilkår for forskningsbevillinger fremgår:

*"Lundbeckfonden forventer endvidere, at værtsinstitutionen og de forskere og andre personer, der er tilknyttet forskningsprojektet, uden unødig ophold skal give offentligheden størst mulig adgang til deres forskningsresultater og datagrundlag med så få restriktioner som muligt."*¹¹

VELUX Fondene

Samlebetegnelse for Willum fonden og Velus fonden. Begge har tilsluttet sig de offentlige forskningsråd og fondes OA-politik, og det er et krav til bevillingshaver, at politikken overholdes for så vidt angår publicering af forskningsresultater. Der stilles ikke krav om tilgængeliggørelse af bagvedliggende data.¹²

Carlsbergfondet

Fondet støtter kun i begrænset omfang enkelte større forskningsprojekter – fokus er på rejsestipendier, postdoc lønninger, indkøb af apparatur, rejser og publikationer mv. Fondet stiller ikke krav om tilgængeliggørelse af de bagvedliggende data, der måtte være resultatet af de støttede projekter.¹³

TrygFonden

Behandling, opbevaring og formidling af forskningsdata er ikke noget fonden stiller krav om. Men fonden accepterer allerede publikationsfees til OA tidsskrifter som del af ansøgningernes budget, og drøfter data spørgsmålet ifm. en igangværende revision af dens forskningsstrategi. Fonden har i flere sammenhænge støttet udviklingen/afprøvning af af datafangstsystemer i fx kommunalt regi, ligesom den påtænker at indlede et samarbejde med Danmarks Statistik mhp. at danne flere og bedre data indenfor det især det sociale område.¹⁴

⁹ http://www.damvad.com/media/46819/resume_fondsanalyse.pdf

¹⁰. Af informationsmaterialet til ansøgere fremgår: "It is assumed that the utilisation of funds granted complies with all applicable laws, rules and standards for trials involving people and animals, as well as for the handling of sensitive data. Furthermore, it is expected that you follow the Novo Nordisk Foundation's standards for good research practice." (sidstnævnte dokument berører ikke emnet behandling, opbevaring og formidling af data). http://www.novonordiskfonden.dk/sites/default/files/guidelines_for_applicants_may_2013.pdf p. 1.

¹¹ http://www.lundbeckfonden.com/media/Generel_information_om_Lundbeckfondens_bevilli.pdf p. 23

¹² <http://villumfonden.dk/oa>

¹³ http://www.carlsbergfondet.dk/~media/pdf-filer/vejledning_engelsk_dansk/Grant 2013 - udgave2.ashx samt mailsvar fra fonden 20.3.2014.

¹⁴ Mailsvar fra fonden 21.3.2014.

Det Obelske Familiefond

Af informationen vedr. behandling af oplysninger står:

*"Vi beder dig også samtykke til vederlagsfrit at stille informationsmateriale i form af tekster, fotografier, levende billeder og lyd mv. til rådighed for Det Obelske Familiefonds eget brug på hjemmeside, i præsentationer mv. Tilsvarende gælder for evalueringssrapporter og andet materiale, som de støttede aktiviteter eller projekter resulterer i, herunder kvalitative og kvantitative data."*¹⁵

Realdania

Reglerne for støttede projekter inkluderer denne paus: " Projektet skal formidle den opnåede viden til relevante interesser". Behandling og opbevaring og formidling af forskningsdata lader ikke til at være noget fonden stiller krav om.

Udenlandske bevillingsgivere af væsentlig betydning for dansk forskning

US National Institutes of Health (NIH)

Iflg. NIHs data sharing policy forventes projekter over en halv mio. USD at adressere spørgsmålet om deling af forskningsdata:

*"[NIH] expects and supports the timely release and sharing of final research data from NIH-supported studies for use by other researchers. "Timely release and sharing" is defined as no later than the acceptance for publication of the main findings from the final data set. All investigator-initiated applications with direct costs of \$500,000 or more (excluding consortium F&A costs) in any single year are expected to address data-sharing in their application."*¹⁶

NIH stiller således krav om en "data sharing plan" fra ansøger – også i de tilfælde, hvor hensynet til forsøgsdeltageres rettigheder begrænser mulighederne for data deling (skal begrundes). Reviewers evaluering af den forskningsmæssige værdi af et projektforslag influeres ikke af dets data sharing plan – det sorterer under NIH programlederne at vurdere og godkende den enkelte plan.

*"The precise content of the data-sharing plan will vary, depending on the data being collected and how the investigator is planning to share the data. Applicants who are planning to share data may wish to describe briefly the expected schedule for data sharing, the format of the final dataset, the documentation to be provided, whether or not any analytic tools also will be provided, whether or not a data-sharing agreement will be required and, if so, a brief description of such an agreement (including the criteria for deciding who can receive the data and whether or not any conditions will be placed on their use), and the mode of data sharing (e.g., under their own auspices by mailing a disk or posting data on their institutional or personal website, through a data archive or enclave [A controlled, secure environment in which eligible researchers can perform analyses using restricted data resources]). Investigators choosing to share under their own auspices may wish to enter into a data-sharing agreement."*¹⁷

Detaljeret gennemgang af, hvad en data sharing plan bør indeholde findes på
http://grants.nih.gov/grants/sharing_key_elements_data_sharing_plan.pdf

Ansøger er velkommen til at lade projektbudgettet afspejle eventuelle ekstraomkostninger som følge data deling og arkivering.

¹⁵ <http://www.obel.com/behandling-af-oplysninger>

¹⁶ NIH data sharing policy http://grants.nih.gov/grants/policy/nihgps_2013/nihgps_ch8.htm - _Toc271264951

¹⁷ NIH Data Sharing Policy and Implementation Guidance http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm - doc

US National Science Foundation (NSF)

NSF stiller krav om udformning af en datamanagementplan i forbindelse med ansøgning om støtte:

"Plans for data management and sharing of the products of research. Proposals must include a supplementary document of no more than two pages labeled "Data Management Plan". This supplement should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results (see [AAG Chapter VI.D.4](#)), and may include:

1. *the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;*
2. *the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);*
3. *policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;*
4. *policies and provisions for re-use, re-distribution, and the production of derivatives; and*
5. *plans for archiving data, samples, and other research products, and for preservation of access to them."*¹⁸

Ift. formidling og deling af data er der ikke tale om krav, men om forventninger:

*"Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants."*¹⁹

UK Wellcome Trust

Alle ansøgere opfordres til at overveje deres tilgang til data management og deling af data. Hvor genereret data vurderes at være af signifikant værdi for et bredere research community, skal ansøgere imidlertid indsende en data management og data sharing plan, der skal godkendes før funding kan udløses.

Der er ikke nogen fast form på, hvordan disse planer skal udformes, men de bør svare på disse syv spørgsmål²⁰:

- *What data outputs will your research generate and what data will have value to other researchers?*
- *When will you share the data?*
- *Where will you make the data available?*
- *How will other researchers be able to access the data?*
- *Are any limits to data sharing required - for example, to either safeguard research participants or to gain appropriate intellectual property protection?*
- *How will you ensure that key datasets are preserved to ensure their long-term value? What resources will you require to deliver your plan?*

Fonden ser data management og data sharing planen – foruden evt. budgetposter relateret til implementeringen af planerne – som en integreret del af ansøgningen.

¹⁸ http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_2.jsp se afsnit C. Proposal Contents, under afsnit 2j.

¹⁹ http://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/aag_6.jsp#VID4

²⁰ <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Guidance-for-researchers/>