

IDK Open Science Politikⁱ

Det overordnede sigte med open science er at sikre åbenhed om forskningsresultaters tilvejebringelse, kvalitet i forskning og evaluering heraf, samt at øge mulighederne for at genanvende forskningsdata i nye forskningsprojekter.

Feltet introduceres i det følgende med afsæt i den danske kodeks for forskningsintegritet (2014) samt FAIR-principper for data management (Wilkinson et al, 2016). Dernæst struktureres beskrivelsen af IDKs open science politik med reference til faserne i en livscyklusmodel for forskningsdata. Derefter kortlægges IDKs nuværende data management-praksis. Dokumentet fokuserer hovedsagligt på data management, herunder inddrages dog tillige anbefalinger vedrørende open access publicering af forskningsartikler. Dette fundament udgør grundlaget for udvikling af en robust open science politik. For nuværende prioriteres temaer vedrørende data management og databeskyttelse over spørgsmål om implementering af FAIR-principper.

I henhold til den generelle beskrivelse af dataadministration i *Den danske kodeks for integritet i forskning* (2014) betragtes primært materiale og data, som:

”ethvert materiale (f.eks. biologisk materiale, noter, interviews, tekster og litteratur, digitale rådata, registreringer, osv.), der udgør grundlaget for forskningen. *Data* er detaljerede optegnelser over det primære materiale, som udgør grundlaget for den analyse, der fører til resultaterne.”

(Den danske kodeks for integritet i forskning, 2014:9)

På den baggrund pointeres det, at:

”ansvarlig forskningspraksis omfatter korrekt administration af primært materiale og data. Hovedformålet med dataadministration er at garantere troværdig og gennemsigtig forskning.”

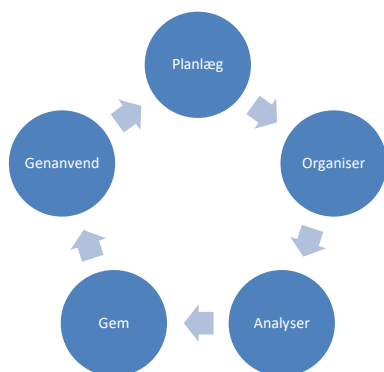
(Den danske kodeks for integritet i forskning, 2014:9)

Tilsvarende søger *guidelines for FAIR Data Principles: Findability, Accessibility, Interoperability, Reusability* (Wilkinson et al, 2016) at sikre kvalitet i frembringelse og evaluering af forskning samt at understøtte genanvendelse af data i andre forskningssammenhænge. Her er udgangspunktet, at indbyrdes relaterede FAIR principper enkeltvis kan karakteriseres som følger:

- a) (meta)data præsenteres på en entydig form, der kan repræsenteres i læsbare formater for mennesker og/eller maskiner. Endvidere beskrives og dokumenteres data fyldestgørende og grundigt med metadata (**F**indability).
- b) (meta)data gøres tilgængelige ved brug af åbne standarder og protokoller, der sikrer at data kan genfindes, identificeres, autentificeres, og at data desuden indeholder procedurer til eventuel autorisation af dataadgang (**A**ccessibility).

- c) (meta)data anvender etablerede standarder for vidensrepræsentation og benytter vokabularer, der følger FAIR principperne for at sikre forenelighed. Interoperationalitetsprincippet kan eksempelvis appliceres på juridiske forhold vedr. ansvar og intellektuel ejendomsret, som illustreret i *The Legal Interoperability of Research Data: Principles and Implementation Guidelines*¹ (Interoperability).
- d) (meta)data offentliggøres, lagres og opbevares med præcise, rige standardiserede beskrivelser, der fremmer genanvendelse (Reusability).

IDKs aktiviteter inden for data management beskrives med henvisning til aktiviteter i nedenstående forskningsdata-livscyklusmodel:



Figur 1: Model - Forskningsdatalivscyklus

I det følgende vil elementer fra denne livscyklusmodel formidle information om IDKs Open Science strategi.

Forskningsprojekter, der behandler persondata, anmeldes til RIO².

IDK open science & data management: organisering og medarbejderkompetencer

IDK understøtter kompetenceudvikling inden for data management og open science i tandem med de generiske kurser, der udbydes fra det Humanistiske Fakultet samt SDU.

¹ Zenodo: <https://doi.org/10.5281/zenodo.162241> (tilgået 01.02.18).

²SDU anmeldelse RIO SDU <https://www.sdu.dk/da/anmeldelse> (tilgået 21.02.18).

IDK sikrer, at nye ansatte introduceres til feltet (ansættelsessamtale, medarbejderkurser, evt. lokale IDK initiativer).

Forskningsmedarbejdere skal have forståelse for vigtigheden af data management og open science. Forskere skal administrere forskningsdata (såsom arkivmateriale, interviewdata, observationsdata, videodata, tekstmateriale, data fra sociale medier) i de forskellige faser i forskningsarbejdet - fra forskningsansøgning til planlægning af dataindsamling, databearbejdning og organisering samt dokumentation, herunder arkivering af data og evt. sikring af mulighed for genanvendelse af data i nye forskningsammenhænge.

Forskningsmedarbejdere skal have kendskab til den databeskyttelsesretlige regulering på området. Desuden skal han/hun have kendskab til etiske spørgsmål i relation til humanistisk forskning. I tilknytning hertil skal han/hun overveje etiske spørgsmål angående privathed, fortrolighed, selvbestemmelse og værdighed i tilknytning til humanistisk databaseret forskning generelt. Såfremt han/hun eksempelvis anvender data fra sociale medier forudsættes kendskab til internet research ethics (Buchanan & Ess, 2008), (Stanford Encyclopedia of Philosophy: Internet Research Ethics).

- ***Data management plan (DMP)***

Forskningsmedarbejdere skal udarbejde en data management plan³ (gennemgang af alle faser i datas livscyklus), herunder have kendskab til online hjælpemidler til udarbejdelse heraf (DMPonline⁴) samt kendskab til formelle krav til DMP i ansøgninger om forskningsmidler (f.eks. Horizon⁵). Ansvarlig for DMP og evt. revision heraf: forskningsprogrammets gruppeleder eller individuel projektejer.

- ***Data organisering, dokumentation og analyse***

Forskningsmedarbejdere skal dokumentere, kontekstualisere og begrunde dataarbejde og analyse (Corti, 2014). Han/hun skal tilse, at data og metadata kan organiseres på struktureret vis til gavn for eget projekt og eventuel genanvendelse og/eller publicering af forskningsdata. I forhold til systematik og sikring af konsistens skal han/hun benytte:

- disciplin-specifikke best practice annoteringsmetoder.
- data og fil-håndtering: formatering, versionering og konventioner for navngivning af datafiler, herunder sikre, at datasubjekter kan re-identificeres og genfindes, hvis de anmoder om dataindsigt el. sletning.

³F.eks. *Digital Curation Centre University of Edinburgh, Data Management plan*: <http://www.dcc.ac.uk/resources/data-management-plans> (tilgået 21.02.18). Samt tjekliste hertil: *Data management plan check list*: <http://www.dcc.ac.uk/resources/data-management-plans/checklist> (tilgået 21.02.18).

⁴DMPonline: <https://dmponline.deic.dk/> (tilgået 21.02.18).

⁵EUs portal for Åbne Data: <https://data.europa.eu/euodp/da/data/dataset/open-research-data-the-uptake-of-the-pilot-in-the-first-calls-of-horizon-2020/resource/c6088e4f-9e49-48f5-9760-18106e3e7822> (tilgået 21.02.18).

- lagringsstrategier, back up/master version af data som analysegrundlag, herunder sikre, at datadokumentation bag publikationer kan genfindes.⁶

- **Lagring og (langtids-)opbevaring af data**

Når der foreligger en udtalelse fra RIO, vil IT-service understøtte forskningsprojektets behov for sikker opbevaring af data. Forskningsmedarbejderen skal have kendskab til internationale standarder for metadata, der kan sikre, at data kan genanvendes og fortolkes af andre i fremtiden. Han/hun skal have kendskab til infrastrukturer (lokale, (inter-)nationale) til lagring af data⁷ samt kunne træffe informeret beslutning om, hvor længe, og hvordan data skal opbevares.

Hvis forskningsmedarbejderen forlader IDK, og ikke ønsker at tage data med sig (når det drejer sig om persondata, ejer forsker data, med mindre andet er aftalt), overdrages data og dataansvar til forskningsprogrammets leder eller projektets leder. Det er institutleders ansvar at sikre, at overdragelsen finder sted, før medarbejderen forlader IDK.

Når projekter er afsluttet kontaktes RIO. Det afgøres, hvorvidt data skal slettes eller anonymiseres (forsker er forpligtet hertil, hvis data ikke kan overgives til Rigsarkivet) eller, hvorvidt de ikke-anonymiserede data kan overføres til Rigsarkivet. Det afgøres endvidere, hvorvidt de anonymiserede data kan publiceres. Det er muligt at bevare data en vis periode efter projektafslutning, f.eks. af hensyn til artikelskrivning eller dokumentation af resultater.

- **Genanvendelse af data**

Forskningsmedarbejderen skal træffe afgørelse om, hvilke data, der på forsvarlig vis kan deles samt, hvornår, og hvordan, de i givet fald kan deles. Han/hun skal tilstræbe at genanvendelse af data fremmes. Han/hun bør orientere sig i guidelines vedrørende og data citation⁸ (viden om, hvor, og, hvordan datasæt kan publiceres).

- **Open Acces Publicering af forskningsartikler**

I overensstemmelse med *Danmarks Nationale Strategi for Open Access*⁹ opfordrer IDK forskningsmedarbejdere til at optimere vidensdeling og anvendelse af forskningsresultater via Open Access Publicering¹⁰. IDK opfordrer til, at forskningsmedarbejdere publicerer artikler i overensstemmelse med de to nedenstående modeller, hvoraf *Green Open Access* dog er den foretrukne, da denne er omkostningsfri for instituttet såvel som for samfundet:

⁶MIT Libraries. Data management and publishing: Organize your files. MIT: <http://libraries.mit.edu/data-management/store/organize> (tilgået 21.02.18).

⁷F.eks.: Forskningsdata opbevaring – Rigsarkivet: <https://www.sa.dk/aflevering-arkivet/private-og-forskere/anmeldelse-og-aflevering-af-forskningsdata> (tilgået 21.02.18), Data DeIC: https://www.deic.dk/da/data_deic_dk (tilgået 21.02.18).

⁸F.eks.: Data Cite – Cite Your Data: <https://www.datacite.org/cite-your-data.html> (tilgået 21.02.18).

⁹Danmarks Nationale Strategi for Open Access: <https://ufm.dk/forskning-og-innovation/samspil-mellem-viden-og-innovation/open-access/billeder-og-filer/danmarks-nationale-strategi-for-open-access.pdf> (tilgået 01.05.18).

¹⁰Information om Open Access publicering ved SDU: <https://www.sdu.dk/da/forskning/forskningspublicering> (tilgået 21.02.18).

Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020:

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

Green Open Access: Herved søges om tilladelse til, at artikler, publiceret i kvalitetssikrede tidsskrifter, kan arkiveres i en Open Access version ("final author version approved") i SDU PURE.

Gylden Open Access: Henviser til tidsskrifter, der udgives som open access med direkte online tilgængeliggørelse af forskningsartikler. Omkostninger til publicering betales af forfatter.

Ved Syddansk Universitetsbiblioteks Open Access Fond¹¹ kan der søges støtte til publiceringsudgifter til Open Access publicering¹². IDK yder ikke støtte til Open Access publicering.

IDK praksis: data management¹³

Forskningsprojekter, der behandler følsomme eller andre fortrolige personoplysninger (forskningsdata), anmeldes til RIO¹⁴ (ved udfyldelse af anmeldesskema), (Efter 25. maj, 2018 anmeldes alle forskningsprojekter til RIOs fortegnelse), der sikrer, at de persondataretlige krav er opfyldt. SDU RIO skal kontaktes på ny:

- når forskningsdata ønskes videregivet til eksterne forskningspartnere.
- når forskningsdata ønskes genbrugt internt til andre forskningsprojekter.
- når forskningsdata overlades til en databehandler.
- når der sker ændringer i projektet, såsom ændring af formål, kontaktperson, personkreds, indhentning af nye typer personoplysninger eller forlængelse af projektperiode. Mindre væsentlige ændringer (f.eks. projektnavn) anmeldes blot til SDU RIO senest 4 uger efter ændringen).

Prototypiske forskningsdata på IDK består af survey data, lyddata eller videoobservationsdata. Der håndteres både personoplysninger og følsomme oplysninger.

- ***Datatyper i Persondataloven (nuværende og nye databeskyttelseslovgivning:25.05.18)***

Pseudonymiserede data: (data, hvor f.eks. identifikator/navn erstattes af nummer, der opbevares/kan genfindes).

Almindelige personoplysninger: identifikationsoplysninger, oplysninger om økonomiske forhold, kundeforhold eller andre lignende oplysninger.

¹¹ Open Access Fond: <https://www.sdu.dk/da/forskning/forskningspublicering/open+access+fond> (tilgået 21.02.18).

¹² Directory of Open Access Journals : www.doaj.org (tilgået 21.02.18).

¹³ Dette afsnit har hovedsageligt fokus på data management i relation til databeskyttelsesregulering og berører dermed kun sporadisk open science strategier for IDK.

¹⁴SDU anmeldelse RIO SDU <https://www.sdu.dk/da/anmeldelse> (tilgået 21.02.18).

Følsomme oplysninger: oplysninger om racemæssig/etnisk baggrund, fagforenings-tilhørsforhold, seksuelle forhold, religion, politisk orientering, filosofisk orientering, genetiske data, biometriske data.

Ikke inkluderet i lovgivningen:

Anonymiserede data: disse er ikke omfattet af databeskyttelseslovgivning (med mindre man med rimelige midler (hacking, el. lovligt) kan re-identificere datasubjekter via datafragmenter.

- **IDK: Det informerede samtykke**

Indsamling, brug, analyse m.v. af forskningsdata, der kun benyttes til forskningsformål, kræver ikke informeret samtykke, men kan ske med hjemmel i persondatalovens § 10, efter den 25/5 2018 med hjemmel i databeskyttelseslovens §10 stk. 1-5. Skal personoplysninger også bruges til andre formål, f.eks. undervisning er samtykke normalt nødvendigt.

Forskningsmedarbejdere på IDK benytter ofte informeret samtykke, eftersom dette er en naturlig indgangsvinkel i mødet med undersøgelsesdeltagere, ligesom der, med et informeret samtykke, kan bedes om tilladelse til at benytte forskningsdata i undervisningssammenhænge samt til fremtidige forskningsprojekter¹⁵.

Det er vigtigt at overveje forsknings- og formidlingsbehov ift. samtykke, så dette ikke afgives i en for snæver kontekst:

- Hvis materialet tillige skal anvendes til undervisningsformål kan forskeren pro-aktivt bede om samtykke til at benytte materiale, der indeholder følsomme oplysninger (f.eks. fra samtalefragmenter i video, audiooptagelser) – her skal man være opmærksom på forbuddet i persondatalovens § 7, stk. 8: *For den offentlige forvaltning må der ikke føres edb-registre med oplysninger om politiske forhold, som ikke er offentligt tilgængelige.*
- Eksempelvis kan materialet benyttes i fremtidige forskningsprojekter. Her bør forskeren, om muligt, specificere disse.

Forskningsmedarbejdere skal udarbejde etisk forsvarlige samtykkeerklæringer, der i videst mulig omfang beskytter undersøgelsesdeltagernes integritet.

- Eksempelvis bør man indhente samtykke, også hvor dette ikke er påkrævet – f.eks. optagelser i det offentlige rum.
- Eksempelvis bør man undlade at bede om samtykke til forhold, der kan udstille deltagere.

¹⁵ Cessda.eric: Consortium of European Social Science Data Archives - European Research Infrastructure Consortium: [Beskrivelse af informeret samtykke samt templates hertil: https://www.cessda.eu/Research-Infrastructure/Training/Expert-tour-guide-on-Data-Management/5.-Protect/Informed-consent](https://www.cessda.eu/Research-Infrastructure/Training/Expert-tour-guide-on-Data-Management/5.-Protect/Informed-consent) (tilgået 23.02.18).

- **Sikker dataopbevaring**

Data lagres for nuværende på PC, lokalt netværksharddisk, Mdrev, Sdrev, mapper, ekstern harddisk, SDU Sharepoint, Dropbox.

Fremover lagres data i overensstemmelse med IT-services' anbefalinger for pågældende datatyper (f.eks. Sharepoint). Der lagres ikke på Dropbox, Google Drev eller lignende.

- **Sikker databehandling**

Data behandles fremover ved brug af godkendte databehandlingsprogrammer, herunder de ressourcer, der stilles til rådighed via IT-service. Eksempelvis har SDU to aftaler med spørgeskemasystemer, der kan håndtere persondata: SurveyXact: <https://wayf.survey-xact.dk/>, hvor forskere selv kan oprette sig, samt Redcap Cloud, hvor oprettelse kan ske ved henvendelse til servicedesk@sdu.dk.

Der benyttes ikke uspecificerede onlineredskaber til databehandling, såsom f.eks. SurveyMonkey. Databehandlingsprogrammer skal registreres og godkendes af IT-service. I sammenhænge, hvor forskningsmedarbejdere anvender værktøjer via institutioner, såsom digital Humanities lab¹⁶, CLARIN¹⁷ og DEIC¹⁸ godkender it-service disse pr. default.

Referencer

- Buchanan, E. A., Ess, C. (2008) Internet Research Ethics: The Field and its Critical Issues. In: K. E. Himma, H. Tavani (Eds.) *The Handbook of Information and Computer Ethics* (1). Hoboken, US: Wiley-Interscience. 273-291.
- Corti, Louise (2014) *Managing and Sharing Research Data: A Guide to Good Practice*. 1st edition. Ch. 3. Thousand Oaks, CA: SAGE Publications.
- Den danske kodeks for integritet i forskning (november 2014) Uddannelses- og forskningsministeriet. <http://ufm.dk/publikationer/2015/den-danske-kodeks-for-integritet-i-forskning> (tilgået 01.02.18).
- Stanford Encyclopedia of Philosophy: Internet Research Ethics. <https://plato.stanford.edu/entries/ethics-internet-research/> (tilgået 01.02.18).
- Wilkinson, M. D. *et al.* (2016) The Fair Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1030/sdata.2016.18.

Ressourcer

(links herunder er tilgået 22.02.18)

Cessda.eric: Consortium of European Social Science Data Archives - European Research Infrastructure Consortium – Expert Tour on Data Management: <https://www.cessda.eu/Research-Infrastructure/Training/Expert-tour-guide-on-Data-Management>

DEIC National HPC Centre, SDU, Abacus 2.0: <https://abacus.deic.dk/>

¹⁶ DIGIHUM: <http://dighumlab.com/>

¹⁷ CLARIN: <https://www.clarin.eu/>

¹⁸ DEIC: <https://www.deic.dk/>

DEIC: <https://www.deic.dk/>

Digital Curation Centre University of Edinburgh, *Data Management plan*: <http://www.dcc.ac.uk/resources/data-management-plans>

Digital Curation Centre University of Edinburgh, *Data management plan check list*: <http://www.dcc.ac.uk/resources/data-management-plans/checklist>

Digital Curation Centre University of Edinburgh, *DMPonline*: <https://dmponline.dcc.ac.uk/>

Digital Humanities in the Nordic countries: <http://dig-hum-nord.eu/projects/>

Digital Humanities Lab Denmark: <http://dighumlab.com/>

Forskerportalen: <http://forskerportalen.dk/?p=1377#Generelt>

Forskningsdata opbevaring – Rigsarkivet: <https://www.sa.dk/aflevering-arkivet/private-og-forskere/anmeldelse-og-aflevering-af-forskningsdata>

Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

Nationale Kulturarvscluster, Statsbiblioteket: <https://www.statsbiblioteket.dk/kulturarvscluster>

Open Access til Data – Uddannelses- og forskningsministeriet <http://ufm.dk/forskning-og-innovation/samspil-mellem-viden-og-innovation/open-access/data>

The EC Digital Single Market Strategy – European free flow of data initiative: <https://ec.europa.eu/digital-single-market/en/economy-society-digital-single-market>

¹Stor Tak til Lars Engelstoft for juridisk kommentering samt tak til Asger Væring Larsen for sparring undervejs. IDKs Open Science Politik er udarbejdet af Anne Gerdes og inddrager materiale fra dele af Gerdes' bidrag til publikationen: *Anbefalinger til udvikling af et Blended Learning-forløb Kursus i Data Management for Ph.d.-studerende - baseret på læringsmål for Humaniora og Sundhedsvidenskab*, af: Gerdes, Anne; Motzfeldt, Hanne Marie; Hansen, Karsten Kryger; Remvig, Kirstin; Hall Andersen, Mette.