

FORVALTNING AF FORSKNINGSDATA I DANMARK – SKER DER NOGET?

De senere år er forskningsdata kommet højt op på den forskningspolitiske dagsorden. Ny rapport tager temperaturen blandt forskningsbibliotekerne herhjemme og ser lidt nærmere på, hvad fremtiden kan bringe.

AF BERTIL F. DORCH, BFD@BIB.SDU.DK, SYDDANSK UNIVERSITETSBIBLIOTEK, FILIP KRUSE, FKR@STATSBIBLIOTEKET.DK & JESPER BOSERUP THESTRUP, JBP@STATSBIBLIOTEKET.DK, STATSBIBLIOTEKET

I de seneste år er forskningsdata kommet højt op på den forskningspolitiske dagsorden. Begreber som 'dataforvaltning', 'data management planning', 'data dissemination' m.m. indgår stadigt oftere i debatten. Bag de forskellige betegnelser gemmer sig aktiviteter som bevaring og kuratering af forskningsdata, sikring af adgang til egne og andres data, og hvorledes disse kan medtænkes allerede fra starten af forskningsprojektet. Denne artikel handler om, hvad der sker – og ikke sker – på danske universiteter på dette område og om nuværende og mulige fremtidige aktiviteter og aktører.

I 2012 undersøgte vi de otte danske universiteters praksis indenfor forvaltning af forskningsdata, især kort- og langtidsbevaring af forskningsdata og muligheder for deling eller andre former for tilgængeliggørelse af disse. Undersøgelsen gennemførtes som et DEFF-projekt med deltagere fra alle universitetsbibliotekerne. Data blev indsamlet ved en række interviews med fast spørgsmålsformat med ledelsesrepræsentanter fra universiteterne, deres IT-afdelinger og universitetsbibliotekerne. Rapporten – *Forvaltning af forskningsdata* – kan downloades fra DEFFs hjemmeside.

Status i Danmark

Hverken korttidsbevaring af forskningsdata – svarende til løbende data-backup undervejs i forskningsprocessen – eller langtidsbevaring ses af universiteterne som deres naturlige opgave. Universiteterne ser klart publicering af forskningsresultater som vigtigere end bevaring og tilgængeliggørelse af data. Korttidsbevaring ses som den enkelte forskers eller instituts ansvar, og udover adgang til universitetsdrev tilbydes der generelt ikke faciliteter hertil. Kun to universiteter råder over tekniske faciliteter til langtidsbevaring.

Det er holdningen, at forskernes behov i vidt omfang bør være bestemmende og samtidig udtrykkes usikkerhed overfor de tekniske krav, også i medfør af de mange forskellige datatyper. Mulige fremtidige fælles løsninger må tage højde for disse forhold.

Nationale dataarkiver som supplement til de eksisterende foretrakkes frem for institutionsbaserede. Begrundelserne er igen de mange forskellige datatyper og bredden i forskernes behov. Desuden anses en national arkivløsning som fremmende for forskningsamarbejdet. Enkelte universiteter overvejer dog at etablere egne arkiver.

De fleste universiteter har aktiviteter i gang – eller planlægger sådanne – for at øge forskernes opmærksomhed på bevaring og tilgængelighed af forskningsdata. Men stadigvæk er publikationer vigtigst. Alt i alt har universiteterne ikke mange aktiviteter i forhold til forvaltning af forskningsdata. Det kan ses som udtryk for passivitet, men også som udtryk for, at forskernes behov bør være bestemmende. Afvisningen af centraliserede løsninger og argumenterne i forhold til diversitet i data, forskningskulturer osv. peger i samme retning.

Udlandet og Danmark

Der er ikke kun tale om en dansk problemstilling. EU arbejder på at indføre krav om adgang til og bevaring af forskningsdata, der er indsamlet for EU-midler, og USA samt en række europæiske lande har indført eller ved at indføre lignende krav. Internationalt har forskningsbibliotekerne allerede en central placering i forhold til forvaltning af forskningsdata, og udviklingen af denne rolle er af stor betydning. Eksempelvis var netop forskningsdata et af de gennemgående temaer på LIBER-konferencen i 2013, hvor blandt andet resultaterne fra ovennævnte undersøgelse blev præsenteret.

For at udvikle en samlet dansk infrastruktur, der kan håndtere forskningsdata på en ensartet måde og dermed undgå risikoen for, at de enkelte universiteter udvikler hver deres system, har DEFF besluttet at støtte projektet *FIF – Fælles Infrastruktur for Forskningsdata*.

Projektet involverer DeIC, DTIC, Statsbiblioteket, Det Kongelige Bibliotek, CBS-bibliotek og Roskilde Universitetsbibliotek. Formålet er, at undersøge de krav en sådan infrastruktur stiller og behovene for løsninger. Derefter skal de involverede parter implementere et sådant system.

Nyt fra Danmark – København og Aarhus

På Københavns Universitet har forskningsdata (organisering, strukturering og opbevaring af disse) indgået i arbejdet omkring regler om god videnskabelig praksis siden 2011. I arbejdet indgår en justering af det eksisterende regelsæt for universitetets praksisudvalg, en mulig ordning omkring "whistleblower" eller "named person" og sidst, men ikke mindst, fremme af en kulturændring på området. En del af baggrunden for



indsatsen er også eksisterende sager omkring overtrædelse af reglerne for god videnskabelig praksis.

Senest er praksisudvalget til bestyrelsen for Københavns Universitet fremkommet med anbefalinger til en politik på området, som vil være grundlag for et nedsat dekanudvalgs arbejde. Disse understreger, at det er forskerens ansvar at beskrive, hvorledes forskningsdata indsamles, opbevares og sikres. Videre, at data skal opbevares, så resultater kan genskabes og, at opbevaringen skal være sikret mod datamanipulation.

Udover det styrkede fokus på det mere overordnede politiske plan, er der samtidig taget initiativer til at øge forskernes opmærksomhed på problemstillingerne omkring forvaltning af forskningsdata. Således arrangerede Københavns Universitets Natur- og Biovidenskabelige Fakultet (SCIENCE) i april 2013 workshoppen *Mega Data* omkring de juridiske, teknologiske, etiske og økonomiske problemer og muligheder. Dette blev et startskud for fakultetet til at iværksætte en proces rettet mod forskerne, der skal sikre, at de er ordentligt klædt på til opgaven. Workshopens resultater er sammenfattet i et white paper, som kan fungere som guide for universitetet i dets fremtidige prioriteringer omkring håndtering af forskningsdata.

Dette white paper behandler bl.a. spørgsmålet omkring de stigende datamængder, som forstærker problemerne omkring sikker opbevaring og herunder uddannelsen af kvalificeret akademisk personale til at varetage bevaring og kuratering. Samtidig kræves det af forskerne, at de stiller resultaterne af offentlig finansieret forskning til rådighed for andre, da resultaterne ses som tilhørende samfundet. Institutionerne må selv støtte forskerne med

de nødvendige retningslinjer, for at de kan imødekomme de nye krav og forventninger. Hertil er der presserende behov for en teknisk infrastruktur. Workshoppen vil senere blive fulgt op med yderligere tre omhandlende publicering, genbrug og deling af data samt metadatering af forskningsdata.

På de indre linjer

Disse aktiviteter eksemplificerer, at der i universitetsregi på de indre linjer arbejdes med forvaltning af forskningsdata. Det er værd at bemærke, at dette arbejde sker på trods af fraværet af en national politik på området og en ligeledes manglende forskningsinfrastruktur.

Også på forskningsbibliotekerne arbejdes der med forvaltning af forskningsdata. Det gælder således LARM, hvor Statsbiblioteket i samarbejde med andre institutioner arkiverer danske radioudsendelser i søgbar form, hvor forskerne kan annotere de eksisterende metadata og dele annotationerne. Langtidsbevaring, genbrug og deling af data er her ikke kun programpunkter, men virkelighed.

Andre eksempler på langtidsbevaring er Netarkivet, hvor Statsbiblioteket i samarbejde med det Kongelige Bibliotek høster det danske internet og Statsbibliotekets eget Mediestream, der samler kulturarven i skikkelse af avis- radio- og tv-samlingerne. Disse er forsynet med metadata og gjort søgbare. Etablering af sådanne digitale forskningsinfrastrukturer kræver emnespecifikt kendskab til samlingerne og til forskernes arbejdsmetoder for at kunne udvikle relevante tekniske løsninger. Forskningsbibliotekerne er således allerede engageret i bevaring af forskningsdata og i samarbejde med forskerne omkring udvikling af tekniske løsninger. ●