

# Forskningsdata Down Under – og måske right here?

ANDS er den australske satsning i håndtering og formidling af forskningsdata. De har midlerne – masser af dem – men spørgsmålet er, hvad de kan bidrage med af visioner og ideer til de danske forskningsbiblioteker.

Af Alfred Heller [ajh@dtic.dtu.dk](mailto:ajh@dtic.dtu.dk), DTU Bibliotek – Danmarks Tekniske Informationscenter og Lars Nondal [ln.lib@cbs.dk](mailto:ln.lib@cbs.dk), CBS Bibliotek

Australien har i flere år investeret store beløb i forskningsinfrastrukturen, og lige nu er forskningsdata et af de højest prioriterede indsatsområder. Det var derfor oplagt, at vi på DEFF's studietur til Australien i oktober 2010 aflagde besøg hos Australian National Data Service (ANDS), som er et nyt initiativ sat i verden for at lette australske forskeres vej til nemt at publicere, finde, genbruge og udgive forskningsdata.

ANDS så dagens lys i 2009 og er nu en organisation med ca. 50 ansatte. Foreløbig har man bevillinger til 2013, og det er ikke småpenge, det drejer sig om: 390 mio. kr. lyder de samlede bevillinger på for perioden 2009 til 2013.

Det er ANDS' erklærede mål at genbruge forskningsdata – "More Australian researchers reusing research data more often", som der står på deres hjemmeside. ANDS vil give australske forskere adgang til data, der giver dem konkurrencefordele i forhold til forskere i udlandet, og samtidig skal det være attraktivt for udenlandske forskere at søge til australske universiteter, mens udenlandske universiteter skal ansøres til at indlede forskningssamarbejder med australske universiteter. ANDS mener desuden – som de fleste andre – at det er nødvendigt med åben adgang til forskningsdata for at kunne se forskerne efter i sømmene og efterprøve deres resultater.

## Markedsføring og formidling af forskningsdata

Overordnet set har ANDS to opgaver: Det er et nationalt markedsførings- og formidlingsorgan for forskningsdata, og det handler om at sørge for, at forskningsdata bliver gjort søgbare og synlige. Men for at kunne formidle data skal de være tilgængelige, og derfor arbejder ANDS intensivt med projekter i forskningensenheder og universiteter, der går ud på at opdyrke kompetencer, hvad angår forskningsdata.

Det betyder, at ANDS primære arbejdsområde er metadata og formidling, mens det tunge arbejde med lagring, storage og langtidsofbevarening af forskningsdata varetages lokalt på de enkelte universiteter. Det er en arbejdsdeling, der begrundes i, at ANDS trods de massive bevillinger endnu ikke er en permanent institution.

På de enkelte australske universiteter arkiveres forskningsdata/metadata som oftest i traditionelle Institutional Repositories, som oprindeligt er bygget op omkring

**“Den store udfordring er at få forskerne og forskningsprojekterne til at aflevere data. Det er sjældent teknikken, der er det største problem, men derimod de lidt blødere aspekter som kultur, psykologi og tradition, der har betydning for, om og hvornår data bliver afleveret.”**

publikationerne. I de tilfælde, hvor data fylder rigtig meget, finder man kun metadata i det institutionelle arkiv, mens rådataene findes i særlige storage-faciliteter.

Og som med publikationerne er en af de store udfordringer at få forskerne og forskningsprojekterne til at aflevere deres data. Det er sjældent teknikken, der er det største problem, men derimod de lidt blødere aspekter som kultur, psykologi og tradition, der har betydning for, om og hvornår data bliver afleveret.

## Hvad høstes?

Der opereres med flere forskellige slags metadata omkring et sæt forskningsdata. Et eksempel er de såkaldte 'experimental metadata', som opstår og fanges umiddelbart i forbindelse med et forsøg, det vil sige ved for eksempel målinger og observationer. Senere tilføjes andre metadata til et sæt forskningsdata, men det er kun det sidste sæt, som kaldes enten 'discovery metadata' eller 'collections metadata', som ANDS importerer eller høster fra universiteterne.

Udover central høstning har de enkelte forskere mulighed for at lade deres data registrere direkte hos ANDS ved hjælp af "Publish My Data self-service/ Register My Research Data", en service, der mest er beregnet for forskere tilknyttet organisationer, hvorfra det ikke er muligt for ANDS at høste de tilsvarende metadata.



Foto: Creative Commons www.flickr.com/photos/36199303@

University of Melbourne er grundlagt i 1853 og er af Times Higher Education World University Rankings bedømt som australiens bedste universitet i 2010.

### Undgå særskilte søgemaskiner

Data samles hos ANDS i en "Australian Research Data Commons" (ARDC), hvis portal har fået navnet Research Data Australia (<http://services.ands.org.au/home/orca/rda/>). Men da man søger at tilpasse sig forskernes sædvanlige søgemønstre og ikke vil introducere alt for mange særskilte søgetjenester, lægger man faktisk mere vægt på at generere et omfattende system af websider, der høstes af de store søgemaskiner (Google, Yahoo osv). Der genereres websider for hver forsker, hvert projekt, hver region/delstat osv.

Udover at sørge for at disse websider høstes af søgemaskinerne, fokuserer man på, at de også eksponeres i de særlige, fagspecifikke databaser eller dokumentationssystemer, der findes indenfor de forskellige fagområder.

ANDS stiller ikke ultimative krav om, at forskningsdata stilles kvit og frit til rådighed for enhver, der finder frem til dem via ANDS-portalen. Men alle datasæt skal være beskrevet med oplysninger om, hvor data findes, og ikke mindst, hvem man skal henvende sig til for at få adgang.

### Kodeks for forskningspraksis

Alle australske universiteter er underlagt *Australian Code for the Responsible Conduct of Research* (<http://www.nhmrc.gov.au/publications/synopses/r39syn.htm>).

Det er et kodeks, som gælder for både den enkelte forsker og det enkelte universitet, og målet er at sikre ansvarlig forskningspraksis – ikke mindst, hvad angår håndtering af forskningsdata og -resultater.

ANDS arbejder intensivt med projekter på de australske universiteter, og man har indtil videre brugt 90 mio.kr på projektarbejdet. Projekterne falder i to kategorier: 'Seeding the Commons' og 'Data Capture'.

'Data Capture'-projekter handler typisk om enkelte forskningsprojekters evne til at indsamle data 'automatisk', dvs. datafangsten finder sted ved målinger eller observationer i forbindelse med et forsøg eller et eksperiment. 'Seeding the Commons'-projekterne finder derimod typisk sted indenfor de blødere discipliner og handler i højere grad om at få opbygget institutionelle infrastrukturer og procedurer til indsamling og organisering af forskningsdata. Derudover drejer det sig om, at få de nødvendige metadata sendt videre til Australian Research Data Commons (ARDC).

### Forskningsbiblioteket som aktør på dataområdet

Men hvilken rolle spiller så de australske universitetsbiblioteker i al dette? Kompetencer til håndtering af metadata, som ANDS jo handler om, er i høj grad til stede på bibliotekerne – i hvert fald hvis vi taler om 'discovery'- eller 'collection'-metadata. Og det er jo også bibliotekerne, der har arbejdet med og ansvaret for de Institutional Repositories, der er centrale omdrejningspunkter i håndteringen af netop metadata.

Arbejdsdelingen på de australske universiteter er derfor skruet sådan sammen, at IT-organisationen tager sig af arkiveringen af de ofte store datasæt, mens bibliotekerne står for metadata og ikke mindst for rådgivning, kontakt, uddannelse, forskerinterviews – og dermed de såkaldte Data Management Plans.



## Data Management Plans

Data Management Plans er i det hele taget en af ANDS' kongstanker, og på hjemmesiden (<http://www.ands.org.au/researchers/manage-data.html>) findes materiale om dette værktøj til forberedelse af dataindsamling, datahåndtering og til beskrivelse, tilgængeliggørelse og arkivering af data.

På Monash University og University of Melbourne er dette arbejde i høj grad placeret hos universitetsbibliotekernes forskerservices/institutservices. På Monash har biblioteket udpeget flere Data Management Specialists, der hjælper institutter og forskere med at få procedurer og kompetencer på plads. Specialisten går i dialog med forskerne og er med til at formulere et projekts Data Management Plan.

Som resultat af arbejdet med Data Management Plans har Monash desuden udarbejdet en Research Data Planning Checklist (<http://www.researchdata.monash.edu.au/resources/datahdrchecklist.doc>), der bruges af bibliotekets dataspecialist, når vedkommende tager et behovsafdækkende interview med forskeren/forskergruppen i starten af et forskningsprojekt.

Ved siden af tilbydes der intern undervisning og uddannelsesforløb for forskerne, f.eks. som led i de ph.d-studerendes introduktionsforløb. Så forskerservice på Monash handler i lige så høj grad om data management som om bibliometri, citationsanalyser, Open Access og informationssøgning. Endelig har Monash University som sådan udarbejdet en foreløbig politik for håndtering af forskningsdata (<http://www.researchdata.monash.edu.au/resources/draftpolicy.html>).

## Centralt datalager og fees

Mens Monash University var de første på området og i øvrigt var med til at starte ANDS, så har University of Melbourne været først om at indføre en egentlig politik på området. De har tilmed opbygget et centralt datalager, som alle fakulteter og forskere kan bruge.

Her opereres i øvrigt med 'data fees', hvor forskningsprojekterne skal betale for at få lagret deres data. Det giver universiteterne mulighed for at etablere økonomisk bæredygtige modeller. Og når f.eks. forskningsråd stiller krav om forsvarlig lagring af data efter projekternes ophør, er det blevet legitimt og almindeligt at inkludere disse omkostninger i projektansøgningerne.

University of Melbourne Library kan ligeledes præsentere en veludviklet efteruddannelsesstrategi for bl.a. data. Der uddannes specialister i data, copyright, data management, lagring, langtidsbevaring osv. Og forskningsprojekterne tilbydes hjælp fra metadataspecialister og dataspecialister fra biblioteket.

## Bibliotekerne som motivationsfaktor

Men uanset indsatsen på Monash og Melbourne, så er det ikke en opgave, der kan løses på kort tid. University of Melbourne giver udtryk for, at det kommer til at tage fem år, før der er nogenlunde styr på tingene. Og som i tilfældet med indsamling af data fra forskere og projekter er det heller ikke denne gang teknikken, der fremhæves som det største problem, men de organisatoriske udfordringer og forskernes adfærd.

Forskerne skal være motiverede, hvis de skal bidrage. Noget af motivationen kommer selvfølgelig 'oppe fra', når der bliver formuleret eksplicite krav om datahåndtering, f.eks. i kraft af Australian Code for the Responsible Conduct of Research.

Men der er også brug for mere positiv motivation. Og her kan bibliotekerne komme på banen. Ikke mindst fordi bibliotekerne har en positiv track-record hos både forskere og institutionsledelse, når det handler om konkrete services som forskningsregistrering, arkivering og publicering – og ikke mindst formidling. I stedet for de administrative byrder, der er forbundet med arbejdet med data, bør der lægges vægt på at opfatte data som et aktiv – for både forskeren og institutionen – der formidlet på den rigtige måde kan give øget synlighed og prestige i en meget konkurrencepræget verden.

## Danske forskningsbiblioteker som datahjælpere

Kan vi tage noget med hjem? Ikke overraskende kan vi i hvert fald konkludere, at forskningsbiblioteker absolut kan spille en central rolle, når man taler om opgaver som håndtering af forskningsdata, herunder lagring, langtidsbevaring og formidling. Og det vil derfor være både naturligt og hensigtsmæssigt, at bibliotekerne får større opgaver og ansvar på dette område, evt. i samarbejde med IT-afdelinger.

En anden og måske mere overraskende konklusion er, at bibliotekerne egentlig føler sig ret komfortable i rollen som datahjælpere. Efter de indledende udfordringer føler de australske forskningsbiblioteker det naturligt at varetage de opgaver, der knytter sig til forskningsdata. De opfatter dem som naturligt sammenhængende med andre forskerstøtteopgaver som citationsmetoder, publiceringsværktøjer, OA og meget mere. Og at det i øvrigt giver gode udviklingsmuligheder og kompetencespring for dele af personalet.

Så det går fint i tråd med den trend, der findes på danske universitetsbiblioteker, hvor fokus i stigende grad rettes mod udvikling af forskerservices.

Danmark har i forhold til Australien en mere veletableret tradition for og infrastruktur til håndtering af forskningspublikationer og videnskabelig kommunikation. Derimod er vejen til både lokale og nationale løsninger på forskningsdataområdet endnu lang.

## Data på dagsordenen

Forskningsdata ligger stadig i mange tilfælde gemt (og glemt?) på usikre steder, og de er ofte utilgængelige for andre forskere. Men forskningsdata er ved at komme på dagsordenen. På det politiske plan sker det i kraft af Videnskabsministeriets Handlingsplan om forskningsinfrastruktur, de faglige panelers rapporter og endelig Open Access-udvalgets rapport fra foråret 2010.

Også DEFF har i den nuværende handlingsplan forskningsdata på dagsordenen. I 2010 har DTIC, AAU, KUBIS, KB i samarbejde med Dansk Data Arkiv (DDA) arbejdet med et projekt, der bl.a. har til formål at belyse en række spørgsmål omkring forskernes incitament til at deltage i arbejdet med indsamling og formidling af forskningsdata. DEFF-projektet skal afrapporteres i begyndelsen af 2011. 