

# Forskningsdata – ett biblioteks läranderesor i de internationella, nationella och lokala landskapen

**Sara Kjellberg & Jonas Fransson**

*Malmö högskolas bibliotek har under flera år gradvis arbetat alltmer med forskningsdatafrågor. Efter en intern fortutbildning inom ansvarigt team på biblioteket har arbetet fördjupats i projektform. En enkät till högskolans forskare gav en nulägesbild, erfarenheter av och syn på forskningsdatafrågor med fokus på öppet tillgängliggörande, samt kunskap om i vilka frågor mer stöd och kunskap behövdes. Biblioteket deltar också i Svensk Nationell Datatjänsts (SND) pilotprojekt. Deltagandet har inneburit kompetensutveckling på flera sätt, genom undervisning, diskussioner och i möten med forskare. Nästa steg är att ta fram förslag på hur biblioteket kan organisera det fortsatta arbetet med stöd till forskningsdatahantering.*

## **Inledning och bakgrund**

Det talas alltmer om forskningsdata i samband med vetenskaplig kommunikation och att såväl öppna upp och sprida forskning som att göra forskningsprocessen mer transparent. I synnerhet har det aktualiserats internationellt genom EU:s senaste ramprogram Horisont 2020 (European Commission, 2015) och nationellt i samband med att ett förslag på nationella riktlinjer för öppen tillgång till forskningsresultat arbetats fram av Vetenskapsrådet (Vetenskapsrådet, 2015).

Forskningsdatahantering är inte något nytt, det har alltid varit en del av forskares arbete med att samla in och bearbeta material av olika slag. Precis som med forskningsresultat som presenteras i form av publikationer kan forskningsdata sägas vara en del av resultaten som kan beskrivas och tillgängliggöras för fler. Det är en underutnyttjad möjlighet som uppmärksammas alltmer i takt med att forskningen blir digital. Det är dock ännu tydligare än vid publicering att forskningsdata är något som är beroende av forskares praktiker i deras discipliner eller epistemiska kulturer (jmf Knorr Cetina, 1989; Borgman, 2015). Dessutom är det inte enbart forskare själva som är involverade i processerna kring datahanteringen i olika skeden. Olika stödfunktioner kan vara viktiga

för att hantera beskrivning, lagring, säkerhet, arkivering och tillgängliggörande, vilket också får betydelse för vad man benämner data. Begreppet data används på flera vis och varierar över tid beroende på när i processerna och vilka funktioner som är inblandade (Haider och Kjellberg, 2016).

Forskningsbiblioteken har under lång tid arbetat med hanteringen av publikationer och synliggörande och tillgång av dessa. Det finns flera aspekter på det arbetet som kan appliceras på forskningsdata, till exempel katalogisering och metadatabeskrivning, samtidigt som det är tydligt att det är ett ännu mer komplext område än vetenskaplig publicering. Formerna på data är långt ifrån homogena och delningskulturerna och hur man arbetar med sin empiri skiljer sig som sagt mycket åt (Borgman, 2015). De olika professionella rollerna som stödjer forskningsdatahantering behöver därför vidareutvecklas och förstärkas (Verbaan och Cox, 2014). Forskningsbiblioteken behöver diskutera de olika perspektiven på forskningsdata, utveckla kompetenserna som omgärdar hanteringen från bibliotekens håll och förankra arbetet tillsammans med de forskare som är aktiva på lärosätet (Tenopir et al, 2014; Nielsen och Hjørland, 2014). På Malmö högskola har vi under 2015 och 2016 fokuserat på bibliotekets roll i förhållande till forskningsdata och forskningsdatahantering på flera sätt. I följande artikel beskriver vi hur vi gått tillväga och de lärdomar och reflektioner som vi har i nuläget om detta sammansatta fält.

## **Omvärldsbevakning och kompetensutveckling**

I ett tidigt skede när vi diskuterade forskningsdata såg vi ett behov av att få en gemensam grund att stå på. Vi hade sedan en längre tid diskuterat forskningsdata på möten i teamet för Publicering och bibliometri och vi beslutade oss för att tillsammans kompetensutveckla oss genom att följa en kurs eller läsa och diskutera ett gemensamt material av något slag. Vi hittade snabbt den kurs som utvecklats vid universitetet i Edinburgh kallad Mantra (University of Edinburgh, 2015) som utöver moduler för forskare hade ett särskilt upplägg för bibliotekarier. Under senvåren och hösten 2015 genomförde vi de självstudier och träffar som Mantra erbjuder. Det innebar ett par timmars inläsning och uppgifter på egen hand kombinerat med diskussioner i hela teamet baserat på övningar och frågor där vi fördjupade oss i fem olika teman: datahantering, organisation och beskrivning av data, datalagring och säkerhet, etik och upphovsrätt samt delande av data. Även om kursen inte var direkt anpassad till svenska förhållanden kunde vi applicera våra diskussioner på redan befintlig kunskap i teamet om praxis i Sverige, bibliotekariekompetenser, forskarerfarenheter och kännedom om vårt eget lärosäte.

I samband med att vi förde diskussioner om forskningsdatahantering på Malmö högskola och bibliotekets roll identifierade vi en viktig ingång till forskningsdata som är en förlängning av redan befintlig verksamhet, nämligen att söka fram existerande forskningsdata som är öppet tillgänglig, beskriven och kan återanvändas. Två personer fick i uppgift att undersöka det området särskilt och utveckla en workshop som kunde genomföras under våren 2016. Tjugo bibliotekarier från olika sydsvenska lärosäten samlades och fick söka och arbeta med olika verktyg för att förstå mer om strukturerna

kring forskningsdata både i form av beskrivningar och av metadata som tillgängliggörande i vissa typer av system, både mer renodlade arkiv som arkivliknande lösningar i likhet med institutionella arkiv.

## Enkät till högskolans forskare

För att få en bättre bild av hur högskolans forskare ser på hantering av forskningsdata genomförde vi under senhösten 2015 en enkät. Enkäten skickades till alla professorer, lektorer och doktorander utifrån information i lönesystemet. Vi fick 161 svar, det vill säga en svarsfrekvens på 27 procent. Professorer besvarade enkäten i högre grad än övriga, kanske för att de reflekterat mer över forskningsdatafrågor i sin roll som forskningsledare. Ett engagemang i frågorna är synligt i kommentarerna som lämnats. Den aktuella bilden är att en majoritet av de som svarat använder öppen eller delad data, likaså delar en majoritet forskningsdata informellt, men bara enstaka har öppet tillgängliggjort data till exempel via ett dataarkiv. Mer kunskap och stöd efterlyses i alla aspekter kring öppen forskningsdata. I kommentarerna framkommer en krock mellan forskningsetiska och juridiska aspekter, till exempel skydd av data gentemot data som allmän handling. Bland svaren och kommentarerna i enkäten märks också en viss oro, både på temat "min data" och för mer administration som kommer att uppta forskningstid. Lagstiftningen är viktig, allt kan inte tillgängliggöras. Öppet tillgängliggörande sker först när data bedömts möjlig att göra öppen. En annan relevant aspekt som inte kommit upp i enkäten är om all forskningsdata är värd besväret att tillgängliggöra, vem kan ha nytta av den? (Fransson et al, 2016).

Tabell 1. Svaren på enkätfråga 17 om önskat stöd och kunskap.

Vad behöver du mer kunskap om alternativt stöd kring för att göra forskningsdatan som du arbetar med öppen?	Andel
Rättigheter och licenser kring öppna forskningsdata	57%
IT-stöd för till exempel databaslösning, lagring eller backup	47%
Tillgängliggörande av forskningsdata (allmänt)	44%
Tjänster för tillgängliggörande av forskningsdata inom ditt ämne	43%
Arkivering av forskningsdata	41%
Tidskrifters krav kring öppna forskningsdata	38%
Forskningsfinansiärers krav (och kommande krav) kring öppna forskningsdata	37%
Stöd kring datahanteringsplaner	35%
Praktiska frågor/stöd för tillgängliggörande av forskningsdata under 2016	31%
Stöd i databearbetning och organisering under forskningsprocessen	27%
Annat [fritextsvar]	11%

I tabell 1 visas den kanske mest centrala frågan för oss i vårt byggande av stödfunktion, de frågor forskarna behöver stöd i (hösten 2015). Svaren på frågan fick oss att inse att stöd behövs i samtliga frågor.

Enkäten till högskolans forskare var en del av ett pilotprojekt som också kortfattat berörde tjänster för tillgängliggörande och roller vid tillgängliggörande. Avrapportering gjordes till forskningsberedningen vid högskolan och till rektors ledningsgrupp. Båda

instanserna tog till sig resultaten och tackade biblioteket, men rapporten ledde inte fram till några beslut. Tillgängliggörande av forskningsdata är ännu ingen viktig fråga, men stödet kring forskningsdata behöver byggas upp i förväg. En av rekommendationerna i rapporten var att tillsätta en arbetsgrupp med uppdrag att arbeta fram en färdplan för stödet, och den rekommendationen kommer aktualiseras igen i det nuvarande projektet kring forskningsdata. Ska olika enheter som arbetar med stöd till forskningen kunna arbeta tillsammans med utveckling måste en färdplan tas fram.

I och med pilotprojektet insåg vi hur stort och komplext området är. Inte minst eftersom data som ska tillgängliggöras bygger på forskning som är komplexa. Vi kommer nog inte kunna hantera forskningsdata på samma sätt som publikationer som har en tydlig form. Insikterna ledde fram till ett fokus på intern kompetensutveckling under 2016. Ett nytt projekt skapades och i projektgruppen ingår även arkivpersonal. En central del i det lokala biblioteksprojektet är att delta i Svensk Nationell Datatjänsts pilotprojekt under 2016.

## **Deltagande SND:s pilotprojekt**

Svensk Nationell Datatjänst (SND) är en forskningsinfrastruktur som tillhandahåller en metadatakatalog och stöd till forskare i datahantering inom humaniora, samhällsvetenskap och medicin och hälsa. De har arbetat med forskningsdatahantering som en stödfunktion sedan i mitten av 80-talet och mycket fokus har varit enkätdata inom samhällsvetenskaplig forskning. Under 2016 har de satsat på att bredda sin verksamhet och också arbeta fram ett koncept där lärosätena blir bättre rustade för att ta hand om forskningsdata och i förlängningen tillgängliggöra data i SND:s söktjänst. I samband med detta kunde lärosäten anmäla intresse för att delta som piloter i utvecklingsarbetet. Med tanke på att vi redan utforskade forskningsdataområdet och ville gå vidare med att utforma det stöd som är lämpligt för Malmö högskolas forskare såg vi SND:s projekt som en möjlighet att ta del av deras expertis och kompetensutveckla oss vidare i dialog med några forskare från Malmö högskola.

### **Utbildning**

Första fasen i projektet tillsammans med SND handlade om kunskapsdelning. SND lade upp ett program med fem träffar där de olika temana var: introduktion till datahantering, metadata för forskningsdata, datahanteringsplan, verktyg och metoder för att dokumentera forskningsdata samt juridiska aspekter. Gruppen som gick utbildningen bestod av bibliotekarier, IT-tekniker och arkivarier. Några av deltagarna hade en dubbel funktion genom att de också är forskare själva. Upplägget var klassiska introduktioner eller föreläsningar till temat med kortare inslag av eget arbete eller diskussion i gruppen. Bredden bland deltagarna var både en fördel och en nackdel. Fördel eftersom olika perspektiv på samma fråga kunde belysas i våra diskussioner och nackdel eftersom förkunskaperna varierade kraftigt. Vårt behov efter att ha skapat den gemensamma basen kring forskningsdatahantering från Mantra handlade mycket om att förstå svenska förhållanden och särskilt vilka metadatastandarder som fungerar för att beskriva data och vad man bör tänka på i mötet med våra forskare.

## Mötet med forskare och forskningsdata

Mötet med forskare, eller kanske är det mer korrekt att kalla det mötet med forskningsdata, var nästa steg i projektet med SND. Vi hade från vår enkät några forskare som visat intresse för att få stöd i hanteringen av data. Vi tog kontakt med flera av forskarna och beskrev att vi skulle vilja träffa dem för att ta del av deras arbete med forskningsdata och att de i sin tur skulle få mer förståelse kring beskrivning och tillgängliggörande av data. Det var svårt att hitta kandidater som hade tid även om vi reducerade tiden för att träffas. En träff med en forskare i odontologi kunde till slut genomföras där två projektdeltagare intervjuade forskaren om dennes data i relation till metadatabeskrivningarna som finns i SND:s formulär för inlämning. Typen av data i detta fall var numerisk data insamlad under kontrollerade materialtekniska experiment. SND närvarade vid mötet för att kunna svara på specifika frågor. Vi kommer att genomföra ytterligare tre tillfällen med "möta forskare med forskningsdata", två med enkätdata och ett fall av intervjudata. Det är oerhört lärorikt att diskutera forskningsdata med forskare och vi kommer att arbeta vidare med olika former av "dataintervjuer" efter projektets slut.

## Läranderesan

Vi har hela tiden sett på forskningsdata som ett område under utveckling och att en första fas för vår del har varit att förstå de olika delarna som ingår i forskningsdatahantering. Vi insåg redan från start hur komplext forskningsdataområdet är och att det är väldigt långt ifrån "one size fits all", vilket också de juniora forskare som skrev ett inlägg till EU angående *Open Science*-arbetet poängterar (Global Young Academy, 2016). Det betyder att vårt angreppssätt också har varit väldigt tydligt inriktat på vårt eget lärande och ganska långt ifrån att implementera tekniska lösningar. Genom att skapa en gemensam bild av forskningsdatahantering och dess olika delar har också vår syn på bibliotekets roll gått igenom olika faser. Från att vi kanske tänkt mer i tekniska termer och liknat forskningsdata vid publikationer har en nyanserad karta skapats som även inkluderar ett slags angreppssätt där till exempel också studenter spelar en roll. Dessutom koncentrerade vi inledningsvis vårt arbete på tillgängliggörandet av data. Det kändes som en naturlig utveckling i relation till *open access* till publikationer och som en del av bibliotekets funktioner. Men det är inte den mest aktuella aspekten av forskningsdatahantering för våra forskare. Dels för att en stor del av forskningen vid Malmö högskola är samhällsinriktad där data berör till exempel elever eller patienter och därmed är svår att öppet tillgängliggöra. Dels för att det generellt inte finns några krav, varken från finansierare eller från lärosätet, att tillgängliggöra data. Därmed lägger man hellre tiden på publicering av forskningsresultat.

Det har dock blivit alltmer tydligt att det inte går att isolera forskningsdata som fenomen, det hör samman med många aktörer och funktioner inom ett lärosäte. En svårighet är hur integrerat det är med metoder och praktiker som kan vara väldigt lokala. Det ställs mot de politiska visioner som finns om hur forskningsdata ska kunna hanteras som mer generell och återanvändbar med liten förankring i hur vetenskapssamhället fungerar, till exempel i förhållande till meritering. Det ska dock inte

hindra forskningsbiblioteken från att ta sig an frågan, men det är oerhört viktigt att närma sig problemområdet från rätt håll och med bra frågeställningar.

Det har också märkts i våra diskussioner att forskningsdata är väldigt abstrakt som koncept men att det förrädiskt nog kan uppfattas som konkret. Särskilt när man rör sig på en mer visionär nivå då det är lätt att föreställa sig data som siffror eller text och något som lätt kan brytas ner och byggas om, men ju mer man pratar med forskare och diskuterar vad forskningsdata är desto mer differentierat blir det. Forskningsdata är hopflätat med forskarna och med deras discipliner (Borgman, 2015). Det är en utmaning för oss som arbetar på biblioteken att närma oss forskningsdata från flera håll och med en beredskap att stödja forskaren i dennes forskningsprocess.

Det är lätt att bara hamna i hur man lagrar stora datamängder, i arkivfrågan för framtiden eller öppen tillgång till forskning. Delarna i forskningsdatahantering är olika aktuell beroende på ämnestillhörighet. Hur gör vi för att gå fram i lagom takt tillsammans med forskarnas behov? I vissa forskningsfält vill man kunna ställa nya frågor till material som är insamlat och inte studerats på alla de upptänkliga sätt som är möjligt, till exempel arkeologiska utgrävningar. I andra ämnen vill man göra metaanalyser nu, till exempel i vissa medicinska områden, i andra handlar det om att återanvända och kombinera andra discipliners empiri för nya frågor, i åter andra handlar det om att kontrollera tidigare studier och undvika fusk. Öppen tillgång till forskningsdata ligger på det sättet långt ifrån öppen tillgång till forskningsresultat i form av en text oavsett om det är en bok, artikel eller något annat format. Det som krävs av återanvändaren av en publikation är läsbarheten som bland annat ligger i genren för texten, disciplinär kunskap och allmän läsförståelse. I jämförelse kan forskningsdata finnas i mängder av upptänkliga fil-format, kräva olika slags programvara för att avläsas, beskrivas med olika typer av metadatastandarder beroende på disciplin och behöver på olika sätt avkodas för att göras användbar igen av någon annan än den som ursprungligen utförde forskningen. Beskrivningen av tillgängliggjord data i form av metadata är viktig för hittbarheten så att den kan återanvändas, vilket är en central bibliotekariekompetens. Dokumentationen av tillgängliggjord data är central för återanvändbarheten, men här är frågan vilket ansvar exempelvis biblioteket ska ta och vilket ansvar som ska läggas på forskaren. Att tillgängliggöra data på ett basalt sätt är enkelt, men att se till att den både blir hittbar och återanvändningsbar med begränsade resurser är en utmaning.

## Nästa etapp

Vi i team publicering och bibliometri upplever nu ett behov av att gå utanför teamet. Forskningsdatafrågan har mognat och kommer sakta involvera allt fler på Malmö högskolas bibliotek som en ny aspekt på befintliga arbetsuppgifter. I och med våra ökade kunskaper kring hantering av forskningsdata så har vi förstått att stöd kring forskningsdata är flera frågor för biblioteket och inte bara en enda. Hanteras den som en enda fråga så blir arbetet snabbt väldigt komplext. Delar av stödet kring forskningsdata har stora likheter med vårt befintliga arbete med registrering i vår publiceringsdatabas och med open access-arbetet. Undervisande bibliotekarier kan komma att undervisa i

hur man hittar och citerar data på samma sätt som de idag hanterar publikationer. Men att stödja forskare i hantering av data under forskningsprocessen är en ny och delikat utmaning. Hantering och bearbetning av forskningsdata är en central del av forskningsprocessen och därmed svårare att närma sig som utomstående än frågor kring publiceringen av forskningsresultat.

Nästa steg på vår resa är att skissa på hur bibliotekets stöd kring forskningsdata kan organiseras. Detta beror delvis på interna förutsättningar, men också på övriga stödfunktioners engagemang i forskningsdatahanteringen. Stödet till forskningen belyses även i andra projekt vid högskolan och som gör att bibliotekets roll tillsammans med övriga stödfunktioner troligen bli tydligare.

Vi kommer åka på studieresa till universitetet i Edinburgh för att ta del av deras erfarenheter av stöd rörande forskningsdata. I Edinburgh finns sedan flera år ett *Research Data Management Programme*. Inom universitetet samverkar flera enheter för att detta ska bli verklighet. En av dessa är universitetsbiblioteket. De har och har haft flera projekt finansierade av JISC som handlar om forskningsdatahantering på olika vis, nu senast om att etablera en säker tjänst för lagring, det så kallade *Data Vault*-projektet. I *Research Data Management Programme* ingår även *Digital Curation Centre (DCC)* som är placerat i Edinburgh och är en nationell stödfunktion med expertis inom forskningsdatahantering för Storbritanniens högre utbildning och forskarsamhälle. De ger expertråd om lagring, hantering, skydd och delande av digitala forskningsdata.

Att börja jobba med forskningsdata är en lång resa för biblioteket. På flera sätt kan forskningsdata ses som en ny dimension i det befintliga arbetet med publikationer. Ur ett annat synsätt kanske det blir ett paradigmskifte för biblioteket. Biblioteket kan komma att stödja mer i själva forskningsprocessen och inte bara stödja publicering och återanvändning av forskningsresultaten i form av publikationer. Oavsett vilket kommer det bli en spännande resa.

## Om författarna

*Sara Kjellberg, filosofie doktor i biblioteks- och informationsvetenskap, arbetar som utvecklare av forskningsstöd vid Malmö högskolas bibliotek. Forskar inom formell och informell vetenskaplig kommunikation, forskares informationspraktiker och forskningsdata.*

*Jonas Fransson, PhD i biblioteks- och informationsvetenskap, arbetar som utvecklare av forskningsstöd vid Malmö högskolas bibliotek. Forskar inom informationssökbeteende och informationssystem och -ekologier, med ett speciellt intresse för relationen mellan hittbarhet och användning.*

## Referenser

Borgman, C. L. (2015). *Big data, little data, no data: scholarship in the networked world*. Cambridge, Massachusetts: The MIT Press.

European Commission (2015). *Guidelines on Data Management in Horizon 2020*.

Hämtad:

[https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)

Fransson, J. , du Toit, M. , Tapia Lagunas, P. & Kjellberg, S. (2016). *Tillgängliggörande av forskningsdata : nulägesbeskrivning samt rekommendationer för uppbyggnad av stöd till högskolans forskare*. Malmö högskolas bibliotek. <http://hdl.handle.net/2043/20266>

Global Young Academy (2016). *Position Statement on Open Data by the Young Academies of Europe and the Global Young Academy*.

<https://globalyoungacademy.net/wp-content/uploads/2016/04/Position-Statement-on-Open-Data-by-the-Young-Academies-of-Europe-and-the-Global-Young-Academy.pdf>

Haider, J., & Kjellberg, S. (2016). Data in the making: Temporal aspects in the construction of research data. In J. Rekers & K. Sandell (Eds.), *New Big Science in focus: Perspectives on ESS and MAX IV*. Lund: Lund Studies in Arts and Cultural Sciences.

<http://portal.research.lu.se/portal/files/3051583/8852248.pdf>

Knorr-Cetina, K. D. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press.

Nielsen, H.J., & Hjørland, B. (2014). Curating research data: The potential roles of libraries and information professionals. *Journal of Documentation*, 70(2), 221–240. doi:10.1108/jd-03-2013-0034

Tenopir, C., Allard, S., Birch, B., & Sandusky, R. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library And Information Science Research*, 36(2), 84-90. doi:10.1016/j.lisr.2013.11.003

University of Edinburgh (2015). *MANTRA: Research Data Management Training*. <http://datalib.edina.ac.uk/mantra/>

Verbaan, Eddy & Cox, Andrew M. (2014). Occupational sub-cultures, jurisdictional struggle and third space: Theorising professional service responses to research data management. *Journal of Academic Librarianship*, 40(3–4), 211–19.

Vetenskapsrådet (2015). *Förslag till Nationella riktlinjer för öppen tillgång till vetenskaplig information*. <https://publikationer.vr.se/produkt/forslag-till-nationella-riktlinjer-for-oppen-tillgang-till-vetenskaplig-information/>