

PÅ VÄG MOT ETT FORSKARSTÖD KRING HANTERING OCH TILLGÄNGLIGGÖRANDE AV FORSKNINGSDATA

Slutrapport från biblioteksprojektet
Tillgängliggörande av forskningsdata 2016

Jonas Fransson
Sara Kjellberg
Karin Sedvall
Ola Tengstam
Pablo Tapia Lagunas
Madeleine du Toit
Jenny Widmark
Tomas Holmström
Hanna Höie



Innehållsförteckning

1. Projektbeskrivning, resultat och rekommendationer	3
2. Om rapportens struktur och innehåll.....	12
3. Definitioner av begrepp	14
4. Forskningsdata – ett biblioteks läranderesor i de internationella, nationella och lokala landskapen.....	15
5. Poster och paper: Developing integrated research data management support in close relation to doctoral students’ research practices.....	22
6. Poster presenterad vid IFFIS 2016	30
7. Vägen från policy till praktik – ett studiebesök i Edinburgh universitetsbiblioteks forskningsdatauniversum	31
8. SND-pilotprojekt - beskrivning, erfarenheter och slutsatser	35
9. Sammanfattning av journal club	38
10. Beskrivning av Lärosäten Syd-workshoppar.....	39
11. Beskrivning av undervisning för kriminologistudenter	40
12. Beskrivning av doktorandworkshop	41
13. Metadata för forskningsdata - krav och ambitioner	42
14. Arkivets synpunkter.....	46
15. Om lagring av forskningsdata.....	47
16. Samlade referenser	49

1. Projektbeskrivning, resultat och rekommendationer

Bakgrunden till projektet är Vetenskapsrådets förslag kring öppen tillgång till forskningspublikationer och forskningsdata (Vetenskapsrådet, 2015) inför forskningspropositionen 2016 (Prop. 2016/17:50). Biblioteket genomförde ett pilotprojekt under 2015 (Fransson et al, 2016), och detta projekt är en fortsatt fördjupning i ämnet tillgängliggörande av forskningsdata. För en kort översikt om öppet tillgängliggörande av forskningsdata i de nordiska länderna se Nordforsk rapport (Nordforsk, 2016).

Projektbeskrivning

Projektets syfte och mål

Projektets syfte har varit att öka förståelsen för och kunskapen om hur biblioteket kan stödja forskare i hantering av forskningsdata. Biblioteket arbetar med beskrivning i form av metadata vid katalogisering, likaså att tillgängliggöra samlingar via söksystem. Båda verksamheterna behöver fördjupas i förhållande till forskningsdata. Biblioteket har också en pedagogisk roll inom informationshantering (medie- och informationskompetens). Denna roll behöver utvecklas när det gäller forskningsdata, både för tillgängliggörande av data från högskolan och användning av externa forskningsdata i högskolans forskning och undervisning (Malmö högskolas bibliotek, 2016).

Samtidigt har Malmö högskola deltagit i Svensk nationell datatjänsts (SND) pilotprojekt under 2016, vilket syftat till att testa en ny modell för SND, överföra kunskap till lärosätenas stödfunktioner, samt skapa möjliga modeller för organiserandet av lokal stödverksamhet.

Formulerade mål i projektdirektivet (Malmö högskolas bibliotek, 2016) är att vid projektavslut finns:

- Ett par tänkbara modeller för bibliotekets stödverksamhet gällande tillgängliggörande av högskolans forskningsdata, samt grundläggande information om resursåtgång för stödet.
- Kunskap och kompetens som är tillräcklig för ett basalt stöd från biblioteket.
- Grundläggande informations- och undervisningsmaterial kring användande och tillgängliggörande av forskningsdata.

Genomförande av projektet

Projektet startade den 1 mars och avslutades 1 december. De första månaderna fokuserades på deltagande i första fasen av SND:s pilotprojekt med utbildningsträffar i Lund (se kapitel åtta om SND:s pilotprojekt).

Projektgruppen sattes samman genom projektdeltagarnas olika roller och kompetenser, från bibliotekets båda avdelningar, Medier och vetenskaplig publicering (MVP) och Pedagogisk verksamhet och lärandemiljö (PVL). Rollerna/kompetenserna som eftersöktes var:

- Arbete med open access och registrering av forskningspublikationer
- Handledning och undervisning av studenter
- Upphovsrätt
- Katalogisering och metadata
- Forskarstöd

Fokus för projektet har varit projektgruppens kompetensutveckling rörande hantering av forskningsdata med en inriktning på tillgängliggörande av data. Inom ramen för projektet har en mängd aktiviteter genomförts, vilka presenteras på olika sätt i rapporten (se kapitel två om rapportens struktur och innehåll).

Projektresultat och diskussion

Tillgängliggörande av forskningsdata

Forskningsdata kan öppet tillgängliggöras på olika sätt, exempelvis:

- deposit in a specialist data centre, archive or thematic repository
- deposit in an institutional repository
- submitting to a journal to support a publication
- publish in a data journal
- dissemination via a project or institutional web site
- self-publish via a cloud-based system such as figshare (Corti et al, 2014, p. 197)

De olika tillvägagångssätten har olika förutsättningar för hittbarhet, kurering och långsiktig tillgång till tillgängliggjorda data. Genom mer strukturerade sätt att tillgängliggöra, som via ett dataarkiv eller en forskningsdatatidskrift (*data journal*), så blir datasetet i större grad synligt i traditionella söksystem. Öppet tillgängliggjorda data kan, om beslut tas, registreras i MUEP, högskolans publikationsdatabas, för ökad hittbarhet oavsett hur den är tillgängliggjord. Publicering på webbplats eller i specifika webbtjänster kan på kort sikt vara attraktiva lösningar, men hittbarheten kan begränsas till webbsökmotorer och långsiktigt omhändertagande av materialet blir i stort sett omöjligt för stödfunktioner (Fransson et al., 2016).

Ambitionsnivå när det gäller metadata är en central fråga som är kopplad både till resursåtgång, till tillgängliggörande i externa söksystem och i viss grad till återanvändbarhet. Att kontrollera och berika metadata, i detta fall ursprungligen registrerad av forskare, är en självklar del av tillgängliggörandet av forskningsdata för att metadata ska hålla en hög kvalitet. Däremot är ambitionsnivån något som måste diskuteras, speciellt i relation till SND:s föreslagna nationella infrastruktur (SND, 2016) och resursåtgång lokalt.

En annan fråga är ambitionsnivå när det gäller dokumentationen av tillgängliggjorda forskningsdata. För att data ska vara återanvändningsbar behövs beskrivningar utöver de metadata som krävs för till exempel arkivering eller söksystem. Antingen kan utförlig dokumentation medfölja data eller så kan data beskrivas i en dataartikel, där fokus ligger på insamlande, täckning och metod. Den kritiska frågan är vems ansvar detta är, forskarens eller forskarstödet?

I pilotprojektet som genomfördes hösten 2015 drogs slutsatser kring tillgängliggörande (Fransson et al, 2016, s. 33):

Idag kan öppet tillgängliggörande av forskningsdata ske genom externa tjänster. Det finns i nuläget inget skäl att upprätta en tjänst i form av dataarkiv för tillgängliggörande av forskningsdata vid MAH. På kort sikt kan eventuellt forskningsdata tillgängliggöras via publikationsdatabasen MUEP. På längre sikt kan troligen forskningsdata tillgängliggöras via lösningen för elektronisk arkivering (e-arkivet). Enskilda finansiärer eller tidskrifter kan komma att kräva att forskningsdata ska tillgängliggöras på specifika platser även i fortsättningen, även om det finns en institutionell lösning.

Tillgängliggörande i externa tjänster rekommenderas tills vidare. En lokal lösning för tillgängliggörande av data behöver baseras på en omfattande lokal behovsanalys. Att tillgängliggöra eller arkivera forskningsdata kräver även ett visst stöd i forskningsprocessen, till exempel kring dokumentation av forskningsdata, samt utåtriktad verksamhet med informationsmöten och utbildningstillfällen för forskare. Likaså behövs information och stöd rörande datahanteringsplaner inför nya forskningsprojekt och ansökningar.

Utåtriktat arbete

Utifrån projektet och speciellt projektgruppens deltagande i SND:s pilotprojekt rekommenderas att stödet kring hantering av forskningsdata inte fokuserar på enskilda faser, till exempel ansökningsfasen eller slutfasen med arkivering och öppet tillgängliggörande, utan att högskolan erbjuder ett stöd till forskningen före, under och efter forskningsprocessen. Minst tydligt är kanske behovet av stöd under forskningsprocessen då det idag inte matchar stödfunktionernas verksamhet med stöd kring ansökningar, arkivets fokus på arkivering och bibliotekets fokus på tillgängliggörande av data. Ett visst stöd under forskningsprocessen finns idag, till exempel it-lösningar för lagring och erbjuda programvaror för analys (SPSS, NVivo, m.fl.), men ingen användarsupport. Vid aktivt arbete med datahanteringsplaner så krävs till exempel stöd kring datadokumentation och utbildning om datahantering. Ökad dokumentation är också en förutsättning för att forskningsdata ska kunna arkiveras och/eller tillgängliggöras på ett meningsfullt sätt. Ett sätt är att skapa ett datalabb, en plats för stöd i forskningsprocessen med fokus på databehandling och -hantering samt visualisering och metodutveckling.

Även om projektet har fokuserat på tillgängliggörande av forskningsdata så har ett av spåren handlat om hur kunskaperna förmedlas och kommuniceras med berörda aktörer såsom exempelvis forskare eller studenter. De pedagogiska ansatserna har varit av två slag i projektet. En ansats riktad till forskare i form av samtal med forskare under faserna 2 och 3 i SND:s pilotprojekt där stödenheternas perspektiv förklarades. Forskaren fick ta del av hur tillgängliggörande i externa dataarkiv kan gå till och vilka metadata som är central för sökning. Den andra ansatsen handlar om att hitta tillgängliggjorda data. Detta har i projektet riktats till bibliotekspersonal genom workshops samt ett undervisningstillfälle för studenter. Våren 2017 planeras också en workshop för doktorander, som kommer att kombinera tillgängliggörande och att hitta tillgängliggjorda data.

Ett exempel på hur man arbetar med öppna data finner man på Digital Social Science Lab (DSSL) på Köpenhamns universitetsbibliotek. Förutom att erbjuda programvaror för analys och arbete med data beskrivs DSSL som "an intelligent space for education and events focusing on digital tools in social sciences" (Københavns Universitetsbibliotek, 2016). Förutom att erbjuda olika typer av programvaror för databehandling hjälper man här också till med att hitta dataset. Att hitta och ge tillgång till material är något som bibliotekarier arbetar med och denna kompetens kan utvecklas för en ny informationstyp som dataset (Lauersen, 2016).

Arbetet med att hitta och återanvända data kräver pedagogiskt utvecklingsarbete. Detta gäller inte minst om vi ska vända oss mot studenter. All undervisning måste förankras i en kontext och anpassas efter ämne och studenternas förkunskaper, och detta gäller i högsta grad för undervisning om forskningsdata. Precis som annan undervisning i grundutbildningen är det viktigt med ett nära samarbete med undervisande lärare. Syftet med att använda öppna data måste formuleras och studenternas förtrogenhet med olika program för analys av data är en fråga som måste undersökas vidare. I dagsläget erbjuds inte heller något samlat stöd för exempelvis arbete med statistik.

Oavsett hur vi utvecklar vårt stöd för forskningsdata kommer de pedagogiska ansatserna att vara fortsatt viktiga. Mycket arbete kommer att behöva läggas på kommunikation och undervisning av olika slag och detta kommer att innebära många verksamheter i biblioteket. Det behövs en allmän kompetensutveckling för alla bibliotekarier. Kontaktbibliotekarier beskrivs ofta som väl lämpade att sprida kunskap och medvetande om forskningsdatastöd, då de både har kontakter på fakulteterna och kunskap inom ämnet (Krier & Strasser, 2014; Pryor, 2012). På Malmö högskolas bibliotek avskaffade vi kontaktbibliotekarier och har nya former för vårt kontaktarbete där undervisande och handledande bibliotekarier spelar en stor roll.

Ökat tillgängliggörande av forskningsdata kan också komma att förändra vad man inbegriper i medie- och informationskompetenser. Vi behöver utöka vår kunskap som handlar om att hitta och citera publikationer till att också gälla data förankrat i användarnas situation och behov. Det gäller såväl studenter som forskare och doktorander. Stödet måste grundas i de lokala behov våra användare har.

Framtida organisering vid biblioteket samt för stödet vid Malmö högskola

Idag ansvarar team publicering och bibliometri för arbetet med forskningsdatahantering inom biblioteket. Flera saker påverkar hur det fortsatta arbetet med stöd till hantering av forskningsdata bör organiseras: forskningspropositionens formulering om att det främst kommer att handla om att tillgängliggöra forskningsdata som ligger till grund för en publikation; möjligheterna för stödet att få en resursförstärkning vid högskolan; i vilken grad stödet vid högskolan hålls samman eller de olika aspekterna i forskningsdatahantering fördelas på till exempel ansökningsprocessen, säker lagring under forskningsprocessen och arkivering/tillgängliggörande; om konsortiet lett av SND kommer till stånd och om Malmö högskola deltar; i vilken grad stödet blir forskardrivet eller baserat på uttalade behov.

Det kan finnas en fara att i för stor grad utgå ifrån bibliotekets och övriga stödverksamheters befintliga verksamhet och organisation som utgångspunkt för utvecklande av stödfunktioner för forskningsdata. Risken finns att fokus inte blir att fylla verkliga behov för forskningen i tillräckligt stor utsträckning. Oavsett hur vi organiserar arbetet internt på biblioteket kommer det behövas ett forum med representanter från de olika inblandade funktionerna, jämte representanter från forskningen för att hantera det samlade stödet när det gäller forskningsdata. Representanterna bör ha mandat att arbeta strategiskt med utvecklingen av stödet kring hantering av forskningsdata före, under och efter forskningsprocessen (se även rekommendation i Fransson et al, 2016). Detta kan utformas på olika sätt antingen som en ny grupp eller att frågan placeras inom en redan befintlig grupp där forskningsdata kan passa för att diskuteras på tvärs mellan olika stödfunktioner. En möjlighet kan vara en grupp under forskningsberedningen.

En naturlig fortsättning på forskningsdataprojekten vid biblioteket vore att dedicera resurser till utvecklingsarbetet och skapa ett arbetslag. Arbetslaget bör knytas nära bibliotekets publiceringsstöd då verksamheterna än så länge är likartade med ett troligt fokus på datahanteringsplaner och tillgängliggörande utifrån finansieringskrav. Utåt sett bör stöd till hantering av forskningsdata och publiceringsstöd (inklusive OA) inte sammanföras då det är skilda frågor för forskarna. Sammansättning av arbetslaget bör sättas i relation till bibliotekets prioriteringar rörande datahanteringsstöd, det vill säga vilka delar som biblioteket satsar på under 2017.

Ur ett användarperspektiv kan en samlad forskningsdatastödenhet vara ett bättre alternativ än ett spritt stöd på många specialiststödfunktioner om man ser till hela högskolan. Det avgörande för hur stödet bör organiseras är i vilken grad forskarpraktikerna kommer ändras utifrån kommande krav, djupgående förändringar kan kräva ett samlat stöd kring datahantering, medan ytligare förändringar kan hanteras i ett decentraliserat stöd. Ett alternativ vore en gemensam ingång till stödet (en *first line*-funktion) som svarar på enklare frågor och som fördelar övriga till rätt stödinstant. Biblioteket kan stå för en första samlad ingång till stödfunktionerna (*first line*) för att hänvisa vidare vid behov till rätt expertis (*second line*). Var en sådan *first line* placeras i verksamheten beror på hur vi hanterar andra frågor in till biblioteket. Det finns redan idag funktionsadresser, en gemensam e-postadress för allmänna biblioteksfrågor och användning av ärendehanteringssystem. Inledningsvis är möjligen antalet frågor inte så många att de behöver tas om hand i alla kanaler utan ett tydligt synliggörande av arbetslaget är tillräckligt.

I nuläget kan vi ge stöd till enskilda forskare som vill tillgängliggöra data via externa tjänster. För ett storskaligt tillgängliggörande krävs troligtvis dels ett lokalt system där arkivering och tillgängliggörande kan ske i samma process, dels krävs gallringsbeslut för forskningsdata, ställningstagande rörande statistiksekretessens räckvidd, samt tydliga riktlinjer för dokumenthantering inklusive forskningsdata.

I nedanstående tabeller visas olika inriktningar på en utbyggnad av stödet. Indelningen är gjord för att tydliggöra möjligheten att göra olika strategiska val, till exempel så är det möjligt att satsa på att utveckla undervisningen gentemot studenterna (nr 4) samtidigt som man avvaktar med utvecklandet av stöd för forskarna. Den uppskattade resursåtgången ska bara ge en grov bild av hur projektgruppen uppfattat tidsåtgången utifrån erfarenheterna i projektet och den troliga utvecklingen inom området med allt fler krav på tillgängliggörande av forskningsdata.

Tabell 1. Utveckling av stöd, generella ansatser.

Nummer	1	2a	2b	2c
Benämning	Grundläggande stöd	Utåtriktad ansats	Stöd i forskningsprocessen (datalabb)	Fördjupat tillgängliggörande
Erbjudet stöd	datahanteringsplaner, tillgängliggörande i externa tjänster + metadata i MUEP, hitta forskningsdata, dataintervjuer	informera och utbilda forskare och doktorander, skapa informationsmaterial (tryckt och digitalt), guider och riktlinjer, självstudiematerial	dokumentera data, bearbeta data, visualisering, analysverktyg (t.ex. Nivio, SPSS), referenshantering, datahantering inkl lagring	tillgängliggörande på fördjupad nivå (ej metadata på miniminivå), tillgängliggörande via lokalt system eller lokalt e-arkiv?
Förutsätter		1 samt gärna 2b eller 3b	1	1
Resursåtgång	1 heltidstjänst	+1 heltidstjänst	+1 till 3 heltidstjänster	+1 till 2 heltidstjänster
Aktörer	biblioteket	biblioteket	biblioteket, it, m.fl.	biblioteket, arkivet
Kommentarer			expansion av Medieverkstan?	

Tabell 2. Utveckling av stöd, specifika ansatser.

Nummer	3a	3b	4	5
Benämning	Data Access Unit (DAU) enligt SND	Data Access Unit (DAU) (utan stöd 2)	Undervisa studenter	Forskningsarkivarie
Erbjudet stöd	tillgängliggöra/lagra i nationell tjänst enligt SND:s riktlinjer, även variabelnivå och utökad dokumentation	tillgängliggöra/lagra i nationell tjänst enligt SND:s riktlinjer, även variabelnivå och utökad dokumentation + utbilda forskare och ge stöd i forskningsprocessen kring dokumentation och metadata	pedagogisk utveckling med ämnesanpassning	En arkivarie inriktad mot att stödja hela forskningsprocessen samt bevaka arkivmetadata-standarder för forskningsdata.
Förutsätter	1 och delar av 2a & 2b	1	1	
Resursåtgång	+2 heltidstjänster	+4 heltidstjänster	+minst 1 heltidstjänst	+1 heltidstjänst
Aktörer	biblioteket	biblioteket, it	biblioteket	Arkivet
Kommentarer		2a och 2b finns ej		

Under livscykeln för arbetet med forskningsdata finns olika specialistroller. Dessa roller innefattar specifik kompetens, i jämförelse med kompetensen och rollen bibliometriker krävs djup kunskap inom ett komplext verksamhetsområde. Två förekommande specialistroller är databibliotekarie och *data science librarian*, där den sistnämnda ger stöd i forskningsprocessen rörande databearbetning och dokumentation (se 2b i Tabell 1 ovan). Databibliotekarier kan arbeta mer utåtriktat med undervisning och handledning, men också arbeta med tillgängliggörande i form av metadatabeskrivning (se även Fransson et al, 2016).

Viss del av arbetet kan förläggas i anslutning till bibliotekets undervisning. Kompetensutveckling är dock besvärligt eftersom området både är brett och komplext, och en förutsättning är att man har en grundförståelse för den generella forskningsprocessen och gärna någon specifik forskningspraktik. I projektgruppen har det varit tydligt att medlemmarna med forskarutbildning har en helt annan förståelse än icke-disputerade projektdeltagare. Viss kunskap behövs på bredden inom hela biblioteket. Kompetensutveckling bör därför ske över tid inom hela organisationen för ökad förståelse.

Slutsatser och rekommendationer

Att förstå vilket stöd man behöver ge och vilka krav man kan ställa måste vara grundat i en förståelse för forskningsprocessen och i en fördjupad behovsanalys. Både kompetensutvecklingen och stödet som utvecklas bör vara kopplade till den lokala forskningens behov, och inte en generell stödverksamhetsdiskurs, där man automatiskt kopiera hur stödverksamhet utvecklas i omvärlden (Krier & Strasser, 2014; Pryor, 2012). Vi har också förstått att det är svårt att grunda arbetet med forskningsdatahantering i mjuka policies som inte följs upp, men att dessa kan användas som redskap i dialog med forskare som påbörjar sitt arbete med forskningsdata. Fokus måste därför vara bredare än att beskriva och tillgängliggöra forskningsdata i olika system. Det handlar om ett samarbete mellan flera olika funktioner på ett lärosäte.

Det är viktigt att dimensionera tilldelade resurser efter vald ambitionsnivå. Forskningsdataområdet kommer att fortsätta kräva omfattande utvecklingsarbete. Kompetensutveckling är fortsatt en nödvändig prioritering samt att det eventuellt kan kräva rekrytering av specifika kompetenser. Samordning med övriga stödfunktioner inom högskolan är av största vikt för att få ett samlat stöd för de olika aspekterna på forskningsdatahantering före, under och efter projekt. Allra viktigast är dock att förankra arbetet med forskningsdatahantering i de discipliner och deras behov som finns på Malmö högskola.

Föreliggande projekt är ett biblioteksprojekt och därför riktas nedanstående rekommendationer till bibliotekets ledningsgrupp. De rekommendationer som rör verksamhet utanför biblioteket antas bibliotekets ledning ta upp i lämpliga sammanhang.

Projektet rekommenderar att Malmö högskolas bibliotek:

#1 Inrättar ett arbetslag under team publicering och bibliometri på biblioteket

På biblioteket bör arbetet 2017 organiseras i ett arbetslag underordnat team publicering och bibliometri med avsikt att klara av det grundläggande stödet och omfattar minst en heltid uppdelat på minst två personer. Om arbetsinsatserna och resurserna senare utökas kan arbetslaget omformas till ett eget team (eventuellt SND:s motsvarande DAU) vid behov. Arbetslagets uppgifter är relaterade till de prioriteringar vi identifierat i projektet samt arbetslagets resurssättning. Arbetslagets kompetenser bör bestå av djup kunskap om forskningsprocessen och forskningspraktiker, kunskap om datahanteringsplaner samt en specialistnivå på kunskaper vad gäller forskningsdataområdet som helhet.

#2 Satsar på fortsatt kompetensutveckling för hela biblioteket

I takt med att forskningsdata blir ett nytt verksamhetsområde för högskolebiblioteken så behövs en allmän förståelse och kunskap hos alla anställda på biblioteket, vilket kräver kompetensutveckling över tid. Fokus bör i första hand vara på att hitta och återanvända data i linje med annan informationssökning och användning av informationsresurser.

#3 Arbetar med stödet för forskningsdatahantering kopplat till alla delar i forskningsprocessen

Arbetslaget arbetar med det grundläggande stödet som relaterar till alla delar i forskningsprocessen det vill säga omvärldsbevakning av policies samt arbete med: datahanteringsplaner; hitta forskningsdata; dataintervjuer; tillgängliggörande i externa tjänster samt eventuellt metadata i MUEP (se tabell 1). Undervisning av doktorander är prioriterat och relaterar till samordningen av doktorandstöd i arbetslag FU. Biblioteket bör också verka för en utbyggnad av stödet under forskningsprocessen, kanske i form av ett datalabb. Stöd kring analysverktyg, databearbetning, dokumentation och liknande.

#4 Gör en fördjupad behovsanalys med hjälp av dataintervjuer

Det är viktigt att diskutera vad forskningsdata är för olika vetenskapliga discipliner. Forskare vill veta vad de kan få hjälp med, inte hur forskningsdataprocessen modelleras. Dataintervjuer är ett bra redskap för att skapa mer ämnesbaserad kunskap och en förståelse för de lokala behoven.

#5 Tar fram kriterier för att bedöma externa tjänster för tillgängliggörande av data

Det finns ett antal externa tjänster forskare kan använda sig av redan idag för deponering av forskningsdata och tillgängliggörande. Riktlinjer bör utarbetas som stöd för val av externt dataarkiv i likhet med kriterier som finns som stöd för val av tidskrifter eller förlag vid publicering.

#6 Arbetar vidare med kriminologi som fallstudie för undervisning om forskningsdata

Arbeta vidare med kontakterna med kriminologi för att inte förlora momentum när det gäller studentundervisning samtidigt som det inte kan prioriteras högt. Stöd till studenterna ingår inte i arbetslagets primära uppgifter utan är förlagt till lärarlaget HUM/SAM i avdelningen PVL med samarbete med arbetslaget för forskningsdata.

#7 Medverkar i samordningen för högskolans stöd kring hantering av forskningsdata

Forskningsdatahantering är ett komplext område, därför bör stödet vara samordnat mellan olika stödfunktioner. Arbetet måste vara förankrat och resurssatt både inom forskningen och av administrationen. Samordningen bör ske i en form som synliggör forskningsdata samtidigt som det ligger i linje med hur övrig stödverksamhet organiseras. En viktig aspekt är de etiska och juridiska frågorna som kräver insatser från högskolejurist och forskningsetisk expertis.

#8 Medverkar i nationella samarbetsprojekt

Mycket utveckling sker parallellt just nu inom biblioteksvärlden och nationellt. För att bättre utnyttja resurser och erfarenhetsutbyte är det viktigt med ett ökat nationellt samarbete mellan forskningsbibliotek som ligger långt fram i utvecklingen av stödverksamhet kring forskningsdata. Under 2017 bör åtminstone ett samarbetsprojekt initieras med ett eller flera bibliotek utöver att följa den nationella utvecklingen. En möjlighet är att söka utvecklingsmedel från biblioteksforeningen höstterminen 2017.

#9 Medverkar i utredning av e-arkiv med avseende på tillgängliggörande av forskningsdata

I forskningspropositionen hänvisas till att Vetenskapsrådet i samråd med Riksarkivet ska hjälpa till att utveckla lärosätenas arbete med långsiktigt bevarande och arkivering av forskningsdata. Att införa ett e-arkiv på högskolan är prioriterat för Gemensamt verksamhetsstöd under 2017. Under 2018 kommer den nya dataskyddsförordningen att träda i kraft vilket också har betydelse för hur lärosätena tar hand om data. Tillgängliggörande genom användning av e-arkiv är en möjlig väg som behöver diskuteras i samband med övriga e-arkivfrågor.

#10 Deltagande i forskningsinfrastrukturen Svensk Nationell Datatjänst (SND)

Vid deltagande i SND:s kommande forskningsinfrastruktur, the Swedish Research Data System (SND, 2016), bör erforderliga resurser tilldelas stödverksamheten. Dessa bör utgå ifrån de krav som ställs på beskrivningsnivån (dokumentation och metadata) (jmf ambitionen som beskrivs i SND, 2016). Lokal styrgrupp bör utses, där högskolans representant i SND:s referensgrupp ingår.

Referenser

Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2014). *Managing and sharing research data : a guide to good practice*. Los Angeles: SAGE.

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>

Krier, L., & Strasser, C. A. (2014). Starting a new service (Chapter 2), in *Data Management for Libraries : A LITA Guide*. Chicago: ALA TechSource.

Laurson, Christian (2016). Libraries curating open data sources. *The library Lab* [Blogg] 25 augusti <https://christianlaurson.net/2016/08/25/libraries-curating-open-data-sources/>) Hämtad 2016-12-02

Københavns Universitetsbibliotek (2016). Digital Social Science Lab. <http://kub.kb.dk/DSSL> Hämtad 2016-12-02

Malmö högskolas bibliotek (2016). Projekt Tillgängliggörande av forskningsdata 2016. Projektdirektiv, version 1.0. <http://mah.se/Bibliotek/Om-Biblioteket/Projektkontor/Pagaende-projekt/Tillgangliggorande-av-forskningsdata-2016/>

Nordforsk (2016). *Open Access to Research Data – Status, Issues and Outlook. NordForsk Policy Paper 1 –2016.* https://www.nordforsk.org/en/publications/publications_container/open-access-to-research-data-2013-status-issues-and-outlook/download

Prop. 2016/17:50. *Kunskap i samverkan – för samhällets utmaningar och stärkt konkurrenskraft.* <http://www.regeringen.se/rattsdokument/proposition/2016/11/prop.-20161750/>

Pryor, G. (2012). Roles and responsibilities: libraries, librarians and data (Chapter 6), in *Managing Research Data*. London: Facet Publishing.

SND (2016). RFI Application SND ver 313, 2016-12-06.

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/>

2. Om rapportens struktur och innehåll

Denna rapport är utplagd som en antologi. Upplägget är ett sätt att illustrera innehållet i projektet då det i stor utsträckning har varit ett kompetensutvecklingsprojekt. Projektet har resulterat i olika aktiviteter och texter som presenteras kapitelvis. Nedan följer en kortfattad läsanvisning för innehållet i respektive kapitel.

Kapitel 1 - Projektbeskrivning, diskussion, slutsatser och rekommendationer

I kapitel ett har projektets resultat sammanfattats samt presenteras projektets rekommendationer.

Kapitel 2 - Om rapportens struktur och innehåll

Sammanfattande läsanvisningar.

Kapitel 3 - Definitioner och begrepp

Centrala begrepp och koncept beskrivs och definieras för ökad läsförståelse av resterande delar.

Kapitel 4 - Forskningsdata – ett biblioteks läranderesor i de internationella, nationella och lokala landskapen

Artikel som beskriver arbetet med och utgångspunkterna för forskningsdata på Malmö högskola fram till och med sommaren 2016. Publicerad i InfoTrend i december 2016.

Kapitel 5 - Paper och poster från ASIS&T AM16

Postern bygger på ett short paper som också är inkluderat i detta kapitel. Innehållet är en diskussion om riktade insatser för doktorander angående att inkludera stöd för forskningsdatahantering i deras forskarutbildning. Poster presenterad vid ASIS&T Annual Meeting i oktober 2016.

Kapitel 6 - Poster presenterad vid IFFIS 2016

Postern är en infografik över forskningsprocessen och beskriver ett möjligt stöd från biblioteket under 2017. Postern presenterades vid IFFIS (Infrastruktur för forskningsinformation i Sverige) i november 2016.

Kapitel 7 - Vägen från policy till praktik – ett studiebesök i Edinburgh universitetsbiblioteks forskningsdatauniversum

Artikel om erfarenheter och inspiration angående forskningsdatahantering vid Edinburghs universitetsbibliotek. Artikeln är baserad på ett studiebesök i Edinburgh i oktober 2016 med stöd från Biblioteksföreningen. Manuskriptet är inskickat till Revy i december 2016.

Kapitel 8 - SND-pilotprojekt - beskrivning, erfarenheter och slutsatser

Sammanfattning av upplägget för Malmö högskolas deltagande i SND:s pilotprojekt och nyvunna kunskaper och erfarenheter baserat på vår medverkan.

Kapitel 9 - Sammanfattning av journal club

Vid tre tillfällen har projektgruppen haft dedikerade möten där projektet diskuterats utifrån i förväg utvalda och lästa artiklar/bokkapitel. Fokus har inte varit texterna i sig utan frågeställningar som uppkommit i projektet. I kapitlet listas vilka texter som diskuterades.

Kapitel 10 Beskrivning av Lärosäten Syd-workshoppar

Vid två tillfällen har workshoppar anordnats för bibliotekspersonal vid Lärosäten Syd-biblioteken. Fokus har varit hur man hittar tillgängliggjorda forskningsdata.

Kapitel 11 - Beskrivning av undervisning för kriminologistudenter

Ett undervisningstillfälle för studenter har genomförts inom projektets ramar. Texten beskriver upplägg och erfarenheter av att undervisa studenter om att hitta forskningsdata.

Kapitel 12 - Beskrivning av doktorandworkshoppar

Ett förslag på upplägg av kommande workshop om forskningsdatahantering riktad till doktorander. Upplägget är baserat på att passa in i bibliotekets övriga struktur för doktorandstöd.

Kapitel 13 - Metadata för forskningsdata - krav och ambitioner

Beskrivning av och argumentation för projektets rekommendationer för metadatabeskrivning av forskningsdata.

Kapitel 14 - Arkivets synpunkter

Högskolans två arkivariater har varit adjungerade i projektgruppen. Texten är en sammanfattning av arkivets perspektiv på forskningsdata som har beröringspunkter med tillgängliggörande av forskningsdata.

Kapitel 15 - Om lagring av forskningsdata

Sammanfattning om lagring och säkerhet för Malmö högskola grundat på tillgänglig information på webben och samtal med IT-avdelningen

Kapitel 16 – Referenser

Samlad litteraturlista med allt material som använts i projektet.

3. Definitioner av begrepp

Arkivering av forskningsdata – forskningsdata blir vid skapande eller upprättande allmän handling vid högskolan (myndigheten) och ska därför arkiveras enligt högskolans dokumenthanteringsplan. Ett så kallat e-arkiv planeras på högskolan för arkivering av digitala handlingar. Arkivering är ännu skild från lagring och tillgängliggörande.

Dataarkiv – används här för tekniska lösningar för tillgängliggörande av data. Det tar inte hänsyn till långtidsbevarande och lösningarna är inte arkiv i arkivlagens mening. Används istället för datarepositorie.

Datahanteringsplan – dokument där hanteringen av forskningsdata inom ett forskningsprojekt beskrivs. Bör vara ett aktivt dokument under forskningsprocessen som även inkluderar arkivering och tillgängliggörande av data.

Dataintervju – intervju med forskare om hantering av forskningsdata, dennes behov av stöd och identifiering av problem. Kan jämföras med bibliotekets traditionella referensintervju där fokus är att förstå användarnas informationsbehov och hur det kan uppfyllas genom bibliotekets resurser.

Dokumentation – beskrivning av dataset för ökad återanvändbarhet i form av dokument eller relaterade forskningspublikationer, till exempel i metodavsnitt.

Forskningsdata – insamlade data (primära data) eller återanvända data (sekundära data) som ligger till grund för forskningsresultat. Kan benämnas på olika sätt inom forskningsämnen, exempelvis empiriskt material, källmaterial eller statistik.

Hantering av forskningsdata – begrepp som beskriver insamlande, hantering, dokumentation, lagring, arkivering, tillgängliggörande, etcetera av forskningsdata under och efter forskningsprocessen. Högskolans stödfunktioner bistår forskare i hanteringen av forskningsdata.

Interoperabilitet – två eller flera systems förmåga att utbyta information, samt att använda denna information.

Kurering av forskningsdata – att säkerställa och förbättra användbarheten hos data efter deponering i öppet dataarkiv eller lokalt e-arkiv, till exempel genom att berika metadata eller byta filformat.

Metadata – beskrivning av dataset i formulär där varje fält kopplas till en viss typ av administrativa, deskriptiva eller tekniska metadata. Idag registreras metadata för forskningspublikationer i MUEP.

Tillgängliggörande av forskningsdata – publicering av forskningsdata i öppna dataarkiv, på webbplats eller som datapublikation. Syften kan vara möjlighet till reproduktion av forskningsresultat eller ny forskning, men också innovation och demokrati nämns ibland.

4. 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 behövs. 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 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 sensvåren 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 bibliotekarierna, 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 numeriska 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 öppna 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 tillgängliggjorda data i form av metadata är viktig för hittbarheten så att den kan återanvändas, vilket är en central bibliotekariekompetens. Dokumentationen av tillgängliggjorda 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.

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

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/>

5. Poster och paper: Developing integrated research data management support in close relation to doctoral students' research practices

Poster presenterad vid ASIS&T Annual Meeting 2016 i Köpenhamn den 16-18 oktober, samt paper publicerat i ASIST 2016 Proceedings of the 79th ASIS&T Annual Meeting Volume 53, 2016. Se MUEP för pdf-filer: <http://hdl.handle.net/2043/21502>



Developing integrated research data management support in close relation to doctoral students' research practices

Jonas Fransson Malmö University Library, Sweden jonas.fransson@mah.se	Sara Kjellberg Malmö University Library, Sweden sara.kjellberg@mah.se
Pablo Tapia Lagunas Malmö University Library, Sweden pablo.tapia@mah.se	Madeleine du Toit Malmö University Library, Sweden madeleine.du.toit@mah.se

ABSTRACT

The quest for open research data is the driving force behind the development of the whole area of research data management practices. We, as a university library, offer and develop support to researchers and doctoral students. Based on the result of a web survey submitted to all researchers at Malmö University, and the knowledge that doctoral students are on their way of forming their individual research practices, we have made doctoral students our first target group for specific seminars and workshops promoting conscious research data management practices. We will organise these seminars and workshops, which both take into account the general aspects of research data management and the discipline specific practices, so as to develop integrated research data management support in close relation to doctoral students' research practices.

KEYWORDS

Research data management, training, RDM-support, doctoral students, research practices

INTRODUCTION

There are several reasons for the increased interest in making research data openly available. From the political level and from some research funders the cause is primarily driven as a question of resources: the data produced or collected by public funds is seen as a public resource that may be more beneficial to society (OECD, 2007). Another angle is the quality of the research and that it should be easy to review (Corti et al, 2014). Lastly, it would be helpful for those researchers who want to be open and receive feedback during the research, for example as part of open notebook science (Amsen, n.d.). In this poster we discuss how research libraries can get an understanding for and develop their role in relation to these different motives. Particularly important is the focus on the local conditions and the research practices of different disciplines.

BACKGROUND

Status of RDM support in Sweden

Research Data Management (RDM) is working with research data throughout the research process, from planning and making research grant application to finishing and archiving. The largest part of the management responsibility falls to the scientists during their research work (Corti et al, 2014). RDM practices vary between the institutions, disciplines and groups of researchers (Borgman, 2015), and in Sweden no comprehensive approach has been taken on a national level for the best practice

in the management of research data. In addition, universities lack solutions for electronic archiving and only in some disciplines have research data been made openly available.

In the field of RDM there are many stakeholders. Funders push the question of open research data; some publishers have policies concerning research data, publications of descriptions of open data to come (data journals), national data centres promote skill development, educational institutions create and operate data repositories in parallel with their publication repositories, as well as develop various types of support to researchers. In the European Union (EU) there is a willingness on the political level to increase the accessibility of research data and some EU-funded projects have had stipulations about data processing and data availability (European Commission, 2015; OECD, 2007). It is in this evolving field that researchers, research groups and institutions must navigate in order to meet current and future requirements.

Research and support to researchers at Malmö University

Malmö University was founded in 1998 and it is a young, modern and international university. It has about 24,400 students and offers undergraduate and postgraduate education. At the university there are about 500 professors (including Associate Professors) and more than 200 PhD students. Malmö University is a state-governed institution accredited by the Swedish Higher Education Authority.

Malmö University conducts multi-disciplinary research, aiming to contribute to the development of a sustainable society in order to meet the challenges of the future. As a result, our research is frequently pursued in collaboration with partners from outside the university. The researchers cooperate with the business community, with the public sector and with non-profit organisations in various areas such as research on social innovation; sustainable city-planning; biofilm and biological interfaces; and the influence sports have on our society and on public transport of the future.

The university is organised into five faculties:

- Faculty of Culture and Society
- Faculty of Education and Society
- Faculty of Health and Society
- Faculty of Odontology
- Faculty of Technology and Society

All the faculties are multidisciplinary. The range of disciplines spans from history studies to biomedicine.

Support to researchers is organised on different levels – some is close to the departments and some is brought together on a centralised level, for example the grants office and communications office. At the library, research support services constitute core services like access to information resources, information seeking support and guidance, publishing and bibliometric support as well as support in using digital tools in the knowledge production, for example reference managers.

In a study by Tenopir et al. (2014), the most common support from libraries were finding and citing datasets, which can be seen as connected to other types of information seeking support already part of the librarians' area of expertise. However, a range of support services related to research data was identified as part of, or planned for, libraries in the study, such as preparing data management plans and the preparation and description of research data to be deposited in repositories or archives (Tenopir et al., 2014). Nielsen and Hjørland (2014) highlight the potentials for the libraries to play a role in curating data but they also question whether libraries will take the lead in some areas, pointing out the need for very domain specific knowledge in order to select, organise and use

research data. This shows that RDM needs a shared effort and that RDM support to researchers demand various expertise and a range of support functions that are involved in the management of research data: in making it available or archiving it and in long-term preservation (see also Verbaan and Cox, 2014). At Malmö University this means that the archive, the library, the IT department, the university lawyer and the information security officer need to work together.

The quest to develop RDM support

The Malmö University Library's mission includes monitoring and disseminating information on the development of scientific publishing and research information, including RDM. The goal is to develop adequate support in higher education and, in particular, to identify what the library's support to university researchers should contain.

The library's interest in research data management is based on the development of the library's current activities. To describe the research data in the form of catalogue records or metadata is a key part of the library's activities, as well as making available different types of collections. Support for the publication has become an integral part of the research support, and it is partly linked to publishing open access. Accessibility and findability via search systems are central to the reuse of the publications. The same processes apply to the making available of data, and research issues are partially interlinked.

FROM A GENERAL UNDERSTANDING OF RDM ISSUES TO AN INTEGRATED SUPPORT – WORK IN PROGRESS

Our starting point was to increase the knowledge of and understanding for the different aspects of a full-fledged RDM support. As presented above, it is a new area in the library sector overall and to be able to develop the services in relation to researchers, librarians at Malmö University need a common ground. The common store of knowledge was obtained by following an online course developed by University of Edinburgh (University of Edinburgh, 2014).

Learning more about research data and the related aspects, from law to metadata standards and safety issues, made clear that the disciplinary differences and research practices also have to be included in the development of support for the future. From a general knowledge of the phases of research data management a deeper knowledge on a domain specific level has to be provided for.

PhD students are on their way to become researchers following an education at the same time as they are performing research. A project at the library has analysed what needs the PhD students have and RDM was found to be one area needing strengthening, as the PhD students' different practices have to be taken into consideration (Gullbekk, Rullestad, Calvo, 2013). Additionally, to learn more about the specific needs of Malmö University's researchers, a survey of the current state of handling research data as well as attitudes concerning RDM was conducted. The PhD students were treated as faculty researchers in the survey. The results showed the diversity of research practices between the disciplines at Malmö University.

A web survey

A questionnaire was sent to all university researchers (n=601) in December 2015. The response rate was 27% (n=164). A commitment to the issues is visible in the comments submitted. The current state is that a majority of respondents are using open or shared data. In addition, a majority informally shares research data, but only a few have made data openly available, via for example a data archive.

The survey tool Sunet Survey (Survey & Report from Artologik) was used and the survey, as well as the invitation, was bilingual (Swedish/English). Two direct-addressed mailings were made: an invitation to the survey as well as a reminder six days later.

Survey questions

Ten survey questions were designed to cover different aspects of research data to give an overview. At the same time, the goal was to keep the number of questions in the survey down and thus make it easy to answer in order to increase the response rate. In addition to issues surrounding the research data, there was background and follow-up questions to provide the answers with a context.

Results

The answers have shown the broad character of research at the university. Collection methods are plentiful and the respondents often used both quantitative and qualitative methods. A majority see themselves as very central in the data collection on all types of methods, even on typical quantitative methods. It allows one a close connection to the research data at hand. Many have been using open or shared data. 66 per cent have shared research data, mainly informally, but only six per cent have made data available through a data centre or data archive.

In order to make the research data you are working with open, in what areas would you need more knowledge or more support?	Share
Rights and licenses for open research data	57%
IT support for e.g. database solution, storage or backup	47%
Making research data available (general)	44%
Services for making available research data within your subject	43%
Archiving of research data	41%
Journal requirements regarding open research data	38%
Current and coming demands from research financiers on open research data	37%
Support concerning data management plans	35%
Practical questions/support for making available research data in 2016	31%
Support in data processing and organising data during the research process	27%
Other	11%

Table 1. The need of RDM knowledge or support (n=161).

As for the support or the knowledge that is lacking, more than half of the respondents state rights and licenses surrounding open research data, but all possible answers are chosen in to a fairly large extent (Table 1). The conclusion is that we need more information and knowledge about all aspects of making data and about data management at large.

Some of the comments in the survey show the range of opinions in regards to managing research data and making it available openly. One researcher says:

I like the idea and am open to it if ethical.

And another states quite the opposite:

The idea is alien to me.

A third think it is problematic:

When working qualitatively, it is deeply problematic, because the material is strongly connected to the research questions and probably very difficult to interpret for others who do not have the

same ontological and epistemological inputs. Moreover, it is ethically questionable, and therefore no option for my research.

The current situation demands various strategies to meet the researchers' needs in their practices as they are situated in their disciplines. To start with, we have decided to address doctoral students in particular as they are forming their personal research practice and have not yet formed research data management habits.

Approaching doctoral students

We have identified the doctoral students as a target group for the introduction of RDM. They are not only part of a discipline; they are socialised into a cultures that differ for each research area (Peixoto, 2014). But, at the same time, there are some general themes in RDM, and we will attempt to balance the general with the specific for each research area.

The strategy is divided into two parts. The first part is to get to know the field of RDM by way of a seminar on the subject and a workshop on how to find relevant datasets for the individual dissertation project. The second part is to address existing PhD student courses on research methods (quantitative/qualitative) or ethics and integrate elements of RDM in these courses. In Sweden, a PhD student has to take a number of PhD student courses as part of their dissertation project – some courses may be mandatory and others are optional.

As shown by the answers to the survey question in Table 1, there are several RDM issues to address. Potential workshops for the doctoral students are:

- Finding, re-using and citing open research data. Extending practices of searching and citing to include data, and promoting the use of open data as a data source.
- Active data management plans (DMP) as a personal information management tool (PIM) (e.g. Jones, 2008) for increased control, efficiency and quality, for example based on Michener's ten rules for creating a good data management plan (Michener, 2015).
- Structured practices for organising research data during the research process, including data security, naming files, documentation and backup practices.
- Depositing research data to local archives and publishing datasets for dissemination, including data preparation and a final DMP.

As a complement to this strategy containing seminars and workshops, we will give personal support in and supervision of RDM matters. We will develop workshops and other activities continuously. During autumn, we will approach the PhD students taking a research ethics course and offer them a workshop introducing RDM.

Conclusion

The goal is to base the RDM support on the actual needs of the research in the different research practices at Malmö University. No two research practices are the same, as the research areas at the university range from humanities and design to odontology and biomedicine. This presents us with a challenge: How do we support the whole range of doctoral students? The answer must be that we, the staff of the RDM-support, have to have a good general understanding of the different research practices and how to communicate with doctoral students and supervisors in different fields of research.

We will start by focusing on doctoral students as one identified target group, primarily divided into groups based on faculty, in order to address different research practices. Differences have become

clear from the survey about research practices. The need for support has to be realised in relation to distinct research practices and the different phases that PhD students are in. This integrated support will both broaden and deepen as we as give RDM support to the PhD students throughout their dissertation projects.

The situation today with very little formal RDM training and few researchers disseminating research data in data archives is a far cry from the goals of the EU where “all” publicly funded data are open. Our work in developing this integrated research data management support in close relation to doctoral students’ research practices is a small step, but it is an important one for our PhD students and their RDM practices. Our new experiences will increase the knowledge, which will be nationally shared, about how to address PhD students on data management topics. We will also investigate how we, as a library, can work together with other support functions, e.g. IT-department.

REFERENCES

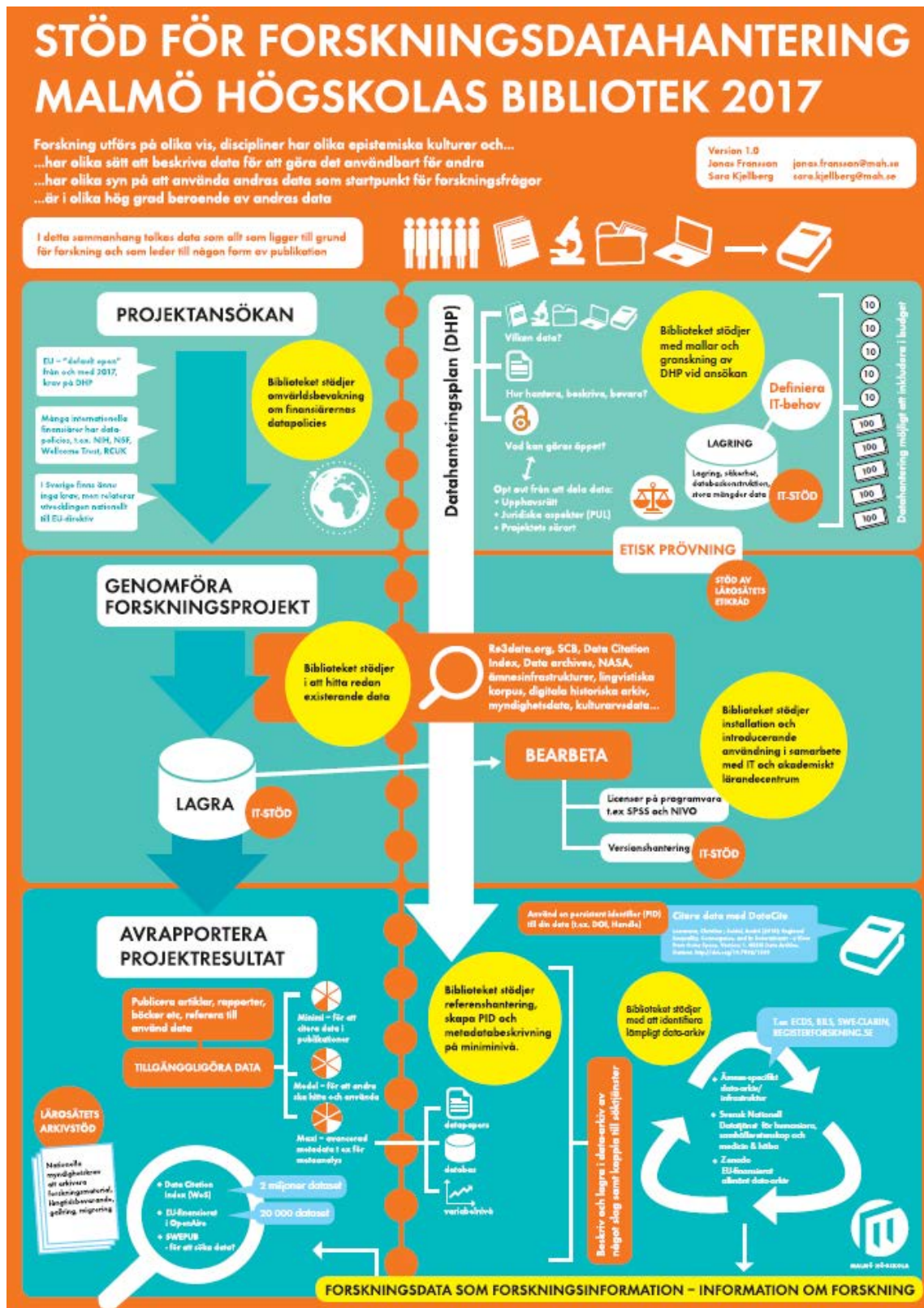
- Amsen, E. (n.d.). *Guide to open science publishing*. F1000Research. Retrieved from <http://blog.f1000research.com/2015/03/06/guide-to-open-science-publishing/>
- Borgman, C. L. (2015). *Big data, little data, no data : scholarship in the networked world*. Cambridge, Massachusetts: The MIT Press.
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2014). *Managing and sharing research data : a guide to good practice*. Los Angeles: SAGE.
- European Commission (2015). *Guidelines on Data Management in Horizon 2020*. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- Gullbekk, E. Rullestad, T. & Calvo, M.C.T. (2013). *PhD candidates and the research process. The library's contribution*. Oslo: University of Oslo.
- Jones, W.P. (2008). *Keeping found things found: the study and practice of personal information management*. Amsterdam: Morgan Kaufmann Publishers.
- Michener, W. K. (2015). Ten simple rules for creating a good data management plan. *Plos Computational Biology*, 11(10).
- Nielsen, H. J & Hjørland, B. (2014). Curating research data: the potential roles of libraries and information professionals. *Journal of Documentation*, 70 (2), 221-240.
- OECD (2007). *Principles and Guidelines for Access to Research Data from Public Funding*. OECD Publications Centre. Retrieved from <http://www.oecd.org/sti/sci-tech/38500813.pdf>
- Peixoto, A. (2014). *De mest lämpade: en studie av doktoranders habituering på det vetenskapliga fältet.[Survival of the fittest – A study of doctoral students’ habituation on the scientific field]* Diss. Göteborg : Göteborgs universitet. <http://hdl.handle.net/2077/35675>
- Tenopir, C., Sandusky, R. J., Allard, S. & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*. 36 (2), 84–90.
- University of Edinburgh (2014). *MANTRA Research Data Management Training*. Retrieved from <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*. Retrieved from <https://publikationer.vr.se/produkt/forslag-till-nationella-riktlinjer-for-oppen-tillgang-till-vetenskaplig-information/>

6. Poster presenterad vid IFFIS 2016

Poster över forskningsprocessen och möjligt stöd från biblioteket under 2017. Postern presenterades vid IFFIS (Infrastruktur för forskningsinformation i Sverige) den 7 november 2016 i Stockholm. Se MUEP för pdf-version av postern: <http://hdl.handle.net/2043/21517>



7. Vägen från policy till praktik – ett studiebesök i Edinburgh universitetsbiblioteks forskningsdatauniversum

Sara Kjellberg, Karin Sedvall och Pablo Tapia Lagunas, Malmö högskolas bibliotek

Det har skrivits mycket litteratur de senaste åren, dels om forskningsdata generellt och dels om hur stödet kring forskningsdatahanteringen byggs upp, inte minst på universitetsbiblioteken. Utmaningarna är många, från de juridiska frågeställningarna till den faktiska hanteringen av olika typer av data. Vilka specialistkompetenser som är centrala för en framgångsrik verksamhet och hur dessa roller samordnas inom universitet är något som Edinburghs universitetsbibliotek arbetar med. Bland annat har biblioteket en policy för forskningsdata sedan 2011 och vi var nyfikna på hur de arbetar med att stödja sina forskare med forskningsdatahantering. På Malmö högskolas bibliotek pågår under 2016 ett projekt för tillgängliggörande av forskningsdata. En studieresa till Edinburgh hösten 2016 kunde genomföras med ett resestipendium från Svensk biblioteks förening.

Under de fem år som gått sedan policyn fastställdes har biblioteket byggt upp en stödverksamhet med databibliotekarier, tagit fram system för lagring och tillgängliggörande och arbetat aktivt med att sprida kunskap och kompetens om forskningsdata ut på universitetet. De delade med sig av sina erfarenheter, vilket gav oss viktiga insikter och inspiration om hur en stödverksamhet för forskningsdata på ett lärosäte kan utvecklas från policy till praktik.

Databibliotekarier, dataintervjuer och data science

I Edinburgh finns flera databibliotekarier med kompetenser inom skilda områden och som arbetar med stöd genom hela livscykeln för forskningsdata. Till exempel erbjuder de stöd för skapande och granskning av datahanteringsplaner, tillgängliggörande och återanvändande av forskningsdata, men forskarna kan också konsultera bibliotekarierna om de behöver leta upp särskilda typer av data och bearbeta och analysera den. En stor del av forskningsdatahanteringen handlar om data curation, det vill säga att behandla, vårda och förvalta, data på olika nivåer. Bland annat kan forskare få hjälp med att beskriva och tillgängliggöra dataset i universitetets institutionella dataarkiv, Datashare. Dessutom erbjuds stöd för att arbeta med beskrivningen av dataset ner på variabelnivå för statistisk analys och återanvändning.

En betydande del av arbetet är de utåtriktade insatser som sker för att sprida kunskap, medvetenhet och intresse för forskningsdata bland universitetets forskare. Detta sker främst genom workshops och besök ute i verksamheterna. I diskussionerna framkom hur viktigt det är med hjälp när frågor och problem uppstår, vilket är svårt att planera i för tid.

”They want it when they need it, you can’t provide it in advance.”

För att möta detta behov måste stödet samlas och byggas på forskarnas behov och sätt att uttrycka detta. Dessutom har det alltmer tydliggjorts att det krävs olika upplägg i olika discipliner, vilket kanske inte är förvånande. Bara ordet data skapar ofta förvirring, till exempel känner forskare inom vissa ämnen sig inte alls bekväma med att prata om sin empiri som data. I Edinburgh har de nyligen valt att skapa en gemensam struktur för allt stöd kallat Research data service. En viktig poäng som

lyfts fram är att forskare inte vill få förklarat för sig hur livscykeln för data ser ut genom olika modeller eller diagram:

”They don’t want to study the data work, they want to do and manage their research.”

Från diskussionerna med Edinburghs databibliotekarier tar vi därför med oss att utgå från användarnas egna uttryck för vad de vill göra före, under och efter ett projekt med sin data.

En viktig aspekt för att kunna öka kontakten med forskarna om forskningsdata är att sprida arbetet och kunnandet om ämnet till fler medarbetare inom stödverksamheterna. I Edinburgh har ett sätt att göra detta varit att låta ämnesbibliotekarier (liaison librarians) genomföra den onlinekurs, Mantra, som de utvecklat. I Edinburgh satsade de dessutom särskilt på att alla bibliotekarier som går kursen ska hålla dataintervjuer. Deras erfarenhet är att dataintervjuerna synliggör mycket av den tysta kunskapen som forskarna besitter om sin empiri och att det tydliggör för bibliotekarierna vilka frågor som kan uppstå när man ska hantera forskningsdata för att beskriva och tillgängliggöra den. Detta är något som vi själva erfarit i liknande intervjuer kring forskningsdata som vi genomfört under vårt projekt. Förhoppningen är att ämnesbibliotekarierna med deras nära kontakt med fakulteter och forskare ska kunna få en bättre förståelse för forskningsdatahantering och vara mer involverade i att stödja sina forskare. Det finns dock en distans mellan att arbeta konkret med forskningsdata som databibliotekarie och med annat forskarstöd på biblioteket. Delvis handlar det om kompetenser som behöver byggas upp och fördjupas för alla bibliotekarier verksamma på forskningsbibliotek. Forskningsdata är trots allt ett relativt nytt verksamhetsområde för oss att ta oss an i biblioteksverksamheten.

I Edinburghs arbete handlade det initialt mycket om själva lagringen av data. I takt med att allt fler systemlösningar har kommit på plats såväl hos dem själva lokalt som på andra håll till exempel EUs olika initiativ, blir det andra delar av forskningsdatahanteringen som blir alltmer tydliga. Nyligen genomförde de Data and software carpentry workshops, det vill säga handfasta aktiviteter i hur man använder olika mjukvaror för statistisk analys, mjukvaru- och databashantering. Det är tydligt att forskningsdatahantering inte stannar vid enkla metadatabeskrivningar på bibliografisk nivå. En sådan utveckling kring databibliotekarierollen går mot data science librarian, vilket i princip betyder att kunskaper om dataanalys behöver finnas inom biblioteket.

Utmaningarna

”The researchers take this as an admin job.”

Den kanske största utmaningen är avsaknaden av en forskardriven agenda för forskningsdata. Många forskare ser det fortfarande som ett administrativt arbete. Inom biblioteksverksamheter som arbetar med att bygga upp forskningsdatakompetenser behöver vi fler fallstudier från olika discipliner och där dataintervjuerna är ett utmärkt redskap. I den forskningsdatapolicy som används i Edinburgh, och de exempel som finns från finansärer, är det fortfarande mjukt formulerade krav med liten eller ingen uppföljning av dem. Möjligheten ligger därför mer i att föra en diskussion kring policyn och tolkning av vad kraven innebär just för den data eller empiri som ska samlas in i ett särskilt projekt än i konkret uppföljning av om man följer policyn.

Att arbeta nära fakulteterna och forskarna tas också upp som recept på framgång och det är viktigt att hitta eldsjälarna eller så kallade data champions, inom fakulteterna. Utvecklingen av ett nytt stöd måste få lov att ta tid och stödet bör utformas i nära samarbete med dem som ska använda det. Vissa svårigheter som uppstod i det inledande arbetet i Edinburgh berodde på att de fick prata om exempelvis lagringstjänster som ännu inte fanns på plats. Där ser situationen annorlunda ut för oss som ska utveckla forskningsdatastöd idag.

För att kunna möta omvärldens nya krav vad gäller hantering och tillgängliggörande av forskningsdata behövs både nya funktioner, kompetenser och samarbeten på tvärs. Även om de i Edinburgh har lång erfarenhet och god kompetens inom området uttrycker de svårigheter med att förankra arbetet på universitetet. Biblioteken har en roll att spela tillsammans med andra stödverksamheter som IT, arkiv, kommunikationsavdelning och grants office. De utmaningar och vägval som forskningsbiblioteken står inför är många. I och med de nya EU-kraven om forskningsdatahanteringsplaner från och med 2017 blir frågan aktuell för de forskare som söker medel därifrån. Biblioteken kommer med stor sannolikhet att bli inblandade i att stödja och erbjuda hjälp vid tolkningen av de policies och krav som utvecklas just nu.

Databibliotekarier har många skilda uppgifter och vilka av dessa som vi på Malmö högskola ska fokusera på är en fråga som just nu diskuteras flitigt. Behöver vi möta upp behovet av statistiksupport i likhet med de diskussioner som finns i Edinburgh just nu kring att ha en helpdesk för statistikfrågor? Samtidigt är det viktigt att komma ihåg att statistisk data ju bara är fråga om en typ av data, den så kallade fyrkantiga och kvantitativa. Det finns en mångfald av empiri i olika discipliner. Från bibliotekets håll kan vi identifiera att det finns en lucka att fylla som gäller att stödja forskningsdatahantering före, under och efter forskningsprojekt och att behovet av stöd kommer öka, men var det ska organiseras inom lärosätet är en större fråga. Även om biblioteken skulle satsa enbart på ett tillgängliggörande perspektiv, att stödja beskrivning och lagring i system som Datashare, skulle det inte räcka med mer kunskap och förståelse om metadata. Det kommer också att krävas förståelse för förutsättningarna för att återanvända, bearbeta och använda data grundat i forskarnas ämnestillhörighet och olika forskningspraktiker.

Tack till Tony Mathys, Cuna Ekmekcioglu, Stuart MacDonald, Robin Rice, Laine Ruus och Dominic Tate som tog sig tid att prata med oss under ett intensivt och lärorikt studiebesök i oktober 2016.

Tio tankar om forskningsdata-stöd

1. viktigt att diskutera vad forskningsdata är för olika vetenskapliga discipliner
2. svårt med mjuka policies som inte följs upp men använd dem för dialog
3. dataintervjuer är ett bra redskap för att skapa mer ämnesbaserad kunskap
4. forskare vill veta vad de kan få hjälp med, inte hur forskningsdataprocessen modelleras
5. fokus måste vara vidare än att beskriva och tillgängliggöra forskningsdata i olika system
6. flera olika funktioner på ett lärosäte måste samarbeta
7. rikta insatser mot forskarstuderande som har möjlighet att arbeta in nya rutiner i sin forskning
8. bra att ha med sig tydliga exempel på data champions som kan beskriva hur de arbetar
9. Single Point of Contact för stödet underlättar för forskarna
10. ha system och visst stöd på plats innan det marknadsförs

Lästips:

Research Data Management Training, MANTRA (online course):

<http://datalib.edina.ac.uk/mantra>

Edinburgh Research data policy:

<http://www.ed.ac.uk/information-services/about/policies-and-regulations/research-data-policy>

Data Carpentry & Software Carpentry workshops

<http://datablog.is.ed.ac.uk/2016/08/25/carpentry-workshops/>

Upplägg kring dataintervjuer från Edinburgh:

<http://datalib.edina.ac.uk/mantra/libtraining.html#profiles>

Om Malmö högskolas forskningsdataprojekt:

Fransson, J. och Kjellberg, S. (2016). Forskningsdata – ett biblioteks läranderesä i de internationella, nationella och lokala landskapen. InfoTrend, vol 69, nr 1/2.

Tillgänglig: <http://www.sfis.nu/ojs/index.php/infotrend/index> (kommer inom kort)

Fransson, J. och Kjellberg, S. (2016). Stöd för forskningsdatahantering, Malmö högskolas bibliotek 2017, poster på IFFIS, Stockholm 7 november:

Tillgänglig: <http://dspace.mah.se/handle/2043/21517>

8. SND-pilotprojekt - beskrivning, erfarenheter och slutsatser

Beskrivning av SND:s pilotprojekt

Bakgrunden till projektet är att SND ska utveckla sin roll mot ett nationellt kunskapscentrum för datahantering och skriva en ansökan till Vetenskapsrådet till utlysningen om infrastrukturer under 2017.

"Syftet med pilotprojektet är att utveckla en modell för att hantera de utmaningar som svenska lärosäten och SND kommer att ställas inför i samband med de ökande kraven på tillgängliggörande och långtidsbevarande av forskningsdata i Sverige. Denna modell baseras på forskarnas behov och omfattar arbetsformer för överföring av kompetens inom datahantering och -kurering från SND till de funktioner (Forskningsstödsenheter, FSE) på lärosätena som i framtiden kommer att arbeta med dessa uppgifter för att leva upp till kraven på tillgängliggörande av forskningsdata." (SND, 2016a, s. 1)

Efter genomgången utbildning ska deltagarna (SND, 2016b, s. 1):

- Ha kännedom om olika typer av forskningsdata och filformat
- Förklara innebörden av metadata, metadataprofiler och metadatastandarder
- Kunna redogöra för väsentliga delar i en datahanteringsplan och stödja andra i arbetet med att utforma en sådan
- Känna till verktyg och metoder för att dokumentera data
- Känna till riktlinjerna för Open Access och veta hur forskningsdata kan tillgängliggöras för sekundärforskning
- Kunna värdera vilka metadata som krävs för att data ska vara användbara för sekundärforskning
- Vara bekanta med grundläggande juridiska aspekter som påverkar hantering av forskningsdata
- Vara kompetenta att ge råd kring ovanstående områden på en grundläggande nivå.

Pilotprojektet har haft fyra faser:

- Fas 1: SND-ledd utbildning
- Fas 2: dataintervju med forskare tillsammans med SND
- Fas 3: dataintervju med forskare
- Fas 4: bearbetning av inlämnade data inför tillgängliggörande

Utformningen för fas 2 och 3 bestämdes av respektive lärosäte. Fas 4 blev ett tillägg i projektet. Fas 1 följde en utbildningsplan med fem utbildningstillfällen, där varje tillfälle tog ca en halv dag. Tema för respektive tillfälle:

1. Introduktion till datahantering, arkivering
2. Metadata för forskningsdata
3. Datahanteringsplan
4. Verktyg och metoder för att dokumentera data
5. Juridiska aspekter, summering, fortsättning till fas 2

Genomförande vid Malmö högskola

Hela projektgruppen från Malmö högskola deltog i SND:s utbildning i pilotprojektets fas 1. Utbildningen genomfördes i Lund tillsammans med Lunds universitets pilotprojektgrupp från HT-fakulteterna och Samhällsvetenskapliga fakulteten. Fas 1 genomfördes under mars till maj.

I fas 2 träffade två från projektgruppen tillsammans med en representant från SND en av högskolans forskare. Mötet var upplagt som en diskussion kring en tänkt inlämning av forskningsdata genom SND:s inlämningsformulär på webben. Fas 2-mötet genomfördes i juni. Det hade förberetts genom förmöten med forskaren där vi bland annat diskuterat syftet med projektet och en genomgång av SND:s datahanteringsplan i förhållande till den data som fas 2-mötet skulle utgå ifrån.

Fas 3 genomfördes under september och bestod av möten med två forskare på samma sätt som under fas 2. Olika typer av data diskuterades i förhållande till en tänkt inlämning av data i SND:s inlämningsformulär. Mötet leddes av en projektmedlem som samtalsledare men till skillnad från fas 2 så deltog hela projektgruppen för att ta del av "dataintervjun" och det efterföljande samtalet.

Fas 4 var ett tillägg i pilotprojektet om kurering av inlämnade data inför ett tillgängliggörande. Exempeldata studerades och vanliga problem och dess lösningar diskuterades. Det genomfördes som ett webbmöte i oktober där tre från projektgruppen deltog.

Projektgruppens erfarenheter och slutsatser

Planen var god (SND, 2016a). Tematiskt är upplägget bra, men i praktiken blev det kvalitetsmässigt ojämnt. Kanske berodde det på SND:s parallella metaperspektiv för att samla in data i pilotprojektet samtidigt som man genomförde undervisningsaktiviteter. Oavsett varför så var det ett problem under fas 1 eftersom den pedagogiska situationen blev lidande. Vi såg det som centralt att få ta del av SND:s expertis på djupet vid dessa tillfällen. Några pass blev för översiktliga och vi saknade konkreta och djupa metadatudiskussioner, till exempel om SND:s metadataschema. Fas 1 var på för basal nivå, och byggde inte vidare på MANTRA-kursen (University of Edinburgh, 2015) som vi redan genomfört i tillräckligt stor utsträckning.

Faserna 2 och 3 var givande med ökad förståelse kring forskningsdata, den bästa delen i projektet enligt flera i projektgruppen. Bra att komma i närkamp med forskningsdata (även under fas 4). För icke-forskare kan forskningsdata lätt bli ett abstrakt begrepp och för forskare kan synen på begreppet variera stort beroende på ämnestraditioner.

Vi har insett att SND:s modell av tillgängliggörande av forskningsdata behöver ställas i relation till andra modeller, till exempel data i institutionella repositorer som därmed kan bli sökbara i Swepub. Men även hur SND:s katalog fungerar i den internationella kontexten är av stor vikt för att förstå helheten, till exempel hur metadata skickas vidare till internationella söktjänster.

Genomgående har konceptet *forskningsstödsenhet* (FSE) använts av SND. Utifrån processen att tillgängliggöra forskningsdata är det logiskt (SND:s perspektiv), men det är problematiskt som begrepp då det är svårt att relatera till övrig stödverksamhet gentemot forskningen på högskolan.

Deltagandet i pilotprojektet har fungerat som en katalysator, vi har inte fått så många svar utan främst fler frågor. Vi behöver mer diskussion kring ambitionsnivå, resurser och kompetenser. Kraven på kompetens kommer att bero på nivån på metadatabeskrivningen och kravnivån på dokumentation. Pilotprojektet har också tydliggjort behovet av lokala riktlinjer och juridiska tolkningar vid det lokala lärosätet. På en basal nivå är arbetet med att tillgängliggöra forskningsdata likt det arbete vi idag utför rörande hantering av publikationer, forskarstöd och administration av MUEP.

Referenser

SND (2016a). *Projektplan för pilotprojekt Bibliotek-Arkiv-SND 2016*, version 2.

SND (2016b). *SND utbildningsplan pilotprojekt*, version 009.

University of Edinburgh (2015). *MANTRA: Research Data Management Training*.

<http://datalib.edina.ac.uk/mantra/>

9. Sammanfattning av journal club

Vid tre tillfällen har projektgruppen haft dedikerade möten där projektet diskuterats utifrån i förväg utvalda och lästa artiklar/bokkapitel. Fokus har inte varit texterna i sig utan frågeställningar som uppkommit i projektet.

Tillfälle 1 - metadata för forskningsdata

Krier, L., & Strasser, C. A. (2014). Metadata (Chapter 5), in *Data Management for Libraries : A LITA Guide*. Chicago: ALA TechSource.

E-bok MAH:

<http://search.ebscohost.com.proxy.mah.se/login.aspx?direct=true&db=nlebk&AN=759506&site=ehost-live>

Tillfälle 2 - skapa stödfunktion

Krier, L., & Strasser, C. A. (2014). Starting a new service (Chapter 2), in *Data Management for Libraries : A LITA Guide*. Chicago: ALA TechSource.

E-bok MAH:

<http://search.ebscohost.com.proxy.mah.se/login.aspx?direct=true&db=nlebk&AN=759506&site=ehost-live>

Pryor, G. (2012). Roles and responsibilities: libraries, librarians and data (Chapter 6), in *Managing Research Data*. London: Facet Publishing.

E-bok MAH:

<http://proxy.mah.se/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=558459&site=eds-live>

Tillfälle 3 - metadata med fokus på DDI

Block, W. & Thomas, W. (2003). Implementing the Data Documentation Initiative at the Minnesota Population Center, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 36:2, 97-101, DOI: 10.1080/01615440309601219

Wong, E. Y. (2016). Data Documentation Initiative, *Technical Services Quarterly*, 33:1, 93-94, DOI: 10.1080/07317131.2015.1093852

10. Beskrivning av Lärosäten Syd-workshoppar

Under projektet genomfördes två workshoppar för bibliotekspersonal inom ramen för samarbetet Lärosäten Syd. De genomfördes den 30 mars och den 26 oktober på Malmö högskolas bibliotek. Deltagare kom från Högskolan i Halmstad, Lunds universitet, Malmö högskola och SLU Alnarp.

Fokus var på att hitta data samt att förstå vad forskningsdata kan vara och hur data görs öppna.

Lärandemålen var att efter genomförd workshop ska deltagarna:

- förstå hur forskningsdata tillgängliggörs på olika sätt
- kunna hitta relevanta dataarkiv och dataset för de ämnesområden deltagarna verkar inom

11. Beskrivning av undervisning för kriminologistudenter

Inom ramen för projektet genomförde tre av projektmedlemmarna ett undervisningspass om att hitta forskningsdata. Deltagarna var studenter från kriminologiprogrammets femte termin och målet var att förbereda studenterna inför kandidatuppsatsskrivandet nästkommande termin. Inom kriminologi används ofta olika typer av data som andra samlat in, till exempel myndighetsstatistik eller stora enkätundersökningar där trygghetsfrågor är inkluderade.

Passet hölls efter önskemål från kursansvarig på kriminologiprogrammet och tanken var att också ha ett tillfälle för mastersstudenterna, vilket tyvärr aldrig genomfördes på grund av schemakrockar. Upprinnelsen till passet var en önskad visning av *Data citation index* (DCI), men vi kände att ämnet behövde mer introduktion och kontext för att studenterna skulle kunna ta till sig området. Vi försökte fokusera på data producerade och tillgängliggjorda av forskare, även om gränserna är flytande. Undervisningstillfällets två timmar blandade också genomgångar med egna sökningar. Vid den tidpunkt vi mötte studenterna hade de börjat fundera på uppsatsämnen, men få hade hunnit välja, vilket förmodligen försvårade de egna sökningarna.

Efter en allmän introduktion med avstamp från öppna myndighetsdata (ett exempel var London Datastore) exemplifierade vi forskningsdata genom att visa SND:s söktjänst och en utvald post och dess data (som var tillgänglig för nedladdning). Därefter visades Re3data.org och studenterna fick söka runt efter ämnesrelevanta dataarkiv. Efter en paus tittade vi på Data Citation Index (DCI) i *Web of Science*. Under passet var kursansvarig lärare med, vilket gjorde att vi kunde få en återkoppling efteråt och studenterna kunde fråga om uppsatsskrivande. I samtal med läraren framkom också att det finns en önskan att få fler studenter att arbeta med kvantitativa data, och att återanvändandet av forskningsdata kan vara en del av detta.

Forskningsdata som område är komplext och mångfasetterat och därmed även undervisning i ämnet. För att öka förståelsen kring området krävs att vi hittar fler sammanhängande trådar så fler moment kan bindas samman. Precis som för annan undervisning som biblioteket erbjuder krävs att vi arbetar nära lärarna för att identifiera studenternas behov. Ett område som forskningsdata ställer också nya krav på kunskaper hos den undervisande bibliotekarien. Biblioteket kan stödja i att hitta data och dataarkiv men stödet för att bearbeta och använda redan existerande dataset måste ges någon annanstans på högskolan, kanske som en del av den ämnesspecifika metodundervisningen.

Detta första tillfälle kan ses som början på ett explorativt utvecklingsarbete, där varken målet (vad som ska läras ut) eller vägen (hur man gör det på ett bra sätt) är tydliga ännu. Om vi ska jobba vidare med undervisning kring att hitta data kommer det kräva ett systematiskt utvecklingsarbete så att undervisningen kopplar samman områdets olika delar och att studenternas lärande därmed blir effektivt.

12. Beskrivning av doktorandworkshop

Doktorandworkshopen för hantering av forskningsdata är en del av stödet som biblioteket erbjuder doktorander. Denna workshop ges i samarbete med högskolans arkivfunktion. I dagsläget ligger tillfället "to be announced" på bibliotekets sida för doktorandstöd och kommer att erbjudas under våren 2017. I likhet med övriga aktiviteter som organiseras kommer Eventbrite användas för beskrivning och anmälan. Tänkt upplägg är en tre timmar lång workshop indelad i två delar; hitta och hantera forskningsdata, och forskningsdata ur ett myndighetsperspektiv.

Hitta och hantera forskningsdata

Hantering av forskningsdata är relevant för alla doktorander. Vilken typ av data man arbetar med beror på ämne och metodval. Tillfället ger en överblick över forskningsdatahanteringens olika delar och fördjupar sig dels i att hitta tillgänglig forskningsdata och del i datahanteringsplaner. Allt fler forskningsfinansiärer kräver datahanteringsplaner som också kan vara ett effektivt redskap för en enskild forskare eller en forskargrupp i datahanteringen.

Lärandemål:

- Få en översiktlig bild av forskningsdatahanteringens olika delar under forskningsprocessen
- Förstå hur forskningsdata tillgängliggörs på olika sätt
- Kunna hitta relevanta dataarkiv och dataset för de ämnesområden deltagarna verkar inom
- Ha en grundläggande förståelse för vad en datahanteringsplan är och hur den används

Forskningsdata ur ett myndighetsperspektiv

Vilka regler omgärdar bevarandet och tillgängliggörandet av forskningsdata? Liksom all verksamhet vid svenska myndigheter omfattas även forskningen vid svenska lärosäten av de regelverk som gäller för myndighetens handlingar. Även insamlade forskningsdata ska därför, i enlighet med arkivlagstiftningen, bevaras genom arkivering och om inga juridiska hinder finns även kunna begäras ut.

Därutöver ställer beslutsfattare och forskningsfinansiärer ökade krav på öppen tillgång till vetenskaplig information. I detta inkluderas tillgång till forskningsdata. Öppna forskningsdata med tillhörande metadata innebär att forskningsdata kostnadsfritt, inom ramen för gällande lagstiftning och med hänsyn till eventuella juridiska restriktioner, kan sökas och lokaliseras via internet. Det innebär att allt fler forskningsfinansiärer ställer krav på att forskare som söker och får anslag ska planera hur insamlade data ska tas om hand, dokumenteras och i vilken form de kan göras tillgängliga för framtida forskning.

Lärandemål:

- Kunskap om insamlad forskningsdatas relation till bestämmelserna om allmän handling och arkivering
- Kunskap om när i forskningsprocessen insamlade forskningsdata bör arkiveras
- Kunskap om vilken forskningsdata som kan omfattas av sekretess
- Kännedom om den personuppgiftsproblematik som kan vara aktuell vid tillgängliggörande av forskningsdata
- Kännedom om den upphovsrättsliga problematik som kan vara aktuell vid tillgängliggörande av forskningsdata

13. Metadata för forskningsdata - krav och ambitioner

Utgångspunkter

Metadata för forskningsdata är ett stort och komplext område. Metadata kan på ett systematiskt sätt användas både i forskningsprocessen och för beskrivning av själva dataseten i ett dataarkiv.

Projektets utgångspunkter angående metadata i nedanstående text är följande:

- Metadata bör finnas i MUEP, högskolans institutionella arkiv.
- Fokus på metadata för dataset och inte för högre nivåer som studie eller publikation.
- Forskningsdata i form av dataset tillgängliggörs i annan (extern) tjänst, men kan nås via länk från MUEP.
- Metadata ska uppfylla Swepubs krav.
- Ambitionsnivå när det gäller antal fält bör vara rimlig i förhållande till arbetsinsats och resurser.
- Ambitionsnivån ska uppfylla grundläggande behov, både enligt FAIR-principerna och interna behov vid Malmö högskola.

Huruvida metadata för tillgängliggjorda forskningsdata ska registreras i MUEP bör beslutas av ansvarig instans inom biblioteket. En central fråga är om ekosystemen för forskningspublikationer och forskningsdata ska blandas ihop; vad skulle det innebära för söktjänsterna och för användarupplevelsen vid sökning att publikationer och data är blandade?

Behov och ambitionsnivå

FAIR-principerna är skapade för att säkerställa grundläggande hittbarhet, åtkomst, interoperabilitet och återanvändbarhet.

The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

Tabell 13-1. FAIR-principerna från Wilkinson et al (2016).

Ambitionsnivå utifrån behov vid Malmö högskola: Det mest grundläggande behovet är att ett dataset ska kunna hittas, nås och citeras. Forskningspropositionen 2016 betonar tillgängliggörande av data som ligger till grund för forskningspublikation (Prop. 2016/17:50). Utöver dessa grundläggande behov så är ambitionen att beskrivningen i MUEP ska ge en översiktlig bild av innehållet och dess användbarhet. Av denna ambitionsnivå följer att vi uppfyller kraven för FAIR samt tillgodoser de interna behov vi ser på Malmö högskola genom registrering i MUEP, där systemet uppfyller ett flertal FAIR-krav såsom interoperabilitet.

Frågan om återanvändbarhet kan också lösas med data papers, korta vetenskapliga artiklar som beskriver tillgängliggjorda forskningsdata, på ett annat sätt än ett omfattande antal fält. Data papers eller annan publicerad dokumentation kan länkas till eller refereras till ifrån den bibliografiska posten i MUEP för forskningsdata.

Vilka fält krävs?

Nästa fråga är vilka fält som krävs för att uppnå denna ambitionsnivå. Under alla omständigheter krävs de fem obligatoriska fält som ingår i miniminivån för Data Cite (Data Cite Working Group, 2014):

- Identifier
- Creator
- Title
- Publisher
- Publication Year

Utöver dessa fält måste vi också ha med ämneskategori/klassning ("subject" är ett rekommenderat fält i Data cite). Detta ökar sökbarheten även för den forskare/student som inte söker på korrekt titel.

- Ämneskategori (till exempel SCB:s ämnesklassning)
- Ämnesord (okontrollerade)

Eftersom dataseten är tillgängliggjorda i externa tjänster så behövs metadata för åtkomst för att dataseten ska kunna nås, förslagsvis:

- Plats/lokation (datarkiv, webbplats, etc.)
- Unik identifierare, fält för till exempel DOI och Handle
- Webbadress (vid avsaknad av unik identifierare)

Dessutom ser vi fördelar med följande fält:

- Sammanfattning/abstrakt
- Licens (rörande återanvändning av datasetet)
- Relaterade publikationer, såsom data papers eller citerande publikationer
- Finansier

En anmärkningsvärd sak är att dessa fyra fält inte alls ingår i Data cite metadataschema, varken i kategorierna *Mandatory*, *Recommended* eller *Optional*. Däremot ingår dessa fält i de scheman som används vid bland annat Imperial College och Cambridge University (Brown, 2016; Reimer, 2016).

Administrativa metadata kring forskningsprojektet kan vara av intresse, speciellt för framtida kopplingar till andra interna system med forskningsinformation, förslagsvis:

- Forskningsprojektets namn
- Projekt-id (från beviljande finansiär)
- Diarienummer vid Malmö högskola

Bortsett från frågan om vilka fält som ska finnas tillkommer frågan om huruvida vissa fält bör vara kontrollerade.

Vi tror att det skulle vara en uppenbar fördel om åtminstone ämneskategoriseringen hämtas ur en kontrollerad vokabulär för ämne, vilket är en självklarhet till exempel vid katalogisering av monografier där SAO används. Förslagsvis bör SCB:s ämneskategorisering användas som används i Swepub och i andra sammanhang där svensk forskningsinformation ska beskrivas.

Kontrollerad vokabulär ingår redan i DDI (DDI Alliance, 2016) och kommer sannolikt att bli vanligare i metadatastandarder, och även bli mer efterfrågat av både forskare och studenter ju mer forskningsdata delas och används. (Wong, 2016).

Vi skulle även kunna ha en högre ambitionsnivå och ha metadata på variabelnivå (alltså metadata för varje variabel i ett dataset). Dessutom skulle vi kunna sträva efter att beskriva hela livscykeln för ett dataset. Det finns åtminstone en metadatastandard som snarare än att bara beskriva ett dataset, beskriver en livscykel där datasetet är en produkt. Det är *DDI Lifecycle*, men de flesta metadatastandarder är underförstått en beskrivning av datasetet.

I vår ambitionsnivå i detta sammanhang väljer vi emellertid att varken ta med kontrollerad vokabulär, metadata på variabelnivå eller metadata för hela livscykeln.

Förslag till rekommendation

Vi föreslår att följande fält är obligatoriska om tillgängliggjorda forskningsdata börjar registreras i MUEP:

- Identifier
- Creator
- Title
- Publisher
- Publication Year
- Ämneskategori (t.ex. SCB:s ämnesklassning)
- Ämnesord (okontrollerade)

Följande rekommenderas:

- Plats/lokation (datarkiv, webbplats, etc.)
- Unik identifierare, fält för t.ex. DOI och Handle
- Webbadress (vid avsaknad av unik identifierare)

Vidare anser vi att följande fält är intressanta att diskutera vidare:

- Sammanfattning/abstrakt
- Licens
- Relaterade publikationer, såsom data papers eller citerande publikationer
- Finansiär
- Forskningsprojektets namn
- Projekt-id (från beviljande finansiär)
- Diarienummer vid MAH

Våra motiv till att vi inte har en högre ambitionsnivå med kontrollerade vokabulärer, metadata på variabelnivå eller för livscykel är:

- Det kräver (för) mycket arbetskraft
- Vi vet inte hur vi bäst ska göra det med tanke på forskningens ämnesbredd vid högskolan, och eventuellt skulle vi ändå tvingas ändra snart
- Vi vet inte hur efterfrågat det är, det vill säga resursåtgången i relation till nytta
- Vi skulle (nästan) vara pionjärer och inte kunna dra nytta av andra, vilket skulle göra att det skulle kräva extra mycket (tanke)arbete och därmed resurser

Vidare föreslår vi att forskare informeras om data papers. Att skriva en kort artikel med fokus på ett dataset ökar dess återanvändbarhet avsevärt samtidigt som forskaren får en publikation att visa upp.

Referenser

Brown, C. (2016). Research Data Discovery: How much metadata is enough? In *Research Data Discovery Service*. <https://rdds.jiscinvolve.org/wp/2016/03/18/how-much-metadata-is-enough/>

DataCite Metadata Working Group (2014). *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data. Version 3.1*. DataCite. <http://doi.org/10.5438/0010>

DDI Alliance (2016). Controlled Vocabularies - Overview Table of Latest Versions. <http://www.ddialliance.org/controlled-vocabularies>

Prop. 2016/17:50. *Kunskap i samverkan – för samhällets utmaningar och stärkt konkurrenskraft*. <http://www.regeringen.se/rattsdokument/proposition/2016/11/prop.-20161750/>

Reimer, T. (2016). Less is more? A metadata schema for discovery of research data. In *Open Access and Digital Scholarship Blog*. <http://wwwf.imperial.ac.uk/blog/openaccess/2016/02/19/less-is-more-metadata-schema-discovery-research-datary-of-research-data/>

Wilkinson et al (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3. doi:10.1038/sdata.2016.18 <http://www.nature.com/articles/sdata201618>

Wong, E. Y. (2016). Data Documentation Initiative, *Technical Services Quarterly*, 33:1, 93-94, DOI: 10.1080/07317131.2015.1093852

14. Arkivets synpunkter

Grunden för all publicering är att det som skall publiceras finns bevarat. Det kan verka extremt grundläggande, men eftersom en mycket liten mängd forskningsdata hittar vägen till högskolans arkiv är det i den ändan vi behöver börja arbetet med publiceringsfrågan.

Själva publiceringen är en fråga för biblioteket, men hur forskningsdata hanteras är ytterst en fråga om hur högskolan hanterar allmänna handlingar och därmed en fråga för högskolans arkivarier. Högskolans arkivfunktion kan bidra med kompetens grundläggande för publicering genom att stödja hantering-, bevarande- och gallring (förstörning) av forskningsdata.

I dagsläget är högskolans stöd kring hantering av forskningsdata bristfälligt. Något som också påpekas av forskare. Arbetet med att stödja hanteringen av forskningshandlingar är komplext och forskningsdata är endast en del av de forskningshandlingar högskolans arkiv måste hantera. Följande beskrivning rör de delar av hanteringen som utgör en grund för publicering av forskningsdata.

För att underlätta hanteringen är det viktigt att forskningsdata är strukturerad med tanke på både bevarande och publicering. Det är också viktigt att sådant som innehåller personuppgifter och/eller sekretess skiljs från det som skall publiceras.

Som med all hantering av digital information spar det både tid och arbete om dessa hänsyn tas redan när informationen skapas. Forskare bör initialt, eller i samband med att data samlas in, fundera över informationens livslängd och om den innehåller personuppgifter eller sekretessbelagd information. Detta gäller oavsett om forskarna anser att forskningsdata kan och skall publiceras eller inte.

Där det är aktuellt bör man strukturera informationen på ett sådant sätt att delmängder som inte kan publiceras skiljs från sådant som kan läggas ut och att sådant som skall förstöras faktiskt förstörs när gallringsfristen går ut. Där det är aktuellt att avidentifiera data inför publicering måste högskolan också säkerställa att avidentifieringen skett på ett tillfredsställande sätt.

Vi ser ett behov av stöd kring denna sorts datahantering, något högskolan i dagsläget inte erbjuder på ett strukturerat sätt

Frågan om förvaring av forskningshandlingar under pågående projekt är i dagsläget i bästa fall en fråga som landar hos IT. Vi ser behov av ett stöd utöver det rent tekniska kring förvaring. Vi tror arkivet skulle kunna bidra med kunskap och rutiner som skulle underlätta både för pågående forskningsprojekt och för eventuell publicering av forskningsdata. Här ser vi en lucka i det stöd högskolans arkiv med rådande bemanning har möjlighet att erbjuda.

Från arkivets sida ser vi det som grundläggande för publiceringsfrågan att högskolan utarbetar och erbjuder stöd till forskare när det gäller hantering av forskningsdata och att detta stöd skall finnas tillgängligt från början av forskningsprojekt och under projektets hela livslängd.

Det kan vara svårt för den enskilda forskaren att själv avgöra exakt vilken sorts datahanteringsstöd det enskilda forskningsprojektet har behov av. Ordet datahantering kan innefatta allt från lagringsutrymme till hur man bäst skapar struktur eller hur man på bästa sätt analyserar en större mängd data. Vi ser därför att högskolan bör erbjuda ett samlat stöd kring hantering av forskningsdata och att detta stöd bör erbjudas i samverkan mellan bibliotek, IT och högskolans arkivfunktion.

15. Om lagring av forskningsdata

Till viss del handlar mycket av IT-perspektivet avseende lagring av forskningsdata om säkerhet. Följande bygger på samtal med Jesper Wokander som är informations- och IT-säkerhetsansvarig på högskolan, samt information från IT-avdelningens och högskolans webbsidor. Det är en översiktlig bild så som vi i projektet har uppfattat nuvarande status för IT-säkerhet och lagring av forskningsmaterial. För att få en mer detaljerad bild bör IT-avdelningen konsulteras vidare.

På Malmö högskola erbjuds ett antal olika lagringslösningar idag bland annat:

- Share N:\ för delad tillgång och personlig hemkatalog M:\ på högskolans servrar - <http://mah.se/medarbetare/For-ditt-arbete/IT-service-for-personal/Vara-tjanster/Hemkatalog/>
- Box - <http://mah.se/medarbetare/For-ditt-arbete/IT-service-for-personal/Vara-tjanster/Hemkatalog/box/>
- Office 365 - ingen dokumentation på webben
- Sharepoint Online - ingen dokumentation på webben
- Det går att beställa inköp av serverutrymme om det behövs för ett särskilt projekt - <http://mah.se/medarbetare/For-ditt-arbete/IT-service-for-personal/Vara-tjanster/Inkop-och-konfiguration/RFC/Bestallning-av-server/>

De olika lagringsytorna har olika för- och nackdelar. I Box kan man i likhet med Dropbox dela arbetsytan mellan flera deltagare, även personer utanför Malmö högskola inom svenska högskolevärlden. Det är alltid möjligt att be om mer diskutrymme för Share eller hemkatalogen om man har större mängder data som behöver lagras. Det går att nå Share och hemkatalog hemifrån via webbgränssnitt. IT rekommenderar i första hand Share och hemkatalog om man ska ha hög säkerhet, till exempel om man hanterar känsliga data eller personuppgifter.

Säkerhetsaspekter kommer bli än viktigare med den dataskyddsförordning som träder i kraft i EU den 25 maj 2018. En enkät till all personal har gått ut för att kartlägga hur man lagrar olika typer av information idag. Det finns redan ett ledningssystem för informationssäkerhet (LIS) som gäller för Malmö högskola. Dokumenten kring informationssäkerhet finns på flera ställen: på sidan med Malmö högskolas styrdokument; på sidorna om information om Datoridentitet för personal; på sidan om informationssäkerhet kopplat till kvalitetsarbete. Att informationen finns på flera ställen och är beskriven på olika vis gör den något svår att hitta och omsätta till specifika förhållanden.

Alla uppmanas av IT att gå den online-utbildning som finns tillgänglig för att förstå hur man ska hantera informationssäkerhet som anställd. Det är dock oklart hur många som är medvetna om detta. Själva dokumentationen är inte anpassad till olika målgruppers data/information och en forskare har troligen svårt att leta sig fram till underlagen och göra dem begripliga för sina behov.

Utbildning om informationssäkerhet finns i högskolans lärplattform som nås via länk till *It's learning* på informationssidan om informationssäkerhet kopplad till kvalitetsarbete - <https://www.mah.se/informationssakerhet> [Observera att denna sida för tillfället inte verkar vara länkad inom MAH-strukturen?]

Som forskare vill man ofta komma åt filer varsomhelst ifrån och veta var man kan lägga vad samt att det finns en bra backup. Rekommendationer från IT är att aldrig lagra något lokalt på din hårddisk. När det gäller versionshantering är det ofta upp till forskaren att skapa en sådan, till exempel vid användning av hemkatalogen. Använder man Sharepoint eller Box finns delvis versionshantering, eller snarare ändringshistorik, för att kunna gå bakåt i tiden.

En aspekt som ännu har låg prioritet för IT är hur man på enkla och samtidigt säkra vis delar filer med flera personer inblandade både på Malmö högskola och utanför. Sannolikheten att internationella samarbeten ökar när Malmö högskola blir universitet är hög och då behövs möjligen nya former av (säkra) lagrings- och arbetsytor. Det är inte alltid lagring är samma sak som att det finns bra möjligheter att arbeta med sitt material under forskningsprocessen på ett flexibelt och enkelt vis. Behov av digitala redskap i forskningsarbete är troligen något som skulle behöva undersökas vidare.

När det gäller långtidslagring efter projektslut hänvisas till ett kommande e-arkiv där man deponerar data och kan "hämta ut" den vid behov och begära sekretessprövning när det är fråga om känsliga data. Ett sådant system anses vara relativt okomplicerat att införa om det enbart handlar om arkiveringsmöjligheten och något som Malmö högskola måste tillhandahålla inom kort som myndighet.

Framtida frågor relaterat till lagring och informationssäkerhet som vi ser behöver diskuteras vidare med IT utifrån projektets erfarenheter är:

- Det finns liten eller nästan ingen alls anpassning till målgrupp av information kring lagring och informationssäkerhet för vad som gäller forskningsdata, vilket gör det svårt för forskare att hitta informationen och omsätta den i sin vardag.
- Lagring är också relaterat till bearbetning och analys av data. När man analyserar och bearbetar kan det vara fråga om mer komplexa systemlösningar såsom databaser eller användning av analysprogramvara som behöver hanteras tillsammans med forskningsdata.
- Bearbetning och analys sker ofta tillsammans med andra, möjligen i andra länder. IT nämner att det finns lösningar för detta, men det framgår inte någonstans i IT:s information i dagsläget.
- I inledningsfasen av ett forskningsprojekt eller vid forskningsansökan behöver man diskutera sina IT-behov så att man kan dimensionera rätt typ av lagring, vilket fortfarande känns relativt frikopplat från ansökningskrivandet.
- Långtidsbevarande och arkivering genom ett e-arkiv är möjligen svåränvänt som tillgängliggörande om man vill kunna citera ett dataset. Det kan behöva göras på ett enklare sätt än att man behöver "hämtar ut" från ett e-arkiv.

16. Samlade referenser

- Amsen, E. (n.d.). *Guide to open science publishing*. F1000Research. Retrieved from <http://blog.f1000research.com/2015/03/06/guide-to-open-science-publishing/>
- Block, W. & Thomas, W. (2003). Implementing the Data Documentation Initiative at the Minnesota Population Center, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 36:2, 97-101, DOI: 10.1080/01615440309601219
- Borgman, C. L. (2015). *Big data, little data, no data: scholarship in the networked world*. Cambridge, Massachusetts: The MIT Press.
- Brown, C. (2016). Research Data Discovery: How much metadata is enough? In *Research Data Discovery Service*. <https://rdds.jiscinvolve.org/wp/2016/03/18/how-much-metadata-is-enough/>
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2014). *Managing and sharing research data: a guide to good practice*. Los Angeles: SAGE.
- DataCite Metadata Working Group (2014). *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data. Version 3.1*. DataCite. <http://doi.org/10.5438/0010>
- DDI Alliance (2016). *Controlled Vocabularies - Overview Table of Latest Versions*. <http://www.ddialliance.org/controlled-vocabularies>
- 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
- 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>
- Gullbekk, E. Rullestad, T. & Calvo, M.C.T. (2013). *PhD candidates and the research process. The library's contribution*. Oslo: University of Oslo.
- 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>
- Jones, W.P. (2008). *Keeping found things found: the study and practice of personal information management*. Amsterdam: Morgan Kaufmann Publishers.
- Knorr-Cetina, K. D. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press.
- Krier, L., & Strasser, C. A. (2014). *Data Management for Libraries : A LITA Guide*. Chicago: ALA TechSource.

- Københavns Universitetsbibliotek (2016). *Digital Social Science Lab*. <http://kub.kb.dk/DSSL> Hämtad 2016-12-02
- Laursen, Christian (2016). Libraries curating open data sources. *The library Lab* [Blogg] 25 augusti <https://christianlauersen.net/2016/08/25/libraries-curating-open-data-sources/>) Hämtad 2016-12-02
- Malmö högskolas bibliotek (2016). *Projekt Tillgängliggörande av forskningsdata 2016. Projektdirektiv, version 1.0*. <http://mah.se/Bibliotek/Om-Biblioteket/Projektkontor/Pagaende-projekt/Tillgangliggorande-av-forskningsdata-2016/>
- Michener, W. K. (2015). Ten simple rules for creating a good data management plan. *Plos Computational Biology*, 11(10).
- 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
- Nordforsk (2016). *Open Access to Research Data – Status, Issues and Outlook. NordForsk Policy Paper 1 –2016*. https://www.nordforsk.org/en/publications/publications_container/open-access-to-research-data-2013-status-issues-and-outlook/download
- OECD (2007). *Principles and Guidelines for Access to Research Data from Public Funding*. OECD Publications Centre. Retrieved from <http://www.oecd.org/sti/sci-tech/38500813.pdf>
- Peixoto, A. (2014). *De mest lämpade: en studie av doktoranders habituering på det vetenskapliga fältet.* [Survival of the fittest – A study of doctoral students' habituation on the scientific field] Diss. Göteborg : Göteborgs universitet. <http://hdl.handle.net/2077/35675>
- Prop. 2016/17:50. *Kunskap i samverkan – för samhällets utmaningar och stärkt konkurrenskraft*. <http://www.regeringen.se/rattsdokument/proposition/2016/11/prop.-20161750/>
- Pryor, G. (2012). Roles and responsibilities: libraries, librarians and data (Chapter 6), in *Managing Research Data*. London: Facet Publishing.
- Reimer, T. (2016). Less is more? A metadata schema for discovery of research data. In *Open Access and Digital Scholarship Blog*. <http://wwwf.imperial.ac.uk/blog/openaccess/2016/02/19/less-is-more-metadata-schema-discovery-research-datary-of-research-data/>
- SND (2016a). *Projektplan för pilotprojekt Bibliotek-Arkiv-SND 2016, version 2*.
- SND (2016b). *SND utbildningsplan pilotprojekt, version 009*.
- SND (2016). *RFI Application SND ver 313, 2016-12-06*.
- Tenopir, C., Sandusky, R. J., Allard, S. & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*. 36 (2), 84–90. doi:10.1016/j.lisr.2013.11.003
- University of Edinburgh (2014). *MANTRA Research Data Management Training*. Retrieved from <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/>

Wilkinson et al (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3. doi:10.1038/sdata.2016.18. <http://www.nature.com/articles/sdata201618>

Wong, E. Y. (2016). Data Documentation Initiative, *Technical Services Quarterly*, 33:1, 93-94, DOI: 10.1080/07317131.2015.1093852