



LUND UNIVERSITY

Hantering av forskningsdata vid fakulteterna inom Lunds universitet – en lägesbeskrivning hösten 2015

Johnsson, Maria; Lassi, Monica

2016

Document Version:

Manuskriptversion, referentgranskad och korrigerad (även kallat post-print)

[Link to publication](#)

Citation for published version (APA):

Johnsson, M., & Lassi, M. (2016). *Hantering av forskningsdata vid fakulteterna inom Lunds universitet – en lägesbeskrivning hösten 2015*. Universitetsbiblioteket, Lunds universitet.

Total number of authors:

2

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

**Hantering av forskningsdata vid fakulteterna
inom Lunds universitet –
en lägesbeskrivning
hösten 2015**

Rapport

Maria Johnsson och Monica Lassi
Universitetsbiblioteket, Lunds universitet



LUNDS
UNIVERSITET

Februari 2016

Innehåll

Inledning.....	3
Metod	3
Resultat.....	5
Nuläget inom fakulteterna	5
Datahanteringsplan	6
Projektstart.....	7
Datainsamling.....	7
Dataanalys	8
Färdigställa och bevara data	8
Tillgängliggörande av data.....	8
Framtiden	9
Analys	10
Policy och riktlinjer	10
Kulturer och arbetssätt.....	11
Stöd och support	12
Universitetets fortsatta arbete med forskningsdata.....	13
Referenser	15
Bilaga: Intervjuguide.....	16

Inledning

Denna rapport beskriver ett projekt som Universitetsbiblioteket vid Lunds universitet har bedrivit under 2015, i vilket vi utfört en kartläggning av fakulteterna vid Lunds universitet med avseende på deras arbete med hantering av forskningsdata. Studiens mål var att få en bild av var fakulteterna befinner sig i dagsläget, att fånga upp frågor som är aktuella för de olika fakulteterna, och att hitta utgångspunkter för universitetets fortsatta arbete med forskningsdata. Kartläggningen gjordes i form av intervjuer med representanter för fakultetsledningar och fakultetsbibliotek vid Ekonomihögskolan, Humanistiska och teologiska fakulteterna, Juridiska fakulteten, Konstnärliga fakulteten, Lunds tekniska högskola, Medicinska fakulteten, Naturvetenskapliga fakulteten och Samhällsvetenskapliga fakulteten.

Sedan några år tillbaka pågår en diskussion både internationellt och i Sverige kring hantering av forskningsdata. Forskare får i allt större utsträckning krav på sig att tillgängliggöra sina forskningsdata och det pågår många diskussioner om säkra lösningar för långtidsbevarande av forskningsdata. Vetenskapsrådet tog under hösten 2014 fram ett förslag om nationella riktlinjer för öppen tillgång till forskningsdata som nu ligger hos Utbildningsdepartementet (Vetenskapsrådet, 2015). Mot bakgrund av detta, så tyckte vi att det vore intressant att undersöka hur fakulteterna vid Lunds universitet ställer sig till denna fråga.

Målgrupperna för rapporten är beslutsfattare vid Lunds universitet och stödfunktioner inom universitetet, som kan komma att jobba med hantering av forskningsdata framöver.

Rapporten är strukturerad enligt följande: Detta avsnitt följs av en beskrivning av vårt tillvägagångssätt vid datainsamling och analys av materialet. Därefter följer en resultatredovisning, som inleds med en kort redogörelse av fakulteternas nuvarande läge gällande hantering av forskningsdata. Huvuddelen av resultatredovisningen struktureras utifrån de faser som forskningsdata vanligen genomgår under forskningsprocessen (utifrån en guide för datahantering utvecklad vid Svensk Nationell Datatjänst – SND) (SND, 2015). Efter detta följer en analys av resultaten och sist en sammanfattning med förslag på fortsatt arbete med forskningsdata för Lunds universitet.

Metod

Kartläggningen genomfördes med semistrukturerade intervjuer av samtalskaraktär under hösten 2015. Inbjudningar till intervjuer skickades till fakultetsledningar och fakultetsbibliotekarier vid Lunds universitet. Responsen var mycket positiv, och vid de flesta intervjuer medverkade minst två representanter från fakultetsledningen och en representant för fakultetsbiblioteket.

Initialt vid varje intervju berättade vi kort om studien och dess kontext i övrig verksamhet kring forskningsdatahantering inom Lunds universitetsbibliotek. Som utgångspunkt för intervjuerna användes en guide för datahantering som har utvecklats vid Svensk Nationell Datatjänst (SND, 2015). Guiden beskriver detaljerat de olika faserna i forskningsprocessen då forskningsdata skapas, modifieras och berikas på olika sätt: Datahanteringsplan, Projektstart, Datainsamlingsfasen, Dataanalysfasen, Färdigställa och bevara data och Tillgängliggöra data (se bild 1).

SND Svensk Nationell Datatjänst





GÖTEBORGS UNIVERSITET

Sök och beställ data

Beskriv och lämna in data

Forskarstöd

Om oss

Kontakt

Sök

Hem > Forskarstöd > Datahantering

Webbkarta

Datahantering

Datahanteringsplan >

Projektstart

Datainsamlingsfasen >

Dataanalysfasen

Färdigställa och bevara data

Tillgängliggöra data

Datahantering

Med datahantering menas hur forskningsmaterialet hanteras, organiseras och struktureras under hela forskningsprocessen. För att enklare hantera den mängd data som insamlingen genererar och för att undvika tidskrävande arbete i efterhand är det betydelsefullt med tydlig struktur i hur data, material och dokument organiseras. Utförlig och strukturerad beskrivning av data innebär visst merarbete men lönar sig om man skulle behöva gå tillbaka för att verifiera analys och resultat. Det är också helt avgörande för långtidslagring och tillgängliggörande av materialet.

Det brukar gå snabbare att dokumentera ett forskningsmaterial om det görs i nära anslutning till de faser i forskningsprocessen som skapar eller förändrar materialet. Det är även enklare att veta hur forskningsdata ska dokumenteras om man redan i förväg har rådgjort med den institution som i slutändan ska bevara materialet. Därför är det en god idé att kontakta SND tidigt i ett projekt och diskutera dokumentation.

En del forskningsfinansiärer kräver att forskare anger i en datahanteringsplan hur insamlade data ska hanteras, dokumenteras och tillgängliggöras. Detta för att forskningsresultat ska kunna verifieras och för att data, i de fall där det inte föreligger juridiska, etiska eller kommersiella hinder, ska kunna återanvändas i framtida forskning.

SND:s webbguide inleds med Datahanteringsplan, där du hittar information om datahanteringsplaner och dess olika delar, och vad som bör tas upp under varje del. Under Projektstart får du en överblick över områden som behandlas något mer djupgående i guiden under Datainsamlingsfasen, Dataanalysfasen, Färdigställa och bevara data samt Tillgängliggöra data. Webbguiden i datahantering är lanserad i sin första version oktober 2014. Revideringar och tillägg är gjorda i september 2015.

- 1: Datahanteringsplan
- 2: Projektstart
- 3: Datainsamlingsfasen
- 4: Dataanalysfasen
- 5: Färdigställa och bevara data
- 6: Tillgängliggöra data

1

2

3

4

5

6

Bild 1. Första sidan till guide om datahantering av Svensk Nationell Datatjänst (SND),
<http://snd.gu.se/sv/forskarstod/datahantering>, 2016-01-13

I linje med studiens mål att få en bild av nuläget på fakulteterna samt att fånga upp frågor för fakulteternas fortsatta arbete med forskningsdata, så använde vi oss av en intervjuguide (se bilaga 1) vid intervjutillfällena. Vi förväntade oss dock, och var förberedda på, att intervjuerna kunde utvecklas organiskt utifrån vad intervjudeltagarna ansåg vara viktigt att prata om.

Som mest medverkade fem personer vid en intervju och som minst en person. Intervjuerna varade mellan 60 och 90 minuter. Anteckningar togs under intervjuerna, alternativt skrevs ner direkt efteråt, beroende på om vi båda utredare var närvarande vid intervjutillfället. Det första steget i analysen utgjordes av en initial öppen kodning, då vi läste igenom intervjuerna och kodade yttranden fritt, utan kodschema som styrning (Miles & Huberman, 1994). Detta gav upphov till 21 koder. I nästa steg utförde vi axial kodning, i linje med grundad teori, i syfte att hitta relationer mellan de ursprungliga koderna för

att eventuellt föra samman dem (Glaser & Strauss, 1967). Det slutgiltiga kodschemat¹ innehöll fyra huvudkoder, vilka hade underkoder som representerade olika fasetter av varje kods tema. Slutligen skapade vi memos, som sammanfattade de viktiga egenskaperna inom varje kod, inklusive eventuella skillnader och likheter i svar från de olika fakulteterna.

Resultat

Resultaten från intervjuerna visar att det finns en viss medvetenhet kring hantering av forskningsdata vid fakulteterna, även om det finns stora variationer i upplevd relevans och praktisk erfarenhet. Många studiedeltagare hänvisade till att arkivet vid Lunds universitet under 2015 genomförde en inventering av forskningsdata och hade kontaktat fakulteterna för att samla in exempel på forskningsdata. Det fanns även en stor kännedom om att finansiärer allt oftare kräver datahanteringsplaner vid forskningsansökningar, även om få hade praktisk erfarenhet av att skapa sådana planer. På det hela taget är fakulteterna medvetna om att forskningsdatahantering är något de måste förhålla sig till i framtiden, oavsett hur deras verksamhet kring forskningsdata ser ut idag. Vi utgick ifrån SND:s guide för datahantering vilket fungerade mycket bra. Guiden är pedagogiskt upplagd – såväl fakultetsledningarna som bibliotekspersonalen kände igen de olika faserna som presenterades.

Några tydliga mönster inom eller mellan fakulteterna har inte identifierats – variationerna i upplevd relevans och praktiska erfarenheter syns mer mellan olika forskningsområden och forskargrupper. Gängse kulturer och arbetssätt inom olika forskningsområden betyder desto mer för hur hantering av forskningsdata sker.

Nuläget inom fakulteterna

Studiedeltagarna nämnde en mängd olika områden inom vilka universitetets forskare idag jobbar aktivt med hantering av forskningsdata i syfte att tillgängliggöra dessa. Utsagorna presenteras här per fakultet och ger en bild av variationerna som finns inom fakulteterna.

Inom Ekonomihögskolan framhölls demografiområdet, där forskningsdata tillgängliggörs såväl nationellt som internationellt. Inom nationalekonomi kräver flertalet tidskrifter att forskare bifogar de forskningsdata som ligger till grund för ett manuskript.

Humanistiska och teologiska fakulteterna nämnde Humanistlaboratoriet vid Lunds universitet² som en miljö med väl utvecklade arbetssätt för hantering av forskningsdata. I övrigt finns en stor spännvidd inom fakultetens forskningsområden och därmed i sätten att hantera forskningsdata.

Lunds tekniska högskola rapporterade en hög medvetenhet och erfarenhet kring hantering av forskningsdata inom ramen för större EU-projekt. Inom mindre forskningsprojekt, och bland individuella forskare, är det inte lika vanligt att hantera data med tillgängliggörande som mål.

¹ Vid intresse, kontakta författarna för att få tillgång till kodschemat

² Humanistlaboratoriet, <http://www.humlab.lu.se>

Medicinska fakulteten redogjorde för de stora utmaningar som finns när det gäller att tillgängliggöra data, på grund av gränser uppsatta av bland annat ett dataskyddsdirektiv från EU och Personuppgiftslagen. I projekt som görs i samarbete med Region Skåne kan data inte delas lika fritt på grund av den sekretess som råder inom Region Skåne. Data från kohortstudier är viktiga att dela med sig av, och kan tillgängliggöras, under förutsättning att de inte begränsas av de lagar som styr sekretess.

Naturvetenskapliga fakulteten talade sammantaget om att forskarna är långt framme när det gäller hantering och tillgängliggörande av forskningsdata. De exemplifierade med en pan-europeisk forskningsinfrastruktur för hantering av klimatdata som är under utveckling, ICOS (Integrated Carbon Observation System) Carbon Portal³

Samhällsvetenskapliga fakulteten nämnde statsvetenskap och kulturgeografi som exempel på områden där forskarna samlar in och strukturerar data, som sedan tillgängliggörs i egenutvecklade databaser. Även inom psykologi och etnografi rapporterades att forskare är vana vid att dela med sig av data.

Två fakulteter uttryckte att forskningen inom deras fakulteter var av sådan karaktär att hantering av forskningsdata med målet att tillgängliggöra kanske inte är så applicerbart som för övriga fakulteter. Juridiska fakulteten beskrev att deras forskning främst utgörs av arbete med öppna källor, exempelvis lagar och förarbeten. Konstnärliga fakulteten svarade att forskningen i regel inte genererar forskningsdata, men att de forskningsområden som i någon mening skulle kunna tala om forskningsdata är musikpedagogik och fri konst.

Datahanteringsplan

SND rekommenderar i sin guide för datahantering att forskare upprättar en datahanteringsplan i samband med att en ansökning om forskningsanslag görs (SND, 2015). Datahanteringsplanen kan utgöra riktlinjer för hur forskaren hanterar sina data under resten av forskningsprojektet, fram till dess att data arkiveras och tillgängliggörs.

Studiedeltagarna var väl medvetna om att datahanteringsplaner krävs av somliga finansiärer, och det fanns en beredskap för att sådana kan komma att krävas i större omfattning i framtiden. Det är ofta inom större forskningsprojekt, till exempel EU-projekt eller andra internationella projekt som det skapas datahanteringsplaner på fakulteterna i nuläget. Inom dessa projekt finns det också mer resurser till datahantering, ibland till att inkludera en så kallad "data manager" som stöttar forskarna under pågående projekt. Det var dock få av studiedeltagarna som hade någon egen erfarenhet av att upprätta en datahanteringsplan. Några hade erfarenhet då Vetenskapsrådet under ett år införde krav på datahanteringsplaner i samband forskningsansökningar.

I SND:s guide betonas även vikten av att tidigt bestämma hur forskningsdata skall beskrivas och vilken standard för metadata som ska användas. Vi frågade hur medvetenheten ser ut beträffande metadata och beskrivande av data inom fakulteten. De studiedeltagare som forskar hade i regel inte tänkt så mycket på metadataaspekter, medan det för biblioteken var naturligt att arbeta med metadata redan, till exempel i samband med registrering av publikationer i universitetets publikationsdatabas. Det rådde

³ ICOS Carbon Portal, <https://www.icos-cp.eu/>

enighet om att kunskap om och kompetens kring metadata kommer att bli viktigt i framtiden, bland annat för att kunna hitta relevanta data samt för att kunna avgöra en datamängds kvalitet och applicerbarhet för den egna forskningen. Samtidigt, uttryckte en studiedeltagare, kan det bli stora krav på forskaren att beskriva forskningsdata så att det blir användbart och läsbart för någon annan. Vikten av standarder för metadata ur användar- och systemperspektiv betonades också, såväl för läsbarhet som för att datamängder och tillhörande metadata ska vara maskinläsbara och kunna inkorporeras i olika infrastrukturer.

De juridiska och etiska aspekter som SND föreslår är viktiga att ta ställning till inkluderar: etikprövning innan projektstart, informationssäkerhet och skydd för personlig integritet under projektets gång, upphovsrätt samt arkivering i samband med att projektet avslutas. Vi frågade om det finns särskilt stöd inom fakulteterna för juridiska och etiska aspekter av hantering av forskningsdata. Vid de fakulteter där detta var mer aktuellt, så hade vissa rutiner inrättats, men studiedeltagarna önskade utöver detta ett bra juridiskt stöd centralt på universitetet för rådgivning om exempelvis Personuppgiftslagen, sekretess och forskningsetik.

Projektstart

När forskaren har fått finansiering och skall starta sitt forskningsprojekt råder SND att datahanteringsplanen ska konkretiseras ytterligare och att forskaren skall bestämma hur datainsamlingen skall ske under projektet. Ingen fakultet rapporterade om några särskilda rutiner just vid projektstart. Den lokala IT-supporten, biblioteket och i vissa fall även jurister nämndes som viktiga stödfunktioner vid projektstart.

Datainsamling

Under fasen datainsamling råder SND forskare att säkra datakvaliteten bland annat genom att strukturera och namnge sina filer på ett genomtänkt sätt. En välorganiserad mappstruktur för filerna är av betydelse för kvalitetskontrollen av forskningsdata och bidrar till att göra forskningsarbetet effektivare. Vid namngivning av filer är det viktigt att det är begripligt, inte bara för forskaren själv, utan även för en utomstående person som vill ta del av filerna. Vi frågade om fakulteterna hade några särskilda riktlinjer för detta. Av svaren framgår att tillvägagångssätten vid datainsamling ser väldigt olika ut, och att utarbetade metoder för namngivning och strukturering av filer är långt ifrån givet.

Under datainsamlingsfasen pekar SND också på vikten av datasäkerhet, till exempel hur data bör lagras i stationära datorer respektive molntjänster, vilket blir allt vanligare. Vid lagring av personuppgifter så gäller särskilda regler för datasäkerhet, och SND hänvisar där till Datainspektionen och till lärosätets lokala personuppgiftsombud. Gällande rutiner för datasäkerhet vid datainsamling hänvisade flera fakulteter till den lokala IT-avdelningen. Vid vissa fakulteter är lagring av personuppgifter vanliga och då är kraven på datasäkerhet vid hantering och lagring av data stora. Några riktlinjer eller rutiner för datasäkerhet kring personuppgifter, som gäller genomgående för en hel fakultet, kan vi dock inte skönja i materialet. Tillvägagångssättet beror antagligen mycket på vilket ämnesområde som ett forskningsprojekt bedrivs inom.

Dataanalys

I dataanalysfasen är det viktigt att hålla reda på olika versioner av filer och att dokumentera förändringar i filerna. SND råder forskare att förse varje version av filerna med tydliga versionsnummer och datum för uppdateringen. Varje fil bör därtill ha en variabellista eller kodbok med beskrivningar av variablerna och hur de är kodade, vilka också bör förse med versionsnummer och datum. Med tydliga versionsnummer på såväl filer som variabellistor och kodböcker blir det lättare att matcha dokumentationen till rätt fil och därmed ha kontroll på vad som har gjorts, varför, av vem och vid vilken tidpunkt.

Vi frågade studiedeltagarna om det inom fakulteten finns några särskilda rutiner för versionshantering av filer. Av svaren framgår att versionshantering, återigen, beror främst på ämne och huruvida forskningen bedrivs individuellt eller i grupp. I forskningsprojekt där flera forskare är involverade blir versionshanteringen helt klart viktig, och i projekt där det finns särskilt stora behov av strukturerade arbetssätt. En fakultet betonade vikten av versionshantering i dataanalysen, eftersom forskningsdata kan genomgå olika faser under analysen och berikas allt eftersom analysen fortskrider. Då är det viktigt att känna till vilken eller vilka versioner man arbetar med, och att man kan gå tillbaka till sina rådata vid behov.

Färdigställa och bevara data

När ett forskningsprojekt är på väg att avslutas rekommenderar SND att forskningsdata förbereds för långtidsbevarande och arkivering, bland annat genom att filer sparas i ett format som är (eller tros komma att vara) läsbart i framtiden. Digital arkivering ställer speciella krav på underhåll, till exempel att filerna regelbundet konverteras till nyare filformat och att filerna migreras, det vill säga flyttas över från ett medium till ett annat, vid behov. SND tar även upp viktiga aspekter vid hantering och lagring av personuppgifter, såsom informerat samtycke vid forskning om människor och etikprövning vid behandling av känsliga personuppgifter. Vi frågade fakulteterna om de har något särskilt stöd eller särskilda rutiner för arkivering och långtidsbevarande av forskningsdata, vilket samtliga svarade nekande på. De forskare som hanterar personuppgifter i sin forskning hade i viss mån berört bevarandefrågan, ofta i samband med etikprövning. En fakultet påpekade att forskare ofta har en oklar uppfattning om skillnaden mellan arkivering och lagring. Detta visar på att det finns behov av vägledning och kompetensutveckling för forskare när det gäller arkivering och långtidsbevarande.

Tillgängliggörande av data

SND ger i sin guide en del råd och anvisningar hur forskare kan tillgängliggöra forskningsdata, till exempel i internationella dataarkiv, genom publicering i datatidskrifter, via SND eller andra forskningsinfrastrukturer. SND lyfter fram fördelarna med att tillgängliggöra data via en resurs som förser data med unika identifierare, till exempel DOI (Digital Object Identifier) eller PID (Persistent Identifier). Detta underlättar både sökbarheten och citeringen av forskningsdata. Vi frågade studiedeltagarna vilka erfarenheter som finns bland fakulteternas forskare när det gäller tillgängliggörande av forskningsdata, och även i denna fråga varierade svaren mycket. Inom vissa forskningsområden fanns erfarenhet av, och en positiv syn på, att tillgängliggöra forskningsdata, till

exempel inom epidemiologi, demografi, nationalekonomi och språkforskning. Flera forskare som ingår i internationella forskningsprojekt, till exempel inom EU, rapporterade en viss erfarenhet av att tillgängliggöra forskningsdata. Inom humaniora lyftes vikten av att tillgängliggöra *levande* data, det vill säga data som fortsätter att uppdateras efter arkiveringen och tillgängliggörandet. En annan faktor som ansågs viktig vid tillgängliggörandet av forskningsdata är att infrastrukturerna där forskare skall deponera sina data är tillförlitliga och förtroendeingivande. Flera studiedeltagare tryckte också på vikten av tid. Forskare behöver få tillräckligt med tid på sig att analysera och använda sig av sina data innan de släpps fria – jämför med embargotid för Open Access-publicering av vetenskapliga publikationer.

SND lyfter fram ”datacitering” som viktigt i sin guide, enligt en liknande princip som för citering av publikationer. Citering till en datamängd bör innehålla tillräckligt med information för att det skall vara möjligt att hitta och nå den korrekta versionen av datamängden. Detta kan underlättas genom användning av unika identifierare, till exempel DOI, som ska kunna säkerställa tillgången till en datamängd. Vi frågade studiedeltagarna hur känt begreppet datacitering är på fakulteterna. Det visade sig att det är ett ganska okänt fenomen för de flesta, och att det inte är särskilt vanligt bland forskare att citera forskningsdata. En studiedeltagare framhöll att det torde ingå i den så kallade tredje uppgiften, det vill säga samverkan med det omgivande samhället, att dela med sig av forskningsdata, och därför torde inte en praktik kring citering av data behövas. Någon annan reflekterade kring att citering nog kommer att bli viktigt på sikt, då det kan öka forskarnas synlighet.

Gällande incitament för att motivera forskare att tillgängliggöra sina forskningsdata svarade flertalet att det antagligen kommer att krävas tvingande policyer och gärna ett samspel mellan ”morot och piska”. Någon nämnde att det vore bra om forskningsprocessen kunde underlättas genom att använda infrastrukturer för att dela med sig av data. Stöd i projektadministration skulle också kunna underlätta processen för att tillgängliggöra forskningsdata. Att bidra till sitt ämne eller forskning i stort – till en större helhet – som forskarna själva också kan ta del av, nämndes också som incitament.

Framtiden

Vid de sista frågorna i intervjuerna lämnade vi SND:s guide och ställde några frågor om universitetets arbete med forskningsdata framöver. Vi efterfrågade studiedeltagarnas reflektioner och visioner kring hantering av forskningsdata i framtiden. Flera pekade på den kommande forskningspropositionen 2016 som viktig för att bestämma riktningen för lärosätena. Vetenskapsrådet ses också som en viktig aktör. Flera nämnde i detta sammanhang att universitetet borde ha en policy för forskningsdata och att universitetsledningen aktivt borde arbeta för detta. Flera påpekade också att det är viktigt att bygga vidare på de befintliga, lokala stödfunktionerna inom fakulteterna. En fakultet uttryckte viss oro för om det kommer att finnas tillräckligt med incitament för mindre forskningsprojekt att tillgängliggöra data, då resurserna för datahantering antagligen är mindre än i de stora projekten. Kopplingen mellan forskningen och undervisningen nämndes också som en viktig fråga. Undervisningen skulle kunna få en extra dimension om studenterna fick tillgång till forskningsdata att arbeta med. Att bli lite mer öppen som forskare och kanske dela med sig av data under pågående forskningsprojekt föreslogs också som ett intressant nytt arbetssätt i framtiden.

Ett explicit behov, som flera studiedeltagare gav uttryck för, gällde att få stöd i forskningsdatarelaterade frågor genom hela forskningsprocessen, vilket spänner över flera organisatoriska enheter, exempelvis biblioteken, Forskningsservice och juridiska avdelningen. Stödet föreslogs vara utformat så att forskaren inte behöver veta exakt vilken organisatorisk enhet hen bör kontakta i en specifik fas av forskningsprocessen, utan att stödfunktionen skulle vara sömlös utifrån ett användarperspektiv. LUCRIS, det system för forskningsinformation som kommer att ersätta Lund University Publications (LUP) från och med våren 2016, nämns som ett tänkbart verktyg för tillgängliggörande av forskningsdata och för de behov av administrativt stöd som till exempel kan ligga på biblioteken ⁴. Biblioteken kan spela en viktig roll, bland annat för stöd i beskrivning av forskningsdata med hjälp av lämpliga metadata.

Analys

Analysen av våra data presenteras i tre teman: policy och riktlinjer, kulturer och arbetssätt, samt stöd och support. Dessa tre teman diskuteras nedan i relation till trender i omvärlden av betydelse för öppen tillgång till forskningsdata.

Policy och riktlinjer

Som vi nämnde i resultatavsnittet, så efterlyser flera fakulteter en policy för tillgängliggörande av forskningsdata, i synnerhet att Lunds universitets ledning är drivande i utarbetande av en sådan policy. Även om det lär dröja ett tag tills det finns en nationell policy för öppen tillgång till vetenskaplig information i Sverige, så är diskussionen i gång i olika sammanhang. Policyfrågan har funnits på dagordningen nationellt sett ett tag, och blev aktuell i samband med att Vetenskapsrådet presenterade ett förslag till riktlinjer för öppen tillgång till vetenskaplig information för regeringen i januari 2015 (Vetenskapsrådet, 2015). Detta förslag finns för närvarande för beredning hos Utbildningsdepartementet och förväntas ingå i forskningspropositionen 2016. Två medarbetare från Universitetsbiblioteket närvarade vid den hearing om förslaget till riktlinjer, som hölls 16 december 2015⁵. Förslaget gäller för forskning som helt eller delvis finansierats med offentliga medel, med målet att all forskningsdata skall öppet tillgängliggöras så snart som möjligt. Vid hearingen och i Vetenskapsrådets förslag betonades vikten av samordning mellan lärosätena vad gäller arkivering och långtidslagring av forskningsdata, exempelvis i lärosätenas val av digitalt arkiv, medan en teknisk plattform för tillgängliggörande föreslås ligga på nationell nivå.

Som medlem i EU har Sverige även flera EU-rekommendationer att förhålla sig till, exempelvis inom ramen för Horisont 2020⁶ där datahanteringsplaner krävs vid ansökning. Vetenskapsrådets förslag inkluderar, som en första åtgärd under åren 2015-2020, att pilotprojekt ska utlysas, vilka kommer att kräva en datahanteringsplan i linje med den som används inom Horisont 2020. Datahanteringsplanen

⁴ LUCRIS, <http://www.medarbetarwebben.lu.se/forska-och-utbilda/stod-till-forskning/lucris-system-for-forskningsinformation>

⁵ För att ta del av en sammanfattning av hearingen, kontakta Jette Guldborg Petersen eller Monica Lassi vid Universitetsbiblioteket.

⁶ Horisont 2020, <https://ec.europa.eu/programmes/horizon2020/>

kommer sedan att användas under hela projektet som en vägledning för hantering av forskningsdata, ända fram till att projektet avslutas och data skall tillgängliggöras.

Även om det nationellt sett går långsamt i policyfrågan, så hindrar det inte att lärosäten tar egna initiativ till lokala policyer kring forskningsdata. Lärosätens internationella forskningssamarbeten och krav från forskningsfinansiärer är ytterligare faktorer som påverkar dem att upprätta lokala policyer och riktlinjer för forskningsdata.

Kulturer och arbetssätt

I det tidigare avsnittet "Tillgängliggörande av data" redogjorde vi för vilka erfarenheter och praktiker det finns bland fakulteterna gällande tillgängliggörande av forskningsdata. Vi kan konstatera att det är stora skillnader mellan forskningsområden och att attityderna skiljer sig en del mellan olika forskare. De flesta forskare följer en gängse kultur inom sina respektive ämnen, där man i mångt och mycket följer praxis för hur forskning bedrivs och publiceringsmönster för i vilka kanaler man publicerar sina forskningsresultat. För de flesta forskare är den egna ämneskulturen och forskningsgemenskaperna mycket viktiga – i vissa fall viktigare än det lokala nätverket hemma på lärosätet. Det blir därför naturligt att följa de praktiker som gäller inom det egna forskningsområdet framför andra intressenters rekommendationer och policyer som inte är tvingande.

Inom Lunds universitet finns flera intressanta exempel på forskningsområden som tillgängliggör forskningsdata. Själva tillgängliggörandet sker på många olika sätt, men de har ofta en tydlig, uttalad strategi att tillgängliggöra forskningsdata inom ett visst ämne. Vid Humanistlaboratoriet (Humlab) finns webbservern "RWAAI" (Repository and Workspace for Austroasiatic Intangible Heritage) som tillgängliggör data om minoritetsspråk i Sydostasien och Indien. "RWAAI" är ett bra verktyg för samarbeten mellan universitetets forskare och forskare på andra lärosäten. Inom ekonomisk historia och demografi finns databasen "SEDD" (The Scanian Economic Demographic Database) som innehåller data om den demografiska utvecklingen i Skåne från sent 1800-tal och framåt. "SEDD" är ett exempel på registerbaserad forskning, där registret utgör forskningsmaterialet, som forskaren rensar och bearbetar utifrån sin forskningsfråga. Inom epidemiologi finns projektet "Skånes Metadatabas för Epidemiologi", i vilken metadata, det vill säga beskrivningar av forskningsdata, från ett antal epidemiologiska studier vid Lunds universitet och Region Skåne tillgängliggörs. Syftet med metadatabasen är att forskningsdata från dessa resurser skall kunna återanvändas i andra studier. På grund av sekretess är dessa data dock inte öppet tillgängliga, utan intresserade hänvisas att kontakta projektets forskare för vidare information. Det finns många fler exempel på forskningsområden som tillgängliggör forskningsdata vid Lunds universitet, detta är endast ett axplock.

Att utvecklingen går mot att forskningsdata skall delas och tillgängliggöras är tydligt – det finns många fördelar både forskningsmässigt och samhällsekonomiskt med detta, och det stämmer väl in på trenderna i samhället. Framförallt främjar det forskares samarbeten med externa forskargrupper och lärosäten. Å andra sidan, så finns det även en skepsis och tvekan hos många forskare till att dela med sig av forskningsdata. Det finns många anledningar till varför tillgängliggörande inte sker, och inom litteraturen diskuteras underliggande faktorer till varför det ser ut som det gör i forskarvärlden (Fecher et al., 2015; Goodman et al., 2014; Lassi, 2014; Tenopir et al., 2011). Faktorer som rädsla för konkurrens,

skyddande av studiedeltagare, brist på tid och resurser, brist på stöd från organisationen, okunskap om fördelarna med tillgängliggörande är exempel på hinder för tillgängliggörande av forskningsdata enligt dessa källor. Vi ser en del liknande tendenser och åsikter i denna kartläggning, även om denna studie haft ett mer övergripande organisationsperspektiv snarare än att undersöka individuella forskares praktiker och attityder.

Våra frågor väckte både intresse och engagemang hos de personer vi intervjuade. Många kände igen sig i intervjuernas teman och ämnet forskningsdata kändes aktuellt på flera sätt. Till exempel hade flera hört talas om datahanteringsplaner och någon enstaka hade erfarenhet av att utarbeta sådana. Hantering av forskningsdata är en så central del av forskningsprocessen och utgör i regel själva kärnan i forskningen. Forskningsdata är förknippat med det dagliga arbetet och det forskare jobbar med under stora delar av sin arbetstid. Forskningsdata hänger också ihop väldigt mycket med forskningsmetodik – vid publicering redovisas ofta metod och material tillsammans. Så hantering av forskningsdata är för många det som händer och sker i vardagen. Att gripa in i forskarens vardag och lägga till nya moment i forskningen är kanske den stora utmaningen när lärosäten, finansierare med flera vill få forskare att dela med sig av forskningsdata och att arbeta mer strukturerat med forskningsdata. Sannolikt kommer forskare även i framtiden att följa gängse kulturer och arbetssätt inom det egna ämnet, och då kommer ämnets förmåga att anamma olika sätt att dela sig av forskningsdata att vara avgörande – i de fall det inte finns krav från någon intressent.

Stöd och support

Flera studiedeltagare hade varit i kontakt med olika stödfunktioner i frågor kring forskningsdata, och beskrev att de fått olika mycket hjälp med sina ärenden och ibland fått motstridiga svar på samma fråga. Det finns idag ingen samordning av stödtjänster för hantering av forskningsdata vid Lunds universitet. För Lunds universitet, som är ett förhållandevis stort lärosäte, så finns det en omfattande organisation av stödfunktioner för forskningen. Det finns expertis att rådfråga på nära håll inom universitetet, exempelvis jurister, arkivarier, IT-support, etiker och bibliotekarier. Forskaren skall bara hitta rätt till stödfunktionerna när hen behöver hjälp. För att underlätta för forskarna vore det bra om stödfunktionerna kunde samordnas mer. På så vis kan den kompetens och expertis som redan finns inom universitetet komma till sin rätt.

Några exempel på stödfunktioner inom Lunds universitet som kan vara relevanta för stöd i hantering av forskningsdata är:

- Arkivet
- Biblioteken
- Etikprövningsnämnder
- Forskningsservice
- IT-stöd och -tjänster
- Juridiska avdelningen

Det finns ett tydligt behov av kompetensutveckling gällande hantering av forskningsdata, såsom metadata, juridiska aspekter och hantering av filer för att underlätta åtkomst och läsbarhet för andra

forskare. Forskningsdata behöver därför implementeras i gängse kompetensutvecklingsprogram vid Lunds universitet. Vad gäller forskarutbildningen så kan forskningsdata komma in i flera sammanhang under utbildningens gång. Framtidens doktorander kommer att behöva tänka på datahanteringsplaner på ett tidigt stadium i forskningsprocessen och på tillgängliggörande av forskningsdata när de närmar sig slutet av studierna. Även för seniora forskare behöver det erbjudas kompetensutveckling i hantering av forskningsdata, gärna onlineutbildningar som forskarna kan följa på egen hand. Personal inom stödfunktionerna behöver också utbildning i hantering av forskningsdata och här handlar det mycket om vilka lagar och regler som gäller och vilka tjänster och verktyg som finns för forskningsdata. Det är också viktigt att personalen inom stödfunktionerna känner till, har viss kunskap om övriga stödfunktioner inom universitetet. Flera lärosäten har tagit fram kompetensutveckling och stöd för hantering av forskningsdata, exempelvis University of Edinburgh i Storbritannien som har utvecklat onlineutbildningen "MANTRA". Lunds universitet skulle kunna erbjuda något liknande för sin personal.

Många forskare har uppfattningen att de själva äger sina forskningsdata, när det i själva verket är lärosätet eller myndigheten som står som ägare (Vetenskapsrådet, 2015). Detta leder ofta till att forskare behandlar sina data som sina egna och tar dem med sig vid flytt till annat lärosäte, ett fenomen som även beskrevs av en studiedeltagare i denna studie. Det kan bli en pedagogisk utmaning för lärosäten att förklara för forskare att forskningsdata är att betrakta som en allmän handling vilken kan begäras ut av allmänheten och att forskningsdata tillfaller lärosätet och inte den enskilde forskaren (Vetenskapsrådet, 2015).

Universitetets fortsatta arbete med forskningsdata

Utifrån denna kartläggning kan vi konstatera att det finns tydliga behov av att arbeta vidare med hantering av forskningsdata inom Lunds universitet framöver, även i väntan på nationella riktlinjer kring tillgängliggörande av forskningsdata. Frågan är aktuell i många sammanhang, exempelvis EU:s rekommendationer för forskning och forskningsfinansiärers krav på datahanteringsplaner. Vetenskapsrådet anger i sitt förslag till nationella riktlinjer för öppen tillgång till forskningsdata att långtidsbevarande och arkivering av forskningsdata ligger på lärosätens ansvar, enligt Arkivlagen (Vetenskapsrådet, 2015). Därför föreslår Vetenskapsrådet att lärosätena avsätter medel för de processer och tekniska lösningar som behöver utarbetas under förutsättning att förslaget implementeras.

Lokalt, på Lunds universitet, så hoppas och tror vi att intervjuerna har bidragit till att aktualisera frågan vid fakulteterna, utöver de exempel vi fått på forskningsområden som redan är aktiva när det gäller hantering av forskningsdata i syfte att tillgängliggöra dessa. Detta, tillsammans med andra initiativ kring forskningsdata på universitetet, utgör goda förutsättningar att arbeta vidare utifrån.

Några förslag till åtgärder för Lunds universitet som kan underlätta arbetet med forskningsdata är:

- Öka samordningen av stödfunktionerna för hantering av forskningsdata, exempelvis med gemensam information på universitetets webbplats, som tar utgångspunkt i forskningsprocessen och inte i vilket stöd som erbjuds av specifika organisatoriska enheter.
- Starta en arbetsgrupp som tittar på framtida policy och riktlinjer för hantering av forskningsdata.

- Skapa kompetensutvecklingsprogram inom hantering av forskningsdata för forskare, forskarstuderande och personal inom stödfunktionerna.

På längre sikt bör meriteringssystemet för forskare och doktorander ses över, och anpassas så att forskningsdata inkluderas i meriteringen av forskare. Hänsyn kan exempelvis tas till huruvida en person tillgängliggör sina forskningsdata och till andra forskares användning och citeringar av forskningsdata, på liknande sätt som citeringar av vetenskapliga publikationer.

Det pågår flera projekt inom Lunds universitet som kan relateras till hantering av forskningsdata. Som vi redan har nämnt så arbetar universitetets arkiv med en inventering av forskningsdata. Inventeringen av forskningsdata är en del i en större inventering som även omfattar administrativa och undervisningshandlingar upprättade till och med år 2013. I ett första steg, som utfördes under 2015, fokuserade arkivet på fysiskt material, och i nästa steg, under 2016, kommer fokus att ligga på digitalt material⁷. Ett annat projekt som har startat upp 2016 är ett samarbete mellan Humanistiska och teologiska fakulteterna och Samhällsvetenskapliga fakulteten om hantering av forskningsdata.

Inom Universitetsbiblioteket kommer vi under 2016 att arbeta vidare med forskningsdatahantering. Vi kommer bland annat att genomföra en studie där forskare intervjuas om deras erfarenheter kring hantering av forskningsdata med hjälp av ett standardiserat verktyg för att skapa forskningsdataprofiler, "Data Curation Profiles Toolkit". Studien blir av fördjupande art jämfört med kartlägningsstudien, med fokus på individuella forskare i stället för på organisationen i stort. Studien förväntas bidra med ytterligare vägledning om vilka stöd forskare behöver för hantering av forskningsdata i olika delar av forskningsprocessen. Vi kommer även att testa ett onlineverktyg för datahanteringsplaner, "DMP Online" som är utvecklat av Digital Curation Centre i Storbritannien, i vilket forskare kan skapa planer utifrån forskningsfinansiärers mallar (DCC). Att bygga upp en tjänst kring datahanteringsplaner ligger i linje med att allt fler finansiärer kräver sådana vid ansökningar om anslag, och att Vetenskapsrådet i sina förväntade pilotutlysningar kommer att kräva det. Därtill fanns en stor medvetenhet kring finansiärers krav, och en tro om att datahanteringsplaner kommer att krävas i allt större utsträckning, bland alla fakulteter inom universitetet.

Sammanfattningsvis, så ser vi att det finns goda förutsättningar att arbeta vidare med hantering av forskningsdata inom universitetet, utifrån den medvetenhet och de behov som uttrycktes av fakulteterna, samt utifrån de krav från externa intressenter som förväntas öka med tiden. Ett samlat grepp i det fortsatta arbetet på universitetet kan skapa goda förutsättningar för att bli ett Open Science-universitet.

⁷ Uppgift från Åsa Berglund, Avdelningschef, Avdelningen Dokumenthantering vid Sektionen Juridik och dokumenthantering

Referenser

- Data Curation Profiles* [Online]. Purdue University Libraries. Tillgänglig: <http://datacurationprofiles.org/> [Hämtad 2016-02-02].
- DMP Online*. [Online]. Digital Curation Centre. Tillgänglig: <https://dmponline.dcc.ac.uk/> [Hämtad 2016-02-03].
- Fecher, B., Friesike, S. & Hebing, M. 2015. What Drives Academic Data Sharing? *PLoS ONE*, 10, 1-25.
- Glaser, B. G. & Strauss, A. L. 1967. *The discovery of grounded theory; strategies for qualitative research*, Chicago,, Aldine Pub. Co.
- Goodman, A., Pepe, A., Blocker, A. W., Crosas, M., Di Stefano, R., Kashyap, V., Siemiginowska, A., Borgman, C. L., Cranmer, K., Hogg, D. W., Gil, Y., Groth, P., Hedstrom, M., Mahabal, A. & Slavkovic, A. 2014. Ten Simple Rules for the Care and Feeding of Scientific Data. *PLoS Computational Biology*, 10.
- Lassi, M. 2014. *Facilitating collaboration : exploring a socio-technical approach to the design of a collaboratory for library and information science*, Borås, University of Borås.
- MANTRA - Research Data Management Training* [Online]. EDINA, University of Edinburgh. Tillgänglig: <http://datalib.edina.ac.uk/mantra/> [Hämtad 2016-02-03].
- Miles, M. B. & Huberman, A. M. 1994. *Qualitative data analysis : an expanded sourcebook*, Thousand Oaks, Sage Publications.
- RWAAI - Repository and Workspace for Austroasiatic Intangible Heritage* [Online]. Humanistlaboratoriet, Lunds universitet. Tillgänglig: <http://hdl.handle.net/10050/00-0000-0000-0003-66A4-2@view> [Hämtad 2016-02-03].
- SEDD – the Scanian Economic Demographic Database* [Online]. Centre for Economic Demography. Tillgänglig: <http://www.ed.lu.se/databases/sedd> [Hämtad 2016-02-03].
- Svensk Nationell Datatjänst - SND. 2015. *Datahantering* [Online]. Tillgänglig: <http://snd.gu.se/sv/forskarsupport/datahantering> [Hämtad 2016-02-03].
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M. & Frame, M. 2011. Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE*, 6, 1-21.
- The Scania Metadatabase for Epidemiology* [Online]. Medicinska fakulteten, Lunds universitet. Tillgänglig: http://www.med.lu.se/bbmri/the_scania_metadatabase_for_epidemiology [Hämtad 2016-02-03].
- Vetenskapsrådet. 2015. *Förslag till nationella riktlinjer för öppen tillgång till vetenskaplig information* [Online]. Tillgänglig: <https://publikationer.vr.se/produkt/forslag-till-nationella-riktlinjer-for-oppen-tillgang-till-vetenskaplig-information/> [Hämtad 2016-01-13].

Bilaga: Intervjuguide

Kommentar: röd text är SND:s benämningar

0: Introduktion

- Berätta kort om UB:s projekt
- Aktuellt kring datahantering
- SND:s modell för datahantering
- Hur ser typiska forskningsdata ut hos er/vid er fakultet? Kan ni ge tre exempel på forskningsdata?
- Har ni diskuterat hantering av forskningsdata på ledningsnivå inom fakulteten eller i andra sammanhang?
- Är biblioteket involverat i ovanstående diskussioner? Om ja, på vilket sätt?

Steg 1: Projektansökan (**Datahanteringsplan**)

- Datahanteringsplaner – Är det något som används vid er fakultet?
- Juridik och etiska aspekter – Finns det särskilt stöd eller rutiner för att hantera detta vid er fakultet?
- Metadata och beskrivande av data – Hur ser medvetenheten ut om detta vid fakulteten?

Steg 2: Projektstart (**Projektstart**)

- Finns det något särskilt stöd eller särskilda rutiner vid projektstart?

Steg 3: Datainsamling (**Datainsamlingsfasen**)

- Struktur och begriplighet – Finns det särskilda riktlinjer för hur man namnger och strukturerar dataset/datafiler? Eller stöd?
- Datasäkerhet – Finns särskilda rutiner för detta?
- Organisering –
- Dokumentation – Finns särskilda rutiner för detta? Används loggböcker, labböcker etc?

Steg 4: Dataanalys (**Dataanalysfasen**)

- Versionering av data och dokumentation – Finns särskilda rutiner för hur man håller reda på olika versioner av dataset?
- Säkerhetskopiering och backup – Finns särskilda rutiner för säkerhet och backup?

Steg 5: Dataförberedelse (Färdigställa och bevara data)

- Arkivering och långtidslagring – Finns särskilt stöd eller särskilda rutiner?
- Hantering av personuppgifter – Finns särskilt stöd eller särskilda rutiner?
- Filformat för långtidslagring – Finns särskilt stöd eller särskilda rutiner?

Steg 6: Tillgängliggörande av data (Tillgängliggöra data)

- Tillgängliggöra data – Finns det särskilda krav från era forskningsfinansiärer?
- Kultur, rutiner, erfarenheter bland forskare – Hur ser rutinerna ut bland era forskare för tillgängliggörande av forskningsdata?
- Data citering – Hur känt är detta begrepp inom er fakultet? Tror ni att det kommer att bli viktigt för forskares karriärer i framtiden, t.ex. att vara en faktor vid anställningar?
- Incitament för att dela med sig av data – Vad tror ni är viktigt för forskare, utifrån fakultetsperspektiv och individperspektiv?

Avslutande frågor

- Har ni några önskemål på tjänster eller stöd för hantering av forskningsdata från centralt håll inom LU?
- Känner ni till någon forskare eller forskargrupp som är lämplig för deltagande i en pilotstudie för hantering av forskningsdata?
- Vad tror ni om framtiden när det gäller hantering av forskningsdata, reflektioner, tankar?