MALMÖ HÖGSKOLA

Tillgängliggörande av forskningsdata

Nulägesbeskrivning samt rekommendationer för
uppbyggnad av stöd till högskolans forskare

Malmö högskolas bibliotek

Februari 2016

Sammanfattning

Biblioteket har genomfört ett internt projekt med syftet att få en bättre bild av nuläget kring forskningsdatahantering för att utveckla ett adekvat stöd till forskningen och forskarna. Rapporten innehåller en redovisning av en enkät till högskolans forskare, en utblick kring dataarkiv för tillgängliggörande av data och en diskussion kring roller vid tillgängliggörande av data. Rapporten avslutas med rekommendationer kring stödet till högskolans forskare baserat på ett antal identifierade frågeställningar. Tillgängliggörande av forskningsdata är komplext. Inte minst eftersom de data som ska tillgängliggöras bygger på forskning som är komplex. Högskolans forskares syn på frågan är central för utvecklingen av bibliotekets verksamhet.

Områdets komplexitet innebär en rad frågor som framförallt behöver diskuteras på fakulteterna och mellan forskare. Bakgrunden är bland annat Vetenskapsrådets förslag till nationella riktlinjer för öppen tillgång till vetenskaplig information (Vetenskapsrådet, 2015a) och högskolans inspel till forskningspropositionen 2015 (Jönsson, 2015).

En enkät skickades ut till högskolans samtliga forskare (ca 600) i december 2015. Svarsfrekvensen blev 27 % (n=164). 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 öppna 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?

Det råder också en viss begreppsförvirring i svaren, kring tillgängliggörande kontra arkivering och data kontra publikationer. Men också kring vad forskningsdata innebär för olika forskare. För somliga är begreppet helt naturligt, andra använder källmaterial eller empiri. En del samlar inte in material utan genererar data utifrån beräkningsmodeller. I bred bemärkelse är data mer än det som samlas in, det kan också vara annan output från forskningsprocessen än rena publikationer, som enskilda illustrationer eller en presentation.

Tjänster för tillgängliggörande och roller vid tillgängliggörande har i rapporten översiktligt beskrivits, men behöver fördjupas och formaliseras. Högskolans kommande e-arkiv kan utgöra en central funktion vid tillgängliggörande, speciellt om en nationell plattform för öppna forskningsdata skapas. Tillgängliggörande och arkivering av forskningsdata är två likartade, men skilda processer. I vissa fall kan arkiverad data tillgängliggöras rakt av, till exempel data från viss naturvetenskaplig forskning, i andra fall kan data behöva anonymiseras eller på annat sätt anpassas. Alla data kanske inte lämpar sig för tillgängliggörande. Datahanteringsplaner kan vara ett centralt redskap, både före, under och efter avslutat forskningsprojekt. Väl kurerad data är lättare att arkivera och tillgängliggöra. Vad krävs för att data ska bli användbar för andra? Ett svar är att en långsiktighet krävs och att resurser för datahantering behöver inkluderas i ansökningar och projektplaner.

Vetenskapsrådets förslag till nationella riktlinjer för öppen tillgång behöver spridas mer eftersom det kommer påverka alla forskare. Hur information och kommunikation om det ska läggas upp är beroende av om och i så fall hur förslaget kommer att se ut i regeringens kommande

forskningsproposition. Men med tanke på att detta inte är en nationell fråga, utan internationell, och att EU är starkt pådrivande i öppenhetens riktning så kan högskolan troligen räkna med att någon form av krav på öppen tillgång till forskningsdata kommer från regeringen.

Ett antal behov har identifierats som resulterar i rekommendationer för vad som behöver byggas upp i stödet kring tillgängliggörande av forskningsdata.

Under 2016 kommer ett pilotprojekt tillsammans med Svensk Nationell Datatjänst (SND) att genomföras vid högskolan. SND är en nationell forskningsinfrastruktur som kommer förändra sin roll och där lärosätesbiblioteken kommer att bli noder i stödet för tillgängliggörande av forskningsdata. Biblioteket kommer i projektet tillsammans med bland annat arkivet och ett par forskare utforska de processer som krävs för tillgängliggörande av forskningsdata.

Rekommendationer

- *Publicera information på mah.se på de teman som är efterfrågade (biblioteket).*
- *Ta fram riktlinjer kring val av externa dataarkiv för tillgängliggörande (biblioteket).*
- *Ta fram en kortare manual med riktlinjer och lathundar som stöd till forskningsdatahantering (biblioteket).*
- *Utreda e-arkivets roll och resurser med tanke på att det kan bli en central del i en eventuell nationell lösning för tillgängliggörande.*
- *Formalisera det existerande informella nätverket mellan arkiv, bibliotek, forskarservice, juridik och IT.*
- *Forskningsberedningen tillsätter en arbetsgrupp med uppdrag att arbeta fram en färdplan för stödet.*
- *Stimulera en fortsatt högskoleövergripande diskussion om frågor kring forskningsdata där bland annat krocken mellan juridiska aspekter och det forskningsetiska perspektivet tas upp.*
- *Satsa på forskning och utveckling kring hantering av forskningsdata på högskolan. Området behöver beforskas för att skapa bättre förutsättningar och ökad förståelse för olika typer av forskning. En möjlighet kan vara att inkludera forskningsdataaspekter i framtida forskningsansökningar.*

En kvalitativ uppföljning av enkäten behöver också göras för att identifiera ytterligare frågeställningar i dessa komplexa frågor tillsammans med forskare vid högskolan.

Rapporten är framtagen av team publicering och bibliometri, biblioteket. Huvudförfattare är Jonas Fransson. Projektgrupp: Madeleine du Toit, Pablo Tapia Lagunas, Sara Kjellberg och Jonas Fransson.

Innehåll

1	Inledning.....	7
1.1	Tendenser i omvärlden.....	7
1.2	Hantering av forskningsdata som fält	7
1.3	Bibliotekets roll.....	8
1.4	Projektets syfte.....	8
1.5	Disposition.....	8
2	Enkäten.....	10
2.1	Genomförande	10
2.2	Resultat bakgrundsfrågorna (11-16)	11
2.2.1	Fråga 11: Fakultetstillhörighet.....	11
2.2.2	Fråga 12: Forskningscentrum	12
2.2.3	Fråga 13: Ämne.....	12
2.2.4	Fråga 14: Forskarroll.....	13
2.2.5	Fråga 15: Forskning i grupp eller individuellt	14
2.2.6	Fråga 16: Externfinansiering.....	14
2.3	Resultat forskningsdatafrågorna (1-10)	15
2.3.1	Fråga 1: Typ av data.....	15
2.3.2	Fråga 2: Datainsamling	16
2.3.3	Fråga 3: Involvering i datainsamling.....	17
2.3.4	Fråga 4: Användning av öppen eller delad data.....	18
2.3.5	Fråga 5: Tillgängliggjort eller delat forskningsdata	19
2.3.6	Fråga 6: Datahanteringsplan	21
2.3.7	Fråga 7: Krav på öppna data.....	21
2.3.8	Fråga 8: Vilja att öppet dela forskningsdata.....	22
2.3.9	Fråga 9: Arbete inför öppet tillgängliggörande	24
2.3.10	Fråga 10: Incitament för ökat tillgängliggörande	24
2.4	Resultat Uppföljningsfrågorna (17-20).....	26
2.4.1	Fråga 17: Stöd och kunskap.....	26
2.4.2	Fråga 18: Återkoppling från biblioteket	28
2.4.3	Fråga 19: Utlottning.....	28
2.4.4	Fråga 20: Kommentarer.....	28
2.5	Diskussion och slutsatser.....	30
3	Relevanta dataarkiv.....	31
3.1	Olika typer av tillgängliggörande.....	31
3.2	Dataarkiv och deras kontext.....	31

3.3	Kriterier för bedömning av dataarkiv	32
3.4	Relevanta dataarkiv för MAH	32
3.4.1	Generella dataarkiv för allmänt tillgängliggörande.....	32
3.4.2	Lokalt tillgängliggörande av forskningsdata	33
3.4.3	Kommande nationell lösning för tillgängliggörande	33
3.5	Diskussion och slutsatser.....	33
4	Roller vid tillgängliggörande	34
4.1	Olika typer av datahantering.....	34
4.2	Stödfunktioner.....	35
4.3	Organisering av stödet	35
4.4	Stöd på MAH.....	36
4.5	Kommande pilotprojekt	36
5	Identifierade frågeställningar	37
5.1	Avslutande diskussion	37
5.2	Frågeställningar för biblioteket.....	38
5.2.1	Representation i bibliotekssystem	38
5.2.2	Information på webben.....	38
5.2.3	Riktlinjer kring externa dataarkiv	38
5.2.4	Studera <i>best practice</i>	38
5.2.5	Kompetens för att kunna beskriva forskningsdata	38
5.2.6	Rekommendationer.....	38
5.3	Övriga frågeställningar kring stöd för tillgängliggörande.....	38
5.3.1	E-arkivets roll	38
5.3.2	Lokalt nätverk.....	38
5.3.3	Rekommendationer.....	39
5.4	Övergripande frågeställningar.....	39
5.4.1	Ägandeskap och kontroll	39
5.4.2	Skydd av informanter	39
5.4.3	Merarbete	39
5.4.4	Data som begrepp	39
5.4.5	Vilka data ska tillgängliggöras.....	40
5.4.6	Rekommendationer.....	40
6	Referenser	41
7	Bilagor.....	42
7.1	Inbjudan.....	42
7.1.1	Svensk version	42

7.1.2	Engelsk version	42
7.2	Enkäten på svenska	43
7.2.1	Introduktionstext.....	43
7.2.2	Frågor om forskningsdata.....	43
7.2.3	Bakgrund respondent.....	45
7.2.4	Uppföljning	46
7.3	Enkäten på engelska.....	47
7.3.1	Introduction to the survey.....	47
7.3.2	Research data questions	47
7.3.3	Bakgrund respondent.....	50
7.3.4	Uppföljning	51
7.4	Tackmejl.....	51

1 Inledning

1.1 Tendenser i omvärlden

Det ökade intresset för öppet tillgängliggörande av forskningsdata har flera orsaker. Från politiskt håll och från vissa forskningsfinansiärer drivs frågan främst som en resursfråga. Data framtagen eller insamlad med offentliga medel ses som en offentlig resurs som kan komma fler till nytta. En annan ingång är kvalitet i forskningen och att det ska vara lätt att granska forskningsresultat. En tredje är forskare som vill vara öppna och få feedback under forskningsprocessen. I Vetenskapsrådets (VR) förslag på riktlinjer nämns följande argument för öppet tillgängliga data (Vetenskapsrådet, 2015a, p. 16):

- **Demokrati och transparens.** *Forskningsresultat bekostade av offentliga medel bör av princip vara tillgängliga för medborgarna. Tillgång till de data som använts för att påvisa ett resultat som presenteras i till exempel en artikel är nödvändig för att kontrollera att slutsatserna stämmer.*

- **Forskning.** *Öppen tillgång till forskningsdata kan leda till nya forskningsmöjligheter, särskilt inom tvärvetenskapen. Med öppen tillgång till forskningsdata behöver experiment i vissa fall inte upprepas, utan tidigare resultat kan återanvändas. Data kan också användas inom nya forskningsfält och genom att kombinera olika datamängder kan forskare adressera nya frågeställningar. Vidare underlättas utvecklandet av nya analysmetoder om forskare har tillgång till experimentella data. För dataintensiv forskning (populärt kallat big data) där olika datakällor integreras för att se nya samband är det en förutsättning med tillgång till stora datamängder.*

- **Innovation och användning utanför forskningen.** *Lättillgänglig data har också potential att kunna användas utanför forskarsamhället till exempel av företag, privatpersoner och andra myndigheter.*

- **Citering.** *Forskare som återanvänder forskningsdata ska referera till den som ursprungligen tagit fram forskningsdata. Det kan ske genom referens till en publikation som beskriver data eller direkt till datamängden. Öppen tillgång till forskningsdata kan därmed leda till ökat antal citeringar, vilket då blir meriterande för forskargruppen som producerat den.*

1.2 Hantering av forskningsdata som fält

Hantering av forskningsdata (*Research Data Management, RDM*) är arbetet med forskningsdata under hela forskningsprocessen, från planering och ansökan till avslutande och arkivering. De största delarna av hanteringen åligger forskarna under forskningsprocessen och praktikerna varierar mellan lärosäten, ämnen och grupper av forskare. I Sverige har inga samlade grepp tagits kring *best practice* i hantering av forskningsdata, lösningar för elektronisk arkivering har saknats på lärosätena och bara i enskilda discipliner har forskningsdata öppet tillgängliggjorts. I till exempel Storbritannien har man under en längre tid arbetat med RDM.

Inom fältet finns det många aktörer. Forskningsfinansiärer driver på frågan kring öppna forskningsdata, vissa vetenskapliga tidskrifter har policies för forskningsdata, tidskrifter för beskrivningar av öppna data tillkommer (så kallade *data journals*), nationella datacenter utvecklar och kompetensutvecklar, lärosäten skapar och driver dataarkiv (så kallade *data repositories*) parallellt med sina publikationsdatabaser, samt utvecklar olika typer av stöd till forskare. I EU finns

en vilja på politisk nivå att öka tillgängliggörandet av forskningsdata och vissa EU-finansierade projekt har haft krav kring datahanteringen och tillgängliggörandet. Det är i detta föränderliga fält som forskare, forskargrupper och institutioner måste navigera för att uppfylla nuvarande och kommande krav.

Öppet tillgängliggörande av forskningsdata är en del av forskningsdatahantering och kan omfatta bara vissa delar av det samlade materialet. På det sättet är öppet tillgängliggörande skilt från arkiveringen där all insamlad data ses som offentlig handling och ska arkiveras. Sker ett löpande strukturerat arbete med att dokumentera hanteringen av forskningsdata blir både arkivering och tillgängliggörande av data enklare.

1.3 Bibliotekets roll

I bibliotekets uppdrag ingår att bevaka och sprida information om utvecklingen av vetenskaplig publicering och forskningsinformation där tillgängliggörande av forskningsdata ingår. Enkäten är en del av detta arbete och syftar till att ge ökad förståelse för forskningens och forskarnas behov av stöd. Målet är att utveckla ett fullgott stöd i högskolan och att särskilt identifiera vad bibliotekets stöd till högskolans forskare ska innehålla

Bibliotekets intresse kring forskningsdatahanteringen grundar sig på en utveckling av bibliotekets nuvarande verksamheter. Att beskriva objekt i form av katalogposter eller metadata är en central del av bibliotekets verksamhet. Likaså att tillgängliggöra olika typer av samlingar. Stöd till publicering har blivit en viktig del av forskningsstödet, och är delvis kopplat till Open Access för publikationer. Tillgång och upptäckbarhet via söksystem är centralt för återanvändning av publikationer. Samma processer gäller för tillgängliggörande av forskningsdata och frågorna hänger delvis ihop.

1.4 Projektets syfte

Text från projektplanen:

Bakgrund: Praktiken kring forskningsdata är i förändring i vetenskapssamhället. Att dela och återanvända öppna forskningsdata ses både som en god forskningspraktik och som ett gott ekonomiskt tänkande. Internationellt kräver allt fler forskningsfinansiärer att forskningsdata ska göras tillgänglig. Det kan också i vissa publikationer ingå att man ska ge tillgång till data som forskningsresultaten bygger på.

För högskolans del kan öppen tillgång till forskningsdata stärka utvecklingen för nyttiggörande och innovation för samhällsutveckling.

Syfte: Projektets syfte är att skapa en utgångspunkt för det fortsatta arbetet med tillgängliggörande av forskningsdata vid Malmö högskola. Syftet är också att sätta tillgängliggörandet i relation till närliggande aktiviteter som stöd vid upprättande av datahanteringsplaner för forskningsprojekt, datahantering under forskningsprocessen, samt arkivering av forskningsdata och -dokument.

Aktiviteter: Några viktiga aktiviteter är att undersöka hur forskarna ser på och vilka erfarenheter de har av öppna forskningsdata och delning av data, att kartlägga de externa tjänster som idag finns för tillgängliggörande av forskningsdata, samt att identifiera centrala frågeställningar som resulterar i ett antal förslag för det fortsatta arbetet med forskningsdata.

1.5 Disposition

Rapporten är disponerad i några tydliga delar. Kapitel 2 behandlar resultaten av enkäten och ger en bild av nuläget vid Malmö högskola. Kapitel 3 behandlar tjänster för de forskare som vill tillgängliggöra data idag, men även e-arkivets kommande roll berörs kort. I kapitel 4 diskuteras roller

vid tillgängliggörandet och organisering av stöd på MAH. Rapporten avslutas med kapitel 5 med en avslutande diskussion, samt presentation av identifierade frågeställningar och rekommendationer. Enkätfrågor och utskick kring enkäten återfinns som bilagor.

2 Enkäten

2.1 Genomförande

Tio enkätfrågor utformades för att täcka in olika aspekter kring forskningsdata för en översiktlig bild. Samtidigt var målet att hålla ner antalet frågor i enkäten och därmed göra den lätt att besvara för att öka svarsfrekvensen. Utöver frågorna kring forskningsdata så lades bakgrund- och uppföljningsfrågor till för att ge svaren en kontext i det vidare arbetet. (Se bilagorna 7.2 och 7.3 för frågorna på svenska och engelska). Enkätfrågorna har varit allmänt formulerade för att passa in på alla typer av forskning. Därmed har frågorna ibland blivit uppfattade på olika sätt av respondenterna, något som troligen är svårt att komma ifrån i en kortfattad och övergripande enkät.

Forskare (inklusive anställda doktorander) valdes ut baserat på kategorier i personalsystemet Primula. Inget urval gjordes i gruppen. De aktuella grupperna i Figur 1 valdes för att säkerställa att utskicket gick till personal med forskningstid. Det innebär att potentiella respondenter i andra personalkategorier valdes bort. Utskicket med inbjudan till enkäten sändes till 601 forskare med inlagd och aktiv e-postadress (se 7.1 för inbjudan).

A screenshot of a list of person categories from the Primula system. The list is displayed in a white box with a thin grey border. The categories are listed in a standard sans-serif font, with some entries containing double slashes (//) to indicate that other categories have been excluded. The categories are: Doktorand, Forskare, Lektor, Postdoktor, Professor, Professor, adjungerad, Professor, biträdande, Professor, gäst, Professor, kallad, Professor/Övertandläkare, Universitetslektor, Universitetslektor // Forskare, Universitetslektor // Postdoktor, Universitetslektor, adjungerad, Universitetslektor, befördrad universitetsadjunkt, and Universitetslektor/Övertandläkare.

- Doktorand
- Forskare
- Lektor
- Postdoktor
- Professor
- Professor, adjungerad
- Professor, biträdande
- Professor, gäst
- Professor, kallad
- Professor/Övertandläkare
- Universitetslektor
- Universitetslektor // Forskare
- Universitetslektor // Postdoktor
- Universitetslektor, adjungerad
- Universitetslektor, befördrad universitetsadjunkt
- Universitetslektor/Övertandläkare

Figur 1. Personalkategorier från Primula inkluderade i undersökningen (där // förekommer har andra skiljetecken tagits bort).

Enkätverktyget *Sunet Survey* (*Survey & Report* från Artologik) användes. Två direktadresserade utskick gjordes, en inbjudan till enkäten (se bilaga 7.1) samt en påminnelse sex dagar senare. Båda utskicken var tvåspråkiga (svenska/engelska). Direktmejl skickades via enkätverktyget för att ha kontroll över vilka som bjöds in till enkäten samt vilka som svarat. Enkäten var aktiv tio dagar och de flesta svaren kom i anslutning till e-postutskicken.

Totalt 164 svar, det vill säga svarsfrekvens 27 %. (se avsnitt 2.2.1 och 2.2.4 för mer information). Alla svar var anonyma. I enkäten var det frivilligt att ange e-post för att bli kontaktad under våren eller för att man ville delta i utlottningen av luncher. Angivna e-postadresser har inte och kommer inte relateras till enkätsvaren. Den engelskspråkiga versionen av enkäten användes av 6 procent av respondenterna.

2.2 Resultat bakgrundsfrågorna (11-16)

Bakgrundsfrågorna kom efter frågorna kring forskningsdata i enkäten för att enkäten presenterades som "tio snabba frågor kring forskningsdata". Här presenteras svaren på bakgrundsfrågorna först som en kontext till svaren på frågorna 1-10.

2.2.1 Fråga 11: Fakultetstillhörighet

Vilken fakultet tillhör du huvudsakligen?

Figur 2. Resultat fråga 11: Fakultetstillhörighet.

Underlaget från Primula för utskicket innehöll inte fakultetstillhörighet, men antalet respondenter per fakultet kan ställas mot antal årsarbetskraft per fakultet redovisad i högskolans årsredovisning (Malmö högskola, 2015) för en indikation på svarsfrekvensen. Det ska noteras att fakulteterna kan ha olika fördelning mellan olika personalgrupper, till exempel lektorer i förhållande till adjunkter, och det påverkar hur många från varje fakultet som inkluderats i enkäten.

	Antal respondenter	Årsarbetskraft 2014	Andel respondenter ifht. årsarbetskraft
HS	37	218	17%
KS	47	279	17%
LS	50	276	18%
OD	16	221	7%
TS	12	97	12%

Tabell 1. Antalet respondenter per fakultet i förhållande till talet för årsarbetskraft 2014 (Malmö högskola, 2015).

Tillhör du något forskningscentrum?

Forskningssentrum	Antal respondenter
Biofilmer – Research...	~5
Centrum för profess...	~3
Centrum för sexolog...	~4
Centrum för tillämp...	~3
Institutet för stud...	~1
Internet of Things ...	~5
Malmö Institute for...	~7
Nej	~115

21 % av respondenterna angav att de tillhör något av MAH:s forskningscentrum.

Inom vilket ämne/ämnena forskar du?

2.2.4 Fråga 14: Forskarroll

Vilken forskarroll har du?

Figur 5. Resultat fråga 14: Forskarroll

Frågan ger en grov indelning på senioritet som forskare. Dels för att kunna bryta ner enskilda frågor kring forskningsdata på kategorierna. Dels för att veta mer om vilka som svarat på enkäten. Svarsfrekvensen är högre i gruppen professorer. Antalet doktorander som har fått inbjudan till enkäten (baserat på information i Primula) är betydligt lägre än antalet som är identifierade i bibliotekets doktorandstödsprojekt¹. Det kan bero på att forskarstudier bedrivs vid högskolan men att personen har en anställning någon annanstans.

	Inbjudna	Svarat	Svarsfrekvens
Professor	95	36	38%
Lektor, postdoktor, eller annan forskartjänst	393	100	25%
Doktorand	113	28	25%
	601	164	27%

Tabell 2. Svartsfrekvens per forskarroll.

¹ <http://mah.se/Om-Malmo-hogskola/Om-Biblioteket/Projektkontor/Pagaende-projekt/BITdoktorandstod/>

2.2.5 Fråga 15: Forskning i grupp eller individuellt

Forskar du främst i grupp eller individuellt?

Figur 6. Resultat fråga 15: Forskning i grupp eller individuellt.

2.2.6 Fråga 16: Externfinansiering

Är din nuvarande forskning externfinansierad?

Figur 7. Resultat fråga 16: Externfinansiering.

Externfinansiering bör indikera en större grad av forskningstid. Forskningstiden som är inkluderad i respektive tjänst är starkt begränsad, förutom för till exempel doktorander och postdoktorer.

2.3 Resultat forskningsdatafrågorna (1-10)

Målet med enkätfrågorna kring forskningsdata har varit att få en uppfattning om vilka erfarenheter av och syn på tillgängliggörande av data det finns bland Malmö högskolas forskare. Bibliotekets fokus på tillgängliggörande har färgat enkäten och vissa frågor har upplevts som svåra att svara på.

2.3.1 Fråga 1: Typ av data

Vilken eller vilka typer av data arbetar du med? (flera val möjliga)

Figur 8. Resultat fråga 1: Typ av data.

Textuell data är den mest frekvent använda typen av forskningsdata. Många arbetar med flera olika typer av data, i genomsnitt angavs 2,5 datatyper per respondent.

Att begreppet forskningsdata kan vara problematiskt framgår av följande kommentar:

Ser inte det material jag huvudsakligen arbetar med som "data" - oerhört problematiskt begrepp. Väljer trots det att svara på frågeställningarna utifrån en tolkning av begreppet i bred bemärkelse som källmaterial.

Men det behöver inte heller vara insamlad data eller empiri, tre svar angivna under "annat (ange nedan)":

Akademisk litteratur: böcker och tidskrifter.

workshop and performances (embodied data)

Alstrar egen data från teori

2.3.2 Fråga 2: Datainsamling

Vilken typ av metod för datainsamling använder du huvudsakligen?

Figur 9. Resultat fråga 2: Datainsamling.

Ett flertal respondenter kommenterar att de använder flera typer av metoder och då både kvalitativa och kvantitativa. Spännvidden är stor vid högskolan när det gäller insamlingsmetoder, här citeras några axplock bland de angivna metoderna:

Etnografisk metod som innehåller analys av dokument, dolda observationer, deltagande observationer, intervjuer

Data mining, regression, machine learning

Alstrar egen data från teori via beräkning på dator

register data och frågeformulär för analys med kvantitativ design samt svar som fri text för kvalitativ analys

Begreppshistoria

Samlar även in språkliga data från webben.

Registerdata

Jag använder både kvalitativ och kvantitativ i lika stor utsträckning. Analys av text enligt innehållsanalys

Abduktiv, praktikbaserad metod på kvalitativ grund, med stöd av statistik (kvantitativ metod), visuellt material samt litteraturstudier.

Regressions analys och gis [Geografiska informationssystem]

Olika typer av enkäter

Filmade fokusgrupper

2.3.3 Fråga 3: Involvering i datainsamling

I vilken grad är du som forskare en del av själva datainsamlingen?

Figur 10. Resultat fråga 3: Involvering i datainsamling.

80 % anser sig i stor grad vara en del av själva datainsamlingen. Svaren kan relateras till fråga 2 där 58 % svarade att de huvudsakligen använder kvalitativ metod för datainsamlingen. Det innebär att även ett antal forskare som i huvudsak använder icke-kvalitativ metod ser sig vara en del av datainsamlingen i stor grad. Det belyser forskarnas roll i forskningsprocessen, det vill säga någon som inte är lätt utbytbar och som har ett nära band till de data som samlas in.

Noteras bör också att bara 6 % svarar *ingen grad*, medan i fråga 2 så anger 12 % att de inte samlar in data själva.

2.3.4 Fråga 4: Användning av öppen eller delad data

Har du i din forskning använt öppen data? (Öppen data i meningen att den är allmänt och enkelt tillgänglig.) (flera val möjliga)

Figur 11. Resultat fråga 4: Användning av öppen eller delad data.

74 % har använt öppna eller informellt delade data (från forskare eller från myndigheter och organisationer) i sin forskning. Frågan kan dock tolkas på olika sätt, vad man lägger i begrepp som öppna data från myndigheter eller informellt delad data skiftar. Tendensen är dock tydlig, det är inte ovanligt att använda data som man inte samlat in själv.

	HS	KS	LS	OD	TS
Jag har använt öppna data från myndigheter och organisationer	38%	70%	68%	31%	55%
Jag har använt öppen data från andra forskare	16%	47%	34%	13%	36%
Jag använt informellt delad forskningsdata	22%	34%	18%	19%	27%
Jag har använt data från myndigheter och organisationer som inte är öppen	30%	34%	14%	25%	9%
Jag har inte använt någon av ovanstående typer av data	43%	6%	26%	44%	27%

Tabell 3. Användning av öppen eller delad data, svar per fakultet. Observera att det är ett litet antal svar från OD och TS.

	Professor	Lektor, postdoktor, eller annan forskartjänst	Doktorand
Jag har använt öppna data från myndigheter och organisationer	50%	61%	50%
Jag har använt öppen data från andra forskare	31%	33%	25%
Jag använt informellt delad forskningsdata	28%	24%	21%
Jag har använt data från myndigheter och organisationer som inte är öppen	33%	21%	21%
Jag har inte använt någon av ovanstående typer av data	28%	25%	29%

Tabell 4. Användning av öppen eller delad data, svar per forskarroll.

2.3.5 Fråga 5: Tillgängliggjort eller delat forskningsdata

Har du delat eller tillgängliggjort forskningsdata? (flera val möjliga)

Figur 12. Resultat fråga 5: Tillgängliggjort eller delat forskningsdata.

66 % har delat eller tillgängliggjort forskningsdata. De olika svarsalternativen kan kanske ses som en skala där många delat med sig av data informellt medan få har tillgängliggjort data via ett datacenter eller ett dataarkiv. En tredjedel har inte delat eller tillgängliggjort forskningsdata.

	HS	KS	LS	OD	TS
Ja, jag har delat med nära forskarkollegor (informellt)	59%	62%	54%	56%	64%
Ja, jag har delat forskningsdata efter direkt förfrågan	16%	23%	24%	6%	36%
Ja, jag har tillgängliggjort data på webbplats (forskningsprojektets eller personlig)	8%	15%	8%	19%	36%
Ja, jag har tillgängliggjort data via datacenter eller dataarkiv (om ja, ange gärna i vilken tjänst nedan)	3%	6%	8%	0%	9%
Nej	38%	32%	36%	44%	9%

Tabell 5. Tillgängliggjort eller delat forskningsdata, svar per fakultet. Observera att det är ett litet antal svar från OD och TS.

	Professor	Lektor, postdoktor, eller annan forskartjänst	Doktorand
Ja, jag har delat med nära forskarkollegor (informellt)	58%	61%	54%
Ja, jag har delat forskningsdata efter direkt förfrågan	22%	22%	14%
Ja, jag har tillgängliggjort data på webbplats (forskningsprojektets eller personlig)	8%	12%	21%
Ja, jag har tillgängliggjort data via datacenter eller dataarkiv (om ja, ange gärna i vilken tjänst nedan)	6%	6%	4%
Nej	36%	31%	39%

Tabell 6. Tillgängliggjort eller delat forskningsdata, svar per forskarroll.

Fritextkommentarerna visar att frågan inte är helt lätt att svara på. Data är ibland väldigt tydligt inkluderad i publikationer och därmed anser en del respondenter att data har tillgängliggjorts via till exempel MUEP (MAH:s publikationsdatabas). Flera nämner att data används i undervisningssyfte, men tillgängliggörandet sker i publikationer:

Genom artiklar! Använder data aktivt i min undervisning!

Att öppet tillgängliggöra forskningsdata kan vara problematiskt av olika skäl som belyses i nedanstående fritextkommentarer:

Filmerna kan inte delas av etiska skäl

Menar ni att man kan vara generös med sin insamlade data? Risken om man lägger ut sin rådata är ju att andra använder den och publicera studier före mig som jag tänkt publicera.

Arbetar med kommersiella data som är företagskritiska och därmed inte är öppna

Ett sätt att säkerställa insyn och transparens bland andra som därmed kan kontrollera min _professionalitet_ i hanteringen av materialet – och det är vad som ska säkerställas, inte att exakt vad som finns i mina källor är tillgängligt öppet, det vill säga utan min som projektansvarigs kontroll. Det senare vore direkt oprofessionellt av mig att tillåta, framför allt ur etiskt perspektiv. Att hantera material är att hantera människor och deras säkerhet. Min forskning skulle inte kunna bedrivas i det sammanhang som den här enkäten implicit refererar till. Det vill säga frågeställningar och material skulle inte kunna beforskas och det gäller generellt inom mitt område. Etiken jag hänvisar till är inte begränsad till nytta för akademien (vilken i princip kan överskrida en hel del då gränsen går vid vad akademiska auktoriteter önskar beforska), utan rör akademins eget existensberättigande som del av ett öppet samhälle. Det senare kan inte skyddas eller främjas av möjligheten till kontroll (i öppen data) från vem som helst över vad någon specifik person har sagt ang. viktiga samhällsfrågor. De jag intervjuat hade inte kunnat vara med i studien och därmed hade kunskapen som nu vi genererar förblivit otillgänglig. Öppen data innebär alltså otillgänglig kunskap och demokratiskt omotiverad kontroll.

Viss forskning kan vara lätt att göra öppen eller har designats för att vara öppen:

I am just starting my data collection, but I plan to make everything open on my personal website. I would be happy to contribute to a repository!

Det kan också handla om traditioner som skiljer sig åt mellan lärosäten eller ämnen:

När jag arbetade vid Lunds universitet fanns tradition att tillgängliggöra data. Detta har jag aldrig sett möjligheten till vid MAH/LS.

Tjänster för tillgängliggörande som nämns i kommentarerna är bland annat EUDO Citizenship Observatory (European Union Democracy Observatory on Citizenship) och NIST (National Institute of Standards and Technology).

2.3.6 Fråga 6: Datahanteringsplan

Har du sökt externa medel där det krävts en datahanteringsplan vid ansökan?

Figur 13. Resultat fråga 6: Datahanteringsplan.

Datahanteringsplaner kan vara ett värdefullt verktyg både under forskningsprocessen och för arkivering och tillgängliggörande. För forskningen kan datahanteringsplaner bidra till både ökad kvalitet och effektivitet (Eloranta, Johansson, Kristinsson, & Andersson, 2013).

2.3.7 Fråga 7: Krav på öppna data

Har forskningsfinansiärer krävt av dig att forskningsdata ska göras öppen efter avslutat forskningsprojekt?

Figur 14. Resultat fråga 7: Krav på öppna data.

Svaren speglar de krav som funnits, det vill säga att få finansiärer har krävt datahanteringsplaner hittills. I till exempel Horisont2020 är det enbart ett antal pilotutlysningar som krävt datahanteringsplan.

2.3.8 Fråga 8: Vilja att öppet dela forskningsdata

I vilken grad kan du tänka dig att öppet dela den forskningsdata du arbetar med?

Figur 15. Resultat fråga 8: Vilja att öppet dela forskningsdata.

Frågan är ställd i syfte att stämna av viljan att öppet dela den forskningsdata forskare arbetar med. Här är det relativt jämnt mellan svarsalternativen; stor grad, i viss grad och i liten grad. I kommentarerna har framkommit kritik mot att de två sista svarsalternativen överlappar, vilket det kan tolkas som. Tanken var att det sista alternativet skulle väljas om frågan inte var relevant, medan svarsalternativ 1-4 handlar om viljan. Data som innehåller till exempel personuppgifter kan i en del fall anonymiseras och sedan tillgängliggöras.

Figur 16. Vilja att öppet dela forskningsdata fördelat per fakultet.

	HS	KS	LS	OD	TS
Stor grad (allt eller så mycket som möjligt)	8%	36%	24%	6%	27%
I viss grad (större utvalda delar)	27%	19%	26%	31%	36%
I liten grad (mindre utvalda data)	27%	21%	22%	25%	0%
Inte alls	14%	9%	2%	6%	0%
Datan går inte dela pga. sekretess eller liknande	24%	15%	26%	31%	36%

Tabell 7. Vilja att öppet dela forskningsdata, svar per fakultet.

Kommentarerna kring viljan att öppet tillgängliggöra forskningsdata kan delas in i olika teman:

Sekretess och skydd av deltagare:

Konfidentiell data och hanteras så i ök med informanterna.

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

Jag undersöker några få elever och jag kommer endast att dela min data inom det större forskningsprojektet jag deltar. Eftersom det handlar om minderåriga (13-15åriga barn) är det av vikt att behandla deras utsagor och handlingar konfidentiellt.

Forskningsdatan kan inte göras öppen av forskningsetiska skäl då det skulle innebära att brott mot konfidentialitetsprincipen och det löfte om anonymitet som ingår i forskningsansatsen. Vissa delar av materialet är sekretesskyddat, andra delar bara känsligt om än materialet är offentligt. Ett offentligt material (t ex polisförhørsutskrifter eller domar) är inte "offentligt" ur forskningsetiskt perspektiv utan ska behandlas enligt både PUL och olika forskningsetiska principer för hantering av känsligt material och personuppgifter.

Det beror på.... en del är och måste vara sekretessbelagd

Kontroll:

Beror på typ av data och egen användning (dvs t ex hur mycket jag publicerat själv där data varit del)

i use proprietary datasets that cannot be shared

Lite svår fråga. I vissa fall till exempel när det gäller SOM-data får man inte dela denna till andra. När det gäller WVS eller ESS kan vem som helst på in och ladda ner denna, så den behövs inte delas.

This needs to be decided by the researcher on a case-by-case basis given the varying degrees of sensitivity involved in research.

Nyttan för andra:

Arbetar man kvalitativt är det djupt problematiskt, eftersom materialet är starkt knutet till de forskningsfrågor man arbetar kring och troligen mycket svårtolkat för andra som inte har samma ontologiska och epistemologiska ingångar. Det är dessutom etiskt mycket tveksamt, och således inget alternativ för min forskning.

Andra aspekter:

Jag analyserar publicerade texter som redan är tillgängliga för allmänheten.

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

2.3.9 Fråga 9: Arbete inför öppet tillgängliggörande

Hur mycket arbete krävs för att den data som du arbetar med ska kunna göras öppen och användbar för andra?

Figur 17. Resultat fråga 9: Arbete inför öppet tillgängliggörande.

Svaren på frågan belyser det faktum att i de flesta fall så krävs en större eller mindre arbetsinsats för att materialet ska kunna tillgängliggöras och vara användbar för andra. Som tidigare nämnts kan data anonymiseras före tillgängliggörandet, men detta kräver också en arbetsinsats.

2.3.10 Fråga 10: Incitament för ökat tillgängliggörande

Vad skulle få dig att börja, eller i större utsträckning än idag, öppet tillgängliggöra forskningsdatan som du arbetar med? (flera val möjliga)

Figur 18. Resultat fråga 10: Incitament för ökat tillgängliggörande.

I genomsnitt har respondenterna angivit 2,6 svar på frågan, men hela 29% har angivit *vet ej*. Svaren på frågan visar hur komplext öppet tillgängliggörande av forskningsdata är. Det incitament som angivits av flest är efterfrågan av forskare i samma ämne (40%). Att avsatt tid är det näst mest frekventa svaret tyder på att det finns en vilja (se fråga 8) men att det kräver extra arbete (se fråga 9). Ur detta perspektiv blir det en resursfråga hur den tillgängliga forskningstiden bäst används.

Knappt en tredjedel svarar att om tillgängliggörande av forskningsdata skulle ses som en merit så skulle det fungera som drivkraft. Lika många svarar att krav från finansierare skulle vara ett incitament, medan en lokal policy inte ses lika motiverande.

Kommentarer i urval:

Om jag fick ett stort anslag som gjorde att jag kunde hålla på med forskning i flera år och ha tid att själv ordentligt bearbeta materialet skulle jag gärna dela med mig. Ofta är dock medlen man får små och arbetsinsatsen stor. Då vill man inte ge den till andra utan att själv ha arbetat ordentligt med materialet.

Med anledning av att mitt arbete är direkt kopplat till näringslivsparter kommer det aldrig vara tal om att dela datan, annat än genom artiklar som publiceras och före detta granskats för känsliga uppgifter. Vad jag kan säga är att policy vid MAH är direkt galet, med anledning av att detta ställer krav som går emot vad finansierare (inklusive näringslivsparter vi arbetar med). Det måste vara lätt att göra forskning på MAH, inte bli svårare!!!

Detta är en komplex fråga med många svar - alla möjligheter ovan måste beaktas och det finns förmodligen mer att tänka på

The idea is alien to me.

När man arbetar med kvalitativ metod och djupintervjuer skulle det vara etiskt problematiskt att tillgängliggöra sina data.

Att tillgängliggöra forskningsdata får inte krocka med rådande lagstiftning, exempelvis PUL och Lagen om etikprövning av forskning som avser människor. Inte heller får den krocka med Datainspektionens anvisningar.

[...] Om det fanns krav om "öppenhet" (dvs kontroll) av material skulle jag söka ett annat jobb. Jag som projektansvarig med många års utbildning inom forskning måste få vara betrodd att avgöra vad och hur något innebär "sekretess eller liknande". Kontrollen av god forskningssed och hantering av "data" (material) kan göras ändå - och _det_ är vad akademien och andra som ska ta del av forskningen behöver (mer av).

Inget kan få mig att öppet tillgängliggöra rådatan för forskningen då det skulle innebära att jag utsätter informanter och andra observerade för integritetskränkning.

Gör det redan, alternativen passar inte därför har jag satt vet inte - felaktigt formulerad fråga

Svårt att tillgängliggöra existerande datamängder eftersom de inte är öppna. Krav från finansiärer och ökad impct skulle göra det intressantare att arbeta med öppen data.

Har inte alls någon avsikt att dela forskningsdata till fler än de personer jag samverkar med och då behövs ingen plattform för detta.

Att öppet tillgängliggöra empiri som samlats in via kvalitativa forskningsintervjuer kräver stor eftertänksamhet, då innehållet kan vara av mycket konfidentiell art. Ofta anges dessutom konfidentialitet för att få informanter och att endast utdrag/citat som inte kan härledas till den enskilde kommer att publiceras, resten presenteras via teman, mönster etc. tillsammans med övrigt insamlat material.

Ej aktuellt pga sekretess av patientjournaler mm.

Lös sekretessproblemet och begränsningen som finns för att dela data med annat land. Borde gå utan att åsidosätta patienternas självklara rätt till sekretess.

Kvalitativt material svårt att dela hur som helst, måste bearbetas och tillgängliggörss genom färdigställt material.

Kommentarerna pekar främst på olika typer av hinder för öppet tillgängliggörande av data. Tillgängliggörande ska naturligtvis följa rådande lagstiftning. Är det personuppgifter som hindrar tillgängliggörande kan materialet anonymiseras.

2.4 Resultat Uppföljningsfrågorna (17-20)

2.4.1 Fråga 17: Stöd och kunskap

Vad behöver du mer kunskap om alternativt stöd kring för att göra forskningsdatan som du arbetar med öppen? (flera val möjliga)

Figur 19. Resultat fråga 17: Stöd och kunskap.

Vad behöver du mer kunskap om alternativt stöd kring för att göra forskningsdata 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	11%

Tabell 8. Svaren på fråga 17: Stöd och kunskap rangordnade efter antal svar.

Svaren på frågan belyser vad biblioteket och övriga stödfunktioner behöver fördjupa sig i och ta fram information kring för att framöver kunna ge ett adekvat stöd. Kommentarer på frågan (några snarlika svar är borttagna):

se ovan om sekretess

ej aktuellt pga patientjournalagen mm

jag är rädd att alla dessa arkiverings aktiviteter kommer att minska min forskningstid ännu mer.

Åter igen, det är inte - och kommer inte - vara relevant med öppen forskningsdata.

Hur förhåller vi oss till etikprövningsnämnderna??

Går inte att göra min data öppen

*Mer kunskaper om etiskt regelverk vid tillskapande av en delvis öppen internationell databas för kliniska data (initialt tillgänglig för deltagare i projektet men på sikt även för utomstående som efterfrågar data).
I en forskarnätverk där jag ingår är vi just nu i färd med att planera för en sådan.*

releasing sensitive data

I do not intend to make my data open.

Att lagstiftningen ändras.

Etiska aspekter

Etik kring att göra forskningsdatan öppen - för att skydda respondenterna.

Återigen en ledande fråga. Enkätens beställare räknar med att det är möjligt att dela forskningsmaterial öppet (dvs möjlig att kontrollera oavsett varför) oavsett att forskare är emot det?

Svårt att veta vilken kunskap som behövs då jag hittills inte alls upplevt krav på att tillgänglighålla empiri som samlats in via intervjuer.

Inget stöd. Vill inte bidra med öppna data.

inget särskilt behov

*blir det ett krav kommer det att påverka den forskning som är möjlig att göra.
redan idag har vi stora problem med kringskuren forskning inom hum-sam*

Inte aktuellt

Kunskap!

2.4.2 Fråga 18: Återkoppling från biblioteket

Om du vill bli kontaktad av biblioteket rörande dina svar i ovanstående fråga ange namn och epost nedan:

14 respondenter önskade att bli kontaktade rörande sina svar på fråga 17. Biblioteket kommer löpande ta kontakt under våren.

2.4.3 Fråga 19: Utlottning

Om du vill delta i utlottningen av luncher ange epost nedan:

Utlottning har genomförts och vinnarna har fått sina presentkort.

2.4.4 Fråga 20: Kommentarer

Övriga kommentarer:

I fritextfältet "övriga kommentarer" i enkäten kom elva kommentarer in. Hänvisningar till tidigare svar är borttagna samt ett undertecknande.

Excellent first step! Happy to meet and exchange with colleagues on this subject!

Deltar gärna i workshop, går på föreläsningar mm om öppna data. Viktigt att vara förberedd om kravet kommer från fler finansiärer och möjlighet att sprida sitt eget arbete, få samarbetspartners mm. Så bra att ni tar tag i detta!

Tillgång till viss öppen data vore mycket bra som ett pedagogiskt utvecklingsmedium särskilt vad gäller undervisning i och övning kring forskningsmetoder, design och analyser för studenter. Lycka till

Comment on question 5: Answering options offered do not include the possibility most common in social science research, namely making research data available through direct quotes from interviews in research-based publications. The way questions 9, 10 and 17 are formulated, leaves me with the uneasy feeling that a policy is emerging, which does not leave a choice of NOT sharing the data openly. Please be so kind as to provide all respondents with the information on the background and the implications of the statement from your letter: "The results will be a base for the Library's continued work with open research data".

Varför dessa frågor? finns en dold agenda att forskningsdata, opublicerade måste göras tillgängliga...?

Det första ni borde göra är att försäkra er om att ert initiativ inte krockar med rådande lagstiftning. Inom väldigt många av högskolans forskningsområden behandlas känsliga personuppgifter respektive biologiskt material från levande människor. Dessa forskningsdata faller under särskilda lagar, fr.a. biobankslagen, PUL, etikprövningslagen och sekretesslagen. Om till exempel Regionala etikprövningsnämnden ger tillstånd att genomföra ett forskningsprojekt kan det finnas upplagor som förhindrar att forskningsdata blir "öppna". Jag har diskuterat detta med Peter Jönsson och jag tror att vi borde ha ett möte kring detta eller så borde ni kontakta Regionala etikprövningsnämnden i Lund och Datainspektionen direkt.

Frågorna är svåra att applicera på mitt forskningsområde, filosofi.

I can also add that in the Living Archives project (VR funded, based at K3) we are in the process of publishing a series of articles on open data that begin by questioning openness and data. See (livingarchives.mah.se) This is a necessary approach to data coming from the Arts and Humanities where we do not necessarily operate with a pre-determined sense of what data might be. From the disciplines I work within, the question is not just how to handle existing data but to question the fundamentals of a data society.

I am an advocate of open data and open access to research. My research is largely qualitative so I have been struggling to set up a good system for opening my data, but I would very much appreciate help in doing so. I would also be happy to lead the way helping others in my department :) Thanks so much for the work that you do, librarians and archivists are the best!!

One of the main issues with making data openly available is that to ensure that the original researcher is kept informed on who is using that data. If it becomes fully open the problem is that it would undermine the possibilities for research

collaboration since a key motivating force behind such cooperation is data sharing, whereas if data is publicly available there is little incentive.

Hej, det känns inte som den här enkäten vänder sig till forskare verksamma inom humaniora. Frågorna är i vilket fall inte helt tillämpliga med avseende på den typ av humanistisk forskning som jag främst ägnar mig åt.

2.5 Diskussion och slutsatser

Enkätsvaren har visat på forskningens bredd vid högskolan. Insamlingsmetoderna är många och ofta används både kvantitativa och kvalitativa metoder. En övervägande del ser sig som väldigt centrala i datainsamlandet vid alla typer av metoder. Det gör att man har en nära koppling till de forskningsdata man arbetar med. Många har använt öppen eller delad data. Många har delat forskningsdata och då främst informellt, väldigt få har tillgängliggjort via datacenter eller dataarkiv.

När det gäller vilket stöd eller vilka kunskaper som saknas så anger över hälften ”rättigheter och licenser kring öppna forskningsdata”, men alla angiva svarsalternativ anges i ganska stor grad. Slutsatsen är att det behövs mer information och kunskap om alla aspekter kring tillgängliggörande av data och datahantering i stort.

3 Relevanta dataarkiv

Begreppet dataarkiv används här för tekniska lösningar för tillgängliggörande av data. Det tar inte hänsyn till långtidsbevarande och lösningarna är inte arkiv i arkivlagens mening. Vi har valt att använda begreppet dataarkiv framför begreppet datarepositorier.

3.1 Olika typer av tillgängliggörande

Material kan öppet tillgängliggöras på olika sätt. Corti et al (Corti, Van den Eynden, Bishop, & Woollard, 2014, p. 197) listar följande sätt:

- 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

De olika tillvägagångssätten har olika förutsättningar för hittbarhet², kurering och långsiktig tillgång till tillgängliggjord 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. Men det gäller också att registrera öppet tillgängliggjord data i MUEP. 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 i stort sett omöjligt för stödfunktioner.

3.2 Dataarkiv och deras kontext

Större infrastrukturer för forskningsdata har funnits länge inom vissa områden, exempelvis inom kärnfysik. Nu kommer det allt fler mindre initiativ av olika slag. Utgångspunkterna är olika, från data som råvara till forskningens integritet och möjlighet till att verifiera resultat till individers önskan att jobba öppet (se även avsnitt 1.2). För att data ska kunna hänvisas till krävs system för tillgängliggörande. Ovan i 3.1 beskrevs olika sätt att tillgängliggöra data. Fokus här ligger på datacenter eller dataarkiv då de ofta bygger på metadatastandarder för att beskriva tillgängliggjord data och även fungerar som kataloger över registrerad data.

För enkel länkning till och citering av data rekommenderas att unika identifierare tilldelas. Det sköts ofta vid registrering/deponering i dataarkiv. Den vanligast unika identifieraren är DOI (Digital Object Identifier) och för data sköter DataCite hanteringen (för publikationer hanteras DOI av Crossref). Tilldelningen av DOI innebär att metadata och URL registreras i DataCites databas och blir därmed sökbara.

Data Citation Index (DCI) är en databas i Web of Science. Den innehåller fylligare beskrivningar av tillgängliggjord data tillsammans med information kring citering av dessa data i artiklar som är indexerade i Web of Sciences Core-databaser (Science Citation Index, Social Science Citation Index och Arts and Humanities Citation Index). Genom DCI är det möjligt att söka efter användbara dataset, att utvärdera enskilda dataarkiv och att mäta spridning genom antalet citeringar.

² Hittbarhet inkluderar aspekter som om objektet går att finna genom sökning eller navigation, hur tillgängligheten är ifht WCAG 2.0-standarderna (för uppläsning i talsynteser och liknande), och om det går att länka till objektet med webbadress (URL) eller en unik identifierare (t.ex. DOI). Hittbarhet förutsätter inte tillgång till objektet, fulltext kan kräva betalning eller licens.

Dataarkiv är speciellt viktiga för dataartiklar, artiklar i vanliga vetenskapliga tidskrifter eller i forskningsdatatidskrifter är ett sätt att göra öppet tillgänglig data mer användbar då insamlingsmetod och omfattning kan beskrivas på ett utförligare sätt. Detta innebär att det blir lättare att bedöma dess potential för återanvändning av andra. Citering av använd data förespråkas även i vanliga vetenskapliga artiklar, även om data inte är öppet tillgänglig.

3.3 Kriterier för bedömning av dataarkiv

Vid val av dataarkiv för öppet tillgängliggörande av data behöver fler aspekter tas i beaktande. Nedan är fem grupper listade. Biblioteket kommer behöva fördjupa arbetet och ta fram någon form av riktlinjer eller rekommendationer.

1. Typ av dataarkiv
 - a. Ämnesområde
 - b. Objekttyp
2. Hemvist
 - a. Geografisk Lokalisering (lagstiftning som gäller, inom eller utom EU)
3. Stabilitet
 - a. Huvudmannaskap
 - b. Finansiering
4. Tillgång
 - a. Licenskrav för uppladdning
 - b. Nivåer av tillgängliggörande (direkt tillgång, hänvisning, endast metadata, osv.)
5. Hittbarhet
 - a. Unika identifierare (DOI samt andra unika identifierare)
 - b. Mängd metadata
 - c. Inkluderad i DCI (arkivnivå, datasetnivå)
 - d. Inkluderad i re3data

Observera att aspekterna i listan inte inkluderar långtidsbevarande. Det är en arkivfråga, se kapitel 4 och 5.

3.4 Relevanta dataarkiv för MAH

Utgångspunkten är att det finns externa dataarkiv att använda av högskolans forskare.

Ämnesspecifika dataarkiv finns för en del ämnesområden. Ett exempel är det vid Universitetet i Tromsø (UiT) framtagna dataarkivet TROLLING³ (The Tromsø Repository of Language and Linguistics.) som skapades på efterfrågan från lokala forskare. Ett annat exempel är Journal of Open Archaeology Data (JOAD)⁴ som driver JOAD Dataverse Repository⁵. Båda är en del DataVerse⁶. Den största fördelen med ämnesspecifika dataarkiv är troligen upptäckt och återanvändning inom det egna ämnesområdet.

3.4.1 Generella dataarkiv för allmänt tillgängliggörande

Svensk Nationell Datatjänst (SND)⁷ lagrar och tillgängliggör data inom humaniora, samhällsvetenskap och medicin och hälsa. Zenodo⁸ är ett EU-finansierat allmänt dataarkiv som drivs av CERN där

³ <http://opendata.uit.no/dvn/dv/trolling>

⁴ <http://openarchaeologydata.metajnl.com>

⁵ <https://dataverse.harvard.edu/dataverse/JOAD>

⁶ <http://dataverse.org>

⁷ <http://snd.gu.se/sv/catalogue>

⁸ <https://zenodo.org>

forskare kan registrera alla typer av data. Figshare⁹ är en annan allmän tjänst för tillgängliggörande av forskningsdata. Den drivs med stöd av ett vetenskapligt förlag, och på samma sätt som Zenodo så är det självregistrering som gäller. Vid SND däremot finns metadataspecialister som hjälper till att beskriva den data som läggs in, vilket är en central skillnad. De generella dataarkiven når troligen en bredare grupp användare, men det beror också på hur väl och omfattande data beskrivs vid registreringen.

3.4.2 Lokalt tillgängliggörande av forskningsdata

Delning av forskningsdata kan också göras lokalt, på webbplatser, genom Box eller i högskolans publikationsdatabas MUEP. Även då är det viktigt att tänka på hittbarhet (i vilka system kan man hitta materialet), nåbarheten (hur stabila är länkarna och publiceringsplatsen över tid), och metadata-beskrivning.

Det kommande e-arkivet kan vara ett sätt att tillgängliggöra data. Troligen kan det göras samtidigt som data arkiveras. E-arkivet kan också komma spela en viktig roll i en kommande nationell lösning (se nedan).

3.4.3 Kommande nationell lösning för tillgängliggörande

SND har börjat talas om en nationell ”yta” för öppen tillgång till forskningsdata. Tanken som presenterats är att data som arkiveras i lokala lärosätesspecifika e-arkiv sedan kan göras tillgänglig i en nationell portal (om öppen tillgång anges vid arkivering) (Petzold, 2015).

Viktigt är att den nationella infrastrukturen för tillgängliggjord forskningsdata kommer vara en del i den internationella metadataekologin för hittbarhet och på sikt användning.

3.5 Diskussion och slutsatser

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.

⁹ <http://figshare.com>

4 Roller vid tillgängliggörande

Utgångspunkten för rapporten är öppet tillgängliggörande av data. Tillgängliggörandet är dock svårt att avskilja ifrån övrig hantering av forskningsdata och därför kommer alla delar i hanteringen att behandlas nedan.

4.1 Olika typer av datahantering

De data som tillgängliggörs skapas och kureras under hela forskningsprocessen. Hantering av forskningsdata kräver en mängd olika kunskaper och färdigheter. Utifrån Figur 20 nedan kan olika typer av datahantering urskiljas:

- Skapande (Data Creator)
- Bearbetning och analys (Data Scientist)
- Strukturell hantering (Data Librarian)
- Administrativ hantering (Data Manager)

De två första typerna av datahantering i listan ovan utförs främst av forskare eller av administratörer i forskargrupper, medan de två senare utförs av olika stödfunktioner som arkiv, bibliotek, forskningsservice och IT.

Figur 20. Core skills for data management (Donnelly, 2008).

Varje typ av datahantering kräver specifika kunskaper och färdigheter. I och med att strukturerad hantering av forskningsdata tidigare främst bara gjorts på forskar- eller forskargruppsnivå kan nya kunskaper och färdigheter behövas hos stödfunktionerna (se till exempel Åhlfeldt & Johnsson, 2015).

4.2 Stödfunktioner

Bilden nedan beskriver stödfunktionernas olika roller baserat på specialistkompetens. Den tydliggör att det är många olika instanser inblandade i hanteringen av forskningsdata, i tillgängliggörande eller arkivering och långsiktigt bevarande. På Malmö högskola kan det röra arkivet, biblioteket, IT, jurist och informationssäkerhetsansvarig.

Figur 21. Roller och kompetenser vid hantering av forskningsdata (Lewis, 2015).

Bilden ovan beskriver rollerna utifrån ett brittiskt perspektiv. Den största skillnaden ur ett svenskt perspektiv är arkivets roll. De svenska lagarna kring arkivering och offentliga handlingar gör att arkivaspekten blir viktigare. Flera av pilarna riktade mot biblioteket i bilden gäller också arkivet, speciellt långtidsbevarande som är en arkivspecifik aspekt.

4.3 Organisering av stödet

Hur själva stödarbetet organiseras skiljer sig mellan lärosäten. Det går att skönja två varianter om man tittar på ledande universitet i Storbritannien. En centraliserad stödfunktion där specialister sammanförs till en ny organisatorisk enhet. Eller en decentraliserad stödfunktion där specialisterna organisatoriskt tillhör olika enheter (till exempel bibliotek, IT eller arkiv). Båda varianterna styrs vanligtvis av en styrgrupp, den decentraliserade kan även ha en projektgrupp för det mer verksamhetsnära arbetet och för att hålla samman de decentraliserade specialisterna.

LERU, League of European Research Universities, rekommenderar bland annat att man på lärosäten skapar en färdplan (*roadmap*) med strategiska mål samt uppgifter och handling att genomföra (LERU Research Data Working Group, 2013, p. 8).

4.4 Stöd på MAH

Biblioteket, IT, arkivet och forskarservice på MAH har idag bra samarbeten. Ett ökat engagemang i hanteringen av forskningsdata kommer troligen att kräva en ökad samordning och någon form av strategisk grupp som kan hantera strategiska frågor.

Utifrån fråga 17 i enkäten (se 2.4.1) om stöd och kunskap kring tillgängliggörande av forskningsdata framgår behov kring många olika aspekter där ansvaret hamnar på olika organisatoriska enheter inom högskolan.

4.5 Kommande pilotprojekt

Malmö högskola kommer att delta i ett SND-projekt under 2016. Detta är en del i arbetet för biblioteket med att utveckla ett stöd till högskolans forskare kring hantering av forskningsdata. I projektet kommer både personal från biblioteket och arkivet samt forskare att delta tillsammans med motsvarande kategorier från andra svenska lärosäten. SND deltar i projektet med sin specialistkompetens.

Projektet syftar dels till att identifiera processfaktorer och ett lokalt arbetssätt, dels till att få fram ett underlag som synliggör vilka resurser som kommer krävas för hantering av forskningsdata, primärt för arkivering och tillgängliggörande. Projektet kommer att utgå från deltagande forskares forskningsdata (en mindre mängd material), som under projektets lopp kommer gå genom hela processen med dokumentation och metadatabeskrivning fram till tillgängliggörande.

Genom projektet kommer förhoppningsvis rollerna i stödet till forskningen vad gäller forskningsdata klargöras och former för samverkan identifieras. Resultaten från projektet kommer att ligga till grund för bibliotekets fortsatta arbete, tillsammans med andra berörda stödfunktioner vid högskolan.

Biblioteket har dessutom sökt pengar från Riksbankens Jubileumsfond för ett infrastrukturprojekt¹⁰ med fokus på ökad användbarhet för öppet tillgängliggjord forskningsdata.

¹⁰ <http://rj.se/For-forskare/Infrastruktur-for-forskning/>

5 Identifierade frågeställningar

5.1 Avslutande diskussion

Tillgängliggörande av forskningsdata är komplext. Inte minst eftersom de data som ska tillgängliggöras bygger på forskning som är komplex. Enkätsvaren visar på den stora bredd som finns vid Malmö högskola avseende insamlingsmetoder för data och typer av data som samlas in. Resultaten visar också att en relativt stor andel av respondenterna har använt öppen eller delad data, liksom att många har delat med sig av forskningsdata och då främst informellt.

Det råder en viss begreppsförvirring, kring tillgängliggörande kontra arkivering och data kontra publikationer. Men också kring vad forskningsdata är för olika forskare. För somliga är begreppet helt naturligt, andra använder källmaterial eller empiri. En del samlar inte in material utan genererar data utifrån beräkningsmodeller. I bred bemärkelse är data mer än det som samlas in, det kan också vara annan output från forskningsprocessen än rena publikationer, som enskilda illustrationer eller en presentation. I tjänster som Zenodo¹¹ och Figshare¹² kan alla dessa typer delas.

Bland svaren och kommentarerna i enkäten märks också en viss oro, både på temat "min data" och för mer administration som kommer äta 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?

Tillgängliggörande och arkivering av forskningsdata är två likartade, men skilda processer. I vissa fall kan arkiverad data tillgängliggöras rakt av, till exempel data från viss naturvetenskaplig forskning, medan i andra fall kan data behöva anonymiseras eller på annat sätt anpassas. Alla data kanske inte lämpar sig för tillgängliggörande. Datahanteringsplaner kan vara ett centralt redskap, både före, under och efter avslutat forskningsprojekt. Väl kurerad data är lättare att arkivera och tillgängliggöra. Vad krävs för att data ska bli användbar för andra? Ett svar är att en långsiktighet krävs och att resurser för datahantering behöver inkluderas i ansökningar och projektplaner.

Vetenskapsrådets förslag till nationella riktlinjer för öppen tillgång behöver spridas mer eftersom det kommer påverka alla forskare. Hur informationen och kommunikation om det ska läggas upp är beroende av om och i så fall hur förslaget kommer att se ut i regeringens kommande forskningsproposition. Men med tanke på att detta inte är en nationell fråga, utan internationell, och att EU är starkt pådrivande i öppenhetens riktning så kan högskolan troligen räkna med att någon form av krav på öppen tillgång till forskningsdata kommer från regeringen.

Slutligen kan hantering av forskningsdata relateras till högskolans inspel till forskningspropositionen 2015 (Jönsson, 2015). Malmö högskolas strategi 2020 uttrycker följande som är relevant:

Vetenskaplig grund, akademisk integritet, kritiskt tänkande och samhällsengagemang utgör fundamenten i vår verksamhet. (Malmö högskola, 2013, p. 6)

För att nå våra mål arbetar vi resurseffektivt och strävar efter ökad systematik och kvalitetssäkrade processer (Malmö högskola, 2013, p. 6)

¹¹ <https://zenodo.org>

¹² <https://figshare.com>

Malmö högskola ska arbeta utifrån ett vidgat innovationsbegrepp inom utbildning och forskning för att möta samhällets utvecklingsbehov. (Malmö högskola, 2013, p. 9)

5.2 Frågeställningar för biblioteket

5.2.1 Representation i bibliotekssystem

Hur och i vilka bibliotekssystem ska metadata för öppna forskningsdata vara sökbara? Frågan gäller både högskolans lokala system och de nationella systemen som LIBRIS och SwePub. Troligen bör öppet tillgängliggjord forskningsdata registreras i MUEP.

5.2.2 Information på webben

Mer information på bibliotekets webb/och eller på andra delar av mah.se bör tas fram under våren. Frågan är i vilken grad generell information är gångbar? En genomtänkt kommunikationsplan behöver tas fram.

5.2.3 Riktlinjer kring externa dataarkiv

Riktlinjer och råd behöver tas fram kring val av externa dataarkiv (med utgångspunkt i kapitel 3). Biblioteket behöver snarast kunna ge handfasta råd kring tillgängliggörande då forskare kan behöva tillgängliggöra data vid till exempel publicering av artikel i tidskrifter med policy kring data.

5.2.4 Studera *best practice*

Biblioteket behöver på ett systematiskt sätt analysera de publikationer som behandlar tillgängliggörande av öppna forskningsdata för ett effektivt arbetssätt. Samt utbyta erfarenheter med andra lärosäten som utvecklar stödverksamhet kring hantering av forskningsdata.

5.2.5 Kompetens för att kunna beskriva forskningsdata

Vilka kompetenser kommer biblioteket behöva kring till exempel forskning och metadatastandarder för att kunna beskriva forskningsdata på ett bra sätt?

5.2.6 Rekommendationer

- *Publicera information på mah.se på de teman som är efterfrågade.*
- *Ta fram riktlinjer kring val av externa dataarkiv för tillgängliggörande.*
- *Ta fram en kortare manual med riktlinjer och lathundar kring stöd till forskningsdatahantering.*

En kvalitativ uppföljning av enkäten behöver också göras för att identifiera ytterligare frågeställningar i dessa komplexa frågor tillsammans med forskare vid högskolan.

5.3 Övriga frågeställningar kring stöd för tillgängliggörande

5.3.1 E-arkivets roll

Det planerade e-arkivet behöver tänkas in i tillgängliggörandet på olika sätt. Dels som en del i processen vid högskolan. Dels hur det kan och bör relateras till den befintliga ekologin med öppna data och dess metadata i olika system.

5.3.2 Lokalt nätverk

Någon form av samordning av stödet kring forskningsdatahanteringen och tillgängliggörande behövs, kanske i form av ett nätverk (som i praktiken redan existerar). För strategiska eller överordnade frågor som rör stöd kring forskningsdata behövs någon form av forum. En färdplan (*roadmap*) kring stödet skulle kunna vara en bra utgångspunkt för det fortsatta arbetet.

5.3.3 Rekommendationer

- *Utred e-arkivets roll och resurser med tanke på att det kan bli en central del i en eventuell nationell lösning för tillgängliggörande.*
- *Formalisera det existerande informella nätverket mellan arkiv, bibliotek, forskarservice, juridik och IT.*
- *Forskningsberedningen tillsätter en arbetsgrupp med uppdrag att arbeta fram en färdplan för stödet.*

5.4 Övergripande frågeställningar

Följande övergripande frågor eller konflikter har blivit identifierade:

5.4.1 Ägandeskap och kontroll

Forskare har en känsla för de data de arbetar med, "min data" är en vanlig uppfattning. Det perspektivet ska ställas mot data som allmän handling. Här krockar två praktiker, forskningspraktikerna mot det juridiska synsättet som åligger en myndighet. Det juridiska läget är beskrivet i (Vetenskapsrådet, 2015b), och merparten av forskningsdata vid ett lärosäte är allmän handling (ofta även offentligt).

Forskningsdata ses också som en egen resurs, *exclusive ownership* nämns i en kommentar i enkäten. Det är en resurs som ska uttömmas eller användas som handelsvara i samarbete med andra forskare eller forskargrupper.

5.4.2 Skydd av informanter

Forskningsetiskt anses det att data kan skyddas eftersom man lovat deltagare konfidentialitet (en forskningsetisk princip) trots att data blir allmän handling där lagstiftning styr om materialet helt eller delvis kan lämnas ut. Se Gillbergfallet vid Göteborgs universitet (Stridsberg, 2014). Det gäller även samarbete med kommersiella aktörer, huvudmannaskapet är avgörande för hur data hanteras i förhållande till lagstiftningen.

5.4.3 Merarbete

I enkäten framkom farhågan att mer arbete krävs för att forskningsdata ska kunna göras öppna och användbara för andra. I normalfallet är inte forskningsdata omedelbart användbara för andra än forskaren eller forskargruppen. I många fall krävs ett dokumenterings- och beskrivningsarbete inför ett tillgängliggörande. Även mer administration befaras, något som äter upp forskningstid. Frågan är om öppet tillgängliggörande av forskningsdata (inklusive dokumentation) bör ses som delvis en ny uppgift och som därmed kommer kräva extra resurser eller avsatt tid.

5.4.4 Data som begrepp

Data är inte ett självklart begrepp för alla. Olika discipliner använder olika begrepp, till exempel källmaterial eller empiriskt material, medan andra använder publikationer som underlag för sin forskning.

Forskningsdata som begrepp är inte heller självklart i förhållande till forskningsinformation eller forskningspublikation. Publikationer ges ut av en utgivare och registreras idag i MUEP. De bygger på och innehåller forskningsdata. Forskningsinformation täcker som begrepp ofta in både forskningsdata och forskningspublikation, samt information om forskningsprojekt.

Om programvara (i form av kod) inkluderas i begreppet data varierar. Ofta inkluderas kod som behövs för att återanvända underliggande data i begreppet. Kod i form av fristående mjukvara kan också ses som en del av den äldre rörelsen för öppen källkod (*Open Source*).

5.4.5 Vilka data ska tillgängliggöras

Ur ett forskarperspektiv är det inte heller självklart vilka data som bör tillgängliggöras. Rådata bearbetas och urval görs under forskningsprocessen, och vad som är användbart för andra beror helt på deras forskningsfrågor (Borgman, 2015; Haider & Kjellberg, n.d.).

5.4.6 Rekommendationer

- *Stimulera en fortsatt högskoleövergripande diskussion om frågor kring forskningsdata där bland annat krocken mellan juridiska aspekter och det forskningsetiska perspektivet tas upp.*
- *Satsa på forskning och utveckling kring hantering av forskningsdata på högskolan. Området behöver beforskas för att skapa bättre förutsättningar och ökad förståelse för olika typer av forskning. En möjlighet kan vara att inkludera forskningsdataaspekter i framtida forskningsansökningar.*

6 Referenser

- 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.
- Donnelly, M. (2008). Research Data Management Forum: RDMF2: Core Skills Diagram. Retrieved February 2, 2016, from <http://data-forum.blogspot.se/2008/12/rdmf2-core-skills-diagram.html>
- Eloranta, S., Johansson, A. L., Kristinsson, S. Y., & Andersson, T. M.-L. (2013). Att strukturera och dokumentera forskningsprojekt. *Läkartidningen*, (8), 416–419. Retrieved from <http://www.lakartidningen.se/Functions/OldArticleView.aspx?articleId=19289>
- Haider, J., & Kjellberg, S. (n.d.). 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.
- Jönsson, P. (2015). Underlag till regeringens forskningspolitik (Dnr. Mahr 69-2015/210). Malmö högskola.
- LERU Research Data Working Group. (2013). *LERU Roadmap for Research Data*. Retrieved from http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf
- Lewis, M. (2015). Slide on roles and competencies in RDM. From a presentation at Malmö University Library 2015-09-23 by Andrew Cox.
- Malmö högskola. (2013). *Strategi 2020 (Dnr Mahr 15-2013/117)*. Retrieved from http://strategi2020.mah.se/wp-content/uploads/2013/06/Final-2020-SV-PDF____.pdf
- Malmö högskola. (2015). *Årsredovisning 2014*. Retrieved from <https://www.mah.se/upload/GemensamtVerksamhetsstod/Ekonomiavdelningen/%C3%85rsredovisningar/%C3%85rsredovisning-2014-webb-final.pdf>
- Petzold, M. (2015). Framtidens SND, Svensk Nationell Datatjänst. Retrieved from <https://www.youtube.com/watch?v=yoRCWiWOJyc>
- Stridsberg, L. (2014). *Juridiska frågor rörande öppen tillgänglighet till forskningsdata (Vetenskapsrådet, dnr 111-2013-9215)*.
- Vetenskapsrådet. (2015a). *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/>
- Vetenskapsrådet. (2015b). *Öppen tillgång (open access) till forskningsdata – en omvärldsbeskrivning 2014*. Retrieved from <https://publikationer.vr.se/produkt/oppen-tillgang-open-access-till-forskningsdata-en-omvarldsbeskrivning-2014/>
- Åhlfeldt, J., & Johnsson, M. (2015). *Research Libraries and Research Data Management within the Humanities and Social Sciences*. Lund. Retrieved from <http://www.lu.se/lup/publication/5050462>

7 Bilagor

7.1 Inbjudan

7.1.1 Svensk version

10 snabba frågor om öppna forskningsdata

Denna enkät är en del i ett pilotprojekt vid biblioteket kring tillgängliggörande av öppna forskningsdata. Syftet med enkäten är att få en nulägesbeskrivning av synen på och erfarenheten av öppna forskningsdata bland Malmö högskolas forskare.

Enkäten är kort och tar cirka 5 minuter att svara på. Den innehåller 10 flervalsfrågor kring forskningsdata samt bakgrundsfrågor. Om du som respondent är intresserad av att bli kontaktad i frågor som rör öppna forskningsdata så ange namn och epost i enkätens slut.

Enkäten skickas ut till följande personalkategorier: professor, universitetslektor/lektor, postdoktor och doktorander. Utskicket baseras på data i lönesystemet.

Begreppet forskningsdata används här för den information som ligger till grund för analys och resultat. Skilda begrepp för detta används inom olika forskningsämnen, ibland används "källmaterial" eller "empiriskt material" om inte data används som begrepp inom ämnet.

Begreppet öppen forskningsdata används här för data som är tillgängliggjord för alla och envar att använda, till exempel via en webbplats för ett forskningsprojekt eller ett dataarkiv.

Resultatet av enkäten kommer att redovisas för forskningsberedningen och publiceras i en rapport. Resultaten kommer vara en utgångspunkt för bibliotekets fortsatta arbete med tillgängliggörande av öppna forskningsdata.

Presentkort på luncher för två kommer lottas ut bland de respondenter som anger sin epost i slutet av enkäten.

Tack för din medverkan!

Biblioteket, Enheten Medier och vetenskaplig publicering,
kontaktperson Jonas Fransson, Jonas.fransson@mah.se

Länk till enkäten:

7.1.2 Engelsk version

10 quick questions on open research data

This survey is part of a pilot project at the library on making research data open available. The aim of the survey is to obtain a status report of the view on and the experience of open research data among the researchers of Malmö University.

The questionnaire is short and takes about 5 minutes to answer. It contains 10 multiple choice questions about research data and some background questions. If you as a respondent are interested in being contacted in matters related to open research data, please fill in your name and email in the end of the survey.

The questionnaire has been sent out to the following categories: professor, senior lecturer/associate professor, postdoctoral researchers and PhD students. The categories are based on data in the salary system.

The concept of research data is here used for the information that forms the basis for analysis and results. Different concepts are used in various research areas, sometimes used "source material" or "empirical data" unless the data is used as a term in the subject.

The concept of open research data is here used for data that is made available for everyone to use, for example, through a site for a research project or a data archive.

The survey results will be reported to the Research Advisory Committee and a report will be published. The results will be a base for the Library's continued work with open research data.

Gift certificate for lunch for two will be drawn among the respondents stating their email at the end of the questionnaire.

Thank you for your participation!

The Library, Department of Information Resources and Scholarly Publishing,
contact: Jonas Fransson,

Link to the survey, change language by clicking on "Svenska" in the upper right corner:

7.2 Enkäten på svenska

7.2.1 Introduktionstext

10 snabba frågor om öppna forskningsdata

Enkäten är kort och tar cirka 5 minuter att svara på. Den är anonym och innehåller 10 flervalsfrågor kring forskningsdata samt bakgrundsfrågor. Om du som respondent är intresserad av att bli kontaktad i frågor som rör öppna forskningsdata så ange namn och epost i enkätens slut.

Begreppet **forskningsdata** används här för den information som ligger till grund för analys och resultat. Skilda begrepp för detta används inom olika forskningsämnen, ibland används "källmaterial" eller "empiriskt material" om inte data används som begrepp inom ämnet.

Begreppet **öppen forskningsdata** används här för data som är tillgängliggjord för alla och envar att använda, till exempel via en webbplats för ett forskningsprojekt eller ett dataarkiv.

Tack för din medverkan!

Biblioteket, Enheten för medier och vetenskaplig publicering
(kontaktperson Jonas Fransson, jonas.fransson@mah.se)

7.2.2 Frågor om forskningsdata

7.2.2.1 (1) Vilken eller vilka typer av data arbetar du med? (flera val möjliga)

- textuell data
- numerisk data
- statistisk data
- bilddata
- ljuddata
- video
- geografisk data

- annat (ange nedan)

+ fritext

7.2.2.2 (2) Vilken typ av metod för datainsamling använder du huvudsakligen?

- Kvalitativ metod (till exempelintervju, aktionsforskning, fallstudie)

- Kvantitativ metod (till exempel mätning, experiment, kvantitativ enkät)

- Samlar inte in data utan använder sekundärdata (till exempelregister, öppen forskningsdata eller litteratur)

+ Ange gärna vilken metod (fritext):

7.2.2.3 (3) I vilken grad är du som forskare en del av själva datainsamlingen

- Stor grad (som forskare är jag en central och aktiv del av datainsamlingen, till exempelgenom deltagande observation eller intervjuer)

- Viss grad (till exempelkvantitativa enkäter)

- Liten grad (använder standardiserade insamlingsmetoder, till exempelblodprov eller psykologiska test)

- Ingen grad (använder sekundärdata, till exempelregister, öppen forskningsdata eller litteratur)

7.2.2.4 (4) Har du i din forskning använt data som du inte själv samlat in? (flera val möjliga)

- jag har använt öppna data från myndigheter och organisationer

- jag har använt öppen data från andra forskare

- jag har använt informellt delad forskningsdata

- jag har använt data från myndigheter och organisationer som inte är öppen

- jag har inte använt någon av ovanstående typer av data

7.2.2.5 (5) Har du delat eller tillgängliggjort forskningsdata? (flera val möjliga)

- ja, jag har delat med nära forskarkollegor (informellt)

- ja, jag har delat forskningsdata efter direkt förfrågan

- ja, jag har tillgängliggjort data på webbplats (forskningsprojektets eller personlig)

- ja, jag har tillgängliggjort data via datacenter eller dataarkiv (om ja, ange gärna i vilken tjänst nedan)

- nej

+ fritext

7.2.2.6 (6) Har du sökt externa medel där det krävts en datahanteringsplan vid ansökan?

- ja

- nej

- vet ej

7.2.2.7 (7) Har forskningsfinansiärer krävt av dig att forskningsdata ska göras öppen efter avslutat forskningsprojekt?

- ja

- nej

- vet ej

+ om ja, i vilket sammanhang (frivillig, fritext)

7.2.2.8 (8) I vilken grad kan du tänka dig att öppet dela den forskningsdata du arbetar med?

- stor grad (allt eller så mycket som möjligt)

- i viss grad (större utvalda delar)

- i liten grad (mindre utvalda data)

- inte alls

- datan går inte dela pga. sekretess eller liknande

+ ev. kommentar

7.2.2.9 (9) Hur mycket arbete krävs för att den data som du arbetar med ska kunna göras öppen och användbar för andra?

- Mycket arbete (den behöver väsentlig bearbetning och/eller uttömmande dokumentation)

- Lite arbete (den behöver smärre bearbetning och/eller viss dokumentation)

- Inget arbete, den går att dela direkt

- Datan går inte dela pga. sekretess eller liknande

7.2.2.10 (10) Vad skulle få dig att börja, eller i större utsträckning än idag, öppet tillgängliggöra forskningsdatan som du arbetar med?

(flera val möjliga)

- efterfrågan från andra forskare i mitt ämne

- mer kunskap om hur data kan delas

- mer stöd

- avsatt tid

- tillgängliggörande av forskningsdata skulle ses som en merit

- mätning av genomslag (impact) för den öppna forskningsdatan

- policy vid MAH för öppna forskningsdata

- krav från finansiärer

- vet ej

+ kommentar:

7.2.3 Bakgrund respondent

7.2.3.1 (11) Vilken fakultet tillhör du huvudsakligen?

- HS

- KS

- LS

- OD

- TS

- annan tillhörighet (ange nedan)

+ kommentar i fritext möjlig som tillägg

7.2.3.2 (12) Tillhör du något forskningscentrum?
(frivillig)

- Biofilmer – Research Center for Biointerfaces

- Centrum för professionsstudier (CPS)

- Centrum för tillämpad arbetslivsforskning och utvärdering (CTA)

- Institutet för studier i Malmös historia (IMH)

- Malmö Institute for Studies of Migration, Diversity and Welfare (MIM)

- Sexologi och sexualitetsstudier (CSS)

- Internet of Things and People (IoTaP–RC)

7.2.3.3 (13) Inom vilket ämne/ämnen forskar du?
- fritext

7.2.3.4 (14) Vilken forskarroll har du?
- Professor

- Lektor, postdoktor, eller annan forskartjänst

- Doktorand

7.2.3.5 (15) Forskar du främst i grupp eller individuellt?
- Främst i forskargrupp

- Främst individuellt

- Blandat, både i grupp och individuellt

7.2.3.6 (16) Är din nuvarande forskning externfinansierad?
- ja

- delvis

- nej

7.2.4 Uppföljning

7.2.4.1 (17) Vad behöver du mer kunskap om alternativt stöd kring för att göra den forskningsdata som du arbetar med öppen?
(frivillig)

(flera val möjliga)

- rättigheter och licenser kring öppna forskningsdata

- tillgängliggörande av forskningsdata (allmänt)
- tjänster för tillgängliggörande av forskningsdata inom ditt ämne
- praktiska frågor/stöd för tillgängliggörande av forskningsdata under 2016
- stöd kring datahanteringsplaner
- IT-stöd för till exempel databaslösning, lagring eller backup
- stöd i databearbetning och organisering under forskningsprocessen
- forskningsfinansiärers krav (och kommande krav) kring öppna forskningsdata
- tidskrifters krav kring öppna forskningsdata
- arkivering av forskningsdata
- annat: (fritext)

7.2.4.2 (18) Om du vill bli kontaktad av biblioteket rörande dina svar i ovanstående fråga ange namn och epost nedan.

(frivillig)

- fritext

7.2.4.3 (19). Kommentarer

(frivillig)

- fritext

7.3 Enkäten på engelska

7.3.1 Introduction to the survey

10 quick questions concerning open research data

The questionnaire is short and takes about 5 minutes to answer. It is anonymous and contains 10 multiple choice questions about research data and some background questions. If you as a respondent are interested in being contacted in matters related to open research data, please fill in your name and email in the end of the survey.

The concept of **research data** is here used for the information that forms the basis for analysis and results. Different concepts are used in various research areas, sometimes used "source material" or "empirical data" unless the data is used as a term in the subject.

The concept of **open research data** is here used for data that is made available for everyone to use, for example, through a site for a research project or a data archive.

Thank you for your participation!

The Library, Department of Information Resources and Scholarly Publishing
(contact: Jonas Fransson, jonas.fransson@mah.se)

7.3.2 Research data questions

7.3.2.1 (1) What kind of data are you working with?

(multiple choice possible)

- Textual data

- Numerical Data
- Statistical data
- Image data
- Sound data
- video
- Geographical data
- Other (specify below)
- + free text

7.3.2.2 (2) What method of data collection do you mainly use?

- Qualitative methods (e.g. interviews, action research, case study)
- Quantitative methods (e.g., measurement, experimentation, quantitative survey)
- I do not collect the data but am using secondary data (e.g. registers, open research data or literature)
- + Please state which method (free text):

7.3.2.3 (3) To what extent are you as a researcher part of the data collection?

- Large degree (as a scientist, I am a central and active part of the data collection, e.g. through participant observation and interviews)
- Some degree (e.g. quantitative surveys)
- Small degree (choose from the standard collection methods, such as blood tests or psychological tests)
- No at all (using secondary data, such as registers, open research data or literature)

7.3.2.4 (4) Have you in your research used data that you yourself have not collected?
(multiple choice possible)

- I have been using open data from agencies and organizations
- I have been using open data from other researchers
- I have used informally shared research data
- I have used data from government agencies and organizations that are not transparent
- I have not used any of the above mentioned types of data

7.3.2.5 (5) Have you shared and made research data available?
(multiple choice possible)

- Yes, I have shared with close research colleagues (informal)
- Yes, I have shared research data on direct requests
- Yes, I have made the data available on a website (research project or personal)

- Yes, I have made the data available through data centers or data archives (if yes, please state in which service below)

- No

+ free text

7.3.2.6 (6) Have you applied for external funding, which requires a data management plan in the application?

- yes

- No

- Do not know

7.3.2.7 (7) Have research funding agencies required that you make the research data open after the completed research project?

- yes

- No

- Do not know

+ If yes, in what context (voluntary, free text)

7.3.2.8 (8) To what extent would you be willing to openly share the research data you work with?

- High extent (all or as much as possible)

- To some extent (large selected parts)

- Small extent (some selected data)

- not at all

- Data cannot be shared due to confidentiality or similar

+ Possible. comment

7.3.2.9 (9) How much work is required on the data that you work with, before it can be made open and useful to others?

- A lot of work (it requires substantial processing and/or extensive documentation)

- Some work (it needs minor processing and / or certain documentation)

- No work, it can be shared directly

- Data cannot be shared due to confidentiality or similar

7.3.2.10 (10) What would make you begin, or to a greater extent than today, to openly make available the research data that you work with?

(multiple choice possible)

- Demands from other scientists in my field

- More knowledge

- More support

- Dedicated time

- That making available research data should be seen as a merit
- Measurement of the impact of the open research data
- An open research data policy at Malmö University
- Demands from research funders
- Do not know

+ Comment:

7.3.3 Bakgrund respondent

7.3.3.1 (11) To what faculty do you mainly belong?

- HS
- KS
- LS
- OD
- TS
- Other affiliation (specify below)

+ Comment ...

7.3.3.2 (12) Do you belong to any research centre?

- Biofilms - Research Center for Biointerfaces
- Centre for Profession Studies (CPS)
- Center for Work Life and Evaluation Studies (CTA)
- Institute for Studies in the history of Malmö (IMH)
- Malmö Institute for Studies of Migration, Diversity and Welfare (MIM)
- Centre for Sexology and Sexuality Studies (CSS)
- The Internet of Things and People (IoTaP-RC)

7.3.3.3 (13) Within what subject/subjects do you research?

- Free text

7.3.3.4 (14) What is your research role?

- Professor
- Associate professor/postdoc - or other research position
- PhD student

7.3.3.5 (15) Do you mainly do your research in groups or individually?

- Mainly in research groups
- Mainly individually
- Mixed, both in groups and individually

7.3.3.6 (16) *Is your current research externally funded?*

- yes

- partially

- No

7.3.4 Uppföljning

7.3.4.1 (17) *In order to make the research data you are working with open, in what areas would you need more knowledge or more support?*

(multiple choice possible)

- Rights and licenses for open research data

- Making available research data (general)

- Services for making available research data within your subject

- Practical questions/support for making available research data in 2016

- Support concerning data management plans

- IT support for e.g. database solution, storage or backup

- Support in data processing and organizing data during the research process

- Current and coming demands from research financiers on open research data

- Journal requirements regarding open research data

- The archiving of research data

- Other: (free text)

7.3.4.2 (18) *If you would like to be contacted by the library concerning the question above, please type your name and email below.*

- free text

7.3.4.3 (19) *Comments...*

7.4 Tackmejl

Utskick till alla som blivit inbjudna till bibliotekets enkät om forskningsdata

Tack för alla svar! 164 svar har inkommit (svarsfrekvens 27 %). Svaren kommer nu att analyseras och sammanställas. Som svar på kommentarer i enkäten kommer här några förtydliganden kring enkäten och bibliotekets arbete kring forskningsdata.

Biblioteket har ett uppdrag från Rektor att bevaka frågan kring tillgängliggörande av forskningsdata. Enkäten är del av detta arbete och syftar till att ge oss ökad förståelse för forskningens behov. Målet är att utveckla ett fullgott stöd, att identifiera vad vi på biblioteket ska inrikta oss på för att ge ett bra stöd till er som forskare.

Den bredare bakgrunden till undersökningen är den internationella rörelsen mot öppen forskningsinformation, både för publikationer och underliggande data, som drivs av forskningsfinansiärer och myndigheter. I Sverige har Vetenskapsrådet lämnat förslag till Regeringen, se 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/). Idag är krav på öppen forskningsdata ett undantag bland svenska forskningsfinansiärer, men troligen kommer krav inom en nära framtid. Vetenskapliga tidskrifter som PLOS ONE har redan idag krav på öppen data vid publicering.

När det gäller vanan att dela eller och fritt tillgängliggöra forskningsdata, skiljer sig traditioner och kulturer mellan olika forskningsområden stort. Data kan tillgängliggöras i olika varianter beroende på tänkt målgrupp. Tillgängliggörande av forskningsdata görs alltid efter gällande lagstiftning. En del känslig data kan anonymiseras och ändå vara intressant att tillgängliggöra.

En sammanställning av enkäten kommer att publiceras på publiceringsservice nya sidor kring forskningsdata (<http://mah.se/Bibliotek/Publicering/Forskningsdata/>) i januari. Där kommer vi fortlöpande skriva om tillgängliggörande av forskningsdata och bibliotekets stöd.

Vänliga hälsningar

Jonas Fransson

Enheten Medier och vetenskaplig publicering, Biblioteket