

Slutrapport från projektet

Tillgängliggörande av forskningsdata

*Lokal funktion för arbetet gentemot
Svensk Nationell Datatjänst*

Jonas Fransson
Aron Lindhagen
Mats Greiff
Zoran Vasiljevic

Skrifter från
Malmö universitetsbibliotek
#10, 2019



Slutrapport från projektet *Tillgängliggörande av forskningsdata - lokal funktion för arbetet gentemot Svensk Nationell Datatjänst*

1. Om projektet

Projektet syftar till att upprätta en lokal stödfunktion vid Malmö universitetsbibliotek för publicering av forskningsdata, främst via Svensk Nationell Datatjänst, SND (se projektdirektiv i bilaga). Projektet har löpt från 180301 till 190531. Mellan 180901 och 190531 har två forskare varit med i projektgruppen, Mats Greiff och Zoran Vasiljevic, dessförinnan endast bibliotekspersonal, Aron Lindhagen och Lena Wennerholm, samt projektledare Jonas Fransson.

Rector Kerstin Tham har varit projektägare. Styrgruppen har letts av bibliotekschef Sara Kjellberg och i övrigt bestått av Thomas Arnebrant (forskarrepresentant) och Emma Nolin (avdelningschef vid UB). En referensgrupp var knuten till projektet med representanter från samtliga fakulteter: Bo Reimer (KS), Marie Torstensson Levander (HS), Tobias Ohlsson (LS), Helena Holmström Olsson (TS) och Per Alstergren (OD).

2. Beskrivning av projektets verksamhet

Projektet har haft två huvudfokus utifrån projektdirektivet. Att utveckla arbetet mot forskningsdatainfrastrukturen SND (Leverans 1) och att ta fram förslag på hur forskarinflytande och forskningsförankring kan ske efter projektslut.

SND konsortiet startade 1 januari 2018. Under projektets gång har SND utvecklats från nystartat konsortium till allt mer praktisk verksamhet, såsom test av nytt system med extern datalagring. SND har fått infrastrukturmedel från Vetenskapsrådet för perioden 2018-2023. Malmö universitet, som betalande nätverksmedlem, är inte inblandat i konsortiets utveckling eller drift. Projektet har följt utvecklingen och deltagit i regelbundna nätverksträffar som SND arrangerat för att informera om infrastrukturens utveckling.

Vi har tillgängliggjort forskningsdata inom ramen för projektet för att få erfarenheter kring systemets funktionalitet och för att testa rollfördelning mellan

forskare och stödfunktion. Malmö universitet har varit med i SND:s pilottest av systemet och personal från SND har besökt Malmö universitet i april 2019. Under pilottestet har en lagringsyta från SNIC¹ använts, men framöver är det tänkt att SUNET ska erbjuda en lagringslösning.

Inledningsvis ser arbetsflödet ut så att forskare lägger in och beskriver sin forskningsdata via ett webbgränssnitt. Därefter granskar och kompletterar den lokala stödfunktionen såväl data som metadata, dvs. den beskrivande posten i SND:s tjänst. Sedan tittar även SND på inlämningen före publicering. Framöver kommer SND:s granskning tas bort, så att den lokala funktionen ansvarar för publiceringen. En viktig del i projektet har varit att titta på ansvarsfördelningen mellan forskare och lokal stödfunktion och hur man får till smidiga arbetsflöden. En slutsats är att det är svårt att hitta en generell uppdelning, utan detta måste hanteras vid varje enskild inlämning beroende på datans karaktär och syftet med publiceringen.

En viktig komponent har varit att arbeta med förankring vid Malmö universitet av frågorna kring forskningsdata under projektet. Projektet eller centrala frågor för projektet har presenterats och diskuterats med:

- Stygruppen
- Referensgruppen
- Beredningen för forskning och forskningsutbildning, BFF, två gånger
- Kommittén för forskningsutbildning, KFU
- Forskningsnämnder vid fakulteter

Dessutom har vi under projektets gång fått frågor kring forskningsdata och datahanteringsplaner från forskare, direkt eller via Forskarservice, som stödfunktionen har hjälpt till att lösa.

I övrigt har projektet genomfört en workshop inom ramen för doktorandkurs i kvalitativ metod inom historievetsenskap tillsammans med kursansvarig Mats Greiff. Universitetsbiblioteket har tidigare utarbetat en workshop för forskarstuderande kring forskningsdatafrågor som har uppdaterats och erbjudits som en del av bibliotekets utbud.

Bibliotekspersonalen i projektet har också varit en del av nätverket för hantering av forskningsdokumentation vid Malmö universitet tillsammans med stödverksamheter vid GV. Nätverket startade höstterminen 2018 som en följd av processen Att stödja forskning.

3. Deltagande forskares reflektioner

Mats reflektion:

¹ Swedish National Infrastructure for Computing, <https://www.snica.se>

I min forskning har jag allt sedan 1980-talet använt mig av kvalitativa metoder och en betydande del av de forskningsdata som samlats in har utgjorts av intervjumaterial, så kallad "oral history". Förutom egna intervjuer i flera olika projekt har jag använt mig av intervjuer som finns i olika slags amerikanska arkiv. Två innehållsrika sådana finns vid Library of Congress i Washington DC och vid Marshall University i Huntington, WV. Erfarenheten av att samla ovärderligt material i dessa arkiv har öppnat ögonen på mig vad gäller betydelsen av att tillgängliggöra forskningsdata, särskilt inom områden som exempelvis "oral history" och antropologi.

Det finns emellertid ett antal principiella frågor att resonera kring, vilket jag bland annat gjort inom "oral history"-kretsar i Sverige och som jag avser att publicera en artikel kring. "Oral history" har en demokratisk forskningsansats i att ge röst åt människor, ofta ur så kallade underordnade grupper i samhället, vars erfarenheter och berättelser oftast inte syns i traditionellt skriftligt källmaterial. I den demokratiska forskningsansatsen ingår också att människor själva äger sina berättelser och att de därför inte, såvida inte andra önskemål finns, anonymiseras. Att öppet tillgängliggöra sådant material reser i relation till den nya GDPR-lagen etiska frågor. Bland annat sådana diskuteras i artikeltexten, men också hur man konkret kan gå tillväga när etiska och andra frågor beaktas.

Vidare kommer sådant forskningsmaterial som samlas in i vinter, transkriberingar av intervjuer om hur Malmös musikliv förändrats under de senaste 30 åren, med beaktande av etiska aspekter, att direkt provas för utläggning digitalt som ett första led i att göra Malmö universitet till en nationell föregångare i att lägga ut insamlat "oral history"-material och tillgängliggöra det för andra forskare. Min uppfattning är att detta är av betydelse för historieforskning såväl i Malmö som nationellt. Jag menar också att det är viktigt att det finns stöd för forskare, både etiskt och tekniskt, i denna process. Ett sådant stöd gör att forskare kan känna större trygghet i att tillgängliggöra sina forskningsdata.

Zorans reflektion:

En central del i projektet har varit att utifrån ett befintligt kvantitativt datamaterial gå igenom hela datahanteringsprocessen, från organisering och beskrivning av datamaterialet till tillgängliggörande. Syfte var att tydliggöra forskarens förutsättningar för tillgängliggörande av forskningsdata och identifiera och klargöra stödbehov. En nödvändig förutsättning för tillgängliggörande är att forskningsdata är i ordnat skick. Utifrån ett forskarperspektiv, så finns ett behov av kompetensutveckling och stöd kring datahantering och hur man handskas med forskningsdata på ett säkert och strukturerat sätt; hur man håller ordning på sina filer, vad som är relevant att dokumentera och vad metadata är, vilka juridiska och etiska aspekter som kan vara viktiga att ta ställning till. Det är viktigt att dokumentationen av forskningsdata börjar så tidigt som möjligt i

forskningsprocessen och att forskaren i ett tidigt skede funderar över vilken information som andra behöver för att kunna förstå och återanvända datamaterialet.

Utvecklandet av gemensamma och ämnesspecifika rutiner kring hur hanteringen av forskningsdata ska kvalitetssäkras är ett annat viktigt utvecklingsområde. Här finns möjligheten att i större utsträckning än vad i dagsläget sker begära in data för granskning vid exempelvis olika hållpunkter i doktorandutbildningen, såsom halvtids- och slutseminarier. Delar av en sådan granskning hade kunnat genomföras med hjälp av stödfunktionen på lärosätet. Implementering av rutiner kring kvalitetssäkring av forskningsdata tidigt i forskningsprocessen minimerar risken för misstag och gör att forskaren känner sig tryggare i att forskningsdata håller tillräckligt hög kvalitet för att så småningom kunna tillgängliggöras och återanvändas. Trots att övergången till öppen tillgång till forskningsresultat är lång är det viktigt att det progressiva arbetet på Malmö universitet fortskrider i den här frågan och arbetar upp en kompetens och beredskap för ett framtida skarpt läge.

Bibliotekets kommentar

För bibliotekets delar har forskarinvolveringen inneburit värdefulla insikter i olika forskningspraktiker och dess datahantering. Forskardeltagandet har också inneburit en legitimering av projektet och en starkare koppling till forskningen vid Malmö universitet.

4. Leverans 1: Etablering av funktion för arbete mot SND

Leverans 1 enligt direktiv: Funktionen DAU² enligt SND för tillgängliggörande av forskningsdata, där data och metadata kvalitetssäkras före uppladdning till en nationell infrastruktur.

Funktionen är upprättad vid biblioteket och uppgiften är att assistera vid tillgängliggörande av forskningsdata, samt att kontrollera och komplettera (kvalitetssäkra) metadata före publicering vid tillgängliggörande via SND. Detta innebär att Malmö universitet i likhet med till exempel Chalmers, GU, och SU har en etablerad lokal funktion i relation till SND-infrastrukturen. Lagringen av data kommer att ske på en lärosätesägd yta på SUNET för att data inte ska lämna myndigheten, vilket krävs när persondata hanteras. Hittills har en lagringsyta på SNIC använts i väntan på SUNET:s lösning.

Funktionen utförs av arbetslag Forskningsdata vid universitetsbibliotekets avdelning Medier och vetenskaplig publicering (MVP). Arbetslaget stödjer även framtagande av datahanteringsplaner och att hitta data för återanvändning. I bibliotekets arbete med stöd och rådgivning kring forskningsdata är nära samarbete med närliggande stödfunktioner vid GV viktigt.

² DAU, Data Access Unit, är SND:s beteckning på de lokala stödfunktionerna som hjälper till vid datapublicering.

Funktionen stödjer forskare som behöver tillgängliggöra forskningsdata på grund av till exempel finansiärers eller tidskrifters krav, men funktionen tar inte ställning i frågan om (öppet) tillgängliggörande av data. Funktionen kommer att utvecklas i takt med yttre krav och inomvetenskapliga behov över tid. Förutom arbetet med att tillgängliggöra forskningsdata så arbetar arbetslaget utåtriktat och besöker gärna handledarkollegier, seminarier och forskargrupper. Även doktorandworkshoppar anordnas vid intresse.

Projektet ser behov av ytterligare specialistkompetenser i kommande utveckling av funktionen, främst kring metadatastandarder och användning av metadata vid hantering av data i forskningsprojekt.

Framöver kommer tillgång till annan expertis krävas, såsom jurist, informationssäkerhetsansvarig, forskningsetiker, arkivarie eller IT. I dagsläget finns informellt samarbete genom Nätverket för forskningsdokumenthantering vid Malmö universitet, detta är viktigt för funktionen och måste säkerställas. Det kommande behovet av arbetsinsatser av olika experter/funktioner beror helt på hur utveckling sker kring datapublicering, finansiärers krav, etcetera.

5. Leverans 2: Förslag kring forskarinvolvering efter projektslut

Leverans 2 enligt direktiv: Förslag på organisation för forskarinflytande och forskningsförankring för stödfunktionen (DAU) efter projektslut.

Att hantera stöd kring forskningsdata är komplext och datafrågorna är en del av forskningens kärna. Verksamheten har också kopplingar till Stödja forskning-processen, internrevisionen rörande dokumenthantering och behovsinventeringen kring forskningsinfrastrukturer (för BFF).

Både forskarinflytande (input från forskare) och forskarförankring (output till forskare) är centralt för verksamheten, speciellt då praktiker och normer kring datatillgängliggörande inte fastställts utan fortfarande är under utveckling för de flesta forskningsämnen.

Med tanke på att inga krav föreligger kring tillgängliggörande av forskningsdata så förordar projektet att fortsatt forskarinflytande/forskarförankring (åt båda hållen) sker genom fakulteternas forskningsnämnder samt Beredningen för forskning och forskarutbildning (plus ev. Kommitteen för forskarutbildning). Funktionen kommer också få forskarinvolvering genom de forskare som ges stöd. Dessutom planeras besök i handledarkollegier. När frågor kring tillgängliggörande av forskningsdata blir mer aktuella så förespråkas en ökad grad av forskarinvolvering.

6. Slutord

Intresset för att publicera data är för närvarande lågt och än så länge finns få krav, vilket gör att det troligen inte kommer att vara någon större aktivitet kring detta i närtid. Däremot ser vi att en funktion som stödjer forskningsdatafrågor i form av bibliotekets arbetslag kommer att fylla en funktion som stöd vid framtagandet av datahanteringsplaner, särskilt eftersom Vetenskapsrådet ställer nya krav på datahanteringsplaner för beviljade projekt.

Bristen på intresse för att publicera eller tillgängliggöra data med begränsningar beror oftast på avsaknad av incitament för enskilda forskare att investera arbetstid. Samtidigt innehåller en stor del av forskningsdata vid Malmö universitet personuppgifter eller annan känslig information, vilket innebär att den aldrig kommer kunna fritt tillgängliggöras. I dessa fall kan SND:s katalog användas för beskrivning och synliggörande av data, medan universitetets e-arkiv kan utnyttjas för lagring och eventuellt utlämnade av data.

Vetenskapsrådet har haft ett utredningsuppdrag kring FAIR-principerna som avrapporterades hösten 2018³. FAIR-principerna innebär att forskningsdata, eller beskrivningar av forskningsdata, ska göras sökbara, nåbara och återanvändningsbara. Detta tillsammans med till exempel EU:s Open Data Pilot inom Horisont 2020⁴ ger en fingervisning om kommande utveckling.

Även om incitamenten för forskare är små att tillgängliggöra forskningsdata bör man arbeta med att ta vara på de positiva sidorna av den pågående utvecklingen mot öppen vetenskap. Samtidigt är det viktigt att utveckling sker på forskarens villkor och att hänsyn tas till de olika forskningsämnena.

³ Kriterier för FAIR forskningsdata, <https://www.vr.se/analys-och-uppdrag/vi-analyserar-och-utvarderar/alla-publikationer/publikationer/2018-12-07-kriterier-for-fair-forskningsdata.html>

⁴ What is the EC Open Research Data Pilot?, <https://www.openaire.eu/what-is-the-open-research-data-pilot>