

Forskningsdatahantering

Inventering 2020

Inledning

SUHF:s nationella arbetsgrupp för forskningsdata har i uppdrag att strategiskt bevaka den nationella och internationella verksamhetsutvecklingen för öppna data. Arbetsgruppen arbetar på uppdrag av SUHF:s samordningsgrupp för öppen vetenskap.

Forskningsdatagruppen har i uppdrag att föreslå konkreta strategier och insatser för att i dialog med relevanta aktörer främja det nationella arbetet med att utveckla och implementera lokala infrastrukturer för öppna data vid de svenska lärosätena.

För tredje gången har en inventering kring hur SUHF:s medlemslärosäten arbetar med hanteringen av forskningsdata genomförts. De andra genomfördes 2018 och 2019.

Den korta enkäten på sju frågor har riktats till dem som har kunskap och kännedom om frågor kopplat till forskningsdatahantering eller någon som arbetar på övergripande strategisk nivå inom lärosätet. Till alla frågor har det funnits möjlighet att lämna kommentarer. Enkäten har besvarats av 33 av 38 SUHF:s medlemslärosäten, vilket ger en svarsfrekvens på 86,8 procent.

Detta dokument är en kort sammanställning av enkätsvaren med några jämförelser med tidigare års enkäter. Sammanställningen går igenom fråga för fråga och avslutas med en kortare sammanfattning.

1. Ange lärosäte, namn på svarande samt funktion/uppdrag inom lärosätet

I fråga ett önskade vi få en bild av vilka som svarat på enkäten. Många svarande arbetar inom högskolans bibliotek och är gärna dess chef, medan andra arbetar inom lärosätenas forskningsstöd eller har andra funktioner. De flesta svaren har lämnats in av flera personer eller av arbetsgrupper.

2. Har ditt lärosäte en forskningsdatapolicy?

Av 34 svarande har fyra lärosäten en forskningsdatapolicy (11,8 procent) (Stockholms universitet, Sophiahemmet högskola, Karlstads universitet och Högskolan Kristianstad), 21 (61,8 procent) lärosäten svarar att de är på gång medan nio svarar nej (26,5 procent).

Flera lärosäten säger att de är på gång att ta fram en forskningsdatapolicy, men att arbetat inte kommit så långt på grund av Coronasituationen. Högskolorna i Borås, Skövde och Väst diskuterar att ta fram en gemensam policy kring öppen vetenskap. Vid Södertörns högskola är en kartläggning på gång kring att upprätta en Data Access Unit (DAU), en del av det arbetet

handlar om att ta fram en forskningsdatapolicy. Vid ett lärosäte är man på gång att ta fram en policy där lärosätets respektive forskarens ansvar ska tydliggöras, preliminärt är även tanken att det arbetet ska bli en del av en publiceringspolicy. Vid andra diskuteras vilken enhet som ska ansvara för framtagandet. Vid Stockholms konstnärliga högskola inväntar en nationell definition på vad forskningsdata är inom forskning på konstnärlig grund. Vid Göteborgs universitet genomförs en intern utredning utifrån ett helhetsgrepp kring hur universitetet ska organisera stödet kring forskningsdata, hur forskarna ska erbjudas rätt verktyg och hur allt ska samordnas.

Jämförelse med tidigare enkäter

Av 2018 och 2019 års enkäter framkom att det bara var ett lärosäte, Stockholms universitet, som hade forskningsdatapolicy. År 2018 hade ett tiotal inlett arbetet och 2019 ökade detta antal något.

3. Vad anser ditt lärosäte är de strategiskt viktigaste frågorna vad gäller hantering av forskningsdata? Vilka är de största hindren/utmaningarna?

Strategiskt viktiga frågor

Många lärosäten lämnade omfattande svar kring vilka som anses vara de strategiskt viktigaste frågorna vid hanteringen av forskningsdata. I centrum står hanteringen och lagringen av data alltifrån datahanteringsplaner till hanteringen under projektets gång och lagring och arkivering både på kort och lång sikt för att på det viset säkra forskningens reproducerbarhet, tillförlitlighet, kvalitet och bevarande. Tillgängligheten och att data ska vara FAIR liksom säkerhetsfrågorna i kombination med transparens lyftes också.

Det finns även önskemål om att det ska skapas en kvalitativ och effektiv infrastruktur för data, liksom samsyn med tydliga mandat och roller på internationell, nationell och lokal nivå för hantering, delning och lagring av forskningsdata. För detta behöver avtalen kring hanteringen ses över utifrån juridiska aspekter som affiliering, samverkan och uppdrag. Flera lyfte att åtgärder behövs för att undvika att lärosäten, myndigheter och organisationer tar fram egna lokala lösningar.

Stödet till forskarna bland genom Data Access Unit (DAU), var en annan fråga som lyftes av flera lärosäten, liksom ökad kunskap bland forskarna och incitament/meriteringsmörötter att säkra, beskriva och tillgängliggöra forskningsdata. Ett lärosäte konstaterade att utmaningarna främst består i avsaknaden av teknisk, juridisk och arkivariekompetens.

Samverkan för att underlätta utbyte av data både inom lärosätet, mellan lärosäten och mellan lärosäten och regioner lyftes också.

Hinder och utmaningar

Vad gäller hinder och utmaningar anses bristfällig nationell samordning vara ett av de största hindren. En tydlig hantering och ansvarsfördelning för forskningsdata under hela datalivscykeln behöver fastställas på nationell nivå som garanterar organisatorisk, ekonomisk, juridisk och teknisk långsiktighet och hållbarhet, och som dessutom visar hur lärosätena ska uppnå öppen tillgång till forskningsdata till 2026, i enlighet med regeringens målbild.

Flera lärosäten lyfter att hitta säkra och kostnadseffektiva centrala lösningar för lagring, arkivering, gallring samt att säkra framtida läsbarhet av data. Det gäller inte minst sekretessbelagda data och infrastruktur, både under pågående projekt (bl.a. för att dela data med externa samarbetspartners) och efter avslutade projekt.

Bristande ekonomiska resurser gör att frågan om data nedprioriteras. Ett annat är att stöd till forskarna inte får ta för lång tid så att det stoppar forskningsprocessen.

En utmaning är att få fler forskare att dela sina data samt att forskarna ofta tror att forskningsdata går under lärarundantaget och inte ska arkiveras. Det vilar stort ansvar för enskilda forskaren att hålla sig och sin data uppdaterad både kort- och långsiktigt, inom ramen för befintlig lagstiftning och hur delning av data kan ingå som del av meriteringssystemet.

Flera påpekar att det kan vara svårt att få deltagare att ställa upp i forskningsprojekt om forskarna inte kan garantera total sekretess (för all framtid) liksom hantera (känsliga) personuppgifter. Detta visar att nivån på kunskapen i dessa frågor behöver höjas både bland forskare och inom förvaltningen.

Konstnärlig och forskningsdata lyfts ur flera aspekter. Det råder oklarhet vad dessa data är och kan vara. Ett problem med dessa data är att den inte är möjlig att upprepa eller falsifiera utan måste förstås utifrån evidensbaserad praktisk erfarenhet där processerna är lika viktiga för förståelsen som själva slutsatsen. En annan aspekt handlar att forska på material som fortfarande är upphovsrättsskyddad, t.ex. musik. Även att det saknas en definition för passivt tillgängliggörande.

Jämförelse med tidigare enkäter

De strategiska frågor som lyfts, liksom vilka största hindren eller utmaningarna är i stort desamma som kom fram i enkäterna 2018 och 2019.

4. Har ditt lärosäte en mall för datahanteringsplan? Om ja, klistra gärna in en länk till den i kommentarfältet nedan.

På frågan om lärosätet har en mall för datahanteringsplan har totalt 33 lärosäten svarat. Elva lärosäten (33,3 procent) svarar ja, 16 lärosäten (48,5 procent) svarar nej och sex lärosäten (18,2 procent) svarar att det är på gång.

Flera lärosäten skriver att de är i process att anpassa sin verksamhet efter Vetenskapsrådets (VR) krav, SND:s mall eller att de helt enkelt anpassar sig efter de olika forskningsfinansiärernas krav. Andra lärosäten använder DMP Tool och DMP Online.

Ett lärosäte lyfter fram att olika forskningsområden inom lärosätet har olika behov och använder sig av olika mallar från finansiärer och ämnesspecifika organisationer. Andra säger att frågan inte är prioriterad då de inväntar VR:s DHP.

Jämförelse med tidigare enkäter

Sedan 2019 års enkät har antalet lärosäten med datahanteringsplan ökat.

5. Finns det en central stödfunktion för hantering av forskningsdata vid ditt lärosäte? Hur ser den ut?

På frågan om det finns en central stödfunktion för hantering av forskningsdata svarar 19 lärosäten (57,6 procent) ja, fem (15,2 procent) och nio (27,3 procent) att det är på gång.

En forskningsdatahanteringsfunktion kan antingen vara en självständig enhet, ett projekt eller en del av t.ex. Grants Office. Kontoret kan organisatoriskt vara en del av förvaltningen, knutet till biblioteket – vilket är vanligast – eller vara en virtuell avdelningsöverskridande funktion. Vid andra är forskningsdatafunktionen kopplat till en forskningsnämnd eller ett forskningsetiskt råd.

Vid flertalet lärosäten finns en s.k. DAU-funktion (Data Access Unit). Den kan bestå av flera kompetenser: arkivarie, bibliotekarie, forskningsrådgivare, jurist, IT och informationssäkerhets-samordnare.

Jämförelse med tidigare enkäter

Antalet lärosäten med en central stödfunktion för hantering av forskningsdata är detsamma som 2019.

6. Hur löser ditt lärosäte lagringsytor för forskningsdata i forskningsprocessens olika skeden? Vilka former och hur?

Vid de flesta lärosätena lagras data lokalt inom ett projekt eller institution. Det finns en stor variation av lösningar. En del av har gemensamma lösningar, andra använder sådant som gemensam mappstruktur, lös lagringsmedia (exempelvis USB-minnen och lösa hårddiskar), DIVA, OneDrive, Dropbox eller Box. Vid något lärosäte utreder man att kunna lagra data hos Sunet eller som molntjänst, andra väntar på den typen av lagringslösningar. Flera påpekar att det är få projekt som har så stora data att man inte kan finna lagringslösningar lokalt.

Vid KI har under våren 2020 införts en ny central lagringslösning med plats för mer statiska data (framförallt som ligger till grund för publikationer och långtidslagring och förhoppningsvis även e-arkivering framöver) och lagring för mer aktiva data som fortfarande används inom forskningsprojekt. Andra rapporterar att upphandlingar är på gång för lagring på såväl kort som lång sikt. Vid Göteborgs universitet finns samarbetsytor lagring implementerat, och vid Uppsala universitet är man på gång med en universitetsövergripande IT-organisation.

Flera lärosäten har datahanteringsriktlinjer eller hanteringsplaner.

Jämförelse med tidigare enkäter

Hur lärosätena löser frågan om lagringsytor för forskningsdata i forskningsprocessens olika skeden är snarlikt det som kom fram i 2019 års enkät. Mycket ansvar ligger på den enskilda forskaren och de flesta lärosätena har någon form av lokala lösningar.

7. Övriga kommentarer

Bland de övriga kommenterarna framkommer att organisationen runt hanteringen av forskningsdata såväl strategiskt som praktiskt har utvecklats sedan förra året och fått tydligare form.

Vidare lyftes att SUHF:s samordningsgrupp för öppen vetenskap inte tillräckligt beaktar de konstnärliga aspekterna av forskning. Vad gäller forskningsdata är detta är mycket komplex fråga inom just konstnärlig forskning. Det anses även "svårt att föra fram de behov som den konstnärliga grunden har för att kunna tydliggöra dels 'vad' som ska beskrivas men även 'hur' detta låter sig göras med hjälp av metadata anpassade efter vetenskapernas beskrivningsformat".

Ett lärosäte lyfter fram att de haft stor nytta av SUHF:s rekommendation avseende styrdokument för forskningsdata som utgångspunkt för deras arbete med forskningsdatapolicyn.

Uppsala universitet ser ett "stort värde i att SUHF:s lärosäten, kanske speciellt de stora, samarbetar och samverkar inom alla aspekter av området hantering av forskningsdata. Ett inte allt för stort land som Sverige har inte råd med att uppfinna ett antal hjul i parallella processer. Dessutom behövs samordning för att kunna samverka i internationella sammanhang, t ex när det gäller European Open Science Cloud. Det finns också ett stort värde i att koppla dessa diskussioner till annan nationell forskningsinfrastruktur, och URFI-gruppen bör vara en central resurs för detta. Det är strategiskt viktigt att lärosäten samarbetar och verkar för att forskarsamfundets repositorer/plattformar för forskningsdata garanteras långsiktig finansiering".

Slutsatser

De slutsatser man kan dra av sammanställningen av enkäten är:

- Utvecklingstakten vid lärosätena vad gäller tillhandahållandet av ett komplett stöd för forskningsdatahantering är fortsatt långsam.
- Lärosätena har kommit olika långt i sin verksamhetsutveckling för lokala infrastrukturer. Det finns risk att den ojämna utvecklingen kan komma att skapa ogynnsamma olikheter för forskarna beroende på vilket lärosäte man tillhör.
- Beroende på lärosätets tillgång och tilldelning av resurser för att utveckla stöd och tjänster för forskningsdatahantering har man olika förutsättningar för att tillhandahålla en komplett central verksamhetsstödande funktion. Risk finns för både kompetensglapp och svårigheter att tillhandahålla tillräckligt stöd till forskarna.
- Utifrån svaren att döma är det tydligt att i centrum står frågan om hur hantering, lagring och arkivering av forskningsdata ska lösas på ett sätt som garanterar forskningens reproducerbarhet, tillförlitlighet och kvalitet över tiden. Detta innefattar även datahanteringsplaner. Flera lyfte att åtgärder behövs för att undvika att lärosäten, myndigheter och organisationer fortsätter att ta fram egna lokala lösningar.
- I svaren lyfts också tillgänglighetsaspekten och att data ska vara FAIR samt säkerhetsfrågor i relation till transparens. Önskemål uttrycks om att det ska skapas en kvalitativ och effektiv infrastruktur för forskningsdata, liksom samsyn med tydliga mandat och roller på internationell, nationell och lokal nivå. För detta behöver även avtalen kring

hanteringen ses över utifrån juridiska aspekter som affiliering, samverkan och uppdrag.

- Vad gäller hinder och utmaningar anses bristfällig nationell samordning vara ett av de största hindren. En tydlig hantering och ansvarsfördelning för forskningsdata under hela datalivscykeln behöver fastställas på nationell nivå som garanterar organisatorisk, ekonomisk, juridisk och teknisk långsiktighet och hållbarhet, och som dessutom visar hur lärosätena ska uppnå öppen tillgång till forskningsdata till 2026, i enlighet med regeringens målbild. På sikt kommer lärosätena även att bli en del av det planerade ekosystemet för EOSC (European Open Science Cloud), vilket ställer krav på en väl fungerande infrastruktur för forskningsdatahanteringen såväl lokalt vid lärosätena som nationell samordning mellan lärosätena och nationella forskningsinfrastrukturer.

Frågor 2020

1. Ange lärosäte, namn på svarande samt funktion/uppdrag inom lärosätet
2. Har ditt lärosäte ett styrdokument för forskningsdata?
3. Vad anser ditt lärosäte är de strategiskt viktigaste frågorna vad gäller hantering av forskningsdata? Vilka är de största hindren/utmaningarna?
4. Har ditt lärosäte en mall för datahanteringsplan?
5. Finns det en central stödfunktion för hantering av forskningsdata vid ditt lärosäte? Hur ser den ut?
6. Hur löser ditt lärosäte lagringsytor för forskningsdata i forskningsprocessens olika skeden? Vilka former och hur?
7. Övriga kommentarer