

Allmänna anvisningar för datahanteringsplaner 2019

Varför är datahantering och att skriva en datahanteringsplan viktigt (data management plan, DMP)?

- Det är god forskningspraxis!
- Risken att förlora forskningsdata minskar.
- I planen kan du redan i förväg ta ställning till eventuella komplicerade frågor som gäller ägande- och nyttjanderätt.
- Planen är ett stöd för öppen tillgång och främjar produktivt framtida samarbete.
- Du bemöter krav från finansörerna.
- Du sparar tid och pengar.
- Datahanteringsplanen visar dina projektledarfärdigheter.

Data är ett brett begrepp som inbegriper "all information som krävs för att upprepa en studie ---, och allt som är av potentiell nytta för andra" – Sarah Jones /Digital Curation Centre.

Datahanteringsplanen ska beskriva hur forskningsdata kommer att hanteras genom forskningens hela livscykel. Datahanteringsplanen är ett levande dokument som bör uppdateras allt efter som forskningsprojektet framskrider.

Rutinerna för hantering av forskningsdata ska följa FAIR-principerna, vilket innebär att forskningsdata ska vara sökbara (Findable), tillgängliga (Accessible), interoperabla (Interoperable) och återanvändbara (Reusable).

Du kan använda dessa allmänna frågor för att strukturera din datahanteringsplan. Besvara de frågor som är tillämpbara på ditt projekt.

Lycka till med datahanteringsplanen!

1. Allmän beskrivning av data
<p>1.1 Vilka slags data baserar sig forskningen på? Hurdana data samlas in, produceras eller återanvänds? Vilka filformat används? Ge även en grov uppskattning av hur stor produktionen/insamlingen av data kommer att vara?</p> <p>Betrakta din datahanteringsplan som en del av forskningsplanen. Datahanteringsplanen behöver inte nödvändigtvis kunna läsas som ett självständigt dokument. Datahanteringsplanen kompletterar forskningsplanen med en beskrivning av den tekniska hanteringen av projektets data. För att undvika upprepningar kan du hänvisa till din forskningsplan i datahanteringsplanen och vice versa.</p> <p>Beskriv kort vilka slags data som samlas in eller produceras. Ange också vilka slags befintliga data som kommer att användas. Till exempel texter, bilder, fotografier, mätningar, statistik, fysiska prover eller koder.</p> <p>Kategorisera dina data på ett sådant sätt att du kan hänvisa till dem senare i planen. Indelningen kan utgöra en allmän disposition för resten av planen, t.ex. A) data som samlas in för detta projekt, B) data som är ett resultat av processen, C) befintliga data som insamlats tidigare och återanvänds i detta projekt, D) projektledningens dokument samt projektrapporter, osv.</p>

Ange filformat för all data. I vissa fall kan formatet som används under forskningsprojektet vara ett annat än det format som används vid arkiveringen av datamaterialet. Ange båda formaten. Filformatet är av central betydelse för framtida tillgänglighet och återanvändbarhet av data.

Rekommendationer för bästa praxis

- Dataanalysen och metodologiska frågor som gäller projektets data ska beskrivas i forskningsplanen.
- Exempel på filformat: .csv, .txt, .docx, .xlsx, .tif
- När du listar vilka filformat som kommer att användas ska du komma ihåg att nämna all slags särskild eller ovanlig programvara som krävs för att läsa eller använda projektets data, i synnerhet om programvaran kodas inom ramen för projektet.
- Använd tabellform eller punktuppställning för att koncist presentera typer av data, filformat, programvara, osv.
- Du kan även uppskatta ökningen i data under projektets gång, t. ex. per vecka: "Projektet producerar / insamlar cirka 100 GB data per vecka."

1.2 Hur kontrolleras datamaterialets enhetlighet och kvalitet?

Beskriv hur metoderna som används för att samla in, analysera och bearbeta data kan påverka kvaliteten på projektets data och hur riskerna ska minimeras.

Kvalitetskontroll av data säkerställer att data inte ändras av misstag och att kvaliteten upprätthålls under hela livscykeln. Kvalitetsproblem kan uppstå i samband med den tekniska hanteringen, konverteringen eller överföringen eller i samband med kontextuell processering och analys.

Rekommendationer för bästa praxis

- Transkriptioner av ljudupptagningar eller videointervjuer ska kontrolleras av en annan person än den som transkriberat.
- Analogt material ska digitaliseras med så hög resolution som möjligt för precision.
- Vid alla konverteringar ska det ursprungliga informationsinnehållet säkerställas.
- Programvara som använder sig av kontrollsummor ska användas.

2. Iakttagande av etiska principer och lagstiftning

2.1 Vilka etiska frågor är relevanta för datahanteringen, t.ex. hantering av sensitiv data, skyddande av deltagarnas identitet, samtycke till delning av data?

Beskriv på vilka sätt hanteringen av forskningsdata håller hög etisk standard och uppfyller relevant lagstiftning. De etiska aspekterna måste beaktas under datamaterialets hela livscykeln.

Exempelvis, till de etiska krav som gäller för så gott som all forskning hör att följa riktlinjerna för hur deltagarna ska informeras. Om det ingår hantering av personliga eller sensitiva uppgifter ska du beskriva hur deltagarnas identitet kommer att skyddas och hur du kommer att anonymisera eller pseudonymisera data.

Rekommendationer för bästa praxis

- Bekanta dig med din organisations etiska riktlinjer och datasäkerhetspolicy och var beredd att följa bestämmelserna i dem.

- Om etikprövningsnämnden ger ett utlåtande om din forskning ska du i datahanteringsplanen ange hur du ämnar följa direktiven (t. ex. hur deltagarnas identitetsskydd tryggas genom att man avlägsnar känsliga eller personliga uppgifter ur data innan de görs tillgängliga.).
- Läs mera om god vetenskaplig praxis t.ex. på [Forskningsetiska delegationens webbsidor](#).
- Läs också riktlinjerna för europeisk forskningsetik: [European Code of Conduct for Research Integrity](#).

2.2 Hur hanteras frågor som gäller äganderätten till datamaterialet, upphovsrätt och immaterialrättigheter? Finns det hinder för användningen eller delningen av data på grund av upphovsrätt, licenser eller andra begränsningar?

Ange vem som äger projektets data och vad som överenskommits beträffande äganderätten. Ange vem som har befogenhet att ge tillstånd till (åter)användningen av datamaterialet.

Rekommendationer för bästa praxis

- Kontrollera din organisations policy i fråga om äganderätt, nyttjanderätt och rätten att sprida data.
- Avtal om äganderätten bör ingås i ett så tidigt skede av projektets livscykel som möjligt.
- Beakta också finansiärens principer för upphovsrätt och immaterialrätt.
- Det rekommenderas att all data och alla koder och all programvara görs tillgängliga för återanvändning t.ex. genom [Creative Commons](#), [GNU](#), [MIT](#) eller en annan ändamålsenlig licens.

3. Dokumentation och metadata

3.1 På vilket sätt dokumenterar du ditt datamaterial så att det är sökbart, tillgängligt, interoperabelt och återanvändbart? Vilka metadatastandarder, README-filer eller annan dokumentation kommer att användas för att andra ska kunna förstå och använda datamaterialet?

Dokumentationen av data möjliggör att andra (både människor och datorer) kan hitta, använda och citera dataset och filer. Metadata är väsentlig information som t.ex. berättar var, när, varför och hur datamaterialet samlades in, bearbetades och tolkades. Metadata kan också omfatta uppgifter om experiment, analysmetoder och forskningsmiljön.

Rekommendationer för bästa praxis

- Beskriv alla dokumentationstyper som kommer att användas (t.ex. README-filer, metadata o.d.) för att andra ska kunna hitta, förstå och återanvända ditt datamaterial.
- När du följer [FAIR](#)-principerna säkerställer du att ditt datamaterial är sökbart, tillgängligt, interoperabelt och återanvändbart.
- Använd sådana forskningsverktyg som automatiskt skapar metadata i standardiserade format. På så sätt kan dina data överföras från en tillverkarens verktyg till ett annat.
- Fundera över hur datamaterialet ska organiseras under projektet. Beskriv t.ex. praxis för hur filer namnges, versionshantering och mappstruktur.
- Identifiera vilken slags information som ska sparas för att andra forskare ska kunna hitta, få tillgång till, tolka, använda och citera dina forskningsdata.
- Datarepositorier kräver ofta att man använder en viss metadatastandard. Ta reda på om det finns ett specifikt metadata-schema eller en standard (dvs. vilka slags metadataelement som efterfrågas) som används inom ditt vetenskapsområde/forskarsamfund eller vid det repositorium där du avser att deponera ditt datamaterial

4. Lagring och säkerhetskopiering under forskningsprojektet

4.1 Var kommer datamaterialet att lagras och hur sker säkerhetskopieringen?

Beskriv var datamaterialet lagras och säkerhetskopieras under forskningsprojektet. Metoderna för bevaring och tillgängliggörande av datamaterialet efter att forskningsprojektet har avslutats presenteras utförligare i avsnitt 5.

Fundera över vem som ska ansvara för säkerhetskopiering och återställande. Om flera forskare är involverade ska du göra upp en plan med dina samarbetspartner och säkerställa en trygg överföring mellan de medverkande.

Rekommendationer för bästa praxis

- Det rekommenderas att du använder dig av en säker förvaringsplats som erbjuds och förvaltas av din egen organisations IT-avdelning.

4.2 Vem reglerar åtkomsten till datamaterialet och hur övervakas skyddad åtkomst?

Det är ytterst viktigt att trygga datasäkerheten, i synnerhet om datamaterialet är sensitivt (t.ex. innehåller personliga eller politiskt känsliga uppgifter eller affärshemligheter). Ange vem som har tillgång till ditt datamaterial, vad de har rätt att göra med det och hur du säkerställer att data tryggt överförs till dina samarbetspartner.

Rekommendationer för bästa praxis

- Åtkomsten till datamaterialet ska regleras och övervakas enligt graden av konfidentialitet.

5. Öppna, publicera och arkivera data efter avslutat forskningsprojekt

5.1 Vilken del av datamaterialet kan göras öppet tillgängligt eller publiceras? Var och när kommer datamaterialet eller dess metadata att göras tillgängliga?

Beskriv huruvida du har för avsikt att publicera datamaterialet eller på annat sätt göra det öppet tillgängligt delvis eller i sin helhet. Om datamaterialet delvis eller i sin helhet inte kan göras öppet ska du motivera varför. Öppen tillgänglighet främjar återanvändningen av forskningsdata.

Rekommendationer för bästa praxis

- Du kan publicera en beskrivning av datamaterialet (dvs. metadata) utan att göra själva datamaterialet öppet tillgängligt. På så sätt kan du begränsa tillgängligheten.
- Publicera datamaterialet i ett datarepositorium eller i en referentgranskad (data)publikation.
- Sök lämpliga repositorer på re3data.org.
- Kom ihåg att kontrollera finansiärens och vetenskapsrådets rekommendationer för datarepositorier och också de nationella rekommendationerna.
- Det rekommenderas att allt datamaterial och alla koder och all programvara som skapats inom forskningsprojektet görs tillgängliga för återanvändning t.ex. genom [Creative Commons](https://creativecommons.org/), [GNU](https://www.gnu.org/), [MIT](https://mit.edu/) eller en annan ändamålsenlig licens.
- Sträva till att använda repositorer eller förläggare som ger permanenta identifikationer (t.ex. DOI, URN) till ditt datamaterial.

5.2 Var bevaras datamaterial av långsiktigt värde och hur länge?

Beskriv kort vilka data som ska arkiveras och hur länge samt vilka data som förstörs efter att projektet har avslutats. Beskriv principerna för det arkiverade datamaterialets tillgänglighet.

Rekommendationer för bästa praxis

- Kom ihåg att kontrollera finansiärens och vetenskapsrådets rekommendationer för datarepositorier och också de nationella rekommendationerna.

6. Datahanteringsansvarsområden och resurser

6.1 Vem ansvarar för specifika uppgifter i datahanteringen under projektets livscykel? Uppskatta även de resurser som krävs för datahantering (t. ex. finansiering, tid och arbetsmängd) .

Några av ansvarsområden har redan beskrivits i föregående frågor. Du kan referera till dem eller sammanfatta alla ansvarsområden och ge en uppskattning av de resurser som behövs.

Gör en uppskattning av huruvida datahanteringen, bevaringen och tillgängliggörandet kräver experthjälp. Fundera över behovet av ytterligare IT-medel och resurser och vilka kostnader som detta medför.

Rekommendationer för bästa praxis

- Kom ihåg att specificera datahanteringskostnaderna i budgeten, enligt finansiärens krav.