



Vägledning till mallen för datahanteringsplaner

Vägledning till mallen för datahanteringsplaner

Dnr 5.2-2018-01016

Swedish Research Council
Vetenskapsrådet
Box 1035
SE-101 38 Stockholm, Sweden

Innehållsförteckning

Vägledning till mallen för datahanteringsplaner	1
Inledning	3
1 Beskrivning av data – återanvändning av befintliga data och/eller produktion av nya data	5
2 Dokumentation och datakvalitet	6
3 Lagring och säkerhetskopiering	7
4 Rättsliga och etiska aspekter	8
5 Tillgängliggörande och långtidsbevarande	10
6 Ansvar och resurser	12
Bilaga 1 Redovisning av anpassningar/ändringar	13

Inledning

Vetenskapsrådet rekommenderar användning av mallen för datahanteringsplaner¹ som grundas i de så kallade baskraven som har tagits fram av Science Europe. Science Europe har även tagit fram en vägledning som kompletterar baskraven med flera vägledande frågor. Detta dokument är en översättning och anpassning av vägledningen från Science Europe. Frågorna som återfinns i Vetenskapsrådets rekommenderade mall är markerade med grå rutor i vägledningen.

Syftet med vägledningen är att underlätta förståelsen av vad de vägledande frågorna i mallen innebär, för att ge ytterligare stöd i arbetet med datahanteringsplaner.

Vägledningen kan användas av forskare för att skapa en bättre förståelse av vad som avses med frågorna i mallen för datahanteringsplanen, men vägledningens huvudsakliga målgrupp är forskningsstödjande funktioner som stödjer forskarnas arbete med datahanteringsplaner. Vägledningen kan även vara till nytta vid framtagning av lokala eller disciplinsspecifika anpassningar.

Utifrån synpunkterna från Vetenskapsrådets arbetsgrupp för nationell samordning av datahanteringsplaner gjordes bedömningen att vissa frågor behöver anpassas eller kommenteras i översättningen. Dessa anpassningar har gjorts för att tydliggöra hanteringen av data, bland annat påverkan av befintlig lagstiftning och rutiner för exempelvis informationsförvaltning och arkivering.

I Bilaga 1 redovisas en sammanfattning av de anpassningar som har gjorts i dokumentet. Frågor i vägledningen är stödjande och formulerade som öppna frågor. Om datahanteringsplanen görs digitalt online rekommenderar vi de forskningsstödjande funktionerna att inkludera tydlig information kring huruvida det finns information som är olämpligt att fylla i online. Det kan också vara värt att överväga att erbjuda flervalsoalternativ för vissa frågor (exempelvis om datavolymer) både för att underlätta för forskarna och för att få högre maskinläsbarhet för svaren.

¹ [Mall för datahanteringsplaner - Vetenskapsrådet \(vr.se\)](#)

För dig som är forskare

Detta dokument är tänkt att användas för att öka förståelsen kring de olika aspekter som efterfrågas i basfrågorna i mallen. Observera att detta är endast stödjande frågor och att datahantering alltid ska utgå ifrån gällande lagstiftning och de riktlinjer och rutiner för informationshantering, arkivering och gallring som gäller vid myndigheten där forskningen genomförs och som påverkar hur en mall för en datahanteringsplan kan vara utformad.

Innan du besvarar frågorna i mallen för datahanteringsplaner rekommenderar vi att alltid först ta reda på de riktlinjer och rutiner för informationshantering som gäller vid lärosätet som är forskningshuvudman samt ifall det vid lärosätet redan finns en särskild anpassning av den gemensamma mallen för datahanteringsplaner som är lärosätesspecifik eller ämnesspecifik. Detta kan du göra genom att ta kontakt med de stödjande funktionerna för datahantering vid ditt lärosäte.

Administrativ information

Förutom den centrala dokumentationen nedan bör en datahanteringsplan även innehålla grundläggande administrativa uppgifter såsom projekttitel, projektledare, diarienummer eller motsvarande, datum och version av datahanteringsplanen.

1 Beskrivning av data – återanvändning av befintliga data och/eller produktion av nya data

Hur kommer data att samlas in, skapas eller återanvändas?

- a) Förklara vilka metoder eller vilken programvara du kommer att använda om nya data samlas in eller produceras.
- b) Ange eventuella begränsningar för återanvändning av befintliga data.
- c) Beskriv hur datans proveniens kommer att dokumenteras.
- d) Ange kort om du har övervägt att återanvända befintliga datakällor, men inte valt att göra det och i så fall varför inte.

Vilka typer av data kommer skapas och/eller samlas in, vad gäller dataformat och mängd/volymer data?

- e) Ange typen av data: till exempel numeriska (databaser, kalkylark), text (dokument), bild, ljud, video och/eller mixed media.
- f) Ange filformat för lagring (framgår ofta av filnamnstillägget, till exempel pdf, xls, doc, txt eller rdf).
- g) Motivera användningen av valda format. Exempel kan vara expertis hos värdorganisationens personal, att man föredrar öppna format, vedertagen standard hos dataförvaringen, utbredd användning inom forskarsamhället, eller den programvara eller utrustning som kommer att användas. Prioritera öppna format och standardiserade format eftersom de underlättar delning och långsiktig återanvändning av data (flera dataförvaringar innehåller listor över sådana ”föredragna format”).
- h) Ange volymen av data (kan anges som det lagringsutrymme som krävs (byte) och/ eller antal objekt, filer, rader och kolumner).

2 Dokumentation och datakvalitet

Hur kommer materialet att dokumenteras och beskrivas med tillhörande metadata om struktur, standarder och format för beskrivningar av innehållet, insamlingsmetod m.m.?

- a) I samråd med relevanta aktörer/funktioner, ange vilka metadata som du tillhandahåller för att hjälpa andra att identifiera och hitta data.
- b) Ange vilka metadatastandarder (till exempel DDI, TEI, EML, MARC, CMDI) som används inom den relevanta disciplinen/ämnesområde/domän. Använd vedertagna metadatastandarder där dessa finns.
- c) Ange hur data kommer att organiseras under projektet, med exempel på konventioner, versionskontroll och mappstrukturer. Konsekventa, välordnade forskningsdata är lättare att hitta, förstå och återanvända.
- d) Överväg vilken annan dokumentation som behövs för att göra återanvändning möjlig. Den kan inkludera information om metoden du använder för att samla in data, information om analyser och processer, definitioner av variabler, måttenheter och så vidare.
- e) Tänk på hur denna information kommer att samlas in och var den kommer att registreras, till exempel i en databas med länkar till varje enhet, en "readme" textfil, filrubriker, kodböcker eller labbanteckningsböcker.

Hur kommer kvaliteten hos data säkerställas och dokumenteras (till exempel repeterade mätningar, validering av datainmatning, m.m.)?

- f) Förklara hur enhetlighet och kvalitet på datainsamlingen kommer att kontrolleras och dokumenteras. Detta kan innefatta processer som kalibrering, upprepade prover eller mätningar, standardiserad datafångst, datainmatningsvalidering, peer review av data eller användning av kontrollerad vokabulär.

3 Lagring och säkerhetskopiering

Hur säkerställs lagring och säkerhetskopiering av data och metadata under forskningsprocessen?

Hur säkerställs datasäkerhet och kontrollerad tillgång till data avseende till exempel hantering av känsliga data och personuppgifter?

- a) Har dessa aspekter säkerställt i samråd med berörda funktioner vid lärosätet, t ex arkivpersonal, dataskyddsombudet eller jurister i din organisation?
- b) Observera att om du bedriver forskning vid ett svenskt statligt universitet eller en högskola omfattas ditt arbete av forskningshuvudmannens krav på arkivering och gallring.

4 Rättsliga och etiska aspekter

Hur säkerställs att data hanteras enligt de rättsregler som gäller till exempel hantering av personuppgifter, sekretess och immaterialrätt?

- a) Säkerställ med de relevanta stödfunktionerna hos forskningshuvudmannen, bland annat jurister, att hantering av personuppgifter, om det förekommer, görs i enlighet med gällande lagar och regler, till exempel Dataskyddsförordningen GDPR. Några aspekter tas upp i punkterna nedan:
- b) Kontrollera med de relevanta stödfunktionerna om det finns behov av att inhämta informerat samtycke för att hantera personuppgifter.
- c) Kontrollera med de relevanta stödfunktionerna om det finns behov av åtgärder för att anonymisera, pseudonymerna eller kryptera personuppgifter som ska hanteras samt hur dessa åtgärder görs på lämpligast sätt.
- d) Säkerställ med de relevanta stödfunktionerna ifall det förekommer behov av att etablera en rutin för reglerad åtkomst för auktoriserade användare av personuppgifter och på vilket sätt.
- e) Säkerställ med de relevanta stödfunktionerna vilka regler och rutiner som finns för tillhandahållande, respektive tillgång till data.
- f) Förklara vilka åtkomstvillkor som kommer att gälla för datan. Kommer data att vara öppet tillgängliga, eller kommer åtkomsten att vara begränsad?
- g) För projekt med flera parter och flera dataägare, säkerställ att tillgång till och användning av data regleras i konsortieavtalet.
- h) Säkerställ med de relevanta funktionerna hos forskningshuvudmannen om data omfattas av immateriella rättigheter. Överväg att använda licenser eller rättighetsmärkningar för att ange villkor för användning och återanvändning.
- i) Om befintliga data har använts inom forskning och blivit en del av forskningsdata eller utgör en ny datamängd kontrollera huruvida villkor för återanvändning av de ursprungliga data påverkar tillgång och användning.

På vilket sätt säkerställs det att data hanteras på rätt sätt utifrån etiska aspekter?

j) Tänk på om etiska frågor kan påverka hur data lagras och överförs, vem som kan se eller använda dem och hur länge de lagras. Visa att du är medveten om dessa aspekter och respektive planering.

k) Följ nationella och internationella uppförandekoder och institutionella etiska riktlinjer, och kontrollera om etisk granskning (till exempel av en etisk kommitté) krävs för datainsamling och annan datahantering inom forskningsprojektet.

5 Tillgängliggörande och långtidsbevarande

Hur, när och var kommer forskningsdata eller information om data (metadata) att tillgängliggöras?

Finns det eventuella villkor, embargo och begränsningar kring tillgång till och återanvändning av data att ta hänsyn till?

- a) Förklara hur data kan visas och delas (till exempel genom deponering i ett certifierat/betrott datarepositorium, indexering i en katalog, användning av en säker datatjänst, direkt hantering av dataförfrågningar eller användning av en annan metod), vilket lämpligen görs i samråd med berörda funktioner vid lärosätet/forskningshuvudmannen.
- b) Har du säkerställt att de gällande rutiner för arkivering och bevarandet följs, i samråd med de berörda funktionerna vid lärosätet?
- c) Ange om och när metadata respektive data kommer att göras tillgängliga (se Vetenskapsrådets rekommendation om öppen tillgång till forskningsdata). Ange om tillgång till eller användning av data kommer att begränsas, och i så fall varför och hur länge, till exempel för att publicera, skydda immateriella rättigheter eller söka patent.
- d) Ange vem som kommer att kunna använda datan. Om det är nödvändigt att begränsa åtkomsten till vissa grupper eller att använda ett datadelningsavtal, förklara hur och varför och hur länge begränsningen gäller. Förklara vilka åtgärder som kommer att vidtas för att lösa eller minimera begränsningar.

På vilket sätt säkerställs långsiktigt bevarande och av vem? Hur kommer urval av data för långtidsbevarande att göras?

- e) Har dessa aspekter säkerställts i samråd med berörda funktioner vid lärosätet, t ex arkivpersonal, dataskyddsombud eller jurister i din organisation?
- f) Observera att om du bedriver forskning vid ett svenskt statligt universitet eller en högskola omfattas ditt arbete av forskningshuvudmannens krav på arkivering och gallring.
- g) Ange de användningsområden (och/eller användare) som kan förutses för data.

Kommer det att krävas särskilda system, programvaror, källkod eller andra typer av tjänster för att kunna förstå, ta del av eller använda/analysera data långsiktigt?

- h) Ange om potentiella användare behöver specifika verktyg för att komma åt och återanvända datan. Tänk på att säkerställa hållbarheten av programvaran som behövs för åtkomst till data.
- i) Ange om data kommer att delas via ett datarepositorium, om dataförfrågningar hanteras via repositoret eller om en annan metod kommer att användas.

På vilket sätt säkerställs användning av beständiga identifierare (PID) till exempel DOI?

- j) Förklara hur data kan återanvändas i andra sammanhang. Beständiga identifierare (Persistent Identifiers, PID) bör användas så att man kan lokalisera och hänvisas till data på ett tillförlitligt och effektivt sätt. Beständiga identifierare hjälper också till att spåra citeringar och återanvändning.
- k) Ange om beständiga identifierare för data kommer att användas. Vanligtvis kommer en betrodd långsiktig repositorium att tillhandahålla beständiga identifierare.

6 Ansvar och resurser

Vem ansvarar för datahanteringen och eventuellt stödjer arbetet med detta under arbetet med forskningsprojektet? Vem har ansvar för datahantering, fortsatt förvaltning, och långtidsbevarande efter projektavslut?

- a) Beskriv rollerna och ansvarsområdena för datahantering/dataförvaltning, till exempel datainsamling, metadataproduktion, datakvalitet, lagring och säkerhetskopiering, dataarkivering och datadelning. Namnge ansvarig(a) person(er) om möjligt.
- b) För samarbetsprojekt, ange hur samordningen av datahanteringsansvar ska ske mellan parterna.
- c) Ange vem som är ansvarig för att implementera datahanteringsplanen (DHP) och se till att den granskas och vid behov revideras.
- d) Tänk på regelbundna uppdateringar av datahanteringsplanen.

Vilka resurser (kostnad, arbetsinsats eller annat) kommer att krävas för datahantering (inklusive lagring, säkerhetskopiering, tillgängliggörande och hantering för långtidsbevarande)? Vilka resurser kommer behövas för att tillse att data uppfyller FAIR-principerna?

- e) Ange hur nödvändiga resurser (till exempel tid) för att förbereda data för delning/förvaring (dataskötsel) har beaktats. Tänk igenom noga och motivera alla resurser som behövs för att leverera data. Dessa kan inkludera lagringskostnader, hårdvara, personaltid, kostnader för att förbereda data inför lagring och kostnader för lagring.
- f) Ange om ytterligare resurser kommer att behövas för att förbereda data för lagring eller för avgifter från datarepositorier. Om ja, ange belopp och hur sådana kostnader kommer att täckas.

Bilaga 1 Redovisning av anpassningar/ändringar

1. Beskrivning av data – återanvändning av befintliga data och/eller produktion av nya data

Hur kommer data att samlas in, skapas eller återanvändas?

1d) Tillägg i slutet: ”och i så fall varför inte” med syftet att undersöka om orsaken kan vara brist på öppen tillgänglighet av forskningsdata.

Vilka typer av data kommer skapas och/eller samlas in, vad gäller dataformat och mängd/volymer data?

1f) ”dataformat för lagring”, ändrat till ”filformat”, då det som anges i exemplet avser filformat och inte bevarandeformat.

1g och 1h) slogs ihop till 1g då frågan om öppna standarder borde komma i samband med kommentaren om att dessa bör prioriteras.

1g) ”vissa format” ändras till ”valda format” pga syftning till de format som har valts.

2. Dokumentation och datakvalitet

Hur kommer materialet att dokumenteras och beskrivas med tillhörande metadata om struktur, standarder och format för beskrivningar av innehållet, insamlingsmetod m.m.?

2a) Tillägg i början: ”I samråd med relevanta aktörer/funktioner”, då det i praktiken kan ske i samråd med andra aktörer, stödjande funktioner eller infrastrukturer.

2b) ”Ange vilka metadatastandarder (till exempel DDI, TEI, EML, MARC, CMDI) du använder” och 2 c ”Använd vedertagna metadatastandarder där dessa finns” slogs ihop till 2b. Formuleringen ”du använder” ändrades till ”som används inom den relevanta disciplinen/ämnesområde/domän”, på grund av att exemplifieringen avsåg den typen av standarder.

Hur kommer kvalitet hos data säkerställas och dokumenteras (till exempel repeterade mätningar, validering av datainmatning, m.m.)?

2f) ”representation med” ändrades till ”användning av” då det är användning av kontrollerade vokabulärer som avses.

3. Lagring och säkerhetskopiering

Hur säkerställs lagring och säkerhetskopiering av data och metadata under forskningsprocessen?

Hur säkerställs datasäkerhet och kontrollerad tillgång till data avseende till exempel hantering av känsliga data och personuppgifter?

Vägledande frågor i detta fall bedömdes ha koppling till informationsförvaltning och arkivering, där lokala rutiner behöver beaktas. Därför föreslås det att stödjande frågor från Science Europe byts ut mot följande:

- a) Har dessa aspekter säkerställt i samråd med berörda funktioner vid lärosätet, t ex arkivpersonal, dataskyddsombudet eller jurister i din organisation?
- b) Observera att om du bedriver forskning vid ett svenskt statligt universitet eller en högskola omfattas ditt arbete av forskningshuvudmannens krav på arkivering och gallring.

4. Rättsliga och etiska aspekter

Den huvudsakliga anpassningen i del 4 avser att frågan ”Hur kommer andra juridiska frågor att hanteras, till exempel immateriella rättigheter och ägande? Vilken lagstiftning är tillämplig?” som i ursprungsdokumentet fanns som en separat fråga utgick och istället anses omfattas av omformuleringen ”Hur säkerställs att data hanteras enligt de rättsregler som gäller till exempel hantering av personuppgifter, sekretess och immaterialrätt?”.

Detta gjordes med hänsyn till att flera olika rättsregler kan vara tillämpliga och frågan borde därmed inte omfatta endast vissa utan alla rättsliga aspekter.

Hur säkerställs att data hanteras enligt de rättsregler som gäller till exempel hantering av personuppgifter, sekretess och immaterialrätt?

- a) ”Se till att följa lagar om personuppgiftsskydd (till exempel GDPR) när du hanterar personuppgifter” har ändrats till ”Säkerställ med de relevanta stödfunktionerna hos forskningshuvudmannen, bland annat jurister, att hantering av personuppgifter, om den förekommer, görs i enlighet med gällande lagar och regler, till exempel Dataskyddsförordningen GDPR. Några aspekter tas upp i punkterna nedan.”
- b) ”Inhämta informerat samtycke för att bevara och/eller dela personuppgifter.” har ändrats till ”Kontrollera med de relevanta stödfunktionerna om det finns behov av att inhämta informerat samtycke för att hantera personuppgifter.
- c-d-e) ”Överväg att anonymisera personuppgifter som ska bevaras och/eller delas (riktigt anonyma uppgifter betraktas inte längre som personuppgifter).” har ändrats till ”Kontrollera med de relevanta stödfunktionerna om det finns behov av åtgärder för att anonymisera, pseudonymisera eller kryptera personuppgifter som ska hanteras samt hur dessa åtgärder görs på lämpligast sätt.” Därmed

ströks de två efterföljande frågorna som omfattade dessa aspekter: ”Överväg att pseudonymisera personuppgifter (den största skillnaden mot anonymisering är att pseudonymisering är kodade uppgifter och är reversibel).” samt ”Överväg kryptering, som ses som ett speciellt fall av pseudonymisering.

(Krypteringsnyckeln måste lagras separat från data, till exempel av en betrodd tredje part).

f) ”Ange om det finns en rutin för reglerad åtkomst för auktoriserade användare av personuppgifter” har ändrats till ”Säkerställ med de relevanta stödfunktionerna ifall det förekommer behov av att etablera en rutin för reglerad åtkomst för auktoriserade användare av personuppgifter och på vilket sätt.”

”Hur kommer andra juridiska frågor att hanteras, till exempel immateriella rättigheter och ägande? Vilken lagstiftning är tillämplig?” ströks (se motivering ovan, i inledningen till 4).

f) ”Ange vem som kommer att äga data, det vill säga vem som har rätt att kontrollera åtkomst” ändrades till ”Säkerställ med de relevanta stödfunktionerna kring vilka regler och rutiner som finns för tillhandahållande, respektive tillgång till data” för att förtydliga att det handlar om vem som ansvarar för tillgång till data, både vid tillfället då data tillgängliggörs och långsiktigt. Här kan flera lagar, regler, policyer och rutiner ha relevans.

g) ”Överväg att använda licenser för dataåtkomst och återanvändning” ändrades till ”Överväg att använda licenser eller rättighetsmärkningar för att ange villkor för användning och återanvändning” och flyttades till j) nedan.

h) ”Förklara vilka åtkomstvillkor som kommer att gälla för datan. Kommer data att vara öppet tillgängliga, eller kommer åtkomsten att vara begränsad? I det senare fallet, vilka? Överväg att använda licenser för dataåtkomst och återanvändning.” ändrades till

”För projekt med flera parter och flera dataägare, ta upp frågan om rätten att kontrollera tillgång till och användning av data i konsortieavtalet”.

i) ”Ange om immateriella rättigheter påverkas (till exempel Databasdirektivet, rätt av sitt eget slag-rättigheter). Om så är fallet, förklara vilka och hur de kommer att hanteras.” ändrades till ”Säkerställ med de relevanta funktionerna hos forskningshuvudmannen om data omfattas av immateriella rättigheter och hur de kommer att hanteras. Överväg att använda licenser eller rättighetsmärkningar för att ange villkor för användning och återanvändning.

j) ”Ange om det finns några begränsningar för återanvändning av tredjepartsdata” har ändrats till ”Om befintliga data har använts inom forskning och blivit en del av forskningsdata eller utgör en ny datamängd kontrollera huruvida villkor för återanvändning av de ursprungliga data påverkar tillgång och användning.”

5. Tillgängliggörande och långtidsbevarande

5a) Tillägg i slutet: ”vilket lämpligen görs i samråd med berörda funktioner vid lärosätet/forskningshuvudmannen.”.

5b) ”Beskriv din plan för att bevara data och ange hur länge data kommer att behållas.” byts ut mot ”Har du säkerställt att gällande rutiner för arkivering och bevarandet följs, i samråd med de berörda funktionerna vid lärosätet?”, med hänsyn till arkivering och bevarande enligt Arkivlagen.

5c) Tillägg i första meningen: ”Ange om och när metadata respektive data kommer att göras tillgängliga (se Vetenskapsrådet rekommendation om öppen tillgång till forskningsdata).

Tillägg/förkortning som omfattar både begränsningar i tillgång och användning, i andra och tredje meningen: ”Ange om tillgång till eller användning av data kommer att begränsas, datan ska användas exklusivt, och i så fall varför och hur länge, Ange om datadelning kommer att skjutas upp eller begränsas, till exempel för att publicera, skydda immateriella rättigheter eller söka patent.

5d) Tillägg i slutet på första meningen ”och hur länge begränsningen gäller.”

5e – h) Frågan är kopplad till arkivering så den justeras till:

e) Har dessa aspekter säkerställt i samråd med berörda funktioner vid lärosätet, t ex arkivpersonal, dataskyddsombudet eller jurister i din organisation?

f) Observera att om du bedriver forskning vid ett svenskt statligt universitet eller en högskola omfattas ditt arbete av forskningshuvudmannens krav på arkivering och gallring.

i) Tillägg i ”hanteras direkt via repositoret”, för att tydliggöra meningen, ”direkt” ströks.

Vetenskapsrådet
Västra Järnvägsgatan 3
Box 1035, 101 38 Stockholm
Tel +46 (0)8-546 44 000
vetenskapsradet@vr.se
Vetenskapsrådet.se