

# Riktlinjer för forskningsdokumentation och datahantering vid Karolinska Institutet

Dnr 1-20/2021

Gäller från och med 2021-01-26



**Karolinska  
Institutet**



**Karolinska  
Institutet**

## Innehåll

1	Inledning .....	3
2	Syfte .....	3
3	Ansvar .....	3
4	Tillstånd och avtal .....	34
4.1	Personuppgifter .....	4
5	Planering och dokumentation .....	4
5.1	Vad ska dokumenteras? .....	4
6	Datahantering .....	5
6.1	Aktiva data .....	6
6.2	Lagring, delning och arkivering av data .....	6
6.3	Tillgängliggöra forskningsdata .....	7
7	Forskarstöd .....	7
8	Länkar .....	7

<b>Diarienummer:</b> 1-20/2021	<b>Dnr för föregående version:</b> -	<b>Beslutsdatum:</b> 2021-01-26	<b>Giltighetstid:</b> Tills vidare
<b>Beslut:</b> Kommittén för forskning		<b>Dokumenttyp:</b> Riktlinjer	
<b>Handläggs av avdelning/enhet:</b> Research Support Office/ Research Data Office		<b>Beredning med:</b> Arbetsgrupp ur kommittén för forskning, referensgrupper med forskare kopplade till Research Data Office samt juridiska avdelningen	
<b>Revidering med avseende på:</b> Nytt dokument			

## 1 Inledning

God forskningssed förutsätter att hela forskningsprocessen, från tanke och idé via insamling och analyser till resultat och spridning av forskningsresultat, dokumenteras och arkiveras på ett korrekt sätt. Det är även viktigt att all forskningsdata som samlas in eller framställs inom ett forskningsprojekt hanteras korrekt under hela processen.

Forskning omgärdas av regelverk (lagar, förordningar, olika kodex och rekommendationer) som är till för att höja och säkerställa god vetenskaplig kvalitet och för att skydda forskningspersoner, försöksdjur, medarbetare och inte minst forskningens integritet och objektivitet. Dessa regelverk reglerar även hur forskningen ska dokumenteras.

Dessa riktlinjer kompletterar styrdokumentet "Riktlinjer för forskning vid KI" och "Sammanhållet kvalitetssystem vid KI", som mer övergripande reglerar forskningen vid Karolinska Institutet (KI).

## 2 Syfte

Riktlinjer för forskningsdokumentation och datahantering vid KI sammanfattar hur forskning ska dokumenteras och hur forskningsdata ska hanteras genom hela livscykeln för data och riktar sig till alla som forskar vid KI.

## 3 Ansvar

KI ansvarar för behandlingen av de forskningsdata som forskare vid KI samlar in, de data som genereras och all tillhörande dokumentation inom ett forskningsprojekt, även om projektfinansieringen är extern. Detta medför att KI ska skydda data från bland annat obehörig åtkomst samt att KI ansvarar för sekretessbedömning och ev. utlämning av data i enlighet med gällande bestämmelser.

För forskningsprojekt som bedrivs i samarbete med andra lärosäten eller aktörer är det viktigt att avtal som reglerar samverkan och ansvar gällande forskningsdata och -dokumentation upprättas. Vid varje institution har prefekten yttersta ansvaret för att forskningen, vilket inkluderar dokumentation, datahantering, lagring, publicering och arkivering, sker enligt gällande bestämmelser.

Den enskilde forskaren ska följa god forskningssed och de regler och riktlinjer som styr forskningen. Om en forskare lämnar KI och vill ha fortsatt tillgång till data, kan forskaren tillsammans med till exempel det nya lärosätet, begära att få data utlämnade. Datautlämningen i kan vissa fall regleras genom avtal.

## 4 Tillstånd och avtal

En rad olika tillstånd och avtal behövs vanligen innan ett projekt kan starta och några av dessa avtal reglerar även datahantering inom projektet.

Krav på hur data ska hanteras kan förekomma i etiktillstånd, biobanksavtal, samarbets- och uppdragsavtal, konsortialavtal (vanligt vid internationella samarbeten), personuppgiftsbiträdesavtal eller i finansierars krav på tillgängliggörande av dokumentation och data inom projekt som de finansierar.

Som forskare är det viktigt att veta vilka avtal som behövs och vad dessa reglerar. Avtalsgranskning görs vanligen av KI:s jurister.

Tillstånd, finansierars villkor och avtal ska diarieföras och arkiveras i enlighet med KI:s dokumenthanteringsplan och enkelt gå att härleda till forskningen som bedrivs.

## 4.1 Personuppgifter

Personuppgifter är information som direkt eller indirekt kan härledas till en levande person. Med känsliga personuppgifter avses uppgifter om ras eller etniskt ursprung, politiska åsikter, religiös eller filosofisk övertygelse, medlemskap i fackförening, hälsa, en persons sexualliv eller sexuella läggning, genetiska uppgifter och biometriska uppgifter för att entydigt identifiera en fysisk person.

Forskningsprojekt som samlar in och bearbetar personuppgifter ska följa kraven i dataskyddsförordningen (GDPR) och anmälas till KI:s dataskyddsombud via ett webbformulär ([registeranmalan.ki.se](https://registeranmalan.ki.se)).

All behandling av personuppgifter inom forskning ska ske i enlighet med gällande bestämmelser för dataskydd samt KI:s riktlinjer för informationssäkerhet. Vid begäran om utlämning av data innehållande personuppgifter ska en sekretessbedömning göras.

## 5 Planering och dokumentation

All forskning som bedrivs vid KI ska dokumenteras, antingen på svenska eller engelska. Medarbetare och andra ska kunna följa och granska forskningen. Det är viktigt att hela forskningsprocessen, från idé och planering till resultat och slutsatser, beskrivs och återges tydligt och detaljerat.

Att hålla forskningen organiserad genom dokumentation är centralt för att kunna redogöra för forskningsprocessen enligt god forskningssed. Forskningsdokumentation behövs också för att säkerställa spårbarhet, reproducerbarhet och återanvändbarhet, skydda upphovsrätt och patent samt underlätta samarbeten.

Forskningsdokumentation vid KI ska ske elektroniskt i godkända system, varför KI tillhandahåller ett centralt ELN-system (elektroniska loggböcker).

### 5.1 Vad ska dokumenteras?

Forskningsdokumentation ska innehålla de intellektuella och praktiska prestationerna av forskningen, och vid behov hänvisa till de administrativa dokument, såsom finansiering, tillstånd och avtal, som rör forskningen.

Mycket av den övergripande informationen kan ges i en projektbeskrivning med tillhörande datahanteringsplan som sedan kompletteras kontinuerligt med aktiv dokumentation av forskningsprocessen där metoderna och datahanteringen beskrivs och eventuella ändringar dokumenteras.

Dokumentationen bör ske löpande och i samband med att forskningen fortlöper, till exempel så dokumenteras analyser och andra bearbetningar av data och/eller material samma dag som de utförs.

Ytterligare information och dokumentation som kan finnas är till exempel korrespondens och minnesanteckningar av betydelse för forskningen samt rapporter och publikationer.

**För den löpande forskningen ska följande dokumenteras:**

- Introduktion (Bakgrund och syfte)
- Material och metod (innefattar även statistik och bearbetning av data)
- Resultat (rådata och bearbetade data)
- Slutsats

För att kunna återanvända forskningsdata bör denna beskrivas med metadata. Metadata kan till exempel inkludera information om metod som använts för att samla in data, definitioner av variabler, måtenheter, eventuella antaganden, format och filtyp för data och programvara som använts för att samla in och/eller bearbeta data.

Forskningsdokumentation ska sparas lika länge som motsvarande data sparas, minst tio år efter publicering eller avslutat projekt, men i ett flertal fall längre än så (om data sparas).

Det är viktigt att de format som används för dokumentation kan läsas under tiden de sparas. Vid behov kan dessa behöva föras över till arkivbeständiga format inför långtidslagring.

## **6 Datahantering**

Datahantering är ett samlingsbegrepp för forskarens arbete med sina data. Viktigt är att forskningsdata beskrivs tydligt, är spårbar, hanteras säkert och att det enkelt går att härleda exempelvis uppgifter i publikationer till underliggande data. All data som samlas in, mäts eller skapas inom ett forskningsprojekt räknas som forskningsdata. Data kan vara numerisk, text, bilder, video- eller ljudinspelningar. Om data inte från början samlas in i ett standardformat, så kan data behöva migreras inför långtidslagring och arkivering.

Det är viktigt att FAIR-principerna (Findable, Accessible, Interoperable and Reusable) tas i beaktande när datahanteringen planeras.

KI rekommenderar att alla forskningsprojekt använder sig av datahanteringsplaner för att underlätta god datahantering. Det är ansvarig forskare som ska ta fram och kontinuerligt uppdatera datahanteringsplanen vid behov. KI erbjuder ett webbaserat system för elektroniska datahanteringsplaner.

En datahanteringsplan innehåller vanligen en beskrivning av data, ansvar och resurser och information om hur forskningen dokumenteras och hur forskaren säkerställer datakvalitet, lagring och säkerhetskopiering samt tar tillvara rättsliga och etiska aspekter. Datahanteringsplanen beskriver också strategier/planer för tillgängliggörande och långtidsbevarande.

## 6.1 Aktiva data

I en tidig fas i projektet bör forskargruppen enas om hur filer ska namnges och hur dessa ska lagras. Sker det i olika mappar bör även namngivningen av mapparna och hierarkin vara enhetlig och konsekvent för att underlätta både spårbarheten och det dagliga arbetet.

Det är viktigt att det tydligt framgår vilken version av en fil som är den senaste samt att det ska vara möjligt att hämta tidigare versioner vid behov. Ändringar kan indikeras i filnamnet och för viktiga dokument kan en versionstabell inkluderas med datum och notering av ändringar för varje versionsnummer.

Det är viktigt att om möjligt välja filformat som är bäst lämpat för långsiktig hållbarhet och tillgänglighet. Helst bör filformaten vara vanligt förekommande och leverantörsoberoende. Vid behov kan de ursprungliga filformaten behöva föras över till arkivbeständiga format inför långtidslagring och arkivering.

### Ett filnamn kan bestå av följande delar:

- Ett prefix som visar vilken typ av dokument en fil är
- En titel: denna ska vara så informativ som möjligt så att det går att förstå dokumentinnehållet från dess titel
- Versionsnumret för dokumentet
- Datumet som visas när det skapades: använd YYMMDD-format
- Dokumentets status: utkast eller slutlig
- Författarens initialer: vem som senast ändrade det

## 6.2 Lagring, delning och arkivering av data

Data ska finnas tillgängliga för behöriga forskare och skyddas mot åtkomst av obehöriga.

Det är viktigt att använda godkända system för att lagra data. En aktuell lista över godkända system finns på medarbetarportalen (avsnitt 8). USB-minnen, privata datorer och externa hårddiskar som det inte görs regelbunden säkerhetskopiering får inte användas inom ramen för KI:s forskningsprojekt.

Vid delning av datafiler ska säkra system användas. Om data som ska delas med samarbetspartners utanför KI innehåller personuppgifter, ska avtal som reglerar tillgång till data upprättas.

Originaldata från genomförandet av forskningsprojektet ska sparas i minst tio år eller längre efter publicering eller avslutat projekt. Prover bör sparas samma tidsperiod om möjligt.

Vägledning kring vilka forskningsdata som ska bevaras och hur länge, samt vad som kan gallras (destrueras) finns beskrivet i dokumenthanteringsplanen för KI.

### **6.3 Tillgängliggöra forskningsdata**

Forskningsdata ska vara så öppna och tillgängliga som möjligt. Detta gäller även andra data som är en del av forskningen, till exempel programkod och skript.

Om data innehåller personuppgifter som direkt eller indirekt kan spåras till en nu levande person så får dessa inte göras direkt öppet tillgängliga. För att även kunna ge tillgång till dessa data så publiceras metadata öppet och sedan krävs det reglerad åtkomst till data innehållande personuppgifter.

Det kan även finnas andra skäl till sekretess än personuppgifter, till exempel kopplat till avtal och/eller patent, som hindrar data från att göras öppet tillgängliga.

## **7 Forskarstöd**

Stöd via Research Data Office (RDO, [rdo@ki.se](mailto:rdo@ki.se)) fokuserar specifikt på forskningsdokumentation och datahantering, till exempel lagring, arkivering och tillgängliggörandet.

## **8 Länkar**

Mer detaljerad information, länkar och hjälp finns på [RDO:s hemsida](#).