

Vedlegg C - Datahåndteringsplan

FoU-prosjekt [Prosjektnavn]

BESKRIVELSE AV DATA

Oppgi hensikten med datainnsamlingen / datagenereringen og hvordan data er knyttet til prosjektets formål.

Her kan du si noe om hvilke problemstillinger, forskningsspørsmål og hypoteser som tas opp i prosjektet, og hvordan disse kan belyses gjennom datainnsamlingen/genereringen.

Skal det samles inn og/eller genereres nye data i prosjektet? JA NEI

Hva er data? Forskningsrådet definerer forskningsgenererte data som "...registreringer, nedtegnelser og rapporteringer i form av tall, tekst, bilder og lyd som genereres eller oppstår underveis i forskningsprosjektet". Se [Forskningsrådets policy for tilgjengeliggjøring av forskningsdata \(PDF\)](#) og [generelle vilkår for FoU-prosjekter \(PDF\)](#).

Skal det brukes allerede eksisterende data i prosjektet? JA NEI

Hva er sekundærdata? Allerede eksisterende data, eller sekundærdata, refererer til data og informasjon som er samlet inn av noen andre enn brukeren. Primærdata er derimot data som er samlet inn av forskeren(e) eller prosjektet som utfører undersøkelsen.

Vær oppmerksom på at data som gjenbrukes ofte kan ha et sett av bruksbetingelser knyttet til seg. Selv om gjenbruk av data i mange tilfeller være uproblematisk, vil det for enhver forskning som innebærer gjenbruk av data samlet inn fra personer og som inneholder personopplysninger, forventes at forskere tar nøye hensyn til den etiske risikoen og oppsøke hjelp fra relevante tilsyn dersom man er i tvil.

Gjenbruk av data som allerede er offentlig tilgjengelig (dvs. publisert hos statistisk sentralbyrå, eller gjort tilgjengelig i bøker, tidsskrift, o.l), og gjenbruk av data der deltakere har gitt samtykke til bruk av data til forskningsformål utover det formålet de opprinnelig gav samtykke til, og/eller som er anonymiserte, er det som regel få eller ingen restriksjoner på å gjenbruke. Eierskapet til de opprinnelige dataene må imidlertid alltid tilkjennevis.

Hvis JA, spesifiser og forklar hvilke eksisterende data som skal benyttes, og oppgi kilden.

Skisser nytteverdien av data; hvem vil ha praktisk nytte av data?

Bruk dette feltet til å si noe om hvem som kan ha nytte av dataene dine. Dette kan være et fagområde eller en fagdisiplin, yrkesgrupper, forskere innenfor spesifikke fagfelt, institusjoner, organisasjoner, etc. Si eventuelt også noe om hvorvidt data vil kunne benyttes, kombineres eller virke sammen med, data fra andre forskere, prosjekt, institusjoner, organisasjoner, e.l.

Spesifiser fagområde(r) for prosjektet ditt.

- Bærekraftig systemutvikling
- Innovativ teknologi
- Smarte Nett
- Annet:

Oppgi relevante tema- og emneord som kan bidra til å beskrive dataene dine.

En gjennomtenkt utvelgelse av søkeord betyr at andre forskere kan finne, lese og sitere forskningsresultatene dine. Ved å optimalisere dataene dine for søkemotorer, vil du øke sjansen for at noen finner dem. Nøkkelord kan for eksempel angi tema/fagområde og/eller datainnsamlingsmetoder.

Oppgi hvilke(t) språk datamaterialet og dokumentasjon vil være på.

- Norsk
- Engelsk
- Annet:

TEKNISK INFORMASJON

Oppgi hvilke metoder (teknikker, prosedyrer) som benyttes for å samle inn eller generere data i prosjektet.

Eksempler på innsamlingsmetoder er eksperimenter, kliniske studier, observasjoner, simuleringer, måling og testing, intervjuer, spørreskjema, opptak/innspilling, etc.

Hvilke typer data vil samles inn / genereres?

- | | | |
|-----------------------------------|--------------------------------|---------------------------------|
| <input type="checkbox"/> Datasett | <input type="checkbox"/> Bilde | <input type="checkbox"/> Video |
| <input type="checkbox"/> Lyd | <input type="checkbox"/> Tekst | <input type="checkbox"/> Annet: |

Kommentar

Dersom mulig, gi et estimat av den antatte størrelsen/omfanget av data.

- | | | |
|--|--|---|
| <input type="checkbox"/> Mindre enn 1 GB | <input type="checkbox"/> 100 GB – 1 TB | <input type="checkbox"/> Mer enn 100 TB |
| <input type="checkbox"/> 1 GB – 100 GB | <input type="checkbox"/> 1 TB – 100 TB | <input type="checkbox"/> Vet ikke |

Kommentar

Oppgi navn på programvare/verktøy og eventuelle filformat som vil gjelde for de aktuelle data. Si gjerne også noe om hvorvidt data vil bli gjort tilgjengelige i åpne eller 'lukkede' (proprietære) format etter prosjektslutt.

Oppgi det tekniske format på data som genereres i prosjektet, f.eks. R, STATA, SPSS, SAS, CSV, TXT, etc.

ETISKE OG JURIDISKE HENSYN

Oppgi hvem som skal ha opphavs- og eiendomsretten til data som skal genereres eller samles inn i prosjektet. Rettighetshaver kan være en eller flere personer og/eller en eller flere institusjoner.

Skal det samles inn eller genereres data om personer? JA NEI

Skal det samles inn eller behandles personidentifiserende data? JA NEI VET IKKE

Hva er personidentifiserende data?

En personopplysning er en opplysning eller vurdering som kan knyttes til enkeltpersoner, slik som for eksempel navn, adresse, telefonnummer, e-postadresse, IP-adresse, bilnummer, bilder, fingeravtrykk, irismønstre, hodeform (for ansiktsgjenkjenning) og fødselsnummer (både fødselsdato og personnummer), eller gjennom en kombinasjon av bakgrunnsopplysninger. Opplysninger om atferdsmønstre kan også regnes som personopplysninger.

Sensitive personopplysninger er opplysninger om rasemessig eller etnisk bakgrunn, eller politisk, filosofisk eller religiøs oppfatning, at en person har vært mistenkt, siktet, tiltalt eller dømt for en straffbar handling, helseforhold, seksuelle forhold eller medlemskap i fagforeninger.

Les mer om personopplysninger hos [Datatilsynet](#) og hos [Personvernombudet for forskning](#).

Hvis JA, har prosjektet ditt vært i kontakt med personvernombud eller andre tilsyn?

Dersom prosjektet skal arkivere personidentifiserbare data, må dataeier dokumentere at de har lov til å oppbevare dataene i henhold til gjeldende regelverk. Dersom dataeier ikke har tillatelse til å oppbevare personidentifiserbare data, må data anonymiseres slik at anonym versjon kan arkiveres før de originale dataene slettes.

- Ja, har vært i kontakt
- Nei, har ikke vært i kontakt
- Nei, men kommer til å ta kontakt

Hvis JA, oppgi hvilke tilsyn prosjektet ditt har vært i kontakt med. Dersom du har fått tildelt saksnummer fra noen av tilsynene, kan det også oppgis her.

- NSD - Personvernombudet for forskning Saksnummer:
- REK - Regionale komiteer for medisinsk og helsefaglig forskning Saksnummer:
- Datatilsynet Saksnummer:
- Andre: Saksnummer:

SIKKERHET, HÅNDTERING OG OPPBEVARING AV DATA

Hvem vil ha hovedansvar for håndtering og oppbevaring av data underveis i prosjektet?

Det bør være én person som har hovedansvar for håndtering og oppbevaring av data i prosjektperioden. Dette trenger ikke være prosjektansvarlig.

Hvordan blir data dokumentert i løpet av forskningsprosessen? Vil det f.eks. bli benyttet noen metadatastandard(er) som er tilpasset din forskningsdisiplin?

Dersom du ikke har tenkt å bruke en metadatastandard, eller det ikke finnes relevante metadatastandarder tilpasset din forskning, si noe om hvilke andre strategier (f.eks. hvorvidt det benyttes fritekstdokument, kode-syntakser eller lignende) som brukes for å dokumentere data. Merk at bruk av en standard strukturerer informasjon i maskinlesbar form, noe som øker synlighet og gjenfinnbarhet for dataene dine i fremtiden.

Hvordan vil tilgangskontroll og datasikkerhet håndteres underveis i prosjektet? Dette gjelder spesielt dersom prosjektdata inneholder personidentifiserbare data eller annen sensitiv informasjon og der dette skal deles mellom flere samarbeidspartnere.

Hvordan vil data oppbevares og sikkerhetskopieres (backup) i løpet av prosjektperioden?

Dersom datamaterialet inneholder personidentifiserbar informasjon må data håndteres og lagres i tråd med samtykket fra de som deltar, samt være i tråd med de tillatelser som er gitt. Det må også være i tråd med virksomhetens bestemmelser om datasikkerhet og i tråd med relevante lovverk.

Hvordan vil filer og mapper organiseres og navngis i løpet av prosjektperioden? Si også gjerne noe om strukturering, håndtering og versjonering av data(filer).

LANGTIDSLAGRING OG DELING AV DATA

Hvilke (deler av) data skal bevares, deles og langtidslagres?

Du bør ha et bevisst forhold til hvilke deler av data som er vurdert til å ha verdi på lang sikt, og hvordan disse skal forvaltes. Det bør samtidig foreligge en plan for hvordan data som vurderes til ikke å ha langsiktig verdi, skal forvaltes, eventuelt destrueres etter en viss tid.

Oppgi hvilke data som skal gjøres åpent tilgjengelige. Dersom enkelte datasett eller deler av data ikke kan deles, forklar hvorfor. Du bør skille tydelig mellom juridiske og kontraktsmessige årsaker, og årsaker til at du selv ønsker å ha tilgangsbetingelser på dataene.

Oppgi hvilke tilgangsbetingelser som vil gjelde i tilfeller der data kan deles, men med begrensninger.

I hvilket arkiv/datasenter skal data og tilhørende metadata lagres etter at prosjektet er avsluttet?

Dataene som ligger til grunn for vitenskapelige artikler bør deponeres ved egnet arkiv og gjøres tilgjengelig så tidlig som mulig, og som regel aldri senere enn ved publiseringstidspunkt. Andre data som kan være av interesse for annen forskning, bør også gjøres tilgjengelig innen rimelig tid, og som regel aldri senere enn tre år etter endt prosjekt. Se også [Forskningsrådets retningslinjer for tilgjengeliggjøring av forskningsdata \(PDF\)](#).

Si noe om hvorvidt det bli brukt persistente identifikatorer (som for eksempel Digital Object Identifier - DOI) for å kunne lokalisere data og metadata etter at prosjektet er avsluttet.

Det er spesielt viktig å være oppmerksom på identifisering og gjenfinnbarhet av data dersom du ikke benytter deg av et datasenter eller en arkiveringstjeneste som tilbyr persistente, standardiserte identifikatorer.

Vil det være kostnader knyttet til tilgjengeliggjøring og deling av data? Si gjerne også noe om den potensielle verdien av å gjøre data delbare og tilgjengelige for gjenbruk.

Øvrige kommentarer til deling og tilgjengeliggjøring av data.