

Innspillsrunde om FAIR forskningdata

Jeg svarer

på vegne av en institusjon/organisasjon

Kontaktinformasjon

Navn på institusjon/organisasjon

Kreftregisteret

Navn på lavere enhet/forskergruppe (hvis relevant)

Ikke besvart

Navn på personen som svarer

mari.nygard@kreftregisteret.no

E-post til kontaktperson

mari.nygard@kreftregisteret.no

Hvis du svarer på vegne av en enhet/forskergruppe eller som enkeltperson: Hvilke(t) fagområde jobber du innen?

Det er mulig å krysse av for flere.

Medisin og helsefag

Om du har oppgitt "Annet", vær vennlig å spesifisere

Ikke besvart

Publisere og presentere innspill

I utgangspunktet har vi ønske om å publisere alle innspill på egen nettside. For enkeltpersoner er det mulig å be om at dette ikke skjer. Ønsker du at ditt innspill skal publiseres?

Ja

Presentasjon på innspillsseminar

Er du interessert i å presentere innspillet i et kort (digitalt) innlegg på innspillsseminaret 19. januar? (3-5 min avhengig av hvor mange som melder seg.)

Ja

Nedenfor er en liste med noen av de spørsmålene vi gjerne vil ha innspill på. Dette er ment som ideer til dere som leser. Dere står fritt til å kommentere både dette og andre sider ved rapporten.

Mangler det viktige elementer i beskrivelsen av dagens tjenester?

Mangler det viktige elementer i beskrivelsen av manglene i dagens tjenester og behovene for å forbedre disse?

Er det andre problemstillinger eller anbefalinger som burde vært løftet?

Er det problemstillinger eller anbefalinger som burde vært fjernet?

Har du/dere forslag til case som burde være med for å belyse problemstillinger knyttet til forskningsdata på en god måte?

Har du/dere forslag til case på miljøer som har gjort litt i retning FAIR, som ikke er super-eksperter,

men som kan være gode forbilder for de som ønsker å komme i gang?

Dere kan gi innspill enten ved å fylle ut i feltene nedenfor, ved å legge ved en fil med kommentarer i (f.eks. Pdf-fil med kommentarer eller Word-fil med «spor endring») eller som en kombinasjon. Har du spørsmål til utfyllingen, send e-post til kristin.braenden@agendakaupang.no

Legg eventuelt ved fil her:

Har du behov for å legge ved flere filer, kan disse sendes på e-post til kristin.braenden@agendakaupang.no

Ikke besvart

Kommentarer til rapporten som helhet (ikke spesifikt til ett kapittel)

- Dette er et VIKTIG tema

- Framstillingen framstår likevel lite spesifikk, hvert fall for helsedata og omfattes noen ganger som «store ord» - uten at det kommer klart fram hvordan dette kan operasjonaliseres. Forslag: omtale problemstillingen knyttet til helsedata separat

- Skille mellom meta-data og data må være mye mer tydelig på tvers av rapporten. Vi må kunne lage bedre metadata og dele dette. Data på individnivå er det derimot et strengt lovverk rundt, og datadeling her må begrenses. Trenden går mot 1) fødererte analyser 2) tilgjengeliggjøring av data ved fjernaksess i en begrenset periode, 3) bruk av syntetiske data.

- Rapporten inneholder lite framtidsrettede elementer som fødererte analyser og bruk av syntetiske data

- Rapporten henviser til at resultater (og ikke rådata) kan gjenbrukes ved såkalte meta-analyse – en viktig del av helseforskning.

- Begrepsbruk/definisjoner

Det er gjort en definisjon av begrepet «forskningsdata» og «åpne data», jf. pkt. 2.5. Ingen av disse begrepene finnes i regelverket, men så lenge meningsinnholdet fremgår, er det forståelig hva som menes. Bl.a. fremgår at «forskningsdata» alltid er resultatet av forskningen og kan inneholde personopplysninger (men behøver ikke gjøre det). Videre at «åpne data» er digitale data osv. (her bør det stå «digitale forskningsdata»). Men begrepsbruken ellers i dokumentet er ikke konsekvent, jf. «sensitive data», «sensitive persondata» «sensitive forskningsdata» og «personopplysninger». Det bør gjøres en opprydding her, og man bør bruke begreper som leser finner igjen i regelverket.

Det som står på s. 16 under «Klassifisering» fremgår flg.:

«Når det gjelder forskningsdata så kan disse kategoriene anvendes for å avgjøre hvor og hvordan data skal håndteres, lagres, prosesseres og arkiveres. Det som gjerne avgjør om forskningsdata havner i den ene eller andre kategorien, er om data inneholder personopplysninger (altså enhver opplysning om en identifisert/identifiserbar enkeltperson) eller ikke. Forskningsdata som ikke inneholder personopplysninger og/eller som er anonyme klassifiseres gjerne som grønne data, mens data med personopplysninger gjerne klassifiseres som gule, røde eller sorte, avhengig av grad av sensitivitet og datamengde.

Det finnes også andre grunner til at data krever ekstra beskyttelse. Eksempler på tilfeller der data ikke kan deles (eller kun kan deles med begrensninger) kan være kontraktsforpliktelser og taushetsavtaler med oppdragsgiver/finansør, interesser knyttet til opphavsrettigheter og patenter, kommersielle interesser, sensitiv informasjon om bedrifter eller organisasjoner, og lignende.»

Begrepet «personopplysninger» bør derfor defineres i pkt. 2.5, og være uttømmende og i samsvar med definisjonen i personvernforordningen art. 4 nr. 1. Definisjonen av personopplysninger under klassifisering i parentes; altså enhver opplysning om en identifisert/identifiserbar enkeltperson, er for enkel, og vil da kunne fjernes.

Det vil være en fordel at dette dokumentet kvalitetssikres med «juridiske øyne» med hensyn på begrepsbruk.

- Mulig feilslutning og feil begrepsbruk

Vi har også merket oss at det på s. 16, nest siste avsnitt, heter: «Dersom dataene er en serie røntgenbilder, kan metadata angi hvilken kroppsdel bildene viser, fødsels- og personnummer på personene det er bilde av osv.» Det er vel litt selvmotsigende at metadata kan angi personidentifikasjon?? (Det heter forøvrig «fødselsnummer» som består av fødselsdato (6 sifre) og personnummer (5 sifre).

Kommentarer til kapittel 2 – Innledning

Ang Bakgrunn:

- UNIT eksisterer ikke lengre og dagens organisering (SIKT) bør redegjøres for lesere
- Hva er konkrete konsekvenser av denne rapporten?
- Bra at man sammenstiller de tre delvis parallelle forskningsdata utvalg.
- Vi savner at det ikke omtales at helseforskningen har et lovverk som hindrer direkte datadeling
- Selv innenfor helseforskningens stramme rammer har vi kommet langt med deling av resultater for å generere ny kunnskap: meta-analyser en viktig del av feltet

Kommentarer til kapittel 3 – Hvilke tjenester og verktøy brukes i dag?

Savner tydelig skille mellom publisering av metadata og publisering av data. Spesielt med tanke på muligheter og tidspunkt når informasjon kan publiseres fordi data vil sannsynligvis bli publisert ved prosjektslutt mens meta-data kan publiseres fortløpende. Insentiver kan komme blant annet fra finansierer.

Datahåndteringsplan burde være MUST og regulert/kontrollert av forskningsinstitusjonen. Retningslinjer bør være tilgjengelig.

Syns også at det er viktig å skille mellom databehandlings tjenester i.e. software og hardware fra samhandlings muligheter MO365, sharepoint etc. Samhandlings software kan absolutt brukes for å beskrive, lagre metadata.

Arkivering og publisering: mangler refleksjoner rundt publisering av resultater som kan brukes til meta-analyser. Definisjon av meta-analyse bør komme tydelig fram. Her er fare for misforståelser.

Kommentarer til kapittel 4 – Hva mangler?

Det var en del : hvordan kan man søke etter metadata? Burde metadata være tilgjengelig på ett sted?

Burde delte data være tilgjengelig på en felles lagringsplass? Hvor skal en slik lagringsplass være?

Metadata er et stort og viktig felt. Innen helse brukes kodeverk som Snomed CT, ICD-10, ICD- O-3, ATC koder osv. En samordnet bruk av koder gjør at vi kan sammenlikne data på tvers av land og vil forenkle fødererte analyser. OMOP initiativet er et eksempel.

Kommentarer til kapittel 5 – Diskusjon

- RDA har laget en oversikt over relevante standarder innen ulike disipliner – men det mangler referanse til helsedata!
- burde man diskutere om mulighet for en Nasjonal lagringstjeneste?
- Forstår ikke tankegangen relatert til at deling av metadata og data er omtalt på samme måte. Deling av meta-data er helt uavhengig av deling av data. Etterlyser bedre struktur av diskusjon/rapporten

Kommentarer til kapittel 6 – Anbefalinger

Ikke besvart