

Innspill fra Sigma2

Sigma2 ser stort potensial i innholdet i rapporten og takker for muligheten for å gi tilbakemelding til selve rapporten samt prosessen og metodikk. Rapporten er en viktig rapport og inneholder mange gode momenter, og i det følgende ønsker vi å påpeke punkter som bør forbedres i den endelige versjonen.

De første tre kapitler fremstår på flere områder som noe overfladiske i nåværende versjon og har derfor begrenset verdi. Vi antar det er fordi disse er ment å være summariske og at det er implisitt at lesere er godt kjent med området (burde nevnes med henvisning til bakgrunnsmateriale).

Vi stiller spørsmål ved representativ deltakelse i spørreundersøkelsen. For eksempel har Sigma2 sin årlige brukerundersøkelse svar som representerer et mye større antall forskere, der det er mange spørsmål som er overlappende med spørreundersøkelsen som har vært gjennomført her. Prosjektet har unnlatt å ta hensyn til enkelte innspill og underlag som er gitt under prosessen, og det er ingen kommentarer knyttet til korrelasjoner eller avvik mellom denne eller andre spørreundersøkelser. Det er heller ikke gjort noen profilering/analyse av spørreundersøkelsen i forhold til deltakelse og representativitet og metodikk. For eksempel en enkel statistikk som viser andel som har svart fra hvilke vitenskapelige områder og roller.

Basert på svar i den begrensede spørreundersøkelsen, fremtrer et klart behov for et større kommunikasjonsarbeid for å informere forskere om hvilke muligheter og løsninger som allerede eksisterer. For eksempel at NIRD Arkivet er åpent og tilgjengelig for alle forskere og uavhengig av størrelse på datasett.

Rapporten er på flere områder ikke representativ for de løsninger og tjenester som allerede er tilgjengelige. Rapporten sier «De infrastrukturene som etableres bør ha et livsløpsperspektiv for forskningsdata, hvor det fra forskerens ståsted er en sømløs brukeropplevelse fra datahåndteringsplaner til arkivering av data — og gjenbruk av disse dataene.», uten å gjøre en vurdering i hvilken grad eksisterende tjenester kan tilby dette, evt. hvilke mangler disse har.

Vi støtter synspunktet til rapporten om at finansiører, f.eks. Forskningsrådet, har en helt sentral rolle i forhold til insentiver for datahåndtering og deling av data. Vi tror også at «behovet for økt datadeling innebærer økt fokus på data og mindre fokus på publikasjoner».

Rapporten unnlater å kommentere utfordringer knyttet til bærekraften (langsiktighet og finansiering) for internasjonale disiplinspesifikke infrastrukturer. Vi får et inntrykk av preferansen for datalagring er:

- 1) Internasjonal disiplinspesifikk lagringstjeneste
- 2) Nasjonal disiplinspesifikk lagringstjeneste
- 3) Nasjonal generisk lagringstjeneste
- 4) Lokal (institusjonell) lagringstjeneste

Dette kan være en god prioritering, men dette bør i så fall være nærmere begrunnet. Man bør også se på hvordan man kan få en god ressursutnyttelse og bærekraftig finansiering for Nasjonale disiplinspesifikke lagringstjenester. Her kan samarbeidet mellom Elixir Norge og Sigma2 med

tjenesten NELS (Norwegian e-infrastructure for Lifescience) være et godt eksempel, hvor den underliggende infrastrukturen står på den nasjonale generiske e-infrastrukturen, og tjenesteleveransen kommer fra ELIXIR.no. Denne modellen kan være anvendbar for mange og bør fremkomme som en mulighet i rapporten.

FAIR står sentralt og i tittelen til rapporten, men leveranser etterspurt i mandatet og selve rapporten fokuserer lite på hva de enkelte F, A, I, R prinsippene innebærer fra et funksjonelt ståsted, og er derfor ikke til stor hjelp for tjenesteleverandører som skal implementere løsninger. For eksempel hvordan maskin håndterbare løsninger skal/bør implementeres.

Sigma2 er enig i anbefalingen "Leverandører bør gi mulighet for disiplinspesifikke og institusjonsspesifikke tilpasninger av maler for datahåndteringsplaner". Sigma2 anser også som avgjørende at mekanismen for å gjennomgå og konsumere informasjon fra DMPer er standardisert, for å støtte bedre flere prosesser: ressurstildeling (finansieringsmyndigheter), kapasitetsplanlegging (tjenesteleverandører) og vitenskapelig produksjon (forskere). Sigma2 ser frem til en nasjonal dialog om interoperabilitet og APIer for DMP-verktøy.

Sigma2 er enig i anbefalingen: "Det bør vurderes å etablere (evt utvide eksisterende) tjenester nasjonalt for arkivering og langtidsbevaring av persondata/sensitive data." og Sigma2 viser til arbeidet som Sigma2 allerede har startet på dette området sammen med sektoren (<https://www.sigma2.no/project-new-research-data-archive>)

Data fra offentlig virksomhet og industrien kan ha stor verdi for forskning, og rapporten sier ikke noe om behovet for å få tilgang til data fra industrien og spesielle krav knyttet til dette. Et eksempel på en slik løsning kan være <https://co2datashare.org/> der Equinor, Sintef og Sigma2 har samarbeidet om dette.

For øvrig, og til tross for noe manglende støtte i underbyggingen (i de første tre kapitlene), konstaterer vi at det mange av anbefalingene som sammenfaller med innspill som Sigma2 har fått fra forskere gjennom våre brukerundersøkelser og fra Sigma2 sin kundebase.