



UNIVERSITETET I OSLO

Digitaliserings- og forvaltningsdepartementet

Per e-post til: postmottak@dfd.dep.no

Dato: 12.12.2024

Deres ref.: 24/3158

Vår ref.: 2024/17877

Høringsinnspill til NOU 2024: 14 Med lov skal data deles

Vi viser til høringsbrev av 19. september 2024 med vedlagt NOU med opprinnelig høringsfrist 9. desember, samt dialog med Silje Johansen om utsatt høringsfrist til 12. desember 2024.

Innledende merknader

Universitetet i Oslo (UiO) ønsker med dette å bidra til høringen om forslagene til datadelingslov og dataforvaltningslov. Som en ledende akademisk institusjon i Norge ønsker vi å fremme synspunkter og anbefalinger som vi mener er avgjørende for å sikre en balansert og hensiktsmessig implementering av åpne data-direktivet (EU 2019/1024), gjennomføringsrettsakten om datasett med høy verdi (EU 2023/138) og dataforvaltningsforordningen (EU 2022/868) i norsk rett.

UiO har gått gjennom utvalgte deler av NOU 2024: 14. Vi merker oss at utredningen har ulike betydninger for universitetet som *bruker* og som *produsent* av forskningsdata.

Viderebruksutvalget trekker fram et meget godt poeng at «det bør fremkomme av formålsbestemmelsen at loven har som formål å fremme forskning», se side 102. Dette støtter UiO. Deling av offentlig data er veldig bra for forskning – som er data *til* forskning.

Dette må imidlertid ikke sammenblandes med deling av forskningsdata – altså data *fra* forskning, som ikke bare er nasjonal, men internasjonal. Som produsent av forskningsdata er UiO bekymret for at de særpreg som gjelder slike data ikke er godt nok ivarettatt i utredningens lovforslag.

Hovedpunkter i UiOs høringssvar

- Akademisk frihet er et premiss for deling av data fra forskning, og forsker må selv kunne styre delingstidspunkt og delingskanal. For data fra forskning må reglene gjøres så fleksible som direktivet tillater.



- For den tekniske gjennomførbarhet, metadatastandarder og prinsippet om innebygd åpenhet vil det være avgjørende at nye løsninger er compatible med eksisterende løsninger og standarder for forskningsdata, slik at systemene kan samhandle med et minimum av manuelle tilpasninger.
- Utforming av en forskrift om lisenser som supplerer datadelingsloven § 9 vil være sentral for delingen av forskningsdata, og UiO ønsker å bidra i dette arbeidet.
- Forventet arbeidsmengde og ressursbruk er etter UiOs mening betydelig, og implementeringen av disse regelverkene må dekkes av øremerkede bevilgninger til oppgradering av infrastruktur, opplæring, mv.

Akademisk frihet som premiss for datadelingsforpliktelsene

Akademisk frihet handler ikke bare om adgangen til fritt å velge forskningsspørsmål og forskningsmetode, men også retten til selv å velge publiseringstidspunkt og egnet publiseringskanal. For UiO er det sentralt at reglene ikke utformes på en slik måte at disse rettighetene påvirkes.

Tidspunktet for tilgjengeliggjøring (utkastet til datadelingsloven § 8)

UiOs forskere *må* selv ha muligheten til å bestemme når forskningsdata skal offentliggjøres, slik det virker som lovutkastet legger til rette for. Det innebærer friheten til å vente med tilgjengeliggjøring av forskningsdata til den tilhørende publikasjonen er publisert. Videre bør UiOs forskere selv kunne vurdere når datasettet er tilstrekkelig kvalitetssikret for tilgjengeliggjøring, eller om hensynet til fremtidig internasjonalt samarbeid gjør det ønskelig å utsette eller begrense tilgjengeliggjøringen.

I enkelte tilfeller må forskningsmiljøene også få anledning til å styre fremtidig bruk av innsamlede datasett for å ivareta forskningsmiljøenes egne interesser og konkurransedyktighet, noe vi mener er i samsvar med hensynet til "legitime forretningsinteresser" i åpne data-direktivet artikkel 10 nr. 2.

Hvor tilgjengeliggjøringen skal skje (utkastet til datadelingsloven § 6)

Hvor forskningsdata arkiveres kan variere avhengig av dataarkiver som er tilgjengelige og hensiktsmessige for de ulike forskningsdataformatene. Ulike fagområder bruker internasjonale forskningsdataarkiv, mens andre benytter institusjonelle arkiv som DataverseNO eller nasjonale dataarkiv som MODANSA. Noen tidsskrifter har egne dataarkiv eller anbefaler spesifikke arkiv.

UiO ser et behov for at lovutkastet tar hensyn til den internasjonale utviklingen innen forskning, og tar i betraktning at forskningsdata ofte er en del av internasjonale prosjekter som organiseres innenfor internasjonale forskningsinfrastrukturer. UiO ser det ikke som hensiktsmessig å indeksere forskningsdata til en nasjonal portal som data.norge.no, ettersom data som er publisert i dataarkiv allerede høstes av relevante internasjonale portaler.



Teknisk gjennomførbarhet, metadatastandarder og prinsippet om innebygd åpenhet og åpenhet som standardinnstilling (utkastet til datadelingsloven §§ 5 (4) og 6, mv)

For at viderebruk og sammenstillinger av (meta-)data skal bli en realitet og ikke bare en rettslig mulighet, er samhandlende datasystemer og felles begrepsapparater helt nødvendig.

UiO vil understreke viktigheten av samhandlende datasystemer ved innføringen av plikter rundt tilgjengeliggjøring av forsknings(meta-)data. Det er avgjørende at disse systemene kan overføre informasjon med minimal bruk av manuelle klikk og at de kan høste eksisterende metadata. Disse systemene må, med andre ord, være kompatible og kunne kommunisere med hverandre.

Persistente identifikatorer som resultatID - DOI, forskerID – ORCID og institusjonsID ROR er identifikatorer som kan tagges på forskningsdata når de arkiveres. På den måten vil det bli enklere å høste forskningsdata til registre som data.norge.no. I tillegg er det viktig at systemene understøtter eksterne finansiører, i særdeleshet EU-kommisjonen, tekniske og formelle krav til åpne data.

UiO fraråder bruken av DCAT-AP-NO på det nåværende tidspunkt. Selv om vi anerkjenner at DCAT er i ferd med å etableres som en høynivå-metadatastandard i Europa, er det fortsatt umodent og ikke tilstrekkelig testet i praksis. UiO foreslår derfor at det gis en anbefaling om bruk av internasjonale metadatastandarder som kan oversettes til domenespesifikke DCAT-standarder i fremtiden.

På UiO mener vi at det bør utvikles datasystemer som tilrettelegger for datadeling. Vi støtter derfor den delen av utvalget som ønsker en konkret henvisning til prinsippet om «innebygd åpenhet og åpenhet som standardinnstilling» i datadelingsloven § 5 fjerde ledd.

Om standardlisenser (utkastet til datadelingsloven § 9)

UiO har kommentarer og innvendinger til drøftelsene rundt valget av lisens for data og datasett, og mener relevante skillelinjer må presiseres tidligere og mer utfyllende enn det er gjort i NOU-ens kapittel 11.8.2, siste avsnitt. For forskere er det sentralt å bli kreditert for data de har delt og å kunne vise til at dere datasett har blitt brukt for å begrunne egen «impact» i videre karriere og tilskuddssøknader. Kreditering for hel eller delvis finansiering av data med stor gjenbruksverdi kan også være sentralt for en finansiør som har finansiert datagenereringen. En slik finansiør forventer å bli kreditert ved videre bruk av dataen, i det minste som en integrert del av metadata.

Etter UiOs mening må det generelt skilles mellom data og databaser som er regnet som åndsverk, og data/databaser som ikke er det. For forskningsdata er det viktig at skillet mellom opphavsrettslig og ikke-opphavsrettslig materiale er tydelig for at både forskere og studenter skal forstå hvilken adgang det er til gjenbruk.

Når det er snakk om åndsverk, kan ikke opphaveren(e) frasi seg navngivelsesretten på generelt grunnlag etter åndsverksloven § 5. For åndsverk anbefaler UiO derfor at CCBY-lisens (forutsatt at man bruker CC-



lisenser), slik at navngivelsesretten gjenspeiles i valget av lisens. Hvis det ikke er snakk om åndsverk, mener UiO at en PDM-merking kan være riktigere.

NOUen viser på side 144 til "Commission Notice, 2014/C 240/01, «Guidelines on recommended standard licences, datasets and charging for the use of documents»." Utvalget skriver "... og at man for fremtiden lisensierer data under en Creative Commons-lisens i samsvar med EUs retningslinjer for anbefalte standardlisenser." Vi kan ikke se at retningslinjene anbefaler bruk av CC-lisenser. CC-lisensene er designet for opphavsrettsbeskyttet materiale. Innad i EU brukes det også andre lisenser. For eksempel European Union Public Licence for programvare og Open Data Commons Attribution License for databaser. Disse bør være inkludert i en utfyllende undersøkelse og vurdering før det konkluderes med hvilke lisenser som bør anbefales.

I tillegg bør det undersøkes om det arbeides på EU-nivå for å utvikle lisenser for forskningsdata. Hvis et slikt arbeid ikke er igangsatt, er det et arbeid som bør initieres. UiO mener at det viktig at bruken av lisenser standardiseres, slik at det er enkelt for både de som lisensierer og de som bruker dataene å forstå hvilken rett det gis til gjenbruk. Slik standardisering må selvsagt også være internasjonal for forskningsdata.

På bakgrunn av dette oppmunttrer UiO en bredere vurdering før det besluttet hvilke lisenser som skal anbefales som standard. Slike vurderinger kan gjerne kombineres med helt andre løsninger, for eksempel at institusjoner og finansiører «kodes inn» som en del av tallrekken i et datasetts DOI.

UiO merker seg at det er gitt hjemmel for nærmere regulering av lisenser i forskrift, og vi bidrar gjerne inn i et slikt forskriftsarbeid.

Forventet arbeidsmengde og ressursbruk

Viderebruksutvalget har ikke hatt anledning til å beregne de budsjettmessige konsekvensene av lovforslagene. Det er foreslått at tilretteleggingen av data skal finansieres gjennom ordinære bevilgninger. UiO mener det er viktig med økte og øremerkede offentlige bevilgninger for å møte det økte behovet for IT-materiell og personell som lovforslagene vil føre med seg. Sistnevnte vil gjelde innen IT, juss, universitetsbibliotekene og administrativt ansatte på fakulteter, institutter og sentre.

Å publisere en oversikt over institusjonens data med tilhørende metadata i en nasjonal datakatalog vil også kreve betydelige ressurser fra en rekke avdelinger ved UiO, særlig de tekniske avdelingene, da nåværende IT-systemer ikke er rigget for å trekke ut og dele slike oversikter. Dette gjelder også ved anskaffelse og utvikling av nye IT-systemer.

Vi stiller spørsmål ved om data.norge.no er egnet til å huse denne infrastrukturen, da arkivering av faktiske forskningsdata her vil kreve mye plass og vedlikehold. En lenking til datasettene i dataarkivene ved bruk av DOI-er gir bedre oversikt og tilgang til kilden for de arkiverte dataene.

UiO har per nå ikke ressurser til å manuelt innhente og registrere informasjon om hvor forskningsdataene er arkivert. Vi peker derfor igjen på at det er sentralt at systemene må kunne samhandle og overføre informasjon med minimal manuell innsats. Dersom de nye lovene vil innebære at UiO selv må publisere metadata i nasjonal datakatalog, er vi tvilende til om dette er *praktisk*, slik åpne data-direktivet artikkel 9 nr. 1 legger opp til.

For forskningsdata som tilgjengeliggjøres, er UiO også svært opptatt av at det må etableres felles kjøreregler for tjenestebelastning (særlig for "bulk download"). Det er veldig enkelt, bevisst eller ubevisst, å overbelaste eller forstyrre et API. Det må derfor åpnes for at en tilbyder kan innføre begrensninger på belastning, og at denne kan granuleres, eksempelvis slik at autoriserte brukere kan konsumere mye data mens anonyme konsumenter ikke kan belaste endepunktet like mye.

Samlet sett mener UiO at det er urealistisk at pliktene som følger av disse regelverkene skal kunne oppfylles uten tilførsel av ressurser til både personell og infrastruktur. Dette må sikres gjennom øremerkede bevilgninger avsatt til implementeringen.

Avsluttende tekniske merknader til lovutkastene

Dynamiske data (utkastet til datadelingsloven § 7)

Å gjøre dynamiske data tilgjengelige vil påføre offentlige virksomheter store kostnader, utover det som følger av de alminnelige delingsforpliktelsene. Vi mener det er nødvendig å sikre at tilgjengeliggjøringen skjer på hensiktsmessige måter, typisk via API og standardiserte dataformater. UiO etterspør derfor en klargjøring av om dynamiske data kan oppfylle vilkårene for å være datasett av høy verdi, og dermed utløse en ubetinget plikt til tilgjengeliggjøring.

UiO vil også påpeke at det for dynamiske data er særdeles viktig at den som skal tilgjengeliggjøre de aktuelle data får avgjørende innflytelse på hvordan dette i praksis skal gjennomføres. Statens Kartverk og Meteorologisk institutt har gjennom mange år gjort slike datasett tilgjengelig på egne plattformer, med stor grad av selvstendig styring.

UiO har så langt ikke denne type datasett eller tilsvarende løsninger, men dersom UiO utvikler slike datasett er vi opptatt av at også våre fagmiljøer får bestemme hvordan slik tilgjengeliggjøring skal skje.

Definisjonen av offentlig finansierte forskningsdata (utkastet til datadelingsloven § 8)

Det er utfordrende å forstå hva "helt eller delvis finansiert av offentlige midler" betyr, basert på NOU-en og lovutkastet, sammenholdt med åpne data-direktivet artikkel 10 nr. 2. Utvalget påpeker at dette ikke er definert, hverken i direktivet eller nasjonalt.

Formålet med delingsforpliktelsene synes å være å maksimere samfunnsnyttene fra offentlig finansierte forskning. Det kan imidlertid stilles spørsmål ved om dette er konsistent med dagens



forskningsfinansieringsmodeller, og om kravet tar tilstrekkelig hensyn til behovene fra ikke-offentlige finansiører.

Spørsmålet om legaldefinisjon av anonymisering

Viderebruksutvalget har valgt å ikke innta legaldefinisjonen av anonymitet fra åpne data-direktivets artikkel 2 nr. 7. Selv om anonymitet ikke er definert i lovteksten, er dette imidlertid en grunnleggende forutsetning for reglene om datadeling. Anonyme data vil man kunne ha en plikt til å dele, mens personopplysninger kan man ha et forbud mot å dele.

Videre er personopplysningsbegrepet i GDPR dynamisk, ettersom man ved vurderingen skal hensynta både eksisterende og fremtidig teknologisk utvikling. Dette innebærer at anonyme data i dag kan være personopplysninger i framtiden. Dette har man eksempelvis sett på det genetiske området, hvor datasett som tidligere har vært tilgjengeliggjort har måttet tilgangsbegrenses i lys av ny kunnskap og teknologi.

UiO foretrekker en nasjonal implementering som tydelig definerer hva som regnes som «anonymt», samt en tydeliggjøring av at man ved offentliggjøring av data som i henhold til direktivet er anonyme på offentliggjøringstidspunktet ikke har ansvar for fremtidig bruk og utvikling.

Nasjonalt prioriteringsråd (utkastet til datadelingsloven § 14)

Et nasjonalt prioriteringsråd foreslås for å vurdere hva slags data og hvilke datasett det skal oppmuntres til at gjøres aktiv tilgjengelig for viderebruk. Det er viktig at institusjoner som besitter den aktuelle dataen får medbestemmelsesrett i disse vurderingene, og UiO forventer at dette tas i betraktning i et eventuelt forskriftsarbeid.

Tilsynsmyndighet etter dataforvaltningsforordningen (utkastet til dataforvaltningsloven §§ 4 og 5)

Viderebruksutvalget anbefaler en mer helhetlig gjennomgang av den nasjonale organiseringen av uavhengige tilsyn på det digitale området, for å utrede hva som er den mest hensiktsmessige løsningen for organiseringen av tilsyn iht. utkastet til dataforvaltningsloven §§ 4 og 5 ref. punkt 15.4.5 og 15.4.6 i NOU-en.

UiO imøteser en mulighet for å gi innspill også til denne prosessen, når en utredning av alternativ er foretatt.

Med hilsen,

Svein Stølen
rektor

Arne Benjaminsen
universitetsdirektør

Dette dokumentet er godkjent elektronisk ved UiO og er derfor ikke signert.

Saksbehandler:
Stine Carlsen
stine.carlsen@admin.uio.no