

Innspill til videre arbeid med åpen tilgang til vitenskapelige publikasjoner og datasett

Vi viser til invitasjon fra UNIT om å gi innspill til hvordan aktører og myndigheter bør arbeide videre for å nå målet om åpen tilgang innen 2024 for alle vitenskapelige artikler som er resultat av offentlig finansiert forskning. Bakgrunnen for invitasjonen er rapporten «Rettighets- og lisensspørsmål ved åpen publisering» skrevet på oppdrag fra Unit av Torger Kielland, førsteamanuensis ved Det juridiske fakultet, Universitetet i Bergen.

Betydning for Nasjonalbibliotekets oppdrag

Nasjonalbibliotekets kjernevirksomhet er forankret i pliktavleveringsloven. Formålet med loven er å sikre levering av dokumenter med allment tilgjengelig informasjon, slik at vitnesbyrd om norsk kultur og samfunnsliv blir tatt vare på og gjorde tilgjengelig som kildemateriale for forskning og dokumentasjon. Nasjonalbiblioteket er dessuten nasjonal infrastruktur for språkteknologisk utvikling og forskning (Språkbanken) og en vesentlig del av infrastrukturen for norsk forskning. Målet om åpen tilgjengelighet for alle vitenskapelige artikler innen 2024 vil ha betydning både for pliktavleveringsoppdraget, språkbankoppdraget og for forskningsinfrastrukturoppdraget.

Åpen publisering vil ha positiv betydning for alle ledd i pliktavleveringsarbeidet: innsamling, bevaring og formidling. Selv om pliktavleveringsloven bare omfatter dokumenter publisert i Norge og dokumenter utgitt i utlandet som er spesielt beregnet på den norske allmennheten, driver Nasjonalbiblioteket også aktiv innsamling av dokumenter publisert av nordmenn eller om Norge i utlandet (norvegica extranea). Vitenskapelige publikasjoner er en viktig kilde til denne delen av samlingen. Digital publisering bak betalingsmur har de senere årene vanskeliggjort innsamling og bevaring av vitenskapelige arbeider publisert av norske forskere i utlandet. Åpen publisering vil fjerne disse hindringene.

Forskrift til åndsverkloven gir begrenset digital tilgang til pliktavlevert materiale. Tilgangen gis ved strømming, er begrenset til to samtidige brukere for hvert objekt, og det er ikke tillatt å laste ned eller videreformidle materialet. Åpen publisering av vitenskapelige artikler vil fjerne begrensningene for digital formidling av denne delen av samlingen.

Nasjonalbiblioteket er en viktig del av infrastrukturen for norsk forskning. I takt med at stadig større deler av samlingen blir digitalt tilgjengelig, utvikler Nasjonalbiblioteket nye redskaper for å utforske den. Mange av redskapene anvendes til tekst- og datautvinning. Åndsverksloven med forskrift legger strenge føringer for formidling og viderebruk av pliktavlevert materiale. Føringene innebærer sterke begrensninger for mange former av tekst- og datautvinning. Åpen publisering av vitenskapelige artikler vil øke mengden tekst som kan gjøres fritt tilgjengelig for alle former for digital analyse. Den andre viktige kilden er i dag offentlig produsert tekst unntatt opphavsrettslig vern (åvl. §14).

Hvordan bør aktører og myndigheter arbeide videre for å nå målet om åpen tilgang?

Nasjonalbiblioteket anser at det kan kreve åpen tilgang til vitenskapelige arbeider publisert av institusjonens egne ansatte. Nasjonalbiblioteket er godkjent av Forskningsrådet som forskningsinstitusjon og har mange ansatte med forskerkompetanse, men ingen ansatte har en fastsatt del av arbeidstiden reservert til forskning, slike beslutninger tas ut fra vurderinger av strategisk relevans for institusjonens øvrige oppdrag.

I kapittel 2 problematiserer rapporten at «bearbeidelser av forskningsartikler medfører risiko for eksempelvis oppstyking, forkorting eller dårlige oversettelser av artikkelen, som kan komme i konflikt med hensyn til forfatterens akademiske integritet» (s. 38). Disse utfordringene er ikke spesifikke for vitenskapelige publikasjoner, men gjelder alle typer publikasjoner som er utstyrt med CC BY eller lignende lisenser, f. eks. offentlige publikasjoner unntatt opphavsrettslig vern. Norske myndigheter arbeider målrettet for videst mulig gjen- og viderebruk av offentlige produsert tekst og data. Den samfunnsøkonomiske betydningen av økt tilgang til åpne data er fremhevet i en rekke offentlige meldinger de senere årene, bla. *Digital agenda for Norge - IKT for en enklere hverdag og økt produktivitet* (Meld. St. 27 (2015-2016)), *Én digital offentlig sektor. Digitaliseringsstrategi for offentlig sektor 2019-2025* (Kommunal- og moderniseringsdepartementet (2019)), og *Nasjonal strategi for kunstig intelligens* (Kommunal- og moderniseringsdepartementet (2020)). Målet om åpen tilgang innen 2024 for alle vitenskapelige artikler bør vurderes i lys av det overordnede politiske målet om videst mulig gjen- og viderebruk av offentlig produsert tekst og data.

Det er i denne sammenheng viktig å skille mellom to typer bearbeiding. Bearbeidingene som problematiseres i rapporten, retter seg mot publikasjonenes kommuniserte innhold. En annen og stadig viktigere form for bearbeiding retter seg mot publikasjonenes språklige strukturer. Det er vanskelig å se at denne siste typen bearbeidinger på noen måte kan true forfatterens profesjonelle integritet. Valget av lisens må balansere hensynet til disse to formene for bearbeiding. I slike vurderinger er det viktig å huske på at behovet for å regulere innholdsrettede bearbeidinger (utdrag, resymeer og oversettelser), også ivaretas av profesjonsetiske krav for journalister, forfattere og forskere og andre yrkesgrupper som benytter seg av dokumentene som informasjonskilder i tradisjonell forstand. Det kan argumenteres for at slike institusjonaliserte normer utgjør en minst like effektiv beskyttelse som lovhjemlete ideelle rettigheter.

Hva gjelder rapportens betraktninger i kapittel 3 om hvilke konsekvenser kravet om åpen publisering kan få for forskningsansattes valg av publiseringskanal, mener Nasjonalbiblioteket at det ikke tas tilstrekkelig hensyn til at Plan S er et internasjonalt initiativ som har som sentralt siktemål å endre dagens publiseringslandskap. Rapportens betraktninger tar utgangspunkt i dagens situasjon hvor mange av de mest prestisjefylte kanalene ikke legger til rette for åpen publisering. Hvis Plan S gjennomføres, vil denne situasjonen nødvendigvis bli endret.

Til slutt vil Nasjonalbibliotekets minne om at også datasett utgjør et resultat av offentlig finansiert forskning som det er viktig å sikre åpen tilgang til på lik linje med vitenskapelige publikasjoner. Rapporten behandler ikke spørsmålet om datasett og databaser, men nevner behandling av vitenskapelige publikasjoner som datasett i forbindelse med tekst- og datamining (s. 30). Med mindre datasett inneholder opplysninger som berører opphavsrett eller personvern, vil de normalt kun være beskyttet av databaserettigheter. Slike rettigheter er langt svakere enn opphavsrettigheter. Blant annet vil mange av problemstillingene knyttet til ideelle rettigheter som rapportens kapittel 2 drøfter inngående, ikke stille seg.

Åpen lisensiering av datasett er viktig av flere grunner. Først og fremst er enkel tilgang til gjenbruk av forskningsdata avgjørende for å sikre etterprøvbarehet av publiserte forskningsresultater. En annen viktig grunn til størst mulig åpenhet er mulighetene for å gjenbruke forskningsdata - enten alene eller i kombinasjon med nye data - for å utforske andre spørsmål og hypoteser enn de som dataene opprinnelige ble samlet inn for å besvare eller etterprøve. En tredje viktig grunn til åpenhet er at forskningsdata kan ha stor verdi også utenfor academia. Dette er en viktig erfaring som Nasjonalbiblioteket har gjort i forbindelse med opprettelsen av Språkbanken i 2010.

Forskningsdata bør ha enda åpnere lisenser enn vitenskapelige publikasjoner for å sikre en mest mulig utstrakt og effektiv viderebruk. Dette kan sikres i form av CC 0 lisenser. Når mange datasett av ulik opprinnelse kombineres, kan sporing og formidling av attribusjoner som følger av CC BY lisenser, ofte omtalt som «attribution stacking», bli veldig krevende. Resultatet blir ofte at kommersielle aktører avstår fra å bruke datasett med attribusjonskrav, selv om lisensen tillater kommersiell bruk. Nasjonalbiblioteket mener derfor at bruk av CC 0 lisens for forskningsdata bør være utgangspunktet for norske forskningsinstitusjoner lisensieringspraksis.

Med vennlig hilsen



Hege Stensrud Høsøien
Avdelingsdirektør
Fag og forskning