

Innlegg på NFR-innspelskonferanse: Muligheter og utfordringer ved bruk av lisenser på forskningsdata

25. mars 2021

Philipp Conzett
Universitetsbiblioteket
UiT Noregs arktiske universitet



Innleiingsvis...

- ❑ Dette er ikkje noko omforeint/samla innspel frå UiT,
- ❑ men eit innspel basert på erfaringane mine frå utvikling og drift av støttetjenester for forskingsdatahandtering på UiT, særleg
 - ❑ DataverseNO (nasjonalt generisk arkiv for forskingsdata) og
 - ❑ TROLLing (the Tromsø Repository of Language and Linguistics) (fagspesifikt, internasjonalt arkiv for språkdata). Begge arkiva er CoreTrustSeal-sertifisert.
- ❑ Eit samla innspel frå UiT kjem i høyringsrunden.

Generelle innspel

- ❑ Samd i og støttar det meste som er sagt i rapporten.
- ❑ Framheva (endå meir) at normalen bør vera å dela data med ingen eller så få restriksjonar som mogleg. Jamvel WHO, som handterer helsedata, skriv i sin datapolicy: «Whenever possible, WHO wishes to promote the sharing of health data.» (World Health Organization u.å.)
- ❑ Framheva (endå meir) det uproblematiske ved deling av data som ikkje har noko lovpålagt vern og der det berre er forskaren sjølv og/eller institusjonen som har eventuelle rettigheter knytte til dataa.

Meir konkrete innspel

«Utvalget ønsker å høre fra dere som til daglig forholder dere til problemstillinger knyttet til deling av data, både når det gjelder **tolkning av regelverk, vurdering av tilgangsbegrensninger og/eller valg av lisenser.**»

Tolking av regelverk

Ufordringar

Det kan vera vanskeleg for forskarar og støttepersonell å navigera i regelverket. Det fører til at ein ofte er usikker på kva ein har lov til å gjera med data ein har gjenbrukt.

Eksempel 1: Åndsverk som definert i Åndsverklova

§ 2: «Med åndsverk forstås i denne loven litterære eller kunstneriske verk av enhver art, som er uttrykk for **original og individuell skapende åndsinnsett**, slik som [...]»

>> Kva skal til for at forskingsdata oppfyller dette kriteriet (= oppnår verkshøgde)?

Ufordringar

Eksempel 2: Databasevern som definert i Åndsverklova

§ 24: «Den som fremstiller en database, som for eksempel et formular, en katalog, en tabell, et program eller lignende arbeid, hvor innsamling, kontroll eller presentasjon av innholdet innebærer en **vesentlig investering**, har enerett til å råde over hele eller **vesentlige deler** av databasens innhold ved uttrekk fra eller gjenbruk av databasen.»

>> Kva inneber ei «vesentlig investering» og «vesentlige deler»?

Utfordringar

Svaret ein som regel får frå juristar, er at slike ting må avklarast i **kvart enkelt tilfelle**, og at ein bør sjå på **rettspraksis**. Men dette er **krevjande** og utgjør eit **unødvendig hinder** for deling og gjenbruk av data.

Jamfør t.d. Doldirina et al. (2018:18): «In short, whether a specific instance of data or metadata is protected by copyright will often require a case-by-case analysis. This lack of legal clarity places substantial burdens on both suppliers and users of data.»

Utfordringar

Eksempel 3: *Er det (tilsikta?) skilnader mellom norsk lovgjeving og EU-lovgjeving?*

Til dømes: Kva slags arbeidsinnsats skal reknast med når ein vurderer «vesentlig investering»?

Jamfør § 24 i Åndsverklova: «Den som fremstiller en database [...], hvor **innsamling**, kontroll eller presentasjon av innholdet innebærer en **vesentlig investering**, [...].»

Ifølgje Doldirina et al. (2018:20f.) seier EU-lovgjeving at **innsamlinga** av sjølve dataa **ikkje** skal inkluderast i vurderinga av om det er gjort ei vesentleg investering:

«In a number of its decisions, the [European] Court [of Justice] has clearly stated that the resources spent on the generation of the contents of a database may not be counted as resources spent on the creation and maintenance of the database itself. Substantial expenditure on the latter is the only basis to acquire the sui generis database protection.»

Moglege løysingar

- ❑ Lovtekstane bør supplerast med **tydelege retningsliner** eller utdjupingar som inneheld **spesifikke kriterium**, slik at det blir lettare å bruka lova på konkrete tilfelle. Retningslinene bør vera utforma på eit **språk** som er **lett tilgjengeleg** også for ikkje-juristar.
- ❑ Eventuelle hòl i lovverket bør fyllast, og lovverket bør i størst mogleg grad vera på line med europeisk/internasjonale lovgjeving.

Vurdering av begrensninger på tilgang og gjenbuk

Utfordringar

- ❑ Sitering og kreditering er eit viktig insentiv for forskarar. Stundom opplever vi at forskarar er redde for manglande sitering/kreditering. Kan bruka lisensar med siteringskrav, men det kan vera til hinder for gjenbruk av data i tilfelle der ein kombinerer data frå mange kjelder.
- ❑ Korleis gje open tilgang til data som i utgangspunktet ikkje kan delast?
- ❑ Forholdet mellom lisensiering og tilgangskontroll gjennom andre verktøy som t.d. avtalar, Data Use Agreement (DUA) og embargo kan ofte vera uklart.

Mogleg løysing

Stol heller på forskingsetiske normer som garanti for sitering/kreditering når ein bruker andre sine data.

Jamfør Doldirina (2018:33): «If attribution is desired by providers for research data, as is usually the case, the citation to parent datasets as **a normative practice in the academic and scientific communities** can fulfill an equivalent role. Data citation may be promoted through development of efficient technological means for facilitating the practice, inclusion of the practice in codes of conduct, and incorporation into the requirements of research publication outlets.»

Men: Kva med gjenbruk av data til andre formål enn forskning, t.d. kommersialisering, journalistikk, osv.? Er det då nok med forskingsetiske normer, eller treng vi meir generelle etiske retningslinjer?

Val av licensar

Utfordringar

- ❑ Restriksjonar på gjenbruk kan gjera det vanskeleg å kombinera datasett frå fleire kjelder til eit nytt datasett (jf. legal interoperability challenges, attribution stacking).
- ❑ Velja rett nivå for lisensiering: fil, datasett, samling?
- ❑ Kva skjer når ein må/ønskjer å endra lisensen på eit publisert datasett?
- ❑ Metadata (for gjenfinning) har ofte ingen lisens.

Moglege løysingar

- ❑ Dersom mogleg, bruk deklarasjonar eller lisensar utan restriksjonar:
 - ❑ Public Domain
 - ❑ Creative Commons Zero (CC0)
 - ❑ Open Data Commons Public Domain Dedication and License (PDDL)
- ❑ Tiltråding om endring av lisensar: Enkelt å gå frå ein restriktiv lisens til ein mindre restriktiv lisens. Andre vegen er verre i og med at dataa då allereie kan ha blitt gjenbrukt med ein mindre restriktiv lisens. Kan vera nyttig å vita når data ikkje kan delast heilt ope frå dag 1.

Referansar

Doldirina, Catherine, Anita Eisenstadt, Harlan Onsrud, og Paul Uhlir. 2018. «Legal Approaches for Open Access to Research Data». LawArXiv.

<https://doi.org/10.31228/osf.io/n7gfa>.

World Health Organization. u.å. «WHO Data Policy». Vitja 24. mars 2021.

<https://www.who.int/about/who-we-are/publishing-policies/data-policy>.