



ARKIVVERKET

KOMMUNAL- OG
DISTRIKTSDEPARTEMENTET

Postboks 8112 DEP
0032 OSLO

Dato	01.07.2022
Din ref.	
Vår ref.	2022/10847
Saksbehandler	HARSOM/STRAKOM

Høringssvar - Gjennomføring av direktiv (EU) 2019/1024 om åpne data og viderebruk av informasjon i offentlig sektor i norsk rett

Arkivverkets samfunnsoppdrag er å bidra til effektiv dokumentasjonsforvaltning og å sikre, bevare og tilgjengeliggjøre et bredt og allsidig utvalg av samfunnets arkiver¹.

Direktiv (EU) 2019/1024 skal sikre åpne data fra offentlig sektor hvor et hovedfokus er «[...] *økning i utleveringen av verdifulle offentlige data til viderebruk, herunder fra offentlige foretak, forskningsutøvende organisasjoner og forskningsfinansierende organisasjoner.*» (fortalepunkt 4).

Arkivverket forvalter historisk arkiv og data fra statlig, kommunal og privat sektor. Den digitale transformasjonen av offentlig sektor krever nye løsninger for å ivareta digitalt skapt dokumentasjon og Arkivverket har derfor fått i oppdrag å utvikle en fellesløsning for mottak, bevaring og tilgjengeliggjøring i Digitalarkivet².

Arkivverkets samfunnsoppdrag innbefatter tilgjengeliggjøring og viktige samfunnsarkiver skal være en datakilde for alle innenfor rammene av lovbestemt taushetsplikt. Arkivverket satser derfor på at nye løsninger skal kunne gjøre data mer tilgjengelig. I Digitalarkivet utforsker vi f.eks. bruk av kunstig intelligens som verktøy for anonymisering og pseudonymisering av konfidensielle data for å kunne legge til rette for viderebruk av offentlig informasjon/data og knytte slike kilder sammen med andre data, f.eks. ved hjelp av APIer.

Arkivverket ser et stort potensial i dette arbeidet for å fremme offentlige og private arkiver som viktige datakilder under prinsippet «så åpne som mulig, så lukkede som nødvendig». Det vises for øvrig til Kulturdepartementets strategi for åpne data 2017-2022 - vedtatt av Kulturdepartementet 03.10.2017.

I arkivloven defineres et arkiv som "dokument som vert til som lekk i ei verksemd." jf. Arkivloven § 2 b. Arkivverket forvalter avleverte arkiv fra offentlig sektor og bevaringsverdig privatarkiv. Offentlig sektors plikt til å holde arkiv fremgår av arkivloven § 6 og statlige arkiv skal leveres Arkivverket jf. Arkivloven § 10. For privatarkiv er det laget en bevaringsplan der bevaringspolitikk, kriterier og samfunnsområder for bevaring er beskrevet³.

Departementet viser under **punkt 2.2 Implementering i norsk rett** til at nødvendige endringer skal gjennomføres gjennom offentliglova og offentligforskrifta og at det skal nedsettes et offentlig utvalg som skal se mer på en mer helhetlig regulering av viderebruk av offentlig informasjon. Arkivverket er av den oppfatning at sekundærbruk av offentlige arkiver og privatarkiv tilsier at arkivloven vil være relevant i den sammenheng.

¹ <https://www.arkivverket.no/om-oss/hva-er-arkivverket?q=samfunnsoppdrag>

² <https://www.digitalarkivet.no/content/733/ofte-stilte-sp%C3%B8rsm%C3%A5l-om-satsingen-p%C3%A5-digitalarkivet>

³ Arkivverkets bevaringsplan for privatarkiv 2020.

Forskningsdata

Arkivverket ser det som positivt at direktivet utvider virkeområdet hva angår forskningsdata. Avlevert materiale til Arkivverket behandles etter offentlighetslova. Arkivverket anser offentlig finansierte forskningsdata som arkiv slik de er definert i arkivloven, som dokument som er blitt til som ledd i offentlig virksomhet.

Direktivets artikkel 2 nr. 9 definerer forskningsdata som *"dokumenter i digital form, unntatt vitenskapelige publikasjoner, som er innsamlet eller produsert ved vitenskapelig forskningsvirksomhet, og som brukes som dokumentasjon i forskningsprosessen, eller som er allment akseptert i forskningsmiljøet som nødvendig for å validere forskningsfunn og -resultater."*

I høringsnotatet slås det fast at forskningsdata faller inn under dokumentbegrepet i offentlighetsloven⁴. Dokumentbegrepet i offentlighetsloven, forvaltningsloven og arkivloven er harmonisert⁵, og er definert som *«ei logisk avgrensa informasjonsmengd som er lagra på eit medium for seinare lesing, lytting, framsyning, overføring eller liknande.»*

I arkivloven § 2 er arkiv definert som *«dokument som vert til som lekk i ei verksemd»*⁶ altså dokumenter som blir til som en del av organets normale virksomhet. Det omfatter alle aktiviteter, oppgaveløsning og saksbehandling. Forskning omfattes av denne definisjonen og forskningsdata er derfor omfattet av arkivloven.

Arkivverket har siden 2014 hatt en avtale med NSD, nå SIKT (Kunnskapssektorens tjenesteleverandør), om at SIKT skal ta imot forskningsdata fra forskere og forskningsmiljøer på vegne av Arkivverket. Avtalen med SIKT legger til rette for at forskningsdata skal lagres og tilgjengeliggjøres for tilpasset analytisk bruk basert på arkivloven og annet relevant regelverk.

Forskningsdata bør i mange tilfeller bevares i et langtidsperspektiv. Forskningsdataarkivenes bevaringsperspektiv strekker seg ofte ikke lenger enn 10 år, mens Arkivverkets bevaringsperspektiv ikke har en tidsbegrensning. Avtalen med SIKT innebærer at Arkivverket skal motta dataarkivene som skal bevares i et langtidsperspektiv. Arkivverket arbeider derfor med å etablere et samarbeid med SIKT om overføring og langtidslagring av bevaringsverdige datasett.

Avtalen med SIKT legger til rette for at forskningsdata skal lagres og tilgjengeliggjøres for tilpasset analytisk bruk basert på arkivloven og annet relevant regelverk.

Dynamiske data, datalager og dataarkiv

Det er positivt at begrepet *dynamiske data* klargjøres i høringsnotatet punkt 6 om Dynamiske data og at det slås fast at også dynamiske data er dokumenter etter offentlighetsloven jf. høringen s. 25. *Dynamiske data* defineres i høringsutkastet som *«dokumenter i digital form som oppdateres hyppig eller i sanntid, særlig på grunn av deres volatilitet, eller at de raskt foreldes; data som genereres av sensorer anses vanligvis å være dynamiske data,»* (Direktivet art. 2 nr. 8 og høringspunkt 6.3.1).

Arkivverket vil bemerke at det innebærer at *dynamiske data*, inkludert sensordata, følgelig også må betraktes som arkiv etter arkivloven. Dette innebærer midlertid ikke at alle dynamiske data skal forvaltes som arkiv (se om dette nedenfor). Dette er i tråd med praksisen i forskningsmiljøene, der det skilles mellom datalager og dataarkiv. Datalager og dataarkiv er betegnelser som er innarbeidet og har vært i bruk lenge når det gjelder lagring av forskningsdata - både nasjonalt og internasjonalt.

⁴ Punkt 5 Begrepsbruk og definisjoner, punkt 5.3 nest siste avsnitt.

⁵ Arkivrett 2020, Herbjørn Andresen, s. 63 «Arkivlovens definisjon av dokument er praktisk talt identisk med definisjonen i forvaltningsloven og offentlighetsloven. [...] De likelydende definisjonene ble introdusert med en endringslov (lov nr. 98/2000), som trådte i kraft 1. april 2001. [...] Formålet med endringen var å oppnå en harmonisert og teknologiavhengig begrepsbruk. Ved en lovendring i 2015 (lov nr. 72/2015) ble samme definisjon av «dokument» også tatt inn i pliktavleveringsloven for avlevering av trykksaker til Nasjonalbiblioteket, i § 3. Formålet med endringen i pliktavleveringsloven var, som omtalt i endringslovens proposisjon Prop. 106 L (2014–2015) s. 47, at definisjonen skulle bli tilsvarende den i arkivloven, forvaltningsloven og offentlighetsloven.»

⁶ Offentleglova § 4 første ledd, forvaltningsloven § 2 litra f, arkivloven § 2 litra a.

Arkivverket støtter departementets forslag om en klargjørende definisjon i offentleglova.

Departementet ber om innspill på bruken av begrepet datalager. Samtidig presiseres det at å bruke et arkivbegrep som skiller seg fra begrepets innhold i arkivlova ikke vil være formålstjenlig (høringsutkastet punkt 4.3 s. 17 siste avsnitt). Ifølge departementet (s. 17 fjerde avsnitt) inneholder et datalager aktive («varme») data som oppstår, innhentes eller brukes underveis i forskningsprosessen.

«Aktive data kan prosesseres, analyseres, endres, deles eller på andre måter brukes operativt underveis i prosjektet. Aktive data kan i så måte bestå av både rådata og prosesserte data. En slik form for aktiv lagring handler også om å oppbevare sikkerhetskopier som kan brukes til gjenoppretting av tidligere versjoner dersom data går tapt eller på andre måter blir ødelagt underveis i prosessen».

Arkivverket vil presisere at bruken av begrepene datalager og dataarkiv ikke skiller seg fra begrepets innhold i arkivloven. Forskningsdata i datalager er arkiv etter arkivloven, men alle dataene skal ikke nødvendigvis forvaltes som arkiv - bare data som har verdi som dokumentasjon eller er gjenstand for saksbehandling skal forvaltes som arkiv. Dette følger motsetningsvis av arkivforskriften § 14. Arkivforskriften § 14 sier at *«Dokument som verken har inngått i saksbehandlingen til organet eller har verdi som dokumentasjon, skal holdes utenfor eller fjernes fra arkivet ved arkivavgrensning»*.

I praksis innebærer det at disse dataene kan slettes når organet ikke lenger har behov for dem. Data som er nødvendige for å dokumentere forskningsprosessen og kunne etterprøve forskningsresultatene skal forvaltes som arkiv. Dette samsvarer som tidligere nevnt godt med forståelsen av *dataarkiv*, slik den er gjengitt i høringsnotatet s. 17:

«Arkiv brukes derimot om langtidsbevaring av data etter at prosjektet er avsluttet eller ved publisering av artikkel. Når data arkiveres så lages det statiske ("kalde") versjoner av data, gjerne basert på et utvalg - for eksempel de dataene som danner grunnlaget for en forskningspublikasjon eller lignende. Sagt på en annen måte: de dataene som velges ut for arkivering anses å ha en viss verdi for (videre) forskning og/eller som grunnlag for validering av forskningsresultater. Data som arkiveres berikes med metadata og annen dokumentasjon, slik at de kan forstås, langtidsbevares og (eventuelt) gjøres tilgjengelige også for andre enn de som deltok i innsamlingen/genereringen av dataene.»

Vi foreslår derfor at det fortsatt differensieres mellom *datalager* og *dataarkiv*.

Arkivverket finner det viktig å poengtere at det i arkivregelverket skilles mellom arkivplikt (arkivloven § 6), journalføringsplikt (arkivforskriften § 9 første ledd, andre punktum) og bevaringsplikt (arkivloven § 9 og 10).

Selv om forskningsdata faller inn under arkivlovens definisjon av arkiv og forskningsdata som har verdi som dokumentasjon (*dataarkiv*) skal forvaltes som arkiv etter arkivloven, innebærer ikke det at alle pliktene i arkivloven gjelder.

Journalføringsplikten og kravet om å følge Noark-standardene gjelder bare for dokumenter som er gjenstand for saksbehandling og har verdi som dokumentasjon (om Noark- se arkivforskriften § 11). Arkivplikten favner videre enn journalføringsplikten, men er mer generell. Arkivloven § 6 sier at «offentlige organ pliktar å ha arkiv» og at disse «skal vera ordna og innretta slik at dokumenta er tryggja som informasjonskjelder for samtid og ettertid». I forslag til ny arkivlov § 4 første ledd utdypes arkivansvaret⁷:

§ 4 Arkivplikt for offentlige organ

Offentlige organ skal arkivere dokument som er skapte eller mottekne som ledd i verksemda deira, og som har inngått i saksbehandlingen, eller som dokumenterer saksbehandling eller oppgaveløysing. Organa skal arkivere strukturert informasjon i databasar og informasjonssystem når den har verdi som dokumentasjon.

⁷ HØYRINGSNOTAT – FORSLAG TIL NY ARKIVLOV av 5.10.2021

Arkiva skal sikrast, ordnast og forvaltast slik at

- a. informasjonsinnhaldet ikkje blir endra*
- b. informasjonsinnhaldet kan nyttast*
- c. opphavet til informasjonen alltid er kjent*
- d. informasjonen blir teken vare på i den samanhengen han inngår i*

De arkivfaglige hensynene handler om at dokumenter (data) som er omfattet av arkivloven skal forvaltes på en måte som gjør at informasjon ikke går tapt, at de forblir anvendbare og at vi kan ha tillit til dataene over tid. Dette samfaller langt på vei med intensjonen og formuleringene i FAIR-prinsippene, slik at ivaretagelse av arkivloven ikke innebærer en tilleggsbelastning for forskningsmiljøene.

Det bør også presiseres at "langtidsbevaring", slik det er brukt i høringsnotatet (side 17 fjerde avsnitt) og bevaringsplikt etter arkivloven ikke er det samme. Bevaringsplikten etter arkivloven har et mye lengre tidsperspektiv (100 år eller mer) enn langtidsbevaring i dataarkiv (som ofte er 10 år). Det er innenfor Arkivverkets arkivmyndighet å fastsette om forskningsarkiv skal langtidsbevares etter arkivloven og leveres til Arkivverket. Dette legges også til grunn i avtalen mellom SIKT og Arkivverket som er omtalt ovenfor.

Lagring av forskningsdata har vært et arbeidsfelt som i stor grad har utviklet seg innenfor universitetene og særlig innenfor universitetsbibliotekene. Etter hvert har også flere direktorater med digitaliseringsoppdrag kommet på banen. Nasjonalarkivene har i liten grad vært aktivt involvert, i motsetning til i Danmark, der Rigsarkivet utfører den samme rollen som NSD/SIKT gjør i Norge. Dette er litt av bakgrunnen for at lagring av forskningsdata ikke er så knyttet til resten av arkivfeltet og har sikkert også betydning for at Arkivverket ofte ikke har blitt tatt med på råd og ikke blir nevnt i offentlige dokumenter som handler om lagring av forskningsdata, selv om dette dreier seg om bevaring av store mengder offentlig skapte data, som er arkiv etter arkivloven. Arkivverket vil fremover ta en mer aktiv rolle på dette området i samarbeid med forskningsmiljøene og gjennom samarbeidet med SIKT.

Datasett med høy verdi (kap. 7)

Departementet skriver under punkt 7.3.2, s. 33 at «*Direktivet er et minimumsdirektiv og innebærer at Norge vil kunne gå lenger enn direktivet mht. å inkludere datasett med høy verdi*» og ber samtidig om hva høringsinstansene mener bør kunne vurderes for en slik liste, «*da dette vil kunne være nyttig informasjon for det forestående utvalget som skal se på en mer helhetlig regulering av viderebruk av offentlige data.*»

Folkeregisterinformasjon og helseregistre er eksempler på «*dokumenter der viderebruk vil være forbundet med viktige fordeler for samfunnet, miljøet og økonomien*» og således er «*datasett med høy verdi*». Arkivverket bevarer historiske helseregisterdata fra spesialisthelsetjenesten (som leveres

Norsk helsearkiv) og historiske folkeregisterdata. Meteorologiske og andre data som viser klimaendringer kan være et annet aktuelt område.

I direktivets foralepunkt 65) fremgår det at «*[...]Biblioteker, herunder universitetsbiblioteker, museer og arkiver har en betydelig mengde verdifulle informasjonsressurser fra offentlig sektor, særlig fordi digitaliseringsprosjekter har mangedoblet mengden digitalt offentlig materiale. Disse kulturarvsamlingene og tilhørende metadata er en potensiell kilde for produkter og tjenester med digitalt innhold, og de har et enormt potensial med hensyn til innovativ viderebruk i sektorer som utdanning og turisme.[...]*».

Arkivverket oppbevarer offentlige data fra mange ulike samfunnsområder som med tilrettelegging kan være egnet for viderebruk. Det er ressurskrevende å utvikle løsninger for viderebruk av digitalt skapte datasett fordi Arkivverket forvalter datasett fra alle samfunnsområder og på svært ulike formater, men også fordi Arkivverket forvalter data i et svært langt tidsperspektiv og bruksformål og brukerbehov endrer seg over tid. Å produsere datasett av data i papirbaserte

arkiver er svært ressurskrevende, men gjør at dataene får mange flere bruksformål, slik at samfunnsnyttene fremover vil være betydelig. Dette er særlig tilfelle når historiske data kan knyttes sammen med samtidsdata.

Et eksempel på bruk av analoge data som siden er skannet og transkribert er Historisk befolkningsregister (HBR-prosjektet), som startet opp i 2014, og nå er forlenget til 2026. I dette prosjektet blir personopplysninger som er tilgjengelige i Digitalarkivet, fra folketellinger, kirkebøker, emigrantprotokoller og andre personkilder, lenket sammen til livsløp for samme individ, med mål å etablere et register over Norges befolkning fra 1801 til 1964. Registret har en «åpen del» fra 1801 til ca. 1920, og en «lukket del» fra ca. 1920 til opprettelsen av sentralt, digitalt folkeregister i 1964.

Departementet skriver i punkt 7.3.4 at oversikten over datasett med høy verdi ikke vil være statisk og at utpeking av datasett med høy verdi vil kunne omfatte datasett som ikke er omfattet av Kommisjonens gjennomføringsrettsakter, og foreslår en forskriftshjemmel om hvilke datasett som har høy verdi.

Arkivverket støtter dette.

Særlig om betaling for innsyn (kap. 10)

Som nevnt over ser Arkivverket et stort potensial i anonymisering og pseudonymisering av konfidensielle offentlige data for viderebruk. Arkiv som langtidslagres etter en bevaringsvurdering, beholdes med alle data som er vurdert å ha en verdi for datasettet. Det er grunn til å bemerke at arbeidet med slik anonymisering for distribusjon/viderebruk og kostnadene ved dette er vanskelig å forutsi per i dag. Arkivverket støtter derfor departementets forslag i punkt 10.3.5 om at «*Også arkiv, bibliotek og museer vil ha adgang til å ta betalt for kostnader til anonymisering, der det er relevant*» og at dette inntas i offentligforskrifta.

Med hilsen

Guri Kaspara Lande
direktør

Harald Sommerstad
seniorrådgiver

Dokumentet er elektronisk godkjent og har ingen underskrifter