



Regulatorisk sandkasse for arkiv, data og offentlighet

Erfaringsrapport #4 med Tolletaten

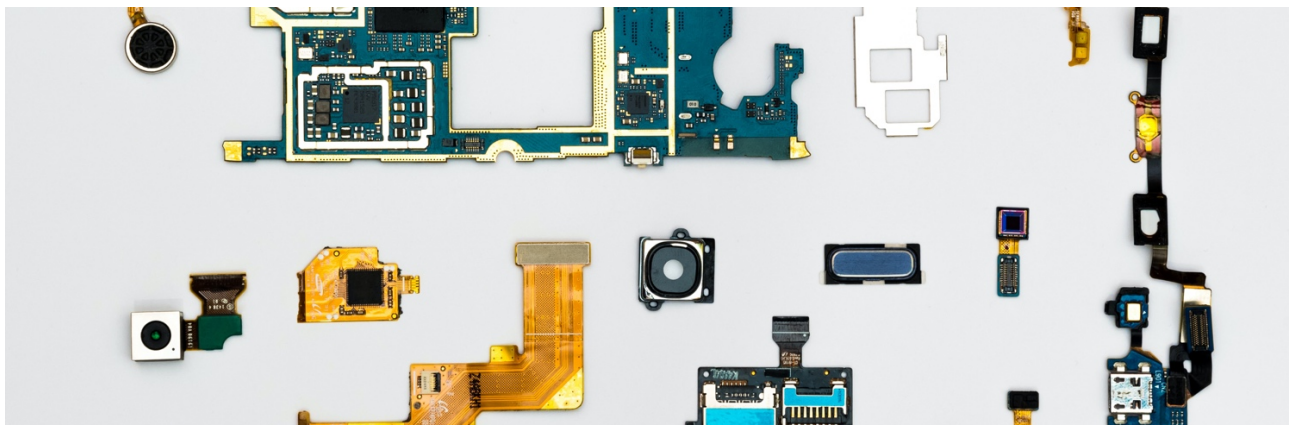
1.12.2022

I rapporten finner du:

- Et lettfattelig sammendrag
- Bakgrunn for prosjektet
- Resultater fra de ulike problemstillingene
- Veien videre og avsluttende refleksjoner

INNHOLDSFORTEGNELSE

<i>Sammendrag</i>	3
1. Om prosjektet	4
1.1. Bakgrunn	4
1.2. Fremtidens arkivregelverk	4
2. Resultater - arkivfaglige hensyn	6
2.1. Utforskning knyttet til dagens regelverk	6
2.2. Utforskning av om standarder til Felles datakatalog kan ivareta arkivfaglige hensyn	9
2.3. Hvilke deler av Noark er Felles datakatalog egnet for å erstatte?	11
3. Resultater - elnnsyn og Felles datakatalog	12
3.1. Kontekst	13
3.2. Resultater fra prosjektet	13
3.3. Hva bør gjøres fremover?	15
4. Veien videre	16
4.1. Gevinster	16
4.2. Læring og regelverksutvikling	17
4.3. Deltagere i prosjektet	18
Vedlegg: Kartlegging av metadata	19



SAMMENDRAG

I perioden februar til juni 2022 har Tolletaten deltatt i regulatorisk sandkasse for arkiv, data og offentlighet. Fra Digitaliseringsdirektoratet har deltagere fra fellestjenestene «elnnsyn.no» og «Felles datakatalog» vært involvert.

«Caset» har vært Tolletatens løsning for grensekontroll, kalt «kontrollstøtteverktøyet».

Mål med sandkasseprosjektet

- Kan arkivfaglige hensyn bli ivaretatt ved bruk av standardene i Felles datakatalog?
- Kan offentlighet bli ivaretatt og informasjon bli publisert på elnnsyn om den beskrives og tilgjengeliggjøres ved bruk av disse standardene?

Konklusjoner

Standardene kan delvis brukes til å ivareta arkiv og offentlighet. Et eksempel er å gi gode beskrivelser av datasettene, noe som kan gjenbrukes ved overføring til Arkivverket og publisering i fellesløsningen Digitalarkivet.

Prosjektet har funnet en tilnærming hvor Tolletaten kan gjøre opplysninger om sine datasett tilgjengelig i Felles datakatalog, og gjøre de samme datasettene søkbare for offentlighet gjennom elnnsyn. Arbeidet har også gått opp grenseflater mellom fellestjenestene. Dette tar oss et steg nærmere en mer helhetlig informasjonsforvaltning.

Veien videre

- ✓ Å lykkes med offentlighet, informasjonsforvaltning og dokumentasjonsforvaltning er overlappende aktiviteter. Det er lurt å tenke på disse tidlig i løpet.
- ✓ Å anvende standarder i Felles datakatalog til arkivformål, vil innebære at den må kombineres med andre standarder som beskriver krav og rutiner for prosesser for håndtering av arkiv.

Dokumentbegrepet i dagens arkivregelverk blir utfordret. Erfaringene gjort i sandkassen medfører en anbefaling om at en samling av data bør falle inn under definisjonen av et dokument. Dette er viktige innspill å ta med i arbeid med nytt arkivregelverk. Forståelsen gjør det mulig for Tolletaten å velge mellom en rekke ulike måter å inndele i dokumenter i kontrollstøtteverktøyet. Erfaringen fra prosjektet er at virksomheter bør stå fritt til å velge en inndeling som er hensiktsmessig for dem.

1. OM PROSJEKTET

1.1. Bakgrunn

Tolletatens samfunnsoppdrag er å sikre at lover og regler for grensekryssende vareførsel etterleves. Etaten skal forvalte og medvirke til at regelverket for dette problemkomplekset etterleves. Tolletaten skal samtidig legge til rette for effektiv handel og vareførsel.

Tolletaten har vært igjennom omfattende prosesser knyttet til digitalisering. De ønsker å se dokumentasjons- og dataforvaltning mer i sammenheng. I søknaden til sandkassen påpekte etaten at: datasett ofte lar seg beskrive og publisere gjennom Felles datakatalog. Tolletaten ønsket å bygge ned det mange har oppfattet å være ett skille mellom dokument og datasett.

Arkivverket ønsket læring om hvordan arkivfaglige hensyn kan ivaretas. Digitaliseringsdirektoratet ønsket å tilgjengeliggjøre og gi innsyn i strukturerte data, og å utforske sammenhengen mellom «elnnsyn.no» og Felles datakatalog.

Prosjektet har undersøkt følgende standarder i Felles datakatalog:

- [Standard for beskrivelse av datasett, datatjenester og datakataloger \(DCAT-AP-NO\)](#)
- [Spesifikasjon for tjeneste- og hendelsesbeskrivelser \(CPSV-AP-NO\)](#)
- [Spesifikasjon for beskrivelse av informasjonsmodeller \(ModellDCAT-AP-NO\)](#)
- [Forvaltningsstandard for begrepsbeskrivelser.](#)

eINNSYN.NO

I denne rapporten omtalt som elnnsyn. Skal gjøre offentlig forvaltning mer åpen, og gjøre forvaltningens dokumentasjon tilgjengelig for innbyggerne. elnnsyn er en søkeportal som publiserer postjournaler og møtedokumentasjon fra forvaltningen.

Brukerne kan søke i metadata og be om å få tilsendt dokumentasjonen, eller i noen tilfeller åpne den ved klikkbar lenke. Dokumentasjonen har typisk form som saksbehandlingsdokumenter og møtedokumentasjon. Tjenesten er gratis og anonym.

FELLES DATAKATALOG

Består av flere katalogtjenester som tilbyr standardiserte beskriver av data slik at andre lettere kan oppdage maskinlesbare data på tvers av virksomheter. I denne rapporten bruker vi Felles datakatalog ikke om løsningen som er tilgjengelig på data.norge.no, men om det settet med standarder som brukes i felleløsningen.

Man henter standardiserte metadata fra virksomheter og sammenstiller disse i en nasjonal tjeneste. Formålet er å legge til rette for gjenbruk av data. I tillegg skal fellestjenesten legge til rette for økt offentlighet.

1.2. Fremtidens arkivregelverk

Definisjon av arkivfaglige hensyn

Strukturert informasjon i løsninger er omfattet av arkiveringsplikten, dersom den har verdi som dokumentasjon eller er gjenstand for saksbehandling. Enkelt forklart innebærer dette at virksomheten, i dette tilfellet Tolletaten, har en plikt til å legge til rette for at vi kan ha tillit til dataene og at de er anvendbare både i dag og så lenge de skal bevares. I forslag til [ny](#)

[arkivlov](#) er dette foreslått regulert i § 4 Arkivplikt for offentlige organ. Sandkassen har sett til disse arkivfaglige kvalitetene, slik de er formulert i § 4, når den har vurdert hvilke beskrivelser av dokumentasjonen som oppstår i kontrollstøtteverktøyet som er nødvendig for å ivareta pliktene etter arkivloven i dokumentasjonsforvaltningsfasen. For enkelhets skyld omtales de i rapporten som arkivfaglige hensyn.

Slik det er formulert i § 4 innebærer de arkivfaglige hensynene at det skal legges til rette for at:

- a) Informasjonsinnholdet ikke skal endres
- b) Informasjonsinnholdet skal kunne brukes
- c) Opphavet til informasjonen alltid er kjent
- d) Informasjonen blir tatt vare på i den sammenhengen den inngår i.

Begrepene i forslag til § 4 og kan minne noe om de internasjonale FAIR-prinsippene, et akronym basert på ordene «findable», «accessible», «interoperable», og «reusable».

Krav i forslag til nytt regelverk som kommer til anvendelse i sandkassen

I dette avsnittet gjennomgår vi de arkivfaglige hensynene som definert ovenfor, og ser på hvilke krav det vil sette til en elektronisk løsning. Det er denne forståelsen som er lagt til grunn i utforskningen. Det er tilnærmingen herfra som er anvendt, mens kontrollstøtteverktøyet er et eksempel for å gjøre diskusjonene mer håndfaste.

Løsningen skal være i stand til å fange dokumentasjon og metadata knyttet til dokumentasjonsforvaltning, og sørge for at brukerne kan søke og hente frem dokumentasjon og metadata. Løsningen skal hente og tilgjengeliggjøre dokumentasjon i et anvendelig format og form, importere dokumentasjon og tilhørende metadata fra eksterne fagapplikasjoner, og integriteten til innholdet og strukturen til dokumentasjonen skal sikres.

Den aktuelle løsningen skal være i stand til å fange og vedlikeholde metadata for dokumentasjon så lenge dokumentasjonen finnes, i tråd med metadatakrav. Løsningen skal ha funksjonalitet for at dokumentasjon kan trekkes ut av løsningen, og på en slik måte at den kan overføres til andre. Dokumentasjon som skal overføres til andre virksomheter eller til depot, skal kunne trekkes ut av løsningen på en slik måte at den fortsatt er satt i sammenheng, er uendret, har kjent opphav og kan brukes.

Løsningen skal opprette og ta vare på metadata om tilgang, bruk og sikkerhet, slik at hvert informasjonsobjekt eller sammenstillinger av informasjonsobjekt har en hendelseslogg som dokumenterer tilgang til og bruk av dokumentasjon.

Løsningen skal være i stand til å konvertere, migrere eller eksportere dokumentasjon med tilhørende metadata og der hvor det er nødvendig, sammenstillinger eller aggregeringer av dokumentasjon til systemuavhengige format for å støtte bevaring utover løsningens levetid.

Dokumentasjon som skal kasseres skal bli slettet i løsningen. Hjemmelsgrunnlag for for eksempel skjerming skal dokumenteres ved overføring. Med overføring mener vi både uttrekk til langtidsbevaring og overføring av data til andre løsninger. Dokumentasjonen skal være tilgjengelig for all berettiget bruk og innsyn. Informasjon med dokumentasjonsverdi

må kunne gjenfinnes, fremhentes, presenteres og forstås innen rimelig tid, slik at alle interessenter har muligheter for etterprøving av ekthet og at informasjonen kan gjenbrukes.

Løsningen skal tilføre identifikatorer til dokumentasjonen og sammenstillinger av dokumentasjon. Løsningen skal være i stand til å knytte dokumentasjon på enkeltdokument eller aggregert nivå til forretningskonteksten over tid.

Løsningen skal sikre at et uttrekk for migrering/eksport er komplett, og inneholde all dokumentasjon og sammenstillinger som skal være med. Den inneholder alle metadata for dokumentasjon og sammenstillinger av dokumentasjon, samt dokumentasjonens hendelseslogg.

2. RESULTATER - ARKIVFAGLIGE HENSYN

Her undersøker vi om arkivfaglige hensyn blir ivaretatt i grensekontrollen ved bruk av standardene som brukes i Felles datakatalog.

2.1. Utforskning knyttet til dagens regelverk

Arkiv- og journalføringsplikt (Arkivforskriften § 9, Offentleglova § 4)

Et aspekt sandkasseprosjektet først måtte avklare, var om løsningen for kontrollstøtte inneholder dokumenter som medfører arkiv- og journalføringsplikt. Det følger av arkivforskriften § 9 at et offentlig organ skal ha en eller flere journaler for registrering av dokument i de sakene organet oppretter. I journalen skal en registrere alle inngående og utgående dokumenter som etter Offentleglova § 4 må regnes som saksdokument for organet, dersom de er eller blir saksbehandlet og har verdi som dokumentasjon.

I den aktuelle løsningen i Tolletaten blir ikke dokumenter sendt inn eller ut av organet. Dokumentene som oppstår, er derfor å regne som organinterne dokumenter. Ifølge arkivforskriften § 9 skal disse registreres i journalen så langt organet finner det hensiktsmessig. Unntaket er organinterne dokumenter som inneholder den endelige avgjørelsen i en sak.

Det er videre slik at dokumentene i Tolletaten sitt kontrollstøtteverktøy ikke inneholder det endelige vedtaket i en sak, med deler av selve saksunderlaget. Det kan altså være arkivplikt, men ikke journalføringsplikt for de strukturerte dataene som har verdi som dokumentasjon. For å øke dokumentasjonsfangsten og mulighet for merinnsyn, har Tolletaten i sandkassen ønsket å teste ut å journalføre og offentliggjøre så mye som mulig av denne dokumentasjonen.

Metadata (Arkivforskriften § 10)

Hvis man ser på den enkelte kontrollrapport som et dokument, kan et datasett som helhet fungere som en journal for kontrollrapportene. Arkivforskriften § 10 definerer hvilke metadata som skal være med ved registrering av dokument i en journal. Samtlige metadatakrav finnes igjen i Tolletatens løsning som del av beskrivelsen av kontrolloppdraget, bortsett fra metadatakravet gitt i § 10 c): navn på avsender/mottaker.

Krav om avsender/mottakeropplysninger kommer ikke til anvendelse for organinterne dokumenter.

Metadata i kontrollstøtteverktøyet

Registreringstidspunkt	ControAssignment.data.created
Identifikator sak og dokument i sak	ControAssignment.data.id (UUID)
Korrespondansepart	<blank>Intern part i saksbehandlingen, ingen verdi, da verktøyet ikke korresponderer med noen
Dato for dokument (for referanse)	Ferdigstilt dato: ControAssignment.data.edited
Knytning til forretningskontekst	Tolls arkivnøkkel (finnes flere).
Tidspunkt behandlingsoppfølging	Ferdigstilt dato av kontrolloppdraget. Ikke i kontrollstøtteverktøyet.
Behandlingsoppfølging	Ikke i kontrollstøtteverktøyet

Betydningen av et utvidet «dokument-begrep» og tidspunktet for registrering i journal (Arkivlova paragraf 2 a og Offentleglova § 4)

Det har vært flere ulike synspunkt rundt hvor stort handlingsrom som ligger i begrepet «dokument». Tolletaten utfordret Arkivverket og Digitaliseringsdirektoratet på definisjonen av begrepene «dokument» og «datasett», og mente at disse var sammenfallende. Arkivloven definerer et dokument som en logisk avgrenset informasjonsmengde som er lagret på et medium for fremtidig lesing, lytting, fremvisning eller overføring. Et datasett defineres i Felles datakatalog som en organisert samling av data, for eksempel i form av en tabell, liste eller en database som kan gjøres tilgjengelig som en nedlastbar fil, og/eller nåes via et Web API.

Sandkassen har lagt til grunn at en samling av data også kan falle inn under definisjonen av et dokument i arkivregelverket. Vurderingen er knyttet til at volum eller størrelse er ikke betinget av «dokument»-definisjonen så lenge informasjonsmengden er logisk avgrenset. Det er dermed ingen forskriftsmessige hindringer slik vi ser det forbundet med å definere et datasett (eller en delmengde av et datasett) som et dokument. Dette kan også bidra til å gjøre flere virksomheter oppmerksom på at deres datasett kan være en del av virksomhetens arkiv og dermed skal ivareta hensynene som forventes av et arkiv. Det bør vurderes videre om det har konsekvenser å likestille disse begrepene.

Arkivloven med forskrifter regulerer ikke tidspunktet for når journalføring skal finne sted. Stortingsmelding nr. 32 (1997/98) *Om offentlighetsprinsippet i forvaltningen* presiserer at det må antas at journalføringen skal skje fortløpende. Også hensynet til offentlighetsprinsippet tilsier dette. Dersom vi skal legge til grunn en utvidet forståelse av «dokument»-begrepet, må det avklares hvorvidt et datasett reflekterer saksbehandlingsprosessen mens den foregår, eller om dette er uvesentlig ettersom datasettet genereres umiddelbart ut fra informasjonen som registreres. Hvis kravet til fortløpende journalføring eller registrering skal kunne opprettholdes, kan det ikke være en vesentlig differanse mellom Tolletatens registrering av data og tidspunktet det publiseres.

I tradisjonell saksbehandling, må et dokument *ferdigstilleles for at offentlighet skal kunne skje*. Rettleiar til offentleglova pkt 3.3.5, sier at et dokument som ikke er ment for utsending er opprettet når det er ferdigstilt. Tanken om at dokument er ferdigstilt gir mening om man legger til grunn at den enkelte kontrollrapport er dokumentet. Hvis man legger til grunn at datasettet som helhet er dokumentet, vil det stadig legges til nye rader, og det vil være vanskelig å fastsette noe tidspunkt dokumentet er ferdigstilt. Merk at Felles datakatalog ikke stiller krav til at et datasett skal være ferdigstilt for at beskrivelsen av det skal offentliggjøres.

En interessant distinksjon finnes også i direktiv (EU) 2019/1024 om åpne data og viderebruk av informasjon i offentlig sektor i norsk rett, hvor det slås fast at dynamiske data vil være dokumenter etter offentlighetsloven jf. høringen s. 25. Følgelig må det også betraktes som arkiv etter arkivloven. 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».

Hva har vi lært?

Så, hvilken betydning får det for Tolletaten at de i sandkassen kan legge til grunn en utvidet «dokument»-definisjon? Og hvilket handlingsrom kan virksomheter potensielt få, dersom en slik justert begrepsforståelse blir innført for alle virksomheter?

Det at definisjonene av dokument og datasett er fleksible, gjør det mulig å fange tre ulike typer dokumenter i kontrollstøtteverktøyet:

Dokumenttype	Nivå 1	Nivå 2	Nivå 3
Beskrivelse	Den enkelte kontrollrapport er et dokument. Dokumentet er en enkelt rad i tabellen over kontrollrapporter, det vil si rapporten fra den enkelte kontroll	En delmengde (granulering) av datasettet er et dokument. Dokumentet er en sammenstilling av data som inneholder alle kontrollrapporter som har visse kjennetegn til felles	Datasettet som helhet kan defineres som et dokument. Dokumentet er da en repeterende struktur av data som inneholder alle kontrollrapporter som skapes, og utvides fortløpende
Eksempel	Rapport fra kontroll av AX79497 ved Ørje tollsted, 8. august 2022 kl. 14.03	Kontrollrapporter for godstrafikk for Ørje tollsted august 2022	Alle kontrollrapporter
Mulige bruksområder (ikke uttømmende)	Bevis på hva som skjedde i den enkelte kontroll. Grunnlag for saksbehandling i f.eks. klagesaker	Produksjon av statistikk. Innsyn som kilde for datadrevet journalistikk	Datadeling med tiltrodde samarbeidsparter. Kvantitativ forskning på grensekontrollvirksomheten
I Noarks Informasjonsmodell	Registrering	Klasse (eventuelt mappe)	Arkivdel (eventuelt klasse)

For mer inngående om hvordan inndelingen (granuleringen) foreslås gjort i kontrollstøtteverktøyet, se kapittel 3 om offentlighet. Siden de ulike dokumenttypene er egnet for ulike formål er erfaringen fra prosjektet at Tolletaten og andre virksomheter bør stå fritt til å vurdere hvilket nivå som er hensiktsmessig for sine behov. Å legge regulatoriske føringer i forskrift eller standardisering for at ett nivå skal gjelde i alle tilfeller, vil kunne medføre en uønsket innsnevring av «dokument»-begrepet. Merk at hvilke data som tas vare på er den samme, uavhengig av hvilket nivå man velger å definere som omfang av dokument.

Når det gjelder vurderingene av om det er journalføringsplikt og arkiveringsplikt er det naturlig å se på nivå 3 i listen over. For øvrige vurderinger står virksomhetene mer fritt til å identifisere hva som er hensiktsmessig.

2.2. Utforskning av om standarder til Felles datakatalog kan ivareta arkivfaglige hensyn

Her tar vi for oss om bruken av standarder fra Felles datakatalog kan brukes til å ivareta arkivfaglige hensyn for beskrivelse av datasett. Arbeidet er inspirert av prosjekt «Metadata og katalog», som pågår i Arkivverket.

Kapittelet bør leses sammen med vedlegget, hvor sandkassen har kartlagt hvordan Felles datakatalog svarer til minimumssettet av metadata som er utarbeidet i prosjekt «Metadata og katalog».

Kort om pågående arbeid med metadata i Arkivverket

Prosjekt «Metadata og katalog» ser på behov for metadata i prosessene for avlevering, langtidsbevaring og tilgjengeliggjøring. Hensikten er å legge til rette for mer gjenbruk av metadata på tvers av disse prosessene i Arkivverket og i fellesløsningen Digitalarkivet. Arkivverket er avhengig av gode beskrivelser av datasett for at de skal være anvendbare i et langtidsperspektiv og fremtidige brukere skal kunne ha tillit til dataene. Maskinelle prosesser, standardisering og gjenbruk av metadata vil også gi bedre datakvalitet, gjøre det enklere å gjenfinne og dele historiske data og å koble arkivmateriale sammen med annen type informasjon.

I 2020 ble det gjennomført en «Proof of Concept», der resultatet var at mange av, men ikke alle metadatabehovene som var identifisert kunne dekkes gjennom standardene som brukes i Felles datakatalog. For mer informasjon se [presentasjon fra arbeidet](#) fra 2020. Sandkassen har bygget videre på disse resultatene.

Metadatabehov i Tolletaten

De fleste metadatabehov er like relevante i dokumentasjonsforvaltningsfasen som etter overlevering til Arkivverket. Dette er typer metadata som også er viktige for å legge til rette for datadeling og offentlighet i dag. Dette gjelder blant annet metadata om formål, hvilken tidsperiode det dekker (yngste og eldste informasjonsobjekt), hvem som eier det og har råderett over dataene, hva slags data det inneholder og hvilke hjemler som regulerer bruk og tilgang.

Tolletaten har i tillegg identifisert noen særlige behov for metadata som Tolletaten har knyttet til dokumentasjonsforvaltning i kontrollstøtteverktøyet (se også kapittel 2.1).

Datakataloger som alternativ måte å ivareta arkivhensyn

Felles datakatalog legger opp til at virksomheter skal ha lokale datakataloger som brukes internt for å ha oversikt over datasett, begrep, informasjonsmodeller og hendelser. Det er noe som kan hjelpe virksomheter med å beskrive sin dokumentasjon med metadata som er nyttig for arkivformål i de lokale datakatalogene. Virksomhetene har da et sted å forbedre sine beskrivelser av datasett, begrep, informasjonsmodeller og annet før de skal publiseres på Felles datakatalog.

De lokale datakatalogene bidrar til bedre beskrivelser av data og arkiv både for dokumentasjonsforvaltning, journalføring og arkivering. Det er derfor viktig å ta i bruk alle delene av lokale datakataloger også med å beskrive begrep, informasjonsmodeller, hendelser og tjenester. Dette er også i tråd med anbefalinger i veileder for orden i eget hus og intern kontroll.

Det er mulig å bruke uttrekk fra lokale datakataloger og Felles datakatalog for å publisere informasjon på offentlig journal og skape arkivuttrekk. Felles datakatalog kan fungere som naturlig del av arkivplanen, særlig for åpne data, og gi oversikt over hva som finnes og hvordan dette skal håndteres.

Arbeidet med å beskrive informasjon i en lokal datakatalog kan ha stor nytte for å ivareta arkivkrav, ved at det skapes bedre oversikt og at det skjer mer kontinuerlig forbedring. Metadata herfra kan inngå i arkivplan og potensielt også i å beskrive arkivuttrekk som skapes.

Hva har vi lært?

Læringen fra utprøvingen viser at det er sammenfallende behov mellom ivaretagelse av arkivfaglige hensyn og hensyn for intern styring av informasjon, deling av informasjon og offentlighet. Dette betyr at det er gevinster i å se samlet på hvordan få til bedre beskrivelser av datasett, og lage felles oversikter over hva som finnes av informasjon, i en virksomhet. Erfaringen fra Tolletaten er at dette gir en helhetlig og tverrfaglig tilnærming til informasjonskartlegging og -forvaltning, gjerne før eller samtidig med at informasjon skapes og er i administrativ bruk.

Videre ligger det verdi i å skape helhet i Felles datakatalog. Bruk av både datasett-, API-, begrep-, informasjonsmodell- og tjenestekataloger i felleskap skaper gevinster som en ikke kan oppnå om disse komponentene skulle brukes hver for seg. Det at det er mulig å både finne datasett, tilhørende API, informasjonsmodeller, begrep, tjenester og hendelser, gjør at det etter sandkassens erfaring er mulig å ivareta aspekter ved arkivfaglige hensyn.

Merk at både datasett, begreps- og informasjonsmodeller kan inneholde sensitive metadata som ikke er egnet for deling. Dette betyr at det kan være lurt å vurdere hvilke type informasjon og datasett en anvender dette på.

For datasett som ikke er åpne, vil beslutningen om å ikke offentliggjøre dem være hjemlet i en lovparagraf. Det er derfor avgjørende for riktig forvaltning av dataene at også disse hjemlene følger med ved overføring. Hjemlene hjelper i å vurderingen rundt å dele, gi

innsyn i og publisere informasjon. Det kan for eksempel være hjemler som anvendes i saksbehandlingen og oppgaveløsningen for Tolletaten. Hjemler er også viktig for å spore sammenheng hvor informasjon har vært skapt og årsaken til at den oppsto. Erfaringen viser videre at metadata er relevant for ulike formål, som gjerne kan sees på i parallell. Det gjeldende sett med metadata fra Arkivverket og som er testet her, er laget for datasett som allerede er avlevert eller deponert. Derfor er ikke alle metadata relevante nå, men behov for metadata som oppstår ved avlevering eller deponering bør likevel hensyntas i kartleggingen mens data er i aktiv bruk.

2.3. Hvilke deler av Noark er Felles datakatalog egnet for å erstatte?

Her adresserer vi om Felles datakatalog er egnet for å erstatte deler av dagens standardisering.

Formål, avgrensninger og bruk

Standardene i Felles datakatalog handler primært om hvordan ulike aspekter ved datasett beskrives. Standardene legger ikke føringer for hvordan enkeltforekomster av informasjonsobjektene i et datasett er beskrevet, men legger til rette for at det skal lages typebeskrivelser av objektene.

Noark-standarden inneholder, i tillegg til krav til metadata også krav til funksjoner i verdikjeden for arkiv, både i opprettelse, overføring, bevaring og publisering. Standarden stiller også krav til beskrivelsene av enkeltforekomstene av informasjonsobjekter. For denne utprøvingen er det opprettelsen av informasjonen som er i fokus, og så lenge den er i aktiv bruk. Overføring og andre prosesser er utenfor prosjektets omfang. Forskjellene i hva standardene inneholder er en naturlig følge av at standardene er utviklet for ulike formål.

Siden datasettet ikke inneholder journalpliktige dokumenter, faller det ikke inn under datasett som skal følge Noark. Men om man legger til grunn at hver enkelt rapport er dokumenter, vil strukturen være så lik det Noark omtaler som en logisk mappestruktur at det faller inn under Noarks beskrivelse av datasett som kan være egnet for Noark (jf. Noark 5 versjon 5 kapittel 1.2.2).

Egenskaper ved objekter

Noarks informasjonsmodell stiller krav til struktur og metadata-beskrivelser for datasett av typen arkiv. En svært forenklet fremstilling av denne er at det finnes ulike nivåer i et hierarki, herunder arkivdel (som en samling av mange dokumenter), klasse (som en avgrensing/gruppering av samlingen av dokumenter) og registrering (som ett informasjonsobjekt i samlingen av mange dokumenter). Kravene til hvilke metadata som skal følge med er til dels ulike for de forskjellige nivåene i hierarkiet.

Hvis man skal identifisere metadata for datasettet kontrollrapporter som helhet, er det beskrivelser på arkivdel- og/eller klasse-nivå som er relevant. Her er det mulig å bruke Felles datakatalog sine felter for å dekke behovene som Noark beskriver. Å bruke Felles datakatalog som erstatning for Noark for dette formålet er dermed mulig, under forutsetning at den som lager datasettbeskrivelsene ikke kun forholder seg til Felles datakatalog sin

definisjon av hvilke felter som er obligatorisk, men at også metadata som dekker behovene for arkiv fylles ut.

Hvis man legger til grunn at den enkelte rapport er et dokument, er det metadatakravene for registrering som er de mest relevante. Disse kan beskrives i Felles datakatalog i de tilfellene de er like på tvers av alle enkeltforekomster i datasettet. Når de er ulike må det stilles tilleggskrav til at metadata på registrering-nivå beskrives som del av dataene i datasettet. Eksempelvis må det for hver rapport finnes metadata for autentisitetstøtte, som at hver registrering har en unik identifikator, og at det er knyttet loggemetaddata til registreringen som blant annet beskriver hvem som opprettet den, når det skjedde og så videre. En slik tilnærming vil kreve at skaperen av datasettet peker til hvilke felter i datasettet som er å anse som de ulike typene autentisitetstøttende metadata.

Konklusjonen er at Felles datakatalog, gitt at det brukes riktig og suppleres med metadata som del av datasettet, kan dekke arkivbehovene for metadata.

Krav til prosess

Noark inneholder en rekke krav til for eksempel dokumentering av prosessen fram til endt saksbehandling, i tillegg til aspekter rundt offentlighet. Prosjektet har med vilje fokusert mest på sistnevnte. Enkelte krav kan også være praktisk vanskelig å verifisere ved å bruke Felles datakatalog eller datasettet alene. Det vil for eksempel kreve tillit til at autorisasjonsregler og logging er implementert på hensiktsmessig vis i løsningen som lager datasettet.

Vi ser med andre ord at standarder i Felles datakatalog, og spesielt DCAT-AP-NO, fremdeles har litt å gå på når det gjelder å ivareta kravene i Noark som gjelder prosess av saksbehandling og håndtering av arkiv. Derfor er det behov for mer utforskning av DCAT-AP-NO og CPSV-AP-NO med tanke på de andre aspektene ved Noark, og eventuelt om det må suppleres med andre standarder.

Hva har vi lært?

Erfaringene fra prosjektet tilsier at standardene i Felles datakatalog har stor overføringsverdi til arbeidet med utvikling av en nye standarder for dokumentasjonsforvaltning på hvordan datasett og informasjon beskrives. Det er behov for mer utforskning for å se hvordan standardene kan ivareta prosesskravene i Noark-standardene.

3. RESULTATER - EINNSYN OG FELLES DATAKATALOG

Her undersøker vi om offentlighet bli ivaretatt og informasjon bli publisert på einnsyn om den beskrives og tilgjengeliggjøres ved bruk av standardene som brukes i Felles datakatalog.

3.1. Kontekst

Tolletaten er underlagt Offentleglova og skal føre offentlig journal som gjøres tilgjengelig på internett, jf. Offentleglova § 10 og Offentlegforskrifta § 6 første ledd. Denne tilgjengeliggjøringen blir gjort på elnnsyn, der overføring av data skjer ved hjelp av en datamodell basert på Noark-standarden.

Imidlertid fører digitaliseringen av forvaltningen til stadig mer saksbehandling i fagsystemer, samhandlingsverktøy og andre systemer som ikke baserer seg på Noark. Det gjør at elnnsyn ikke lenger kan basere seg utelukkende på en Noark-basert datamodell, men må legge til rette for å kunne motta og tilby søkefunksjonalitet i data avlevert fra flere systemer enn de Noark-baserte. Regelverket for offentlighet og behovet for innsyn i forvaltningens saksbehandling er det samme uavhengig av systemer.

Det er også stadig sterkere signal om at offentlige virksomheter bør legge til rette for deling og tilrettelegging av data slik at andre virksomheter, både offentlige og private, skal kunne hente ut og skape verdi fra de delte dataene.

Tolletaten dokumenterer mange av sine oppgaver som strukturerte datasett. I sandkassen bruker vi datasett fra kontrollvirksomheten som case, såkalte kontrollrapporter fra kontrollstøtteverktøyet. Som rapporten viser, er mye av dette ikke journalføringspliktig. Tolletaten ønsker likevel å legge til rette for innsyn og offentlighet i datasettene. Samtidig ønsker de seg en effektivisering av tilgjengeliggjøringen gjennom å avdekke hvilke muligheter som ligger i å benytte Felles datakatalog fremfor en Noark-basert datamodell.

I sandkasseprosjektet har vi utforsker nye måter å tenke offentlighet på ved å se på mulighetsrommet for tilgjengeliggjøring av datasett i et samspill mellom elnnsyn og Felles datakatalog.

Ambisjonen har vært at brukeren skal kunne oppdage dataene både i elnnsyn og Felles datakatalog, mens selve tilgjengeliggjøringen av datasettbeskrivelsene kun skal gjøres i Felles datakatalog. For å imøtekomme ønsket om effektivisering i avlevering fra Tolletaten bør ikke dataene dobbeltpubliseres til begge tjenester. Denne måten å tilgjengeliggjøre datasettbeskrivelsene på vil gi mer og bedre offentlighet for innbyggerne.

3.2. Resultater fra prosjektet

Tolletaten har levert offentlig journal til elnnsyn siden 2010. Før oppstart av sandkasseprosjektet ligger det omkring 13 000 journalposter med tittel «Kontrollrapport» eller varianter av dette fra Tolletaten i elnnsyn. Flertallet er interne dokumenter og i resten er avsender og mottaker i stor grad skjermet.

Det er vanskelig for offentligheten å søke eller finne ønsket informasjon blant 13 000 journalposter hvor alle har samme tittel, men ikke samme innhold. I tilfeller der offentligheten likevel vil søke innsyn i en kontrollrapport vil Tolletaten trolig motta mange flere innsynskrav enn nødvendig fordi brukeren ikke vet hvilke som er relevante og velger å be om innsyn i flere enn nødvendig for å øke sin sjanse for å få ønsket informasjon.

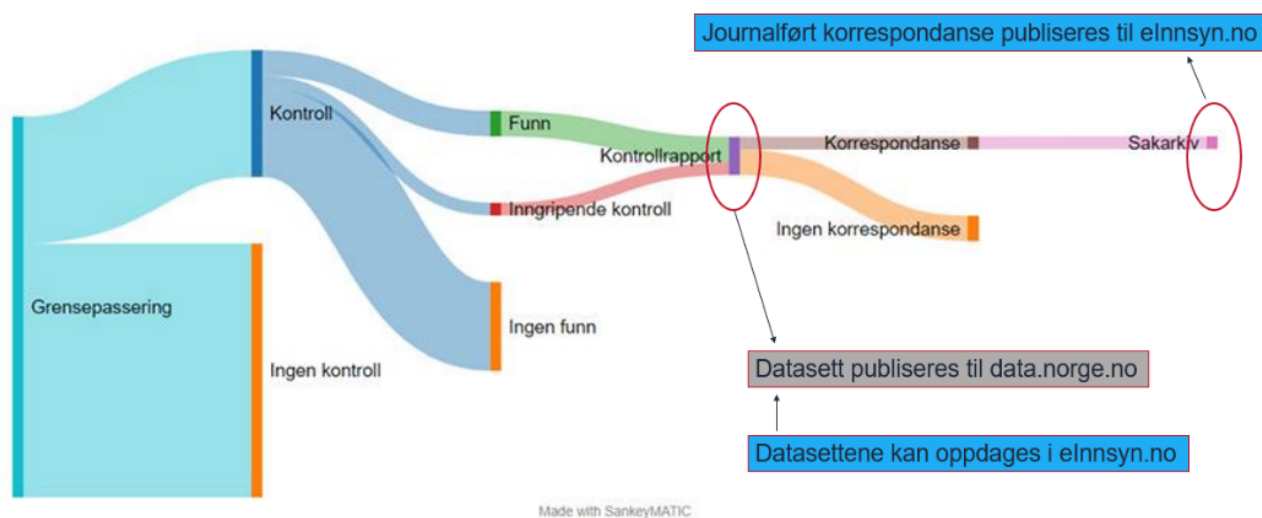
Når et datasett inneholder opplysninger som er unntatt offentlighet, vil verdien ofte være større av å få innsyn i sammenstillinger framfor enkeltrapper. Med dette mener vi nivå 2 eller 3 som fremstilt i tabellen i kap. 2.3. Ved å publisere kontrollrapportene som datasett i Felles datakatalog har Tolletaten ønske om å utvise mer offentlighet og gi innbyggerne enklere og bedre muligheter for innsyn. Samtidig er det et ønske om å effektivisere behandlingen av innsynskrav, ved at innsynskravene har høyere kvalitet og vil være mer målrettede.

Gjennomføring

I Tolletatens kontrollstøtteverktøy blir det generert kontrollrapporter i de tilfeller der kontrollen medfører et funn, eller det er blitt utført en «inngripende kontroll». Disse kontrollrapportene danner til sammen et datasett.

I noen tilfeller etterfølges disse kontrollene av oppfølging og korrespondanse. Denne korrespondansen vil da utgjøre en saksbehandling og som gjennomgangen i kapittel 2 viser, skal denne journalføres og publiseres på Tolletatens postjournal i elnnsyn på vanlig måte.

De resterende kontrollrapportene danner et datasett som ikke trenger å journalføres, og kan overføres Felles datakatalog på standardisert måte for deling av data. Dette er illustrert ved figuren under.



Hvordan få til felles oppdagelse av datasett mellom elnnsyn og Felles datakatalog

En problemstilling i sandkassen har vært å utforske muligheten for mer og enklere offentlighet for innbygger ved å kunne oppdage informasjon på tvers av fellestjenestene elnnsyn.no og Felles Datakatalog. Samtidig skal informasjonen ikke dobbelt publiseres til begge tjenestene. Denne problemstillingen ble etter hvert avgrenset til å kunne søke i elnnsyn etter Tolletatens datasettbeskrivelser som er tilgjengeliggjort i Felles datakatalog.

For å oppnå målsettingen om mer offentlighet og bedre mulighet for innsyn i datasettene er det ikke nok å gjøre det teknologisk mulig. Datasettene må også legges til rette på en måte som gjør at brukeren har nok informasjon til å kunne søke og finne ønsket kontrollrapport.

Sandkassen sitt forslag til løsning:

- Tolletatens hoved-datasett med kontrollrapporter bør deles opp (granuleres) i flere logisk avgrensede datasett, som er tidsavgrenset. Disse bør gis meningsbærende titler med intuitive søkeord som for eksempel «Kontrollrapporter tungbil, Ørje Tollsted, 2021».
- Datasettene tilgjengeliggjøres med datasettbeskrivelser i Felles datakatalog. Relevant informasjon bør også beskrives i Tjenestekatalogen og Begrepskatalogen.
- For hvert datasett oppretter Tolletaten en «dummy-journalpost» med den søkbare tittelen som tittelfelt, og med lenke direkte til datasett-oppføringen i Felles datakatalog. «Dummy-journalposten» overføres elnnsyn og blir søkbar som en del av Tolletatens postjournal.
- Datasettene til Tolletaten oppdateres løpende via de tilgangslenkene som er oppgitt i Felles datakatalog kan fortløpende oppdateres med nye kontrollrapporter uten at det er behov for oppdatering i elnnsyn eller i Felles datakatalog. Det må imidlertid vurderes hvor lang tidsperiode et datasett kan inneholde, sett opp mot at publiseringsdato i elnnsyn kan bli lite intuitiv og at et søketreff etter hvert vil komme langt ned i trefflisten.

3.3. Hva bør gjøres fremover?

Ved publiseringstidspunkt for denne rapporten er løsningsforslaget ikke testet ut i praksis. Det gjenstår for Tolletaten å finne ut hva som vil være en hensiktsmessig oppdeling (granulering) av datasettene, og hva som vil være åpne data. Det må også etableres rutiner og kvalitetssikring for generering av «dummy-journalposter». For å maksimere effektiviseringen er det en målsetting om automatiserte prosesser på dette.

Tolletaten ønsker også å foreta en vurdering på om datasettene kan gi for mye informasjon til publikum slik at kontrollvirksomheten kan bli mindre effektiv.

Både Felles datakatalog og elnnsyn er tjenester som skal understøtte praktiseringen av innsynsretten i offentlighetsregelverket i Norge. elnnsyn har noe mer tilrettelagt funksjonalitet for sluttbruker som Felles datakatalog bør vurdere å gjennomføre. Metadata om dokument skal publiseres på offentlig journal og være tilgjengelig slik at innbyggeren skal ha mulighet til å be om innsyn. Dette gjelder også metadata. I elnnsyn vises dokumenttittelen og hvilken unntakshjemmel dokumentet er vurdert unntatt etter. Selv om dokumentet er unntatt offentlighet kan innbyggeren likevel be om innsyn og organet må gjøre en innsynsvurdering. Slik funksjonalitet er ikke tilrettelagt i dag i Felles datakatalog.

I elnnsyn er det lagt til rette for å enkelt be om innsyn i hvert enkelt dokument via en knapp og bestillingskurv. Sandkasseprosjektet har lagt inn [behov for tilsvarende funksjonalitet](#) som en brukerhistorie i Felles datakatalog sin Github.

Open Ivbachmann opened this issue on 22 Jun • 0 comments



Ivbachmann commented on 22 Jun • edited

Som: Privatperson, presse, næringsdrivende
Ønsker jeg å: Kreve innsyn (etter bestemmelser i Offentleglova) i innholdet i datasett på en effektiv måte (for eksempel via en dedikert knapp).
Slik at: Eier av datasettet kan vurdere om den som ber om innsyn kan få det og eventuelt gi tilgang (gjennom andre kanaler).

Som: Eier av datasettet
Ønsker jeg å: Motta innsynskravet via egnet kanal
Slik at: Jeg kan effektivisere prosesser rundt tilsyn

Merk: Registrert i møte i regulatorisk sandkasse for arkiv, data og offentlighet - med tilslutning fra Digdir, Arkivverket og Tolletaten

Relatert til epos: #731

Lenke til design: Se tilsvarende funksjonalitet i elnnsyn
Lenke til løsningsarkitektur:

Forretningsverdi (Kritisk/Viktig/Mindre viktig/Kjekt å ha):

Assignees
oystein-asnes

Labels
userstory

Projects
Behovsprosjekt
Avklare behov

Milestone
No milestone

Development
No branches or pull requests

elnnsyn jobber også med å utarbeide en generisk metadatamodell for offentlighet/tilgjengelighet som et supplement til Noark (eller fremtidig standardisering). De regulatoriske sandkasseprosjektene har vært gode arenaer for å få innspill og kjennskap til behovene i en slik datamodell.

I dette prosjektet med Tolletaten, så vi at Noark-modellen som brukes i elnnsyn i dag likevel var tilstrekkelig for å tilgjengeliggjøre informasjon om datasettet som ble publisert i Felles datakatalog. Eneste justering som måtte gjøres var å utvide begrepet «dokument» hos elnnsyn til å kunne inkludere lenke til datasettbeskrivelsen i Felles datakatalog. Det ville også være mulig å lenke direkte til datasettet (for eksempel et API), hvis det hadde vært åpen informasjon tilgjengelig. Noe tilpasning må utføres i brukergrensesnittet på elnnsyn for å tydeliggjøre at det her ikke er snakk om et dokument. Dette vil elnnsyn kunne tilpasse ved bruk av «format»-verdien i metadatafeltene i «dokument-objekt»

4. VEIEN VIDERE

4.1. Gevinster

Læringen fra arbeidet kan bli et bidrag til en mer effektiv dokumentasjons- og informasjonsforvaltning. Også andre virksomheter kan på sikt dra gevinster av læringen, og det er også et bidrag til mer og bedre tilrettelagt innhold i elnnsyn og Felles datakatalog.

For Digitaliseringsdirektoratet blir det viktig å bygge videre på de erfaringer som er høstet om relasjonen mellom Felles datakatalog og elnnsyn fremover. Også hvordan elnnsyn skal fungere som innsynsløsning for stadig nye digitaliserte løsninger og dataformater i forvaltningen blir viktig å jobbe videre med. En del av dette arbeidet er utviklingen av en ny datamodell for elnnsyn med arbeidstitel «metadatamodell for offentlighet».

Gevinster for offentligheten

- Mulighet for mer og bedre tilgang på informasjon fra Tolletatens virksomhet.
- Mer relevante søketreff ved søk i elnnsyn, når beskrivelsen av dataene i emnefelt inneholder intuitive titler heller enn tidligere «Kontrollrapport».
- Søketreff i elnnsyn vil peke til et datasett som gir data i kontekst i stedet for enkeltrapper, det vil gi kunne bedre forståelse og merverdi for sluttbruker.

Gevinster for elnnsyn

- Det blir ryddigere og enklere å få relevante treff i elnnsyn når oppføringer med verdiløse titler som «Kontrollrapport» erstattes med oppføringer med relevante, intuitive journalposttitler.
- Mer dokumentasjon som er omfattet av offentligregelverket fra forvaltningen vil bli tilgjengeliggjort gjennom elnnsyn på en brukervennlig måte.
- elnnsyn viser seg som en fleksibel løsning som kan brukes til oppdagelse av et bredere utvalg av informasjon fra forvaltningen enn tradisjonelle journalposter.

Gevinster for Tolletaten

- Mulighet for færre og bedre kvalitet på innsynskrav.
- Mer offentlighet og mer deling av data.
- Effektivisering av overleveringsprosesser til arkiv, offentlighet og datadeling.

4.2. Læring og regelverksutvikling

Prosjektet viser at å lykkes med offentlighet, informasjonsforvaltning og dokumentasjonsforvaltning er overlappende aktiviteter. Det gir mer effektivt, tryggere og bedre oppfyllelse av offentlighetsprinsippet å tenke offentlighet og innsyn sammen med arkivhensyn tidlig i informasjonsforvaltningsløpet.

Det vil videre bli viktig å følge opp prosjektets utprøving av en utvidet forståelse av hva et «dokument», åpne data og «datasett» kan være etter arkivregelverket. Tolletaten en datadrevet og digitalisert etat, er ett av mange eksempel på et tydelig behov for et mer teknologivennlig regelverk. Eventuelle regulatoriske føringer i forskrift eller standardisering bør forsøke å unngå å begrense at for at ett nivå (inndeling) skal gjelde i alle tilfeller. Dette vil erfaringsvis potensielt kunne medføre en uønsket innsnevring av hva som blir tatt var på for ettertiden. Det bør vurderes videre om det har konsekvenser å likestille disse begrepene, for eksempel bør det vurderes om bruk av standardene i Felles datakatalog vil være innenfor arkivlovens krav.

Prosjektet anbefaler at Arkivverket, eventuelt i samarbeid med Kulturdepartementet, i det videre diskuterer grundig om det er ønskelig å harmonisere arkivregelverket med standardene i Felles datakatalog, og eventuelt hvordan dette kan skje. Det bør også være dialog med Justis- og beredskapsdepartementet om offentliglova.

Noark skal ikke videreutvikles, og Arkivverket jobber med behov for ny standardisering som skal komme i kjølvannet av Noark. Det er en viktig læring fra prosjektet at de undersøkte standardene i Felles datakatalog bør vurderes til å inngå som en del av standarder på arkivområdet.

Når prosjektet har sammenlignet Felles datakatalog med dagens Noark, er det tydelig at det er standarder som er ment for ulike formål og med ulik innretning. Felles datakatalog stiller krav til dokumentasjonsprosesser og metadata som dokumenterer kontekst og autentisitet, og kan potensielt brukes til å beskrive arkiv slik de er dannet. Det mangler beskrivelse av funksjonalitet for låsing og frysing metadata og dokumenter på systemnivå. Alternativt kan Felles datakatalog brukes til å beskrive arkiv på et overordnet nivå, og dermed danne grunnlag for arkivbeskrivelse og kataloginformasjon.

Å anvende Felles datakatalog til arkivformål, som den ikke er ment for, vil derfor innebære at den må kombineres med andre standarder som beskriver krav og rutiner for prosesser for håndtering av arkiv. Dette er krav som vil måtte stilles for å ivareta intensjonen i dagens arkivregelverk. Prosjektet anbefaler at Arkivverket følger opp denne læringen gjennom sitt pågående standardiseringsarbeid.

4.3. Deltagere i prosjektet

Navn	Rolle	Virksomhet
André Nergaard	Prosjektkoordinator	Tolletaten
Filip Hansen	Prosjektressurs – informasjonsarkitekt	Tolletaten
Atle Sander	Prosjektressurs – informasjonsarkitekt og ontologi	Tolletaten
Elisabeth Sanneland	Prosjektressurs – juss	Tolletaten
Ingeborg Anna Gaundal	Prosjektressurs – arkiv	Tolletaten
Simon Sjølhagen	Prosjektressurs	Tolletaten
Eirin Hansdatter Krogstad	Prosjektressurs	Tolletaten
Tove Faye-Lund	Prosjektressurs	Tolletaten
Erland Ferrer-Pettersen	Prosjektleder og redaktør	Arkivverket
Marthe Rosenvinge	Ansvarlig for regulatorisk sandkasse og redaktør	Arkivverket
Petr Vasilev	Prosjektressurs – metadata	Arkivverket
Martin Bould	Prosjektressurs – kvalitetssikring	Arkivverket
Olav Hagen-Sataslåtten	Prosjektressurs – juss	Arkivverket
Henry Christoffer Meen	Prosjektressurs – kvalitetssikring	Arkivverket
Mona Danielsen	Prosjektressurs – kvalitetssikring	Arkivverket
Lars Vegard Bachmann	Prosjektressurs – kvalitetssikring	Arkivverket
Jacob H. Handegard	Prosjektressurs – kvalitetssikring	Arkivverket
Stein Magne Os	Prosjektkoordinator	Digitaliseringsdirektoratet
John Eivind Presterud	Prosjektressurs – elnnsyn	Digitaliseringsdirektoratet
Ole Jomar Moen	Prosjektressurs – elnnsyn	Digitaliseringsdirektoratet
Øystein Åsnes	Prosjektressurs – Felles datakatalog	Digitaliseringsdirektoratet
Bodil Therese Cirotzki	Prosjektressurs – kvalitetssikring	Digitaliseringsdirektoratet
Siv Kvåle Sanden	Prosjektressurs – juss	Digitaliseringsdirektoratet

VEDLEGG: KARTLEGGING AV METADATA

Oversikt over behov for metadata som ivaretar arkivfaglige hensyn, hensyn til styring og deling av informasjon og offentlighet:

Minimumssett metadata

Metadatabehov	En forklaring av hva dette er i fritekst for å sikre at vi har felles forståelse av hva det betyr	Formatstandard / Iverksettelse av metadatabehov
Hvilke virksomheter (eller personer) skapte informasjonen (arkiv - ref. definisjon i arkitekturteamet)?	Liste over alle arkivskapere. Det offisielle navnet eller / og orgnr / personnr på den eller de virksomheter eller personer som har skapt arkivet. Unngå forkortelser. Se på samfunnssektor her.	dcat:dataset/dct:creator : foaf:Agent
Hvilken virksomhet eller person skal avlevere arkivet/arkivuttrekket?	Det offisielle navnet eller / og orgnr / personnr på virksomheten eller personen som avleverer datasett. Unngå forkortelser.	dcat:dataset/dct:publisher : foaf:Agent . Den som er ansvarlig for å gjøre datasettet tilgjengelig, kan gjerne være leverandør av løsningen. Muligens trenger vi en annen rolle her. Det er den _eneste_ påbudte egenskapen på datasettene i DCAT-AP-NO. Den er sktrukket et stykke allerede. Det er antagelig creator i tilfelle Toll. Her trenger vi antagelig å beskrive internt eierskap til datasettene.
Hva er arkivets formål og administrative brukssammenheng?	Til hvilket formål er dette datasettet skapt? Hvilke formål skulle det oppfylle? Hvorfor er datasettet skapt (oppgaver og tjenester som den skulle dekke ett behov for)? Dette feltet ligner på dagens arkivbeskrivelse. (Datasett har ikke noe formål)	/dct:description/dcat:distribution : rdfs:Literal. Tjenestekatalogen kan være nyttig her. Kan være referanse til tjenestekatalog, datasett kan være tjenesteresultat. Parallelt med formålsbeskrivelse i registerlov. Formål som særnorsk utvidelse av DCAT-AP. Formålsbeskrivelsen er ofte formulert i en lov / forskrift / annen rettskilde i description. Kan være referanse til arkivplan eller tjenestekatalogen. Her er det forskjell mellom hva er formål med hver enkelt kontroll og samling av kontroller.
Hva er de primære informasjonsobjektene i arkivuttrekket?	Hva handler arkivuttrekket om hovedsakelig? Hvilke objekter eller opplysninger er de viktigste? Eksempelvis sak, hendelse, person, eiendom osv.) En del av objekttypene/opplysningene i arkivuttrekket kan ha blitt identifisert og beskrevet som en del av arbeidet med personvern, informasjonssikkerhet og internt kontroll. For digitalt skapt materiale skal objekttyper være fra informasjonsmodellen til arkivuttrekket. Eksempel: Foto fra Kong Harald Vs besøk i, søknad om naturskadeerstatning.	Forskjell mellom samling av alle kontrollrapporter, og enkelt kontrollrapport. For samling er det enkle kontrollrapporter. For enkelt kontrollrapport: Primære infoobjekter kan kontrolloppdrag, -rapport og -objekt er viktige. Er det mulig å synliggjøre dette med verktøy fra Felles datakatalog? Muligens vi kan bruke begrepskatalog for å beskrive primære infoobjekter? Eller vi kan bruke deldatasett som informasjonsobjekt - informasjonmodellkatalog.
Hvilke hjemler gjelder når arkivet skal behandles i sin helhet	Her lister vi hjemler som gjelder for behandling for hele arkivet eller arkivuttrekket i gjeldende del av verdikjeden. Det er nødvendig å dokumentere endringer i dette feltet når det skjer. Eksempel: Arkiv til Norsk Hydro: Egne hjemler for innsyn i danningen ihht selskapets regler,	cpsv:follows : cpsv:Rule . Vi skiller i tilfelle Tolletatenmed å si at hjemler på samling av kontrollrapporter og enkle kontrollrapporter.

	etter avlevering gjelder avtalen mellom selskapet og Arkivverket.	
Er det spesifikke hjemler som regulerer innsyn som gjelder for alle objektene i arkivet?	Her lister vi hjemler som gjelder for alle objekter i datasett som er IKKE definert som primære. Må beskrive generiske hjemler. Eksempel: Statsministerens konotor: Beskyttelsesinstruksen gjelder for noen filer, og det er eksplisitt definert på disse filene. Instruksen gjelder ikke alle filene, og metadatabehovet som defineres her inneholder hjemmel som gjelder filene uten eksplisitt definert hjemmel. (Her må det kjøres POC)	cpsv:follows : cpsv:Rule . Målet er å ha oversikt over relevante hjemlene, og det som vi skal løse innen denne sandkassen
Hvilket tidspunkt er det eldste informasjonobjekt opprettet?	Startdato til innholdet i datasettet. Dato for det eldste informasjonobjektet i datasettet	dct:temporal / dct:PeriodOfTime. Det er mulig å la være å definere avslutningstidspunkt på periodOfTime. Hva skal publiseres og hva skal ikke? Poenget med elnnsyn i stor grad er å kunne se saken i utvklng.
Hvilket tidspunkt er det yngste informasjonobjekt opprettet?	Dato for det nyeste elementet i datasettet	dct:temporal / dct:PeriodOfTime.
Hvilke geografiske områder er relevante for informasjonen?	Geografisk dekningsområde for datasettet. Landskode (ISO 3166-1 alpha-3), regionkode/fylkesnummer (ISO 3166-2:NO) og/eller kommunenummer (0301, 0411, 2030 osv)	dcat:dataset/dct:spatial : dct:Location. Hvilke relasjoner fra kontrollrapport til lokasjon kan være relevante? Kontrollstedet? Om vi beskriver Tolletatensom organsasjon i CPSV, blir det mulig å koble til tollområdet.
Hvilke naturlige språk er viktige for å tolke arkivet/arkivuttrekket?	Språkene i selve datasettet. Det eller de språk innholdet i datasettet er skrevet på (ISO 639-1)	dcat:dataset/dct:language : dct:LinguisticSystem
Hvordan kan vi entydig identifisere arkiv (arkivuttrekk/arkivserier) (unik identifikator)	Unik identifikator for hele datasettet. F eks UUID, URI, OID. Det finnes anbefaling for identifikatorer for offentlige ressurser på nett: https://www.digdir.no/digitale-felleslosninger/peikarar-til-offentlege-ressursar-pa-nett/1492	dcat:dataset/dct:identifier : rdfs:Literal
Er det kjente lakuner eller manglende informasjon i arkivet (arkivuttrekk/arkivserier) og hva er de?	Med lakuner menes her hele serier eller deler av materialet som mangler. Eksempelvis tidsperioder eller geografiske områder som ikke følger med. Hull i saksnummerrekkefølge er ikke nødvendig å opplyse om her.	dcat:dataset/dqv:hasQualityAnnotation : dqv:QualityAnnotationdcat og dataset/dqv:hasQualityMeasurement : dqv:QualityMeasurement.

Særlige behov – Deling og innsyn

Metadatabehov	En forklaring av hva dette er i fritekst for å sikre at vi har felles forståelse av hva det betyr	Formatstandard / Iverksettelse av metadatabehov
Er det gradert eller fortrolig informasjon i datasettet?	Referanser til hjemler fra sikkerhetsloven, beskyttelsesinstruksen osv. Gradert informasjon defineres alltid kun på datasett nivå. Ref. https://nsm.no/regelverk-og-hjelp/sikkerhetsloven-og-forskrifter/	cpsv:follows : cpsv:Rule med type https://data.norge.no/vocabulary/rule-type#non-disclosure-rule
Er det skjerming som oppheves ved vedtak for hele datasettet?	Lovhjemmel som dekker hele datasettet som helhet. Flere ved behov. Dette inkluderer f. eks. offentlighetslova, personvernlovgivning (GDPR), statistikkloven, helseopplysningsloven.	cpsv:follows : cpsv:Rule med type https://data.norge.no/vocabulary/rule-type#non-disclosure-rule

Er det skjerming som oppheves ved vedtak for elementer i datasettet?	Alle lovhjemler som er nevnt og brukt på elementer i datasettet. Dette inkluderer f. eks. offentlighetslova, personvernlovgivning (GDPR), statistikkloven, helseopplysningsloven.	cpsv:follows : cpsv:Rule med type https://data.norge.no/vocabulary/rule-type#non-disclosure-rule
Er det skjerming som oppheves på dato for hele datasettet?	Lovhjemmel for å unnta offentlighet fram til en bestemt dato når det skal frigjøres. Dette inkluderer f. eks. offentlighetslova, personvernlovgivning (GDPR), statistikkloven, helseopplysningsloven.	cpsv:follows : cpsv:Rule med type https://data.norge.no/vocabulary/rule-type#non-disclosure-rule
Er det skjerming som oppheves på dato for elementer i datasettet?	Alle lovhjemler som er nevnt og brukt på elementer i datasettet og datoer på når det skal oppheve. Dette inkluderer f. eks. offentlighetslova, personvernlovgivning (GDPR), statistikkloven, helseopplysningsloven.	cpsv:follows : cpsv:Rule med type https://data.norge.no/vocabulary/rule-type#non-disclosure-rule
Hvem internt i organisasjon har råderett til datasettet?	Eierskap og forvaltning: hvem i Tolletateneier datasettet. I tilfelle Tolletatendet er divisjonsdirektør (nivå 2 leder). Tolletatenkobler dette til stilling / rolle, og ikke til person. Organisasjonskart beskriver disse stillingene.	dct:datasett/dct:qualifiedRelationship : dcat:Relation
Hva er datasettet del av?	Peker på andre datasett eller deler av katalogen	dcat:dataset/dct:isPartOf

Særlige behov – Tolletaten

Metadatabehov	En forklaring av hva dette er i fritekst for å sikre at vi har felles forståelse av hva det betyr	Formatstandard / Iverksettelse av metadatabehov
Hvem internt i organisasjon forvalter datasettet?	Hvem i Tolletatenforvalter datasettet. I tilfellet Tolletatendet er en er pekt ut av eier av datasett. Faglig forvaltning, peker ofte på divisjonsdirektører på nivå 2. Det kan være flere divisjoner som er ansvarlige for datasett, og det kan brukes som opphavet av data. Det kan være egne forvaltere på enkle distribusjoner av datasettet i begrepsmodell av DCAT-AP-NO.	dcat:Role er custodian - https://standards.iso.org/iso/19115/resources/Codelists/gml/CI/RoleCode.xml
Hvilke hjemler har datasettet vært skapt for?*	formål/hjemmel: hvilket formål dataene ble samlet inn eller skapt for	https://data.norge.no/vocabulary/rule-type#data-processing-rule
Inneholder datasettet personopplysninger?	hvorvidt datasettet inneholder personopplysninger og av hvilken kategori	
Inneholder datasettet særlige personopplysninger?	Først og fremst handler dette om datasettet inneholder særlige personopplysninger.	
Er datasettet en arkiv eller en dokumentasjon?	status som dokumentasjon eller arkiv: dokumenterer datasettet et prosessforløp eller en prosess, eller på annet vis betraktes som som dokumentasjon. Brukes for å skille mellom samling av kontrollrapporter (det er arkiv) og enkle rapporter (dokumentasjon)	
Når skal datasettet slettes?	levetid / tidspunkt for sletting (koblet tett til formål)	
Hva skal til for å bruke datasettet i andre deler av organisasjon?	vilkår for bruk og deling	https://data.norge.no/specification/dcat-ap-no/#klasse-rettighetsutsagn
Hvilke prosesser er datasettet relatert til?*	kobling til prosess: opphavsprosess og andre prosesser datasettet brukes i	https://data.norge.no/specification/dcat-ap-no/#Datasett-bleGenerertVed
Hvilke applikasjoner er datasettet relatert?*	kobling til applikasjon: opphavsapplikasjon og andre applikasjoner datasettet brukes i	https://www.w3.org/TR/2013/REC-prov-o-20130430/#Activity
Hvilke andre datasett er dette datasettet relatert* til?	kobling til andre datasett, opphav/etterkommere / del av/	https://informasjonsforvaltning.github.io/dcat-ap-no/#Datasett-relatertRessurs

Hvilke begrep brukes i dette datasettet?	kobling til begrep. Kobling til internt begrepskatalog.	https://informasjonsforvaltning.github.io/dcat-ap-no/#Datasett-begrep
---	---	---

Behovene markert med asterisk (*) er aspekter som var ment å inngå i “formål og administrativ brukssammenheng” i minimumssettet. Å detaljere dette har dog vært nyttig for å få en bedre mapping mot formatstandarder.