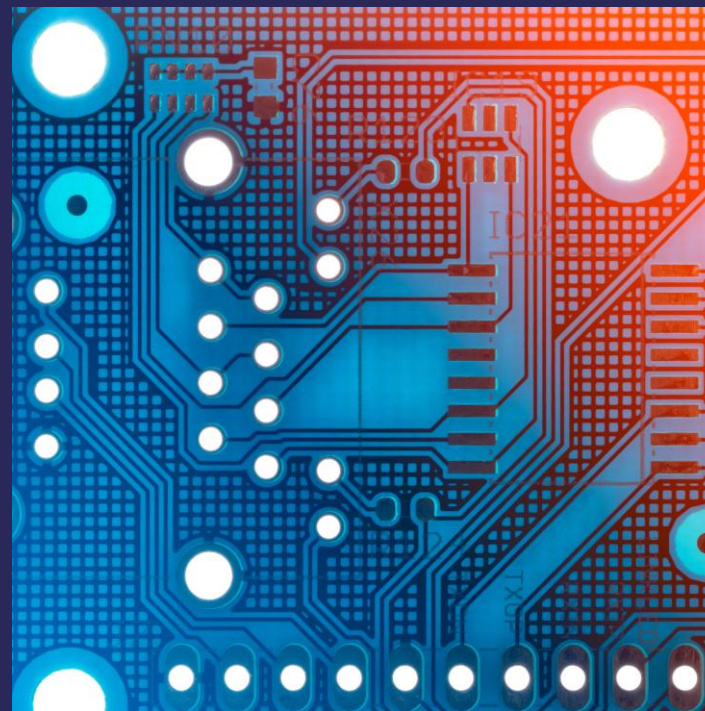


Kravspesifikasjon for innebygd arkivering for informasjonssystemer i Oslo kommune



Med ny arkivlov- og forskrift faller Noark-kravet bort

- ▶ Dette medfører at Byarkivet kan ta en mye mer aktiv rolle inn i prosessene for anskaffelse av nye fagsystemer i Oslo kommune
- ▶ Utfordringen er at vi i mange tilfelle kommer sent inn i prosessene, eller at det ikke er allokert tilstrekkelig midler for å tilfredsstille Byarkivets krav
- ▶ Dette utelukker allikevel ikke at vi kan gjøre en vurdering av systemet slik at det tilfredsstiller APIA-kravene, eller at vi sammen med leverandør og etat kan foreslå noen små justeringer for å komme i mål.

Hva er innebygd arkivering

- ▶ I realiteten er det at et system har funksjonalitet for registrering og lagring av dokumentasjon som sikrer:
- ▶ Autentisitet
- ▶ Integritet
- ▶ Pålitelighet
- ▶ Anvendbarhet

Autentisitet

- Misforstås ofte med «original», en term som ikke gir mening i et digitalt landskap. Det vi i stedet kan si er at påføring av visse metadata kan bevise autentisitet, ofte brutt ned på metadata som beskriver:
 - a) tidspunkt for registrering
 - b) aktør(er)
 - c) hendelse/transaksjon
- En endringslogg kan også bevise autentisitet, og det kan også løses gjennom lås/frys-funksjonalitet knyttet til registrering og informasjonsobjekt, samlet beskrevet som informasjonselement.
- Vi bør også ha i mente at en saksbehandlingsprosess i seg selv er med på å sikre autentisiteten, ved at en saksbehandler i alle ledd kontrollerer eller verifiserer innhold.

Integritet

- ▶ Integritet vil oftest være knyttet til *endring*.
- ▶ Vi kan ikke forhindre endringer knyttet til informasjonselementer i et fagsystem, men vi kan løse det gjennom versjonering, altså overskriving tillatt kun der tidligere versjoner lagres og bevares i løsningen..
- ▶ Vi vil også kunne si at på dette stadiet bør bevarings- og kassasjonsvurdering gjennomføres. I realiteten er dette en identifisering av informasjonsinnholdet som skal skapes i løsningen.

Anvendbarhet

- ▶ Anvendbarhet må i ytterste konsekvens være tilknyttet tilgjengeliggjøring
- ▶ Tilgjengeliggjøring kan både være migrering men også overføring gjennom uttrekk, men det må gjøres i en form som sikrer videre anvendelse av informasjonsinnholdet.
- ▶ Her finner vi overgangen mellom dokumentasjonsforvaltning og informasjonsforvaltning, og områder som semantisk interoperabilitet, datakataloger, datadeling, arkitekturprinsipper osv.
- ▶ Vi skal være klar over at tilgjengeliggjøring er et nedfelt krav i ny arkivlov.

Hva er Innebygd arkivering

- For å forstå behovet for å utvikle innebygd arkivering (IA) må vi fjerne rammen rundt Noark som krav.
- Spissformulering: alle fagsystemer som lagrer informasjon har innebygd arkivering, det vil si at systemet lagrer informasjon, men det arkiverer ikke nødvendigvis ut fra våre prinsipper eller krav.
- Det enkleste er å summere IA opp med APIA
- Noark inneholder en rekke krav som skal sikre lås, frys, journal, struktur og uttrekk.
- Fordi Arkivverket rundt 2021 bestemte seg for å slutte å utvikle Noark-standarden lanserte de konseptet Innebygd arkivering
- Overforenklet kan en si at Innebygd arkivering skal erstatte kravene i Noark som omhandler APIA

Hva Noark krevde, men som vi ikke trenger i fagsystemer

- Den konseptuelle modellen i Noark definerte struktur ut fra *arkiv* på øverste nivå, deretter granulert ned i *arkivdeler*.
- Arkivdeler er ikke en vesentlig komponent for fagsystemer.
- Videre har Noark mappe «klasse» og «mappe», men dette er heller ikke nødvendig for å sikre innebygd arkivering.
- *Klasse* går på klassifikasjon, men det kan være et hvilket som helst klassifikasjonssystem
- *Mappe* er ikke alltid hensiktsmessig, vi kan heller snakke om identifikator som binder informasjon sammen på tvers i en løsning, det kan være personnumre eller et objektnummer for en kommunal bolig.

Planlagt innført:

Reglement 103: 1 Anskaffelse og innføring av nye informasjonssystemer

- ▶ Nye informasjonssystemer som inneholder dokumentasjon skal så langt det lar seg gjøre følge prinsippene for innebygd arkivering, kravene er definert i reglement 103-1: Kravspesifikasjon anskaffelse av fagsystemer.
- ▶ Anskaffelsesprosjektet skal involvere Oslo byarkiv umiddelbart for utforming av kravspesifikasjon. Systemeier og/eller systemforvalter skal ivareta de dokumentasjonsfaglige kravene, slik at informasjonen ikke går tapt.
- ▶ Systemeier og/eller systemforvalter skal sørge for at kravspesifikasjonen inneholder krav som muliggjør:
 - ▶ a) produksjon av en logg eller en journal som inneholder en oversikt over registrerte metadata,
 - ▶ b) lås/frys ved registrering av metadata og lagring av informasjon
 - ▶ c) integrasjon mot statlige eller kommunale fellesløsninger og elektronisk signatur hvis det er behov for det,
 - ▶ d) utforming av et sett av maler for utgående dokumenter som kan ekspederes direkte fra løsningen,
 - ▶ e) produksjon av uttrekk.

APIA i ny arkivlov

- § 5 Arkivplikten
- Organene skal forvalte dokumentasjon som blir til som ledd i virksomheten, som arkiv. Dokumentasjonen skal sikres slik at informasjonen ikke går tapt, og forvaltes slik at:
 - a. opphavet til informasjonen alltid er kjent
 - b. det fremgår i hvilken sammenheng informasjonen er skapt og inngår i
 - c. informasjonen ikke blir endret
 - d. informasjonen er tilgjengelig på en måte som ikke begrenser fremtidig nytte

Dokumentasjonen skal sikres slik at informasjonen ikke går tapt, og forvaltes slik at:

- ▶ **a. opphavet til informasjonen alltid er kjent**

Dette kan være systemproveniens eller organisatorisk proveniens. Dette kan også være prosessproveniens; Identifikator er det sentrale.

- ▶ **b. det fremgår i hvilken sammenheng informasjonen er skapt og inngår i**

Kontekst. Hvis vi etablerer en integrasjon reduserer vi muligheten for å oppfylle dette kravet, fordi vi flytter informasjonen ut av sin opprinnelig kontekst. Fagsystemet har et datasett som er den autoritative kilden eller opphavet; «sammenhengen informasjonen er skapt og inngår i».

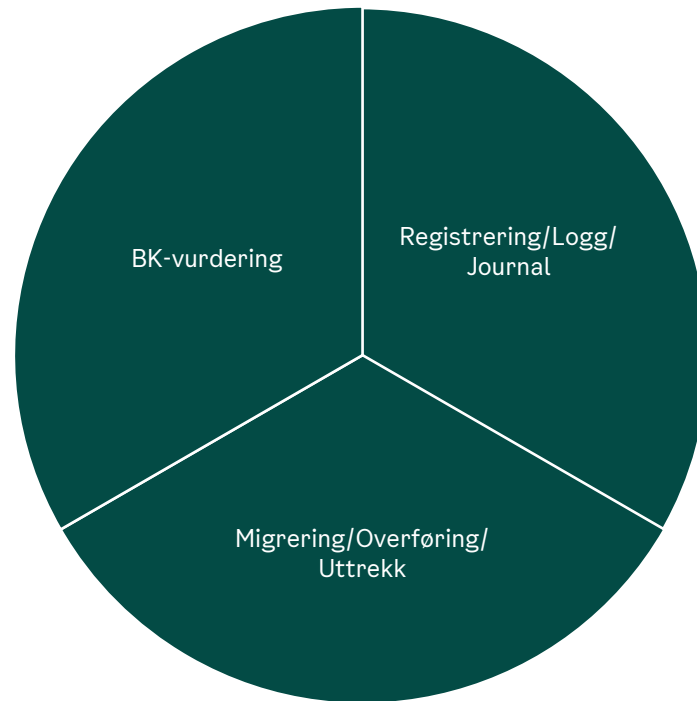
- ▶ **c. informasjonen ikke blir endret**

- ▶ Å flytte informasjonen ut av opprinnelig kontekst via en integrasjon skaper endring. I sin essens er integrasjon et brudd på kravet i arkivloven § 5c.

- ▶ **d. informasjonen er tilgjengelig på en måte som ikke begrenser fremtidig nytte**

- ▶ Fra et tilgjengeliggjøringsperspektiv kan vi si at det å forholde seg til to informasjonskilder (datasett fra fagsystemet og uttrekk fra sakarkivintegrasjonen) begrenser fremtidig nytte fordi en alltid må sammenlikne innholdet fra to uttrekk. Det er ikke bare meningsløst, men også enormt ressurskrevende.

De tre grunnleggende trinnene i prosessen



103 Bevaring og kassasjon

- ▶ Forut for utforming av kravspesifikasjon skal systemeier og/eller systemforvalter ha foretatt en gjennomgang av informasjonsinnholdet som vil bli produsert i systemet.
- ▶ Denne gjennomgangen skal ta utgangspunkt i OKA - Oslo kommunes funksjonsbaserte klassifikasjon med tilhørende bevarings- og kassasjonsbestemmelser.

BK-vurderingen

- Første steg i prosessen er BK-vurdering av innholdet i løsningen
- Be først om møte med etat og leverandør
- Leverandør vil ofte beskrive funksjonalitet som innebærer overskriving og/eller sletting etter en viss tid; vi må sammenholde denne informasjonen med OKA og få dem til å rigge en løsning for uttrekk.
- Et fagsystem mottar noen ganger en feed fra et hendelsessystem. Vi må sjekke om hendelsessystemet eksporterer all bevaringsverdig informasjon til fagsystemet.
- Vi kan også oppleve at det er integrasjon mellom to fagsystemer der det første håndterer svarInn eller har en e-post-handle. Det er ikke noe problem, men konfigurasjonen må dokumenteres.

Case #1: Integrasjon fordi løsningen ikke kan produsere en journal

- Husk: en logg kan være en journal
- De aller fleste fagsystemer kan produsere en logg
- Avklare om dokumentkategoriene er innkomne eller organinterne (notater)
- Gå i dialog med leverandør via etat for å identifisere relevante metadata for å kunne produsere en logg
- Det er ikke annet formkrav til logg enn at den må inneholde registreringer av relevante metadata innenfor en gitt tidsperiode.

Logg i stedet for journal –eller en ny definisjon av obligatoriske metadata

- ▶ Arkivverket kjørte noen prosjekter i Regulatorisk sandkasse med DigDir og utvalgte virksomheter.
- ▶ En del konklusjoner der er viktige for oss:
«Dagens krav til journalføring og metadata er begrensende, og er til hinder for nytenkning. Det er behov for å tenke mer helhetlig rundt fremtidens metadata. Hvordan stilles krav som gir ønsket fleksibilitet, og hvordan kan metadata benyttes som verktøy for å ivareta arkiv og offentlighet».
- ▶ Og:
«Ut fra begrepene benyttet i Noark kan vi omformulere begrepsbruken i arkivforskriften § 10 om metadatakrav til å ha følgende, mer generiske og «åpne» definisjoner:
- ▶
 - Registreringstidspunkt
 - Identifikator sak og dokument i sak
 - Korrespondansepart
 - Tittel eller beskrivelse
 - Dato dokument (for referanse)
 - Knytning til forretningskontekst
 - Tidspunkt behandlingsoppfølging
 - Behandlingsoppfølging»

Arkivverket sa også noe mer som er essensielt:

- » «På alle stegene i prosessene i systemet dannes det logg. Det må vurderes om systemet inneholder tilstrekkelige metadata til at kravet om journalføring blir overholdt.
- » Hvis systemet produserer en logg som inneholder metadata tilsvarende det som er angitt i punktet ovenfor, kan vi vurdere det dithen at dette er tilstrekkelig. Det dreier seg altså om:
 - a) metadata som angir registreringstidspunkt,
 - b) identifikator sak og dokument i sak,
 - c) korrespondansepart (håpløst begrep, tenk en eller annen identifikator som identifiserer opphavet)
 - d) tittel eller beskrivelse,
 - e) dato for dokument,
 - f) knytning til forretningskontekst (kan ofte være metadata om tjeneste eller ytelse)
 - g) tidspunkt for behandlingsoppfølging eller informasjon om selve behandlingsoppfølgingen. (denne kan være vanskelig å løse men kan være kombinasjon av kode for tjeneste/ytelseskategori og timestamp).

Fagsystemer har som regel metadataverdier av følgende art:

- ▶ - Journalføringsdato: `documentDate`
- ▶ - Saks- og dokumentnummer: `pegaCaseID/DocumentID` (`archiveCaseId`)
- ▶ Organ/avsender/mottaker: `caseContacts` eller `recordCreator`
- ▶ - Opplysninger om sak, innhold og emne: `Title` (sak) `Title` (dokument)
- ▶ - Datering på dokumentet: `documentDate`
- ▶ **Alle disse oppfyller metadatakravene**

Formkrav journal/logg

- ♦ I Utkastet til ny arkivlov skriver departementet:
«Heimelen opnar òg for at det kan fastsetjast reglar om andre måtar å registrere naudsynte metadata på enn berre tradisjonell journalføring».
- ♦ Denne fleksibiliteten skal vi utnytte for alt den er verdt.
- ♦ Veldig mye av arbeidet vårt vil ligge i samarbeid og veiledning overfor leverandør for å få dem til å meisle ut en logg som inneholder minimumskravene våre og som kan produseres på forespørsel.
- ♦ Leverandør vil ofte si det er umulig å få løsningen til kontinuerlig å produsere en logg over registreringer. Vi poengterer at det kun skal være på forespørsel. Etat vil si at det hender svært sjeldent at de får forespørsler om kopi av journal.

Case #2: Integrasjon fordi løsningen ikke har brevmaler

- ▶ Dette scenariet krever gjennomgang med leverandør.
- ▶ Ofte erfarer vi at etat ikke har tatt gjennomgangen.
- ▶ Det skyldes automatikken i oppsett med Websak-integrasjon, fordi en opplever dette som «enklere».

Lås/frys – endring og sletting

APIA skal sikre låsing og frysing av registreringer og informasjonselementer

- ▶ APIA er teoretisk og abstrakt
- ▶ Hovedpoenget er å sikre funksjonalitet for administrator og begrensinger i rettigheter for bruker.
- ▶ Autentisitet skires ved at kontekst opprettholdes
- ▶ Konteksten er de «omkringliggende» metadata som beskriver hendelsen
- ▶ Disse metadataene er ofte i strukturert form i fagsystemer, dvs de plottes inn i faste parametere eller felter.
- ▶ De er fagspesifikke og går utenfor våre metadatakrav i arkivforskriften
- ▶ Å sikre at de lagres over tid hvis bevaringsverdige er essensielt (bruk OKA)

Uttrekk

- ▶ Uttrekk defineres av Digitalt depot
- ▶ Det finnes fire mulige oppsett for uttrekk:
- ▶ Noark
- ▶ ADDML
- ▶ CIM
- ▶ Siard

Noark-uttrekk

- ▶ Denne modellen er nedfelt i Noark-standarden, og vil hvis vi har et forvaltningssystem i digitalt depot og en egen tilgjengeliggjøringsløsning gi masse fordeler fordi vi gjennom Noark-uttrekket har en homogen struktur.
- ▶ Sagt veldig enkelt inneholder Noark-uttrekket de opprinnelige relasjonene og den opprinnelige strukturen.

ADDML

- ▶ Standard utviklet av Riksarkivet, XML-basert.
- ▶ Spesielt egnet for fagsystem, i utgangspunktet velegnet for alt som ikke er Noark.
- ▶ Fordrer en beskrivelse av informasjonsinnholdet forut for produksjon av uttrekket.
- ▶ Gir kjempegevinster fordi det har en bevisst tilnærming til informasjonsinnholdet i et fagsystem gjennom at vi identifiserer oppbyggingen, relasjoner mellom registreringer og objekter, strukturen osv.
- ▶ Testes gjennom verktøyet Arkade.

CIM-uttrekk

- ▶ CIM er DSB-CIM, et krise- og beredskapssystem, der Arkivverket i samarbeid med DSB har utviklet en modell som er veldig enkel men som kan utvides.
- ▶ Vi har ingen erfaring med det her i Oslo kommune, men vi har dialog med Arkivverket og vil kunne utvikle varianter for oss som oppfyller behovene våre.

Siard-uttrekk

- ▶ Proprietær standard utviklet av det sveitsiske riksarkivet. Dette er i realiteten en nød- eller brannslukningsløsning som i utgangspunktet bare burde benyttes for å berge et uttrekk, som siste utvei.
- ▶ Det er en databasedump i et format definert av Sveits, men for tilgjengeliggjøringen må en bygge struktur og nøste opp for å danne relasjoner. Dette er masse arbeid som fordrer minst et ukesverk pr uttrekk for å kunne tilgjengeliggjøre innholdet.