

Samfunnsdokumentasjon – hvilken dokumentasjon er det vi ikke trenger å bevare?

Presentasjon for:	Utvalg for bevaring av digitalt skapt dokumentasjon
Av:	Sigve Espeland, Interkommunalt Arkiv i Rogaland IKS

1. Innledende generelle betraktninger om problemstillingen

Samfunnet blir stadig mer digitalisert. Dette reiser selvsagt i økende grad spørsmål om hvilken type samfunnsdokumentasjon man skal bevare og hvilken man kan slette eller bare se bort fra. Og ikke minst hvilken form (teknisk format) har denne dokumentasjonen.

Hvis man skal bevare digital samfunnsdokumentasjon eller mene noe om hvilken digital samfunnsdokumentasjon man bør bevaringsvurdere, ja så må man også ha noen tanker om hva man over hode ikke behøver å bevare og hvordan arkivverdige digitalt materiale konkret kan bidra til å tåkelegge, tilsløre eller vanskeliggjøre det digitale bevaringsarbeidet. Og motsatt, hva kan man gjøre for å synliggjøre hva man vil bevaringsvurdere, f.eks. på et lovregulert fagområde.

Begrepet «å skille skit og kanel» blir nærliggende å ty til for å beskrive hva jeg egentlig søker å belyse i denne artikkelen. Det er «kanelen» vi vil bevare. «Skiten» må i denne sammenheng betraktes som digital informasjon «som ikke inngår i saksbehandling eller har verdi som dokumentasjon» på lang sikt for å bruke en beskrivelse fra dagens regelverk.

Vi bør altså diskutere gode bestemmelser, rammer og arbeidsmetoder som kan fjerne mye av det som tåkelegger og tilslører det digitale arkivmaterialet som egentlig skal/bør bevares. Og da må vi snakke om konkrete ting, konkrete digitale ting, som ikke inngår i en virksomhet eller en saksbehandling og som sådan ikke har verdi på lang sikt. Men som ofte er viktige når et system i det digitale samfunnet er i produksjon.

Mye av diskusjonen må derfor være omkring databaser. Databaser er ikke det som er mest «sexy» i det moderne digitale samfunnet, men de er der nesten overalt og spesielt i systemer som brukes til digital saksbehandling og arkivering. Databaser er og vil bli stadig viktigere i det moderne informasjonssamfunnet, de er motorene og informasjonslagerne i de systemene vi bruker i dag og trolig i fremtiden. Bevaringsvurderinger rundt databaser vil derfor bli stadig viktigere.

I denne artikkelen vil jeg derfor i hovedsak søke å sette søkelyset på hvilke informasjonstyper i moderne databaser vi bør kunne se bort fra – arkivbegrense, og med noen eksempler søke å vise hvor stort datavolum dette kan dreie seg om. Argumentasjonen vil med dette utgangspunktet også i noen grad kunne strekke seg ut over databasebegrensningen.

2. Arkivbegrensning og begrepet *helhetlig samfunnsdokumentasjon* i det digitaliserte samfunn

Begrepet *helhetlig samfunnsdokumentasjon* tar utgangspunkt i arkivlovens §1 hvor det i dag heter: «Føremålet med denne lova er å tryggja arkiv som har monaleg kulturelt eller forskingsmessig verdi eller som inneheld rettsleg eller viktig forvaltningsmessig dokumentasjon, slik at desse kan verta tekne vare på og gjorde tilgjengelege for ettertida».

På dette juridiske grunnlaget har begrepet *helhetlig samfunnsdokumentasjon* blitt definert som å:

«Sikre dokumentasjon om samfunnet som skal kunne brukes til å forstå samtid og fortid, og til å bevare og videreutvikle Norge som rettsstat og demokrati».

Og etter arkivmeldingen utvidet og tolket slik: «Det skal være mulig for ettertiden å forstå hele samfunnet, ikke bare offentlig sektor». Begrepet *helhetlig samfunnsdokumentasjon* blir i forlengelsen av dette summen av de to ovenfor nevnte tolkingene som bør materialiserer seg som et samfunnsoppdrag for arkivinstitusjoner.

Et Google søk på Internett med begrepet «informasjonsflom» gir svært mange treff, f.eks. følgende: «Informasjonsflommen er over oss og den vokser for hver dag. Ikke minst det digitale mediet har skapt uante muligheter for å skape, dublisere, distribuere og lagre informasjon. Dette er vår tids syndeflod. Stephen Hawking har påpekt at mengden av ekstern informasjon kan betraktes som et nytt stadium i menneskets evolusjon, som vil skape problemer fordi vår hjerne ikke henger med i informasjonsflommen». (<https://sites.google.com/site/informasjonsflommen/>)

Hvordan skal samfunnets arkivinstitusjoner sørge for en helhetlig samfunnsdokumentasjon i dette perspektivet? Hvordan skal vi klare å pukke ut representativ helhetlig samfunnsdokumentasjon i dette perspektivet. Hvordan skal vi «fiske» i denne digitale flommen?

Det er i alle tilfeller en ting som er helt klart. Vi må ta utgangspunkt i gamle erkjennelser knyttet til store oppgaver når vi som samfunn skal begynne dette bevaringsarbeidet som Stephen Hawking mener ligger utenfor den menneskelige hjernes kapasitet. Svaret på bevaringsutfordringen ligger med andre ord i det menneskelige fellesskap og teknologi. Hvis vi så tar utgangspunkt i det folkelige begrepet «å skille skit og kanel», så kan vi begynne med å definere hva vi overhode ikke er interessert i å bevare i informasjonsflommen, begrunne hvorfor og forsøke å kvantifisere volumet på denne typen arkivverdige informasjon og dokumentasjon. Det er dette som gjør det aktuelt å løfte opp begrepet *arkivbegrensning* som en sentral del av metoden for å få til å bevare helhetlig samfunnsdokumentasjon i en digital tidsalder. Arkivbegrensning som prinsipp og metode vil definere mye av «maskevidden» på den «trålen vi skal fiske med» i den digitale flommen. Sagt på en annen måte: Med et regelverk for arkivbegrensning som er bedre tilpasset en digital virkelighet vil det være enklere å se hva vi ikke trenger å bry oss om å bevare, først teknologisk så tematisk.

I det følgende vil jeg forsøke å konkretisere, eksemplifisere og kvantifisere hvor viktig arkivbegrensning kan være knyttet til bevaring av digitalt arkivmateriale.

3. Arven fra papirarkiv – begreper og bestemmelser

Ofte greit å ta utgangspunkt i det helt fundamentale når man snakker om bevaring av arkiv. Hva er et arkiv? Det er dokumenter, informasjon eller data som blir til som ledd i en virksomhet hos en arkivskaper og som samles som et resultat av denne virksomheten (også kalt enkeltarkiv). Men det kan også være et oppbevaringssted for arkiv, en organisatorisk enhet som utfører oppgaver knyttet til arkiv (arkivtjeneste). Begrepet brukes også om arkivdepotinstitusjoner.

Elektronisk arkiv blir i forlengelsen av dette: Elektroniske dokument med tilhørende metadata.

I forlengelse av denne definisjonen kan det kobles på 3 sentrale begreper når det gjelder bevaring av arkiv:

Bevaring – påbud om bevaring:

- Ut fra et hjemmelsgrunnlag ta vare på definerte arkiv og typer dokumentasjon.
- For langtidsoppbevaring i arkivdepot
- For senere oppslag, innsyn eller forskning.
- Hjeml: «Forskrift om offentlige arkiv» §15 (<https://lovdata.no/forskrift/2017-12-15-2105/§15>)

Kassasjon:

- Slette eller destruere originalt arkivmateriale eller dokumentasjon med utgangspunkt i et hjemmelsgrunnlag tilpasset lokale forhold.
- Helst innarbeidet i en arkivplan (bevaringsplan).
- Hjeml: «Forskrift om offentlige arkiv» §16 (<https://lovdata.no/forskrift/2017-12-15-2105/§16>)
- I forlengelse av dette har Arkivverket laget dokumentet «[Veiledning til bevarings- og kassasjonsbestemmelser for fylkeskommunale og kommunale arkiv skapt etter 1950](#)».

- Med utgangspunkt i Arkivverkets veileder har vi i Rogaland f.eks. laget dokumentet «[Bevarings- og kassasjonsplanmal for IKA Rogalands eiere](#)». Kommunene har deltatt aktivt i utformingen av dette dokumentet og det er utformet slik at det kan tas direkte inn i kommunens arkivplanen som en bevaringsplan.

For å kunne foreta en arkivfaglig god bevaringsvurdering og foreta kassasjon på en mest mulig oversiktlig og ryddig måte, er og bør *arkivbegrensning* være svært viktig. Man kan gjerne si at arkivbegrensning er måten å «skille skit og kanel» på. Arkivbegrensning kan også beskrives som arkivfagets «slankekur».

Arkivbegrensning:

- Vil si å holde utenfor eller fjerne dokumentasjon fra et arkiv (les også: system) «*som verken har inngått i saksbehandlingen til organet eller har verdi som dokumentasjon*».
- Hjeml: «Forskrift om offentlige arkiv» §14 (<https://lovdata.no/forskrift/2017-12-15-2105/§14>)
- Arkivbegrensning skal gjennomføres seinest ved avlevering heter det videre i §14, det vil i praksis si til arkivdepot.
- Dette betyr også at denne arkivfagets «slankekur» har 3 innsatspunkt:
 - a) I forbindelse med danningen av arkivet som en del av det som ofte kalles for «Post- og ekspedisjonsrutiner».
 - b) Ved bortsetting til nærarkiv
 - c) eller senest ved overføring til arkivdepot.
- a) For elektronisk arkivmateriale kan pkt. b) og c) svært ofte slås sammen fordi denne typen arkivmateriale krever en tidligere avlevering jfr. [§18 i «Forskrift om offentlige arkiv»](#) og [riksarkivarens forskrift §5-5 pkt. 5](#).

Monner arkivbegrensning?

Disse 3 begrepene og arbeidsmetoder utledet av dem har vokst frem i papirarkivtradisjonen. Ved IKA Rogaland har vi f.eks. erfaring for at faglig god arkivbegrensning, vi kaller det for «rensing» av papirarkiv, sammen med bedre pakking av materialet skapt i perioden 1950 til ca. 1990 kan i volum redusere et arkiv med mellom 30 og 40%. Vi har foretatt flere slike «rensinger» for våre eiere før overføring til arkivdepot.

Arkivbegrensning monner med andre ord. Og da blir spørsmålet: Kan regelverket utvikles og tilpasses slik at en slik metodikk også kan anvendes på digitalt skapt arkivmateriale. I det følgende kommer jeg til å legge hovedvekten på arkivbegrensning. For etter god arkivbegrensning trer det arkivmaterialet som skal bevaringsvurderes tydeligere frem og det blir enklere å si hva som skal bevares og hva som kan kasseres.

4. Kan regelverk og prinsipper om arkivbegrensning anvendes på digitalt skapt materiale?

- JA er langt på vei det korte svaret.
- Men det må nok gjøres noen tilpassinger og tilføyelser.
- Arkivloven med forskrift og regelverk utledet av dette er ofte generelt og overordnet.
- Lokale bevarings- og kassasjonsbestemmelser kan derimot være mer konkrete.
- Slikt lokale regelverk er ofte tematisk og prosessinnrettet (jfr. pkt 3). Og som sådan kan de langt på vei også brukes på digitalt skapt dokumentasjon og arkivmateriale. Når slike lokale regelverk er utformet med prosessinnrettet og tematisk arkivering i bakhodet er de på mange måter uavhengig av mediet som bærer dokumentasjonen.
- Når det overordnede regelverket blir forbedret og mer nyansert når det gjelder e-arkiv, kan vi fort tenke oss at det blir aktuelt å bevaringsvurdere hele systemer, spesielt fagsystem, etter slike lokale bevarings- og kassasjonsplaner.
- Fordi et fagsystem ofte er tematisk avgrenset til et fagområde, f.eks. barnevern, PP-tjeneste eller pleie- og omsorgssektor.

- Og i slike sektorer kan man ofte med utgangspunkt i særlov og forskrift definere hvilke typer dokumentasjon et fagsystem inneholder og deretter bevaringsvurdere hele systemet og beskrive dette i arkivplanen.
- I SAMDOK-rapport 2017, delprosjekt kommunale arkiv, oppgave 2: Metodikk for bevaring fra kommunale fagsystem (Pleie og omsorgssektoren), avdekkes det i prosjektrapporten at det er «mellom 40-50 saks- og dokumenttyper i pleie- og omsorgssystemene (PLO-systemene)». Og at det ikke bør «være vanskelig å få på plass bevaringsbestemmelser (BK-bestemmelser) for dette avgrensa antallet saks- og dokumenttyper». Det bør altså være overkommelig å gjøre bevaringsvurdering av hele fagsystem. (<https://www.arkivverket.no/arkivutvikling/utviklingsprosjekter/avsluttede-prosjekter/samdok/samdok-delprosjekt-kommunale-arkiver>)

5. Hvorfor er arkivbegrensning så viktig?

- Fordi det holder utenfor og fjerner informasjon og dokumentasjon fra arkiv som ikke inngår i saksbehandling eller har verdi som dokumentasjon.
- Dette gjør arkiv, uavhengig av medium:
 - Mindre i volum
 - Mer oversiktlig
 - Enklere å bruke i daglig arkivdanning
 - Enklere å bevaringsvurdere
 - Billigere å langtidsbevare
 - Enklere å bruke ved innsyn i arkivdepot

6. Digital dokumentasjon som bør kunne arkivbegrenses - noen forslag

Hovedtyngden av de de forslagene til informasjon- og dokumentasjon typer som bør kunne arkivbegrenses med hjemmel i fremtidig regelverk er knyttet til systemer som i dag bygger på relasjonsdatabaser.

Nr.:	Type:	Beskrivelse/forklaring/begrunnelse:
1	Schema som er knyttet til men som ikke inngår direkte i databasestrukturen til det bevaringsverdige systemet.	<p>En database er en samling av strukturerte data. Dataene i en database er vanligvis organisert på en slik måte at informasjonen er lett tilgjengelig. Et schema (skjema) er i utgangspunktet en formell beskrivelse av hvordan en database er dannet og hvor alt ligger i den. Sagt på en annen måte, hvordan databasen er utformet, hvor data er lokalisert og hvordan data er relatert internt i den.</p> <p>I tilknytning til de svært vanlige og mye utbredte SQL-databasene er et schema en liste av logiske strukturer og data. En <u>databasebruker</u> eier et schema som ofte har det samme navnet som databasforvalteren. Et eksempel: Ved mange kommunale pedagogisk psykologiske tjenester (PP-tjeneste) har det ofte blitt brukt et system som har et schema med brukernavn PPI. (Dette har også gitt navn til systemet i dette tilfellet: PPI.)</p> <p>I et databasebasert system kan det kan det være flere schema. Noen er knyttet til den teknologiske databaseplattformen systemet kjører på, andre til f.eks. det fagsystemet som kjøres på plattformen (og som ofte skal bevares).</p> <p>Hvis et uttrekk for bevaring av f.eks. et fagsystem gjøres med en databasebruker som er en systemadministrator som har tilgang til alt som ligger på databaseplattformen, så får man ofte med seg flere schema enn det som skal bevares. Dette kan være databasebrukere som: ROOT (MySQL), SA (MSSQL), SYSADMIN (Oracle), m.fl. Det bør altså lages et utgangspunkt i regelverket for arkivbegrensning som gir tydelig hjemmel for å arkivbegrenses schema som ikke inngår direkte i databasestrukturen til systemet som skal bevares.</p>
2	Indekser/databaseindekser	<p>En databaseindeks er en datastruktur som forbedrer hastigheten på datainnhentingsoperasjoner på en databasetabell på bekostning av ekstra skrive- og lagringsplass for å opprettholde indeksdatastrukturen.</p>
3	Views (lagrede prosedyrer)	<ul style="list-style-type: none"> • Slike views (visninger) er et konsept innen databaseteori hvor man oppretter en <i>virtuell tabell</i> basert på resultatsettet fra en

		<p>databasespørring. Visninger utvider funksjonaliteten man kan få fra vanlige grunnleggende tabeller på flere måter: De kan representere en delmengde av dataene som finnes i en eller flere tabeller.</p> <ul style="list-style-type: none"> • Views setter ofte sammen innhold fra mange tabeller på en måte som gjør innholdet forståelig for brukeren. Views bør i mange, men ikke alle, tilfeller bevares som dokumentasjon av systemet. • I den grad views fungerer som automatiserte hjelpetabeller som setter sammen data fra ulike tabeller og migrerer resultatet inn i en annen tabell i databasen hvor resultatet lagres bør de kunne arkivbegrenses. • Views som genererer rapporter eller gir visninger som har verdi som dokumentasjon eller inngår i saksbehandlingen bør bevares, f.eks. knyttet til visninger knyttet mot registrerte dokument, presedens, klage, avvik, m.m.
4	<p>Database logger knyttet til:</p> <ul style="list-style-type: none"> • Spørringer som har med systembruk å gjøre. • Systemkjøring • (Annet?) 	<p>Databaselogging lagrer en oversikt over endringer eller bruk av tabeller eller felt i en egen databaseloggingstabell. I et system kan logging f.eks. inkludere: Opprette en revisjonsregistrering av endringer i spesifikke tabeller som inneholder sensitiv informasjon knyttet til kjøring av systemet. Loggtabeller kan også gi oversikt over hvem og hvordan vedkommende har brukt systemet f.eks. knyttet til søk eller spørringer, dvs. logging med tanke på sikkerhet, kontroll og innsyn ut fra et "need to know" prinsipp. Flere av disse loggene kan bli svært store og de kan være svært viktige når et system er i produksjon, men ikke nødvendigvis i forbindelse med langtidsoppbevaring i arkivdepot etter at systemet er tatt ut av produksjon. Da kan det være på sin plass med en hjemmel for å kunne arkivbegrense databaselogger knyttet til:</p> <ul style="list-style-type: none"> • <i>Spørringer som har med bruk av systemet å gjøre</i> – Dette vil i praksis si tabeller som logger alle SQL-spørringene eller SQL baserte søk som kjøres i og mot systemets database. Det kan f.eks. i et pleie- og omsorgssystem være logging av alle standardiserte spørringer eller søk som helsepersonell gjør for å hente frem informasjon fra systemets pasientjournal. I slike logger ligger det ofte informasjon om hvem som har gjort søket og hvilken SQL-kommando som er kjørt mot databasen. Slike loggtabeller kan bli svært store og de har i et mer langsiktig bevaringsperspektiv liten verdi. De bør derfor kunne arkivbegrenses ut fra en hjemmel som vi i dag ikke har. • <i>Systemkjøring</i> – Logger knyttet til: Versjons oversikt, oversikt over oppdateringer, systemfeil og feilretting bør bevares. Øvrige logger bør det være mulig å arkivbegrense. • (Annet?)
5	<p>EDI-data og EDI-logger</p>	<ul style="list-style-type: none"> • EDI - Electronic Data Interchange eller "elektronisk datautveksling" på norsk. Samlebegrep for utveksling av data (integrasjoner). Det er altså her snakk om elektronisk utveksling av data mellom datasystemer på et dataformat partene enes om. • EDI er altså en del av et datasystemet sin innebygde funksjonalitet for elektronisk datautveksling som har til hensikt å sørge for sikker kommunikasjon mellom f.eks. en lege og et laboratorium eller en almenlege og spesialhelsetjenesten. I f.eks. helsesektoren i Norge skjer dette oftest via Norsk Helsenett. • Slik vi vurderer det ved Interkommunalt Arkiv i Rogaland IKS (IKA) er dette dokumentasjon som er viktig å ha på plass så lenge systemet er i produksjon for å ivareta nødvendig sikkerhet, f.eks. i forbindelse med helseopplysninger. Men knyttet til langtidsbevaring i arkivdepot er informasjonen ikke bevaringsverdig. Slike tabelldata er <i>knyttet til transport</i> av melding/dokument og kan ikke direkte knyttes til saksbehandling/klientbehandling. Overføringsmeldinger av denne typen kan derfor med en viss berettigelse arkivbegrenses iht. §14 i arkivforskrifta hvor det heter: "Dokument som verken har inngått i saksbehandlingen til organet eller har verdi som dokumentasjon, skal holdast utanfor eller fjernast frå arkivet ved arkivavgrensing. Arkivavgrensing skal gjennomførast seinast før avlevering [til arkivdepot]". Med dette hjemelsgrunnlaget bør EDI-tabeller kunne fjernes fra f.eks. et siard-uttrekk før det lages en arkivdepotversjon.

		<p>Dette er altså ikke bevaringsverdig dokumentasjon i arkivdepot slik vi har vurdert det ved IKA. Det dreier dessuten svært ofte om pakket, kodet og kryptert informasjon som er sendt fra eller mottatt i systemet og hvor informasjonen høyst sannsynlig ligger utpakket og dekryptert i en annen tabell i systemet. §14 i arkivforskrifta er ikke særlig presis for å foreta arkivbegrensning av denne typen data, men det er den beste bestemmelsen vi har i dag. Arkivbegrensning av EDI data etter §14 i arkivforskrifta må også holdes opp mot §5-16 (2) i Riksarkivarens forskrift hvor det heter: "Sikkerhetsfunksjoner som er mulige i arkivdokumenter, her under passord og kryptering, skal være deaktivert ved avlevering eller deponering til Arkivverket. Arkivdokumenter kan bare overføres i kryptert form når det er gjort særskilt avtale om kryptering. En slik avtale skal omfatte regler om disponeringen av nøkler for dekryptering". De fleste kommunale arkivinstitusjoner følger også dette regelverket.</p> <ul style="list-style-type: none"> • Gitt denne begrunnelsen og forklaringen bør det altså komme på plass muligheter for å arkivbegrense EDI-data og -logge med et mer presist hjemmelsgrunnlag.
6	Produksjonsversjoner av dokument	<ul style="list-style-type: none"> • Produksjonsversjoner (inkl. utkast og arbeidskopier) av dokument som ligger i eller er knyttet til systemet bør kunne arkivbegrenses forutsatt at disse foreligger i arkivversjon i arkivpakken og har en arkivfaglig god kobling til tilhørende metadata. • (Er først og fremst knyttet til arkivpakker med <i>noark-struktur</i>, men er også aktuelt for fagsystem når tilhørende dokument f.eks. ligger som BLOB-element inne i databasen.)
7	Arkivbegrensning av hele system eller deler av det basert på <i>integrasjon</i> med andre system.	<ul style="list-style-type: none"> • Basert på kunnskap om hvilke typer dokumentasjon, informasjon og metadata et system inneholder, gjerne et begrenset fagsystem, bør man i enkelt tilfeller kunne konkludere med at hele systemet (eller store deler av det?) kan arkivbegrenses. Dette fordi data i systemet ikke inngår i saksbehandling eller har verdi som dokumentasjon for virksomheten som bruker det. • ELLER fordi systemet genererer dokumentasjon, informasjon og metadata som isolert sett er bevaringsverdige, men som via integrasjon lagres i et annet system. Her kan det f.eks. i en kommune være snakk om integrasjon mot et annet kommunalt system eller integrasjon mot et statlig system. Det finnes mange slike eksempler. • Dette MÅ i begge de aktuelle tilfellene dokumenteres i forvaltningsorganets eller virksomhetens arkivplan.
8	(Annet?)	(Andre forhold som jeg ikke har erfaring med eller ikke kommer på)

7. Hvor store volumer kan dette dreie seg om – noen eksempler.

Ved IKA Rogaland har vi ofte stilt oss spørsmål om hvor arkivverdig deler av det elektroniske arkivmaterialet vi mottar for oppbevaring i arkivdepot egentlig er. Vi har ofte lurt på om ikke en god del av arkivmaterialet burde ha vært arkivbegrenset før det ble gjort såkalt tabelluttrekk etter f.eks. siard-metoden.

I 2 tilfeller har vi brukt det utilstrekkelige regelverket vi har i dag etter beste evne fordi behovet for å foreta arkivbegrensning har vært påtakelig.

Tilfelle 1: *Visma Profil - Sandnes kommunen*

Generelt:

- Fagsystem brukt i pleie- og omsorgsektoren i Sandnes kommune. Sandnes er en stor kommune i norsk sammenheng med over 81000 innbyggere.

Om arkivbegrensningen i dette tilfellet:

- I denne konkrete bevaringssaken for Sandnes kommune ble det i alt laget 3 siard-uttrekk før vi kom i mål over en perioden på over halvannet år.
- Alle uttrekkene ble produsert med programmet SpectralCore FullConvert (SCFC).

- IKA måtte be Sandnes kommune om å lage nytt uttrekk 2 ganger.
- Det var først det 3. uttrekket vi ved IKA klarte å håndtere med den datamaskin parken og de verktøyene vi har tilgjengelig.
- I det første uttrekket lå dokumentene (CLOB- og BLOB-feltene) inne i sine respektive systemtabeller i den siard-fila som ble produsert.
- IKA klarte ikke å gjøre noe med denne fila.
- Ordnings- og beskrivelsesprogrammet Documaster Decom klarte ikke å lese den på de kraftigste datamaskinene IKA har.
- Vi har også fått vite at det tok nesten to uker å produsere det første siard-uttrekket i Sandnes.
- I det 2. uttrekket la Sandnes kommune helt bevisst etter råd fra IKA dokumentene utenfor siard-fila.
- Her klarte vi ved IKA å gjøre noe.
- Vi klarte å lese siard-fila opp i Documaster Decom, men konverteringen av dokumentene havarete.
- Vi fant videre ut at problemet var knyttet til en bestemt tabell som heter OM_DB_TRANS. En tabell som logger transaksjoner i databasen.
- Denne tabellen logger alle SQL-spørringene som kjøres i databasen.
- Vi fant ut at denne inneholdt over **65 millioner records** og i overkant av **40 millioner CLOB-element**.
- IKA foretok derfor en bevaringsvurdering etter §14 i arkivforskrifta av tabellen OM_DB_TRANS og kom til at de standardiserte SQL-spørringene som her ble logget ikke er bevaringsverdige på lang sikt.
- Ved IKA fant vi også ut at i Sandnes kommune var uttrekket produsert med en bruker som hadde admin-rettigheter mot databasen. Denne brukeren tok også med seg 2 schema som ikke var nødvendige å ta med i uttrekket.
- Ved IKA klarte vi nå å lese dette 2. uttrekket fra Sandnes kommune opp på en MySQL database plattform når vi kun valgte tabeller fra selve Visma Profil schema.
- Vi plukket altså ut kun tabeller som vi viste var sentrale i Profil databasen og vi utelot selvsagt tabellen OM_DB_TRANS.
- For å lage en god "plukklister" for å finne ut hvilke tabeller som skulle med fra Visma Profil schema til Sandnes kommune sammenliknet vi tabellstrukturen i uttrekk fra to andre kommuner - Klepp og Forsand.
- Dette ble gjort på grunnlag av JASON-filer laget i Documaster Decom fra uttrekkene fra disse to kommunene.
- Denne "plukklisten" førte til at det gikk mye greiere å produsere det 3. uttrekket i Sandnes - mellom 4 og 5 timer er vi blitt fortalt.
- Det 3. uttrekket ble også mye enklere å håndtere i Documaster Decom ved IKA.
- Nå var det bare ca **1,16 millioner** CLOB- og BLOB-element i uttrekket mot omkring **42 millioner** ved de to første uttrekkene.
- Ca 1,16 millioner var og er også et mye mer realistisk tall.
- Ved IKA brukte vi 34 døgn på å konvertere disse BLOB- og CLOB-elementene til arkivformat, men det er en annen historie som ikke har noe med arkivbegrensningen å gjøre.

Tilfelle 2: Infodoc - Tysvær kommunen

Generelt:

- Offisielt heter systemet Infodoc Plenario, men i de fleste kommuner kalles det bare Infodoc. Systemet er et EPJ system brukt ved allmennelegekontor, spesialistpraksis, helsestasjon og kommunehelsetjeneste. I Tysvær kommune ble systemet brukt ved et nedlagt kommunalt legekontor. Tysvær er en kommune med ca 11000 innbyggere i Nord-Rogaland.

Om arkivbegrensningen i dette tilfellet:

- IKA begynte å arbeide med dette uttrekket InfoDoc uttrekket fra Tysvær kommune i april 2022. Den 14. april startet vi en konverteringsprosess i Documaster Decom og LibreOffice.

Denne prosessen var ferdig 4. mai 2022. Konverteringsprosessen gikk altså i flere uker og endte opp med følgende resultat:

- IN_PROGRESS: **0**
 - FAILED: **113**
 - SUCCESSFUL: **55558**
 - IDLE: **0**
 - NOT_CONVERTED: **1061790**
-
- Ved IKA fant raskt ut at svært mange av filene som ikke ble konvertert var kryptert, komprimert (zip) og passordbeskyttet. Vi så derfor raskt behov for å finne ut mer om hva som lå bak disse tallene. Det ble derfor gjort en bevaringsvurdering knyttet til de tabellene som inneholdt BLOB- og CLOB-elementer i dette InfoDoc uttrekket. Essensen i denne er presentert i pkt. 5 EDI-data og EDI-logger ovenfor. Vi fant ut at svært mange av de BLOB- og CLOB-elementene som Documaster Decom og LibreOffice ikke konverterte var knyttet til såkalt elektronisk datautveksling (EDI) mye basert på bl.a. det som kalles Advanced Message Queuing Protocol (AMQP) som er en åpen standard for meldingsorientert mellomvare. Hele EDI forholdet går på sikring av automatisert kommunikasjon inn og ut av fagsystemet. Denne teknologien har en tendens til å logge svært mye i systemet og bli svært volumiøs, og den har ikke direkte med saksbehandling eller klientbehandling å gjøre. Det er funksjonalitet som transporterer informasjon på en sikker måte inn og ut av systemet og logger dette. Ved IKA vurderer vi dette som dokumentasjon som kan arkivbegrenses i forbindelse med overføring til arkivdepot.
 - På grunnlag av denne bevaringsvurderingen konkluderte vi at i arkivdepotpakken (AIP) trenger ikke EDI elementene å tas med. De ble derimot liggende i arkivpakken IKA mottok fra Tysvær kommune (SIP). IKA laget derfor en ny siard-file med SpectralCore FullConvert hvor alle EDI tabellene ble sløyfet. Denne ble så behandlet i Documaster Decom og utgjør AIP.
 - Behandlingen av den nye siard-fila i Documaster Decom viser at volumet av BLOB- og CLOB-elementer knyttet til EDI funksjonen til systemet er svært volumiøs. Følgende tall illustrerer dette. I det opprinnelige uttrekket som IKA mottok fra Tysvær kommune var det **1117461** BLOB- og CLOB-elementer. Etter å ha fjernet alle EDI-tabellene fra uttrekket og laget ei ny siard-file startet vi en ny konverteringsprosess i Documaster Decom og LibreOffice. Denne viser at antallet BLOB- og CLOB-elementer er redusert til **261001**. Dette betyr at i dette konkrete uttrekket er **856460** BLOB- og CLOB-elementer arkivbegrenset med utgangspunkt i den ovenfor nevnt bevaringsvurderingen.

8. Fremtidig regelverk – tanker om formuleringer

Jeg vet jo ikke hvordan fremtidens regelverk for arkivbegrensning kommer til å bli, men jeg ser for meg at §14 i arkivforskriften på ett eller annet vis må omformuleres og utvides.

Dagens bestemmelser om arkivbegrensning springer ut av en papirbasert arkivtradisjon og lyder slik:

§ 14. Arkivavgrensing

Dokument som verken har inngått i saksbehandlingen til organet eller har verdi som dokumentasjon, skal holdes utenfor eller fjernes fra arkivet ved arkivavgrensing.

Arkivavgrensing skal gjennomføres seinest før levering.

Følgende dokument skal alltid holdes utenfor arkivet:

- a. trykksaker, rundskriv og annen mangfoldig og allment tilgjengelig materiale. Det organet som har utferd slik materiale, skal likevel arkivere ett eksemplar og de nødvendige forarbeida.
- b. konsept, kladdar, ekstra kopier, interne meldingar o.l. som ikkje har verdi som dokumentasjon. Konsept som har noteringar og påteikningar som er nødvendige for å forstå saka og den samanhengen ho står i, skal arkiverast. Det same gjeld kopi av alle utgåande dokument.

Innledningen til paragrafen kan i fremtiden f.eks. lyde slik:

"Dokument [og informasjonstyper] som verken har inngått i saksbehandlingen til organet eller har verdi som dokumentasjon, skal holdes utenfor eller fjernes fra arkivet ved arkivavgrensing".

For digitalt skapt arkivmateriale ser jeg i dag for meg at bestemmelsene i §14 f.eks. kan utvides slik:

c.	Schema som er knyttet til men som ikke inngår direkte i databasestrukturen til det bevaringsverdige systemet.
d.	Indekser/databaseindekser
e.	Views (lagrede prosedyrer)
f.	Database logger knyttet til: <ul style="list-style-type: none">• Spørringer som har med bruk av/i systemet å gjøre• Systemkjøring
g.	EDI-data og EDI-logger
h.	Produksjonsversjoner av dokument
i.	Arkivbegrensing av hele system eller deler av det basert på integrasjon med andre system.

*Arkivenes Hus i Stavanger den 16. november 2022
Sigve Espeland*