

Digitalt arkivdepot

Forprosjekt i Byarkivet

Innholdsfortegnelse

1 - Innledning.....	3
2 - Gevinster og interessenter	5
2.1 Kommunen som helhet og kommunens virksomheter	5
2.2 Kommunens innbyggere og ulike brukergrupper	5
2.3 Gjenbruk og verdiskaping med utgangspunkt i offentlig skapte data	5
3 - Overordnede krav og behov.....	7
3.1 Lagringsløsningen	7
3.2 Innholdsforvaltning og forvaltningsverktøy	9
4 - Nåsituasjonen.....	11
4.1 Forvaltning av digitalt arkivmateriale.....	11
4.2 ROS-analyse	12
Risiko 1: Manglende mottak av arkivmateriale til digitalt depot.....	13
Risiko 2: Informasjonstap i digitalt depot.....	13
Risiko 3: Lav tilgjengelighet for arkivmateriale i digitalt depot.....	14
5 – Muligheter i markedet og valgte løsninger i offentlig sektor	15
5.1 Depotstyringssystemer.....	15
5.1.1 Archivematica	15
5.1.2 ESSArch	15
5.1.3 Preservica	16
5.2 Lagringsløsning	17
5.3 Kommunearkivinstitusjonenes digitale ressurscenter	17
6 - Elektronisk arkivdepot – alternativer	19
7 – Etterslep i kommunens virksomheter	20
7.1 Fagsystemer.....	20
7.2 Noark-baserte systemer	22
7.3 Mottak av uttrekk – oppgaver og ressursbruk	22
7.4 Prosjekt for innsats med etterslep	23
8 – Forslag til videre arbeid og samlet kostnadsestimat	25
9 – Referanser	27

1 - Innledning

Elektronisk arkivmateriale som skal bevares over tid, krever aktivt vedlikehold og må overvåkes, sikkerhetskopieres og konverteres til nye lagringsenheter og formater med få års mellomrom. Mangel på nødvendig vedlikehold vil føre til informasjonstap, og materialet vil miste viktige egenskaper som integritet, autentisitet og lesbarhet. Etter hvert vil det gå tapt.

Riksrevisjonens undersøkelse fra 2010¹ av arbeidet med å sikre og tilgjengeliggjøre arkivene i kommunal sektor viser at mange kommuner hadde store utfordringer med å sikre og gjøre tilgjengelig bevaringsverdig dokumentasjon. Svakheter ved elektronisk arkiv ga høy risiko for dokumentasjonstap av stor rettslig, forvaltningsmessig og historisk betydning. Riksrevisjonen slo fast at uerstattelig materiale har gått tapt og står i fare for å gå tapt, og at dette svekker innbyggernes rettssikkerhet. Funnene støttes i Arkivverkets rapport *Digitalt skapt materiale i kommunal sektor 1985-2010*². Det understrekes også at elektronisk arkivmateriale vil fortsette å gå tapt så lenge kommunene og fylkeskommunene i liten grad er i stand til å gjennomføre avlevering, vedlikehold og langtidslagring.

Både Oslo kommunes innbyggerdialog og saksbehandling i kommunens virksomheter foregår i større og større grad elektronisk, og vedtak og hendelser dokumenteres etter hvert kun digitalt. Utviklingen forsterkes av kommunens satsing på digitalisering. For at dette skal være forsvarlig, er det nødvendig at kommunen har en løsning for å kunne bevare og holde tilgjengelig arkivverdig dokumentasjon over tid - et såkalt digitalt arkivdepot. Dette i henhold til både lov og forskrift, Oslo byarkivs mandat og kommunens arkivinstruks. Byarkivet er i følge Oslo kommunes arkivinstruks og bystyrevedtak 907/99 Oslo kommunes arkivdepot og arkivfaglige etat. Akkurat som det er Byarkivets oppgave å overta papirmateriale og bevare det over tid, er det Byarkivets oppgave å ta imot og bevare elektronisk skapt materiale over tid. Byarkivet forvalter allerede digitalt skapt arkivmateriale i sitt depot. Det er likevel ikke på langt nær etablert noe egentlig digitalt depot med infrastruktur og metode i henhold til standarder. Man kan ikke si at dagens bevaring av digitalt arkivmateriale i Byarkivets forvaltning innebærer forsvarlig praksis, eller at kommunen gjennom nåsituasjonen tilfredsstiller krav i lov og forskrift til digitalt arkivdepot.

Innsatsen med å få overført digitalt arkivmateriale fra virksomhetene til Byarkivet har videre vært alt for svak gjennom mange år, og restansene er store. Både Riksarkivarens tilsyn med kommunen, innføring av nytt sak/arkiv-system og generasjonsskifte i kommunens fagsystemportefølje vil øke mengden digitalt materiale som overføres til Byarkivet. Det forsterker behovet for så raskt som mulig å få på plass standarder, infrastruktur og rutiner som sikrer forsvarlig forvaltning.

Denne utredningen skal danne et første grunnlag for en prosess med etablering av digitalt depot. Den gir en beskrivelse av nåsituasjonen, inneholder en overordnet ROS-analyse knyttet til

¹ Dokument 3:13 (2009–2010) Riksrevisjonens undersøkelse av arbeidet med å sikre og tilgjengeliggjøre arkivene i kommunal sektor (2010)

² Digitalt skapt materiale i kommunal sektor 1985 til 2010 – kartlegging (2014)

nåsituasjonen, presenterer behov og overordnede krav til et digitalt arkivdepot og skisserer ulike mulige løsninger. Rapporten beskriver kort etterslepet på overføring av digitalt materiale fra virksomhetene, og kommer med forslag til videre arbeid. Det gis også kostnadsoverslag for etablering av infrastruktur og for innsats med etterslep de nærmeste årene.

Utredningen skal danne grunnlag for videre arbeid med anskaffelse av infrastruktur og innsats med etterslep, og for forankring og finansiering av depotet i kommunen.

Det er svært få miljøer i Norge som har digitalt arkivdepot i henhold til regelverk og standarder, og etablering må fortsatt regnes som utviklingsarbeid. Riksarkivaren støttet dette utredningsarbeidet med utviklingsmidler i 2015.

I Kulturretatens tildelingsbrev for 2017 står det at etaten skal arbeide med utredning på arkivområdet, blant annet digitalt depot. Rapporten er et svar på dette oppdraget.

Utredningen har vært gjennomført av en prosjektgruppe i Byarkivet. Gruppen har bestått av Robert Kalleberg (leder), og spesialkonsulentene Mona Thomassen og Petter Svendsen. Byarkivets ledergruppe har vært styringsgruppe for arbeidet. Det har vært dialog med og innhenting av dokumentasjon fra Riksarkivet, Trondheim byarkiv, Bergen byarkiv og Kommunearchivinstusjonenes digitale ressurscenter (KDRS) i forbindelse med arbeidet.

2 - Gevinster og interessenter

Et digitalt depot skal først og fremst sørge for at elektronisk informasjon som dokumenterer kommunal saksbehandling og informasjon av viktig rettslig, kulturell, historisk og/eller annen forskningsmessig verdi, blir bevart for ettertiden i henhold til gjeldende lovverk og standarder. Dette fører til etterlevelse av de gjeldende lov og forskrift (Arkivlova, Offentlighetslova og Forvaltningsloven).

2.1 Kommunen som helhet og kommunens virksomheter

Kommunens virksomheter er viktige behovshavere for det digitale depotet. Når arkivmateriale ikke lenger er i daglig bruk i kommunens virksomheter, skal det overføres til Byarkivet for bevaring og tilgjengeliggjøring over tid. Kommunen er forpliktet etter lov og forskrift til å ha et standardisert depot. Dette gjelder digitalt arkivmateriale så vel som papirbasert arkiv. Langtidsforvaltning av digitalt arkivmateriale krever spesialisert kompetanse, verktøy og infrastruktur. Det er rasjonelt og følger dessuten av lov og forskrift at slik forvaltning skal foregå i et sentralt depot. Ved overgang til nye systemer, kan virksomhetene da overlate forvaltning av materiale fra systemer under avvikling til depotet. Systemer som har gått ut av daglig bruk i virksomhetene kan avvikles og virksomheten behøver ikke selv sørge for ressurser, kompetanse, infrastruktur og verktøy til langtidsforvaltning.

Ettersom dataene sentraliseres, kan også interne arbeidsprosesser standardiseres, forenkles og effektiviseres. Dette vil gi bedre struktur og kvalitet på historisk informasjon og arkiverte data, og det vil være enklere å gjøre dem tilgjengelige for bruk og som den enorme ressursen de er. En økning i kunnskapsnivået hos systemeiere og de enkelte virksomhetene rundt avlevering vil føre til minimering av informasjonstap.

2.2 Kommunens innbyggere og ulike brukergrupper

En viktig oppgave for et arkivdepot er å gjøre arkivmateriale i depotet tilgjengelig for ulike brukergrupper. Det papirbaserte materialet i Byarkivets forvaltning gjøres hvert år tilgjengelig for flere tusen brukere på ulike måter. Det gis partsinnsyn etter skriftlig henvendelse, og materialet hentes frem for blant annet hobby-, studie- og forskningsformål på Byarkivets lesesal. Ett viktig formål er dokumentasjon av den enkeltes rettigheter, for eksempel i forbindelse med skolegang, pensjonsutmåling, og barnevern.

Et eksempel til illustrasjon, er arkivmateriale fra barnevern, som er rettighetsdokumentasjon unntatt offentlighet der tilgangen endres over tid. I tillegg skal arkivmateriale dokumentere hvordan barnevernet utfører sitt arbeid. På sikt vil det være av interesse for forskere og historikere å benytte materialet for å si noe om hvilken samfunnsmessig funksjon barnevernet har hatt. Dersom det er mulig å behandle materialet statistisk, har det forskningsmessig verdi allerede før klausuleringstiden for personsensitivt materiale har gått ut.

Når tilsvarende materiale nå lagres og skapes digitalt, er det stort potensial for tilgjengelighet og bruk for innbyggerne og ulike brukergrupper ved å samle materialet i et digitalt depot og gi samlet, digital tilgang.

2.3 Gjenbruk og verdiskaping med utgangspunkt i offentlig skapte data

Den samlede mengden av kommunens digitale arkivdata utgjør en stor ressurs. Det å samle og forvalte disse i et sentralt digitalt depot med en tilfredsstillende infrastruktur, innebærer også stort

potensial for gjenbruk av offentlig skapt og finansiert dokumentasjon ved at andre aktører kan utvikle tjenester og løsninger som baserer seg på dokumentasjonen.

3 - Overordnede krav og behov

Det er visse prinsipper som ligger til grunn når Byarkivet nå utreder digitalt arkivdepot for Oslo kommune. I følge forskrift av 11.12.1998 nr. 1193 om offentlige arkiv (arkivforskrifta), § 5-1-3, skal kommuner og fylkeskommuner opprette egne ordninger for arkivdepot. Formålet med et arkivdepot for digitalt arkivmateriale er i hovedtrekk å bevare arkivmaterialet på en måte som sikrer mot informasjonstap og utilsiktet tilgang, samt å opprettholde potensialet for bruk. Siden bevaring av digitalt arkivmateriale krever aktiv bevaring ved at formater og strukturer konverteres ved behov og at en hele tiden sikrer at materialet ligger på lagringsteknologi som ikke er utdatert eller i ferd med å forvitte, er det avgjørende å velge løsninger og metodikk som legger til rette for nødvendige skifter i de komponentene arkivdepotet består av.

Byarkivet må ha et robust digitalt arkivdepot som er tilrettelagt for økte datamengder og økt behov for arkivvedlikehold, gjenfinning og tilgjengeliggjøring. Robustheten ligger blant annet i at en minimerer risikoen for tap av data og for utilsiktede og tilsiktede endringer i arkivmaterialet, og tar høyde for teknologiske endringer. En annen sentral egenskap ved et digitalt arkivdepot, er at det må organiseres og forvaltes på en slik måte at det inngir til tillit. Det er essensielt at samfunnet har tillit til hvordan arkivmaterialet hentes inn, bevares og gjøres tilgjengelig. Dokumentert revisjonssporing og logging, samt mekanismer for å fange opp avvik, er med på å bygge en slik tillit.

Sikker lagring av digitalt arkivmateriale og forvaltning på en slik måte at man bevarer sentrale egenskaper ved materialet, unngår informasjonstap og holder materialet tilgjengelig for senere bruk, krever helt spesifikke arkivfaglige metoder nedfelt i etablert arkivfaglig praksis, i standarder og i regelverk. Det digitale arkivdepotet må altså bygge på et sett med prinsipper og krav og ha visse grunnleggende egenskaper, som er utbredte og utprøvd innen digital bevaring.

To sentrale komponenter i et digitalt arkivdepot er den fysiske lagringen av arkivmateriale og verktøy for forvaltningen av det. I det følgende beskrives overordnede krav og føringer for disse to komponentene.

3.1 Lagringsløsningen

Det er et grunnleggende prinsipp at det digitale arkivmaterialet lagres i minimum tre kopier og at arkivmaterialet ligger lagret på minst to geografisk adskilte steder. Dette minsker muligheten for at data går tapt som følge av en katastrofehendelse som for eksempel brann eller oversvømmelse. I tillegg gjør en slik lagringsstruktur det vanskeligere å manipulere informasjonen som bevares.

Det er videre et grunnleggende prinsipp at digitale arkivmaterialet skal lagres med to ulike teknologier. Lagring med to teknologier (p.t. f.eks. disk og tape) er mer robust mot teknologiskifter blant annet ved at kostnader og risikoen for datatap ved skifte av teknologi, kan reduseres. Lagring basert på én teknologi innebærer for eksempel at en er mer sårbar for svingninger i markedet for lagringsløsninger og får økt avhengighet til enkelte produsenter av lagringsteknologi.

Bruk av to teknologier bidrar i tillegg til at arkivdepotet blir mer robust mot innebygde feil i valgte teknologier.

Materialet i det digitale depotet vil vokse etter hvert som materialet overføres. Lagringsløsningen må derfor være skalerbar, slik at det til en hver tid er lagringskapasitet til å håndtere arkivuttrekk.

I lagringsløsningen i et digitalt arkivdepot inngår flere elementer i tillegg til den fysiske lagringen av data. Dette er elementer som sørger for at løsningen er organisert og driftet på en slik måte at informasjonssikkerheten blir ivaretatt og at risikoen for informasjonstap reduseres. Følgende er naturlige elementer i tilknytning til lagring av bestanden i et digitalt arkivdepot:

- avviks- og endringshåndtering
- fullstendige backup-rutiner
- tilgangsstyring
- rapportering til Byarkivet
- katastrofeberedskap
- kontinuitetsplan
- dokumentasjon

Disse elementene må være godt dokumentert. For depot som ivaretas av en ekstern leverandør av lagringstjenester må dette kontraktsfestes.

Byarkivet har behov for en lagringsløsning som fra starten av har en lagringskapasitet som tar hensyn til eksisterende lagringsbehov og til forventet vekst i de nærmeste årene. Det er krevende å estimere kapasitetsbehov i lagringsløsningen. Med utgangspunkt i erfaringer med nåværende bestand og forventede overføringer de nærmeste årene, gjøres likevel et estimat nedenfor.

Usikkerheten har blant annet sammenheng med at Byarkivet ennå ikke har erfaring fra innføring av ny sakarkivløsning (ACOS WebSak) hos virksomheter i Oslo kommune. Videre finnes det en til dels omfattende fagsystemportefølje i kommunen hvor det ikke foreligger en fullstendig oversikt over bevaringsverdig informasjon. I tillegg kan at pågående utvikling i metodikk for bevaring påvirke mengden data som bevares og formatet på dem.

For å beregne lagringsbehovet har Byarkivet innhentet informasjon fra andre arkivdepoter slik som det norske riksarkivet, Bergen byarkiv, KDRS, Trondheim byarkiv og det svenske riksarkivet.

Alle virksomheter i Oslo kommune skal ha et Noark-system til arkivering av saksdokumenter. Byarkivet har så langt mottatt en bare en håndfull arkivuttrekk med dokumentfiler fra Noark-4-systemer, men ingen uttrekk som er egnet for langtidsbevaring fra Noark 5-løsninger. Uttrekkene varierer i størrelse. Med utgangspunkt i seks Noark-4-uttrekk på til sammen ca. 93 GB, kan vi bregne gjennomsnittsstørrelse på uttrekk fra sak/arkiv-systemer i nåværende bestand til 15,5 GB. Disse uttrekkene inneholder til sammen ca. 235 000 dokumentfiler i TIF- og PDF/A-format, noe som gir en gjennomsnittsstørrelse per dokumentfil på 0,4 MB.

Imidlertid er det flere utviklingstrekk som tilsier at gjennomsnittsstørrelsen på Noark-uttrekkene er voksende. For det første vokser antall journalførte dokumenter i kommunen pr år sterkt. I 2015 ble det i følge 2016-kartleggingen utført av Byarkivet³, journalført cirka 1 million dokumenter i Oslo kommune. Dette er en økning på rundt 50 prosent fra 2012, mens økningen fra 2009 til 2012 var på cirka 45 prosent. Dette påvirker størrelsen på uttrekk for en gitt tidsperiode. Videre er det flere faktorer som peker i retning av at gjennomsnittsstørrelsen på dokumentfilene vil øke, blant annet økt fokus på filformater som må spesialbehandles, for eksempel innen 3D-modellering. Byarkivet

³ Arkivsituasjonen i Oslo kommune 2016 – Kartleggingsrapport (2016)

forventer derfor vi at gjennomsnittstørrelsen per Noark-uttrekk vil øke, og legger derfor til grunn en gjennomsnittstørrelse på 30 GB for Noark-uttrekk Byarkivet mottar i de kommende 5 årene.

Lagringsbehovet for arkivuttrekk fra fagsystemer vil variere sterkt. Uttrekk som ikke inneholder dokumentfiler, vil som oftest være mindre enn uttrekk med dokumenter. Periodene som uttrekkene omfatter, vil også variere. Erfaringsmessig vil størrelsene variere fra noen få MB til flere GB. I 2017 forventer Byarkivet med å motta flere fagsystemuttrekk. Ett av disse er anslått å være på cirka 600 GB, men de fleste andre vil være betydelig mindre.

Byarkivets nåværende bestand av digitalt arkivmateriale er på under 1 TB. På bakgrunn av denne bestanden, restansene på overføring fra virksomhetene og voksende størrelse på hvert uttrekk, legger vi til grunn at lagringskapasiteten ved oppstart av det digitale arkivdepotet bør være 8-10 TB. I dette ligger det en margin som skal minimere muligheten for at virksomheter må vente med å overføre uttrekk i påvente av at kapasiteten justeres opp. Anslaget over lagringsbehov avhenger også av hvordan oppstarten av depotet sammenfaller i tid med mottak av uttrekk fra virksomheter med store datamengder og hvor raskt og i hvilket omfang overgangen til nytt sakarkivsystem vil foregå.

Vi planlegger for en årlig økning på 5 TB. Dette er en størrelse som nødvendigvis vil måtte kunne justeres opp eller ned i planarbeidet til Byarkivet alt etter som mengden av avtalte og planlagte mottak varierer. Som en del av forvaltningen av bestanden i det digitale arkivdepotet, inngår vedlikehold og forvaltning, hvorav det for det førstnevnte også inngår konvertering til andre filformater. Det dette betyr, er at den daglige driften av depotet i seg selv vil medføre en økning i lagringsbehov.

3.2 Innholdsforvaltning og forvaltningsverktøy

Det har etter hvert kommet internasjonale standarder for organisering, forvaltning og dokumentasjon innen langtidsbevaring av digital informasjon. Mange depotinstitusjoner i inn- og utland av tatt i bruk disse standardene som grunnlag for å ivareta sine bevaringsoppgaver. Det er med andre ord en stor andel arkivdepoter som legger de samme standardene og prinsippene til grunn for sin forvaltning av digitalt arkivmateriale. I kjølvannet av dette er det i ferd med å bli dannet beste praksiser innen digital bevaring.

Byarkivet legger til grunn at det skal velges en utbredt, utprøvd og standardisert modell som rammeverk for kommende forvaltning av digitalt arkivmateriale, og mener at den svært utbredte OAIS-modellen skal være rammeverk for forvaltning av innholdet i det digitale arkivdepotet. Et høyt antall arkivinstitusjoner i utlandet forholder seg til OAIS. Blant disse finner vi statlige og kommunale arkivinstitusjoner i både Sverige og Danmark.

OAIS (Open Archival Information System) er et rammeverk for å forvalte informasjon for langtidsbevaring organisert i såkalte informasjonspakker. Dette er et rammeverk som opprinnelig ble utviklet av Consultative Committee for Space Data Systems (CCSDS). OAIS ble en ISO-standard i 2003 (ISO 14721:2003). Gjeldende versjon av standarden er ISO 14721:2012⁴.

⁴ Open archival information system (OAIS) -- Reference model (2012)

I juni 2012 kom sluttrapporten fra DIAS-prosjektet (*Digital arkivpakkestruktur*)⁵ som var et samarbeidsprosjekt mellom Riksarkivaren og fire kommunale aktører. Dette prosjektets hovedmål var «å definere en felles og omforent pakkestruktur for bevaring av digitalt arkivmateriale med utgangspunkt i OAIS-standarden».

Resultatet fra DIAS-prosjektet legger relativt sterke føringer på arkivinstitusjonene for hva som er rasjonell forvaltningsmodell for bevaring av digitalt arkivmateriale i Norge. Resultatene fra DIAS-prosjektet, inkludert OAIS-modellen, ligger til grunn for den digitale bevaringen hos de største og mest sammenlignbare arkivinstitusjonene i landet, og for de valgene de har tatt som omhandler forvaltningen av arkivmateriale.

Et digitalt arkivdepot vil normalt være basert på et depotstyringssystem for forvaltning av arkivmateriale. Depotstyringssystemet kan dokumentere bestanden ned på et relativt detaljert nivå gjennom bruk av et antall metadataelementer. Systemet holder rede på arkivuttrekkene blant annet med tanke på hva som er mottatt, langtidslagret og tilrettelagt for bruk, samt at det kan sikre integritet gjennom bruk av sjekksummer. Et depotstyringssystem kan i tillegg bidra til å holde rede på når arkivdokumenter (filer) bør konverteres til andre formater. De mest utbredte depotstyringssystemene innen arkivsektoren er basert på OAIS-rammeverket.

Arkivuttrekk som mottas i Byarkivet fra virksomhetene skal testes for å avgjøre om de er egnet for langtidsbevaring. Utrekkene måles etter kriteriene som er avtalt med virksomheten og om de oppfyller arkivfaglige krav. De måles også etter om tekniske kriterier og krav er oppfylt. Den første typen kriterier og krav kan dreie seg om at uttrekket omfatter riktig periode, arkivdel(er), tabeller og metadataelementer. Dette vil variere ut i fra om det er snakk om fagsystemuttrekk eller uttrekk fra Noark-system. Når det gjelder tekniske krav og kriterier, så ligger blant annet riksarkivarens forskrift og Noark-standardene til grunn for disse. I tillegg har Oslo kommune ved Byarkivet foretatt visse tilpasninger til regelverket for å gjennomføre optimal bevaring for kommunen sett som helhet.

For å teste arkivuttrekk trengs det i mange tilfeller spesialtilpasset programvare. I hovedsak har slik programvare blitt utviklet og distribuert av Arkivverket/Riksarkivet, men det er nå også andre aktører innen digital bevaring som enten utvikler programvaren selv eller som får programvarefirmaer til å gjøre det på bestilling. Det benyttes også standard programvare i arbeidet med å dokumentere og kvalitetssikre arkivuttrekk. Eksempler på dette er editorer som kan behandle store filer og XML, programvare for identifikasjon av filformater, for generering og verifisering av sjekksummer og programvare for filkonvertering. Den spesialtilpassede programvaren som eksisterer i markedet er under utvikling, eller er i liten grad tilrettelagt for distribusjon. Det finnes imidlertid noen versjoner som er egnet for testing av arkivuttrekk, forutsatt at depotinstitusjonen har nødvendig kompetanse på bruken av disse.

⁵ Vel bevart! Rapport fra samarbeidsprosjektet DIAS – digital arkivpakkestruktur (2012)

4 - Nåsituasjonen

Det er altså visse grunnleggende krav til infrastruktur og forvaltning i et digitalt arkivdepot. Hvordan fremstår nåsituasjonen i kommunen og Byarkivet sett i forhold til disse kravene?

4.1 Forvaltning av digitalt arkivmateriale

Selv om Byarkivet har mottatt digitalt arkivmateriale (arkivuttrekk) i flere år, er bestanden samlet relativt liten. Arkivuttrekkene har frem til 2016/2017 blitt lagret på CD-R i to kopier. Begge kopiene har vært oppbevart i Byarkivets magasiner. CD-R har også vært et vanlig medium for overføring av uttrekk fra avleverende virksomhet til Byarkivet. Som nødvendig vedlikehold har det vært vanlig praksis at CD-R-plater kopieres til nye CD-R hvert femte år, eller til andre databærere. I henhold til standarder skal altså det digitale materialet lagres på tre ulike enheter på minst to ulike lokasjoner. Nåsituasjonen innebærer altså en kopi for lite, og lagring på én lokasjon i stedet for to. Dette må betraktes som direkte uforsvarlig. Dessuten innebærer CD-R bruk av teknologi som i flere år har hatt en dalende utbredelse, og dagens praksis innebærer tidkrevende manuelle operasjoner.

Som en midlertidig løsning for å redusere den akutte risikoen for tap av materiale, er Byarkivet i ferd med å kopiere materialet fra CD-R til disk. Det er dessuten anskaffet 1 TB sikker lagringsplass inkludert tilgangskontroll og logging hos Utviklings- og kompetanseetaten, som for tiden bruker lagringstjenester hos Evry. Dette området skal brukes til lagring av en tredje kopi. Dette vil også innebære en lokasjon nummer 2.

Et digitalt arkivdepot vil normalt være basert på et depotstyringssystem for forvaltning av arkivmateriale, jmfør avsnitt 3.2 om innholdsforvaltning og forvaltningsverktøy. Nåværende praksis i Byarkivet er ikke basert på et slikt system. Arkivuttrekk registreres i katalogverktøyet ASTA, som er utviklet for beskrivelse og bestandskontroll for papirarkiver. Dette er ikke i henhold til standard, og innebærer ikke verken hensiktsmessig eller forsvarlig overvåking, kontroll eller gjenfinning av det digitale materialet.

Arkivuttrekk som mottas i Byarkivet fra virksomhetene skal som nevnt testes for å avgjøre om de er egnet for langtidsbevaring. For å teste arkivuttrekk trengs det i mange tilfeller spesialtilpasset programvare. I hovedsak har slik programvare blitt utviklet og distribuert av Arkivverket/Riksarkivet, men det er nå også andre aktører innen digital bevaring som enten utvikler eller får utviklet programvare. Den spesialtilpassede programvaren som eksisterer i markedet er under utvikling. Byarkivet følger disse utviklingsprosessene tett, og deltar også i prosesser rundt noen av disse verktøyene, blant annet Arkade 5 og Documaster Noark 5 Validator, og holder seg oppdatert på aktuell programvare i Norge og i andre sammenlignbare land. Når det gjelder testverktøy, er nåsituasjonen derfor omtrent så god som den kan være.

Behandling av henvendelser om tilgang til elektronisk avlevert arkivmateriale vil bestå av gjenfinning, tilgangskontroll og tilgjengeliggjøring av materialet i et format som er forståelig for sluttbruker. Det materialet som er overført til Byarkivet, er foreløpig i hovedsak deponeringer, der arkivskaper har beholdt sin egen kopi av materialet og selv finner det fram ved behov. Det er foreløpig svært få henvendelser om tilgang til elektronisk avlevert arkivmateriale i Byarkivet. Etter hvert vil deponert materiale regnes som avlevert. Byarkivet overtar da rettighetene til materialet og forpliktelsen til å gjenfinne det ved behov. Foreløpig har Byarkivet bare en enkel beredskap for henvendelser om

innsyn, og ikke hensiktsmessige metoder eller verktøy. Hver enkelt forespørsel om innsyn må behandles for seg.

Et digitalt arkivdepot med lagringsløsninger og verktøy etter standard vil legge grunnlaget for tilpassede innsynsløsninger for avleverte arkivuttrekk som gir en helt annen tilgjengelighet for publikum og også for avleverende virksomhet.

Det preger nåsituasjonen at innsatsen i kommunen med overføring av elektronisk materiale fra virksomhetene til Byarkivet har vært for svakt prioritert. Det er derfor et betydelig etterslep på denne aktiviteten. Byarkivet har nylig kartlagt omfanget. Dette er beskrevet nærmere i kapittel 7.

4.2 ROS-analyse

En beskrivelse av nåsituasjonen viser altså betydelige avvik sammenlignet med gjeldende standarder og overordnede krav. Videre er det påvist et betydelig etterslep i overføring av digitalt skapt arkivmateriale fra virksomhetene til Byarkivet. Byarkivet utførte i mars 2017 en risikovurdering for dagens håndtering av digitalt arkivmateriale i Byarkivet med sikte på å identifisere sentrale utviklingsbehov innen mottak og bevaring av digitalt arkivmateriale. Analysen oppsummeres i det følgende.

Målet som er risikovurdert er: «Byarkivet skal motta og forvalte elektronisk arkiv i tråd med standarder og regelverk.»

Dette målet er forankret i både lov og forskrift, Byarkivets mandat og kommunens arkivinstruks, og er ett av fem særlig prioriterte innsatsområder i Byarkivet strategi de nærmeste årene. Målformuleringen beskriver etter vår oppfatning en minimumsforventning til digitalt arkivdepot i offentlig sektor.

Risikovurderingen tok for seg tre overordnede risikomomenter knyttet til oppfyllelse av målsetningen innenfor avlevering og forvaltning av digitalt arkivmateriale.

Byarkivet er opprettet for å være et sentralisert arkivdepot for langtidsbevaring av viktig dokumentasjon knyttet til enkeltmenneskers rettigheter, Oslo kommunes forretningsdokumentasjon eller dokumentasjon som har kulturell, forskningsmessig eller historisk verdi. Oslo kommunes virksomheter skal derfor i henhold til arkivinstruksen avlevere digitalt arkivmateriale til Byarkivet for bevaring. Manglende overføring av arkivmateriale fra virksomhetene til digitalt depot er identifisert som et sentralt risikomoment knyttet til måloppnåelsen. Arkivmaterialet som mottas og bevares i Byarkivet skal behandles i henhold til gjeldende retningslinjer og standarder innen digital langtidsbevaring, blant annet OAIS. Det andre identifiserte risikomomentet er knyttet til tap av arkivmaterialet når det oppbevares i det digitale depotet. Det er likevel liten grunn til å bevare dokumentasjonen hvis den ikke kan tilgjengeliggjøres ved behov. Det tredje risikomomentet er derfor knyttet til lav tilgjengelighet til arkivmateriale i det digitale depotet.

Nr	Risiko	S	K	R
1	Manglende mottak av arkivmateriale i digitalt depot	4,3	4,2	
2	Informasjonstap i digitalt depot	3,5	4,8	
3	Lav tilgjengelighet for arkivmateriale i digitalt depot	3,7	3,2	

Sannsynlighet	Svært stor					
	Stor			3	1	2
	Moderat					
	Liten					
	Meget liten					
		Ubestydlig	Lav	Moderat	Alvorlig	Svært alvorlig
Konsekvens						

Risiko 1: Manglende mottak av arkivmateriale til digitalt depot.

Sannsynligheten for manglende mottak av arkivmateriale til det digitale depotet er i dag stor. Både Byarkivets kartlegging av arkivsituasjonen i kommunens virksomheter og Riksarkivets arkivstatistikk for 2015, viser at det er et stort etterslep på overføringer fra Oslo kommunes virksomheter til Byarkivet. Byarkivet har hittil i liten grad fått inn digitalt arkivmateriale til bevaring, selv om de elektroniske fagsystemene har vært i bruk i lengre tid.

Identifiserte konsekvenser ved manglende mottak av arkivmateriale omfatter flere områder. Manglende avlevering av digitalt arkivmateriale fra arkivsystemer og andre fagsystemer i strid med gjeldende regelverk. En konsekvens er at virksomhetene selv må forvalte eget arkivmateriale over tid. Forsvarlig forvaltning av digitalt materiale over tid uten tap av viktige egenskaper ved materialet, krever spesialisert kompetanse og infrastruktur, og det er kostnadskreven og lite hensiktsmessig at dette utføres av den enkelte virksomhet. Det fører til en lite helhetlig praksis og fragmentert fagmiljø med svak utvikling. Det vil trolig føre til økte utgifter for kommunen og er lite gjennomførbart på lengre sikt. Dette kan føre til tap av omdømme for Byarkivet, men også Oslo kommune som konsern. Det er også et potensiale for arkivmateriale går tapt eller kommer på avveier, og materialet er dessuten vanskeligere tilgjengelig for andre enn den enkelte virksomheten.

Risiko 2: Informasjonstap i digitalt depot

Sannsynligheten er vurdert som moderat/stor fordi Byarkivet ikke enda har opplevd informasjonstap i arkivmateriale i digitalt depot. Det er likevel slik at dagens løsning vil være uholdbar i løpet av relativt kort tid. Sannsynligheten er ikke vurdert i et evighetsperspektiv, men i løpet av en 5 års periode. Tiltak på området er derfor nødvendig i løpet av en 5 årsperiode fordi sannsynligheten for informasjonstap vil øke betraktelig innen få år.

Konsekvenser ved informasjonstap i digitalt depot er mye av de samme som i risiko 1; manglende mottak av arkivmateriale i digitalt depot. Det vil potensielt være tap av viktig rettighetsdokumentasjon for borgerne og Oslo kommune som rettighetshaver. Dette kan føre til økonomiske tap både ved manglende dokumentasjon på egne forretningsforhold, men også ved manglende bevis på oppfyllelse av lovpålagte forpliktelser og krav mot borgerne. I tillegg kan det være tap av viktig dokumentasjon med stor kulturell og historisk verdi. Tap av informasjon generelt kan føre til tap av omdømme og spesielt siden det er snakk om en manglende oppfyllelse av arkivloven med forskrifter hos en stor kommune som Oslo.

Risiko 3: Lav tilgjengelighet for arkivmateriale i digitalt depot

Sannsynligheten for lav tilgjengelighet for arkivmateriale i digitalt depot er stor fordi det i dag ikke finnes gode løsninger for forvaltning, søk og tilgjengeliggjøring av det digitale materialet. Materialet er der og det er ikke umulig å finne tilbake til det, men det er vurdert som lite tilgjengelig.

Konsekvensen av lite tilgjengelig arkivmateriale er at ansatte i Byarkivet må bruke mye tid på eventuelle innsynsforespørsler. Det er i dag vanskelig både å finne frem til materialet og gi brukeren innsyn til materialet på en praktisk måte. Dette fører til økt saksbehandlingstid og lite brukerrettede tjenester, noe som kan svekke Oslo kommunes og Byarkivets omdømme. Det er også lite økonomisk gunstig for Byarkivet å basere seg på vanskelige manuelle prosesser når mengden arkivmateriale øker i stor skala. Resultatet er dessuten at virksomhetene selv ofte må holde materialet tilgjengelig for søk til tross for at materialet er overført til Byarkivet. Dette er lite hensiktsmessig og unødig kostnadskrevende.

5 – Muligheter i markedet og valgte løsninger i offentlig sektor

Dette kapitlet inneholder en omtale av hvilke løsninger som er tilgjengelige og i bruk for arkivdepoter i offentlig sektor. Både programvare for forvaltning av arkivbestanden og løsning for den fysiske lagringen av den omtales.

5.1 Depotstyringssystemer

Samtlige arkivdepoter i Norge baserer seg på OAIS-rammeverket for organisering av den digitale bevaringen. Det er i praksis bare tre applikasjoner (depotstyringssystemer eller forvaltningsverktøy) i markedet som er aktuelle for å ivareta forvaltningen av innholdet i et digitalt arkivdepot⁶. Alle disse forholder seg til OAIS. De tre applikasjonene er Archivemata, ESSArch og Preservica. Under følger en beskrivelse av disse.

5.1.1 Archivemata

Archivemata er basert på åpen kildekode og åpne standarder og bruker blant annet METS, PREMIS, Dublin Core og the Library of Congress BagIt-spesifikasjonen til å produsere arkivpakker. Archivemata brukes sammen med AtoM (Access to Memory) som er en web-basert åpen kildekode-applikasjon for arkivbeskrivelser og visning av digitalt arkivmateriale.

Archivemata er bygget rundt bruken av micro-services, dvs. at hver oppgave er brutt ned til å kunne utføres av en dedikert og uavhengig tjeneste, i praksis implementert som Perl-kode eller dedikerte programmer.

Archivemata er i utgangspunktet fritt tilgjengelig programvare som kjører under Linux. Effektiv og hensiktsmessig bruk av programvaren fordrer kompetanse på Linux og god kjennskap til prosessgangen i OAIS.

Artefactual er et firma basert i Canada som er ledende innen utviklingen av både Archivemata og AtoM. Firmaet tilbyr tjenester knyttet til disse applikasjonene, blant annet utvikling av tilpasset funksjonalitet.

Archivemata og AtoM benyttes av svært mange større offentlige arkivinstitusjoner i Europa og Nord-Amerika, og har slik en omfattende utbredelse. I Norge er derimot ikke Archivemata i praktisk bruk i arkivdepoter, med de ulemper det innebærer for muligheten for samarbeid om utviklingsarbeid og forvaltningspraksis. Men utprøving av verktøyet er gjennomført av Arbeiderbevegelsens arkiv og bibliotek og i tillegg gjennom en SAMDOK-arbeidsgruppe for bevaring av privatarkiv.

5.1.2 ESSArch

ESSArch er et depotstyringssystem for arkivpakker basert på OAIS. Applikasjonen er utviklet av det svenske firmaet ES Solutions, opprinnelig for det svenske Riksarkivet. Som en fortsettelse av DIAS-prosjektet ble verktøyet for det norske Riksarkivets regning tilpasset kravene og standardene som ble spesifisert gjennom DIAS.

⁶ DIAS sluttrapport del 2. (2017)

ESSArch består av flere programmoduler som brukes avhengig av om brukeren er arkivskaper eller arkivdepot, og hvor i bevaringsprosessen informasjonen befinner seg. Programvaren er åpen kildekode og er basert på Python og JavaScript med web-grensesnitt. Programvaren er i utgangspunktet gratis. Brukere kan inngå avtale med ES Solutions om feilretting, videreutvikling og installasjon/igangkjøring. Kostnadene knyttet til dette ligger på rundt 300 000 initielt og 150 000 årlig.

EssArch er det forvaltningsverktøyet som har klart størst utbredelse i Sverige og Norge. Dette gir gode muligheter for samarbeid med nærstående fagmiljøer. I Sverige er ESSArch i bruk blant annet i Riksarkivet, MKC, SVAR, Folksam, Ericsson, Lantmäteriverket og Sydarkivera. I Norge benyttes det av Arkivverket, Bergen kommune og Kommunearkivinstitutionenes Digitale Ressurscenter (KDRS), og muligens av noen flere arkivdepoter. Utover Sverige og Norge er imidlertid utbredelsen svært begrenset, og ES Solutions må betraktes som et sårbart miljø, med de svakheter det innebærer for utvikling og vedlikehold av løsningen.

DIAS-prosjektet var et prosjekt i regi av Riksarkivaren sluttført i 2012 som ut i fra OAIS-modellen definerte en pakkestruktur for informasjonspakker til bruk ved overføring til og bevaring av arkivmateriale i norske arkivdepot. Denne pakkestrukturen har i høy grad blitt en de facto-standard innen digital arkivbevaring i Norge. Én ting som kjennetegner pakkestrukturen, er tilpasningen av METS- og PREMIS-skjemaer til bruk i DIAS-arkivpakker. Det kun er ESSArch av de tre nevnte depotstyringssystemene som tar hensyn til METS- og PREMIS-skjemaene. Hvorvidt om de to andre applikasjonene er egnet for bruk mot disse skjemaene, og dermed DIAS-strukturen, er uklart.

4.1.3 Preservica

Preservica er programvare for å ivareta langtidsbevaring av digital informasjon. Programvaren implementerer OAIS-rammeverket og inneholder både funksjonalitet for å forvalte informasjonen som arkivpakker og for å administrere selve lagringen av bestanden. Verktøyet utvikles og forhandles av firmaet med samme navn. Preservica er et engelsk firma som har utviklet sentrale deler av sitt bevaringsprodukt i samarbeid med The National Archives i Storbritannia. Firmaet er en stor aktør i markedet og bidrar i arbeidet med sentrale standarder og tilknyttede teknologier som PRONOM, DROID and OAIS.

Preservica kommer både i en versjon som baserer seg på lagring i skyen (Cloud Edition) og i en versjon hvor kunden selv ivaretar lagringen av innholdet (Enterprise Edition). Programvaren er i motsetning til de to andre produktene proprietær og anskaffelse av den innebærer kjøp av årlige lisenser. Per mars 2017 lå inngangssummen på 55 000 GBP per år for en avtale som omfatter minimum 3 år. Preservica har også en «leieboer»-lisens (*tenant*) som tillegg for arkivdepoter som ønsker å la andre bruke deres depotinfrastruktur basert på Preservica.

Preservica har fokus på at selve informasjonen i en *record* skal bli bevart selv om denne blir migrert og konvertert til andre formater (*Active Presentation technology*). Preservica er med tanke på OIAS-implementeringen sentrert rundt bevaring av filer ved at en gruppe sammenhørende filer som skal bevares, kalles en AIP (arkivpakke), og at en del bevaringsinformasjon tilsynelatende befinner seg på utsiden av AIPen. Programvaren har også et sterkt fokus på sikkerhet i følge dem selv.

Preservica er ikke i bruk i arkivsektoren i Norge, med de ulemper det innebærer når det gjelder muligheter for samarbeid om utvikling og praksis. I utlandet er imidlertid utbredelsen relativt omfattende, og det er flere store aktører som baserer arkivdepotet sitt på Preservica.

5.2 Lagringsløsning

Når det gjelder den fysiske lagringen av arkivbestanden, så er det en større variasjon i mulige valg av lagringsløsninger og -leverandører enn hva som er tilfelle for depotstyringssystem.

I Oslo kommune tilbyr UKE (Utviklings- og kompetanseetaten) lagringstjenester for virksomheter i kommunen. UKE benytter lagringstjenester levert av kommersiell underleverandør. Hvorvidt lagringstjenestene UKE leverer egner seg for de helt spesielle kravene som stilles til en lagringsløsning for et digitalt arkivdepot må evalueres systematisk etter en detaljert kravspesifikasjon som utarbeides i henhold til regelverk, standarder og Byarkivets funksjoner og behov.

I det åpne kommersielle markedet i Norge finnes en rekke leverandører av sikker lagring. Hvor mange av dem som tilfredsstillende helt spesielle kravene til en lagringsløsning for et arkivdepot er imidlertid også vanskelig å fastslå uten en detaljert kravspesifisering og påfølgende evaluering av tilbydere.

5.3 Kommunearkivinstitusjonenes digitale ressurscenter

Kommunearkivinstitusjonenes Digitale Ressurssenter (KDRS) er et kommunalt samvirkeforetak eid av medlemmene, som er kommunearkivinstitusjoner fra hele landet. KDRS skal tilby tjenester og kompetanse til sine medlemmer og aktivt bidra til å utvikle og effektivisere langtidslagring av digitale arkiv innenfor kommunal og fylkeskommunal sektor.

KDRS tilbyr en depotløsning som er bygget på OAIS-modellen og de valgene som ble foretatt i DIAS-prosjektet. Som kommunal arkivinstitusjon kan Byarkivet melde seg inn i KDRS, og Oslo kommune får da gjennom medlemskapet tilgang til både lagringsløsning og forvaltningsverktøy (EssArch).

I praksis overføres OAIS-informasjonspakker til og fra depotløsningen som befinner seg i Trondheim (to kopier). Innholdet i løsningen blir speilet hos Riksarkivet (Oslo) og Nasjonalbiblioteket (Mo i Rana). Det er vanlig å la KDRS ta seg av pakkelagringen, mens det finnes en infrastruktur lokalt hos det aktuelle arkivdepotet (KDRS-medlemmet) hvor inn- og utpakking foregår, samt test og tilrettelegging av innholdet i pakkene.

Det er en fordel at KDRS langt på vei leverer en ferdig depotløsning etter standard for sine medlemmer. Videre tyder vurderinger foretatt av Kulturetatens jurist på at man ikke behøver å ta hensyn til anskaffelsesregelverket i offentlig sektor ved innmelding i KDRS. Det må imidlertid betraktes som en ulempe at man drastisk reduserer valgmuligheter og slik sett ikke har den samme muligheten til å tilpasse det digitale depotet til behov og krav i Oslo kommune generelt og Byarkivet spesielt. Videre er KDRS et lite og sårbart miljø, og det er dessuten uklart om løsningen kan tilfredsstillende de behovene en stor kommune som Oslo har for kapasitet. Det er erfart noe treghet når det gjelder behandling av store informasjonspakker (maks. 32GB hittil). Dette er det etter sigende planer om å gjøre noe med, blant annet med raskere brannmurer og linjer. Medlemmer i KDRS får pt.

1TB lagringsplass hos KDRS. Dette er langt mindre enn Oslo kommune har behov for av kapasitet allerede ved etablering av det digitale depotet. KDRS har en ferdig installasjons-/oppsettpakke som følger med. Denne kommer i form av ferdige virtuelle Linux-maskiner med installert forvaltningsprogramvare (ESSArch).

Årsavgiften for å være medlem i KDRS, er i 2017 NOK 305 000,-. I tillegg kommer en innmeldingsavgift. Prismodellen er under vurdering, og det arbeides blant annet med modeller som differensierer etter størrelse på kommuner. I tillegg vil utgifter til nødvendig lokal infrastruktur tilkomme.

6 - Elektronisk arkivdepot – alternativer

En kan se for seg flere alternativer for realisering av et digitalt arkivdepot med et depotstyringssystem som er basert på OAIS og en lagringsløsning som oppfyller minimumskravene til antall kopier, lokasjoner og teknologier.

Siden det er lagt til grunn at Byarkivets digitale arkivdepot skal baseres på OAIS-rammeverket, må lagringsløsninger som skal komme i betraktning derfor fungere med et depotstyringssystem som forvalter arkivuttrekk i henhold til OAIS. Forvaltningssystemet benyttes av ansatte i Byarkivet mot den digitale arkivbestanden som er lagret i løsningen. Ett av prinsippene for Byarkivets digitale arkivdepot, er at arkivmaterialet skal være lagret på minst to lokasjoner. Et annet prinsipp er at det skal være minst tre kopier av arkivmaterialet. Dette betyr at det på én lokasjon kan være flere enn én kopi av bestanden.

Uansett hvordan man organiserer og for øvrig innretter anskaffelse og drift av det digitale depotet, gjelder det at forvaltningen av bestanden skal styres og utføres fra og av Byarkivet. For å kunne sikre forsvarlig forvaltning, inngår det at alle lagringsløsninger må være åpent og fullstendig dokumentert overfor Byarkivet. Med tanke på det som går under betegnelsen «skylagring», er det helt avgjørende at Byarkivet har full kontroll på hvor arkivmaterialet fysisk og geografisk blir lagret. Hvordan disse momentene blir ivaretatt vil ha stor betydning for valg av alternativ og leverandør. Det vil ikke minst ha mye å si for hvordan det digitale arkivdepotet vil fremstå når det kommer i bruk. I dette inngår grad av tillit til at Byarkivet kan ivareta depotoppgavene sine, og faktisk sikring av arkivmaterialet blant annet gjennom opprettholdelse av materialets integritet.

Før man kan gjennomføre anskaffelse av lagringsløsning, er det nødvendig å spesifisere detaljerte krav til denne på en helt annen måte enn de overordnede beskrivelsene av standarder og krav som er gjort i denne rapporten. Det må tas stilling til hvor høy grad av sikkerhet mot uønskede hendelser som datatap, uønsket tilgang og spredning av arkivmateriale arkivdepotløsningen skal ivareta, hvor strenge krav som skal stilles til Byarkivets kontinuerlige tilgang til bestanden osv.

Det er et viktig avklaringspunkt om det bør lagres en kopi av bestanden i Byarkivets lokaler. En slik nærhet reduserer risikoen for manglende tilgang til bestanden, og gir fysisk kontroll med lagringsløsningen på en måte som andre alternativer ikke gir. En ulempe er at det stiller spesielle krav til egnede lokaler i Byarkivet, samt at det må etableres infrastruktur for eksempel i form av lagringsserver i Byarkivet.

Når det gjelder innretning av selve anskaffelsen, kan man også se for seg ulike alternativer. Skal man velge én leverandør av lagringstjenester til tross for kravet om ulike lokasjoner, eller skal man gjøre avtale med to leverandører? Det første alternativet gir fordeler i form av enklere administrasjon av avtaler og kundeforhold og et oversiktlig kostnadsbilde, mens det andre kanskje gir redusert sårbarhet og risiko ved spredning på ulike aktører. Det kan også tenkes etablering av både lagringsløsning og forvaltningsverktøy gjennom én anskaffelse og én totalleverandør, som eventuelt etablerer løsningen i samarbeid med en eller flere underleverandører. En utfordring ved at lagringen utføres av andre enn den Byarkivet forholder seg til som leverandør, er at ansvarsforhold og løsningsmodeller kan bli mer uoversiktlige og mindre transparente. Dette vil kunne være en utfordring også med UKE som leverandør, så lenge UKE benytter seg av underleverandører for lagringstjenester.

Det må vurderes nærmere om det er hensiktsmessig at Byarkivet selv i noen grad skal eie infrastrukturen og drifte lagringsløsningen, eventuelt med tilhørende forvaltningsverktøy. Som et utgangspunkt vurderer vi det uansett foreløpig slik at det ville innebære en uhensiktsmessig oppbygging av infrastruktur og driftskompetanse i et relativt lite miljø dersom Byarkivet selv skulle drifte løsningen, med tilhørende lite rasjonell ressursbruk. Vi tror derfor det er lite hensiktsmessig å basere depotet som helhet på en slik løsning. Imidlertid kan det være en løsning at noe infrastruktur er plassert i Byarkivet, eventuelt med tilhørende driftsoppgaver. Kanskje kan man generelt si at jo færre eksterne leverandører man forholder seg til, jo mer aktuelt er det å vurdere om Byarkivet bør ha en kopi av bestanden i eget hus og kanskje også drifte lagringsløsningen for denne kopien selv, i lys av risiko- og sårbarhetsmomenter knyttet til for eksempel kun én ekstern leverandør.

7 – Etterslep i kommunens virksomheter

I takt med den digitale utviklingen i Oslo kommune de siste tiårene har mengden av digitale arkiver som etter regelverket skulle vært overført til Byarkivet som kommunens arkivdepot økt betydelig. Innsatsen med å få overført og bevart dette i Byarkivet har imidlertid vært alt for svak gjennom mange år, og restansene er store. Både Riksarkivarens tilsyn i kommunen, innføring av nytt sak/arkivsystem og generasjonsskifte i kommunens fagsystemportefølje fører nå til en betydelig økning i antall henvendelser om å få overføre materiale. Dette forsterker behovet for så raskt som mulig å få på plass infrastruktur og rutiner i henhold til standarder, samt kapasitet som sikrer forsvarlig langtidsforvaltning.

Det som finnes av elektronisk materiale hos Byarkivet er i hovedsak uttrekk av Noark-4-baserte journalsystem fra bydelene i forbindelse med bydelsreformen, samt en håndfull uttrekk fra andre virksomheter. Ettersom de fleste av arkivene var papirbaserte, består uttrekkene i hovedsak av journaler, uten medfølgende elektroniske dokumenter. Det var pr utløpet av 2016 overført 49 uttrekk til Byarkivet. Til sammenligning oppga Byarkivet i Bergen på samme tidspunkt å ha 116 uttrekk i sin forvaltning. Den relativt beskjedne mengden elektronisk materiale i Byarkivet tilsier at det er store datamengder ute i virksomhetene. Det er dessverre også rimelig å anta at digitalt bevaringsverdig arkivmateriale allerede har gått tapt.

Blant annet som grunnlag for denne utredningen, foretok Byarkivet i 2016 en kartlegging av arkivsituasjonen i virksomhetene for å få bedre oversikt over restanser på overføring av uttrekk. Denne kan suppleres og sammenholdes med en tilsvarende undersøkelse gjennomført av Byarkivet i 2008. Basert blant annet på disse kartleggingene oppsummerer vi nedenfor omfanget av ulike systemer i bruk i kommunen og estimerer antatt restanse på overføring av bevaringsverdige digitale uttrekk til Byarkivet.

7.1 Fagsystemer

I Riksrevisjonens undersøkelse av arbeidet med å sikre og tilgjengeliggjøre arkivene i kommunal sektor ble utfordringen med fagsystemer fremhevet. I sine vurderinger peker Riksrevisjonen på at «bevaringssituasjonen i kommunal sektor er spesielt kritisk for dokumentasjon i spesialiserte fagsystemer». Dette er fagsystemer som inneholder sentral rettighetsdokumentasjon for enkeltindivider, og behovet for å bevare denne dokumentasjonen er derfor meget stort. Riksrevisjonens undersøkelse og offentlig arkivstatistikk viser at det i svært liten grad er tatt ut og

bevart dokumentasjon fra fagsystemene og andre elektroniske saksbehandlingssystemer i kommunal arkivsektor for å sikre denne dokumentasjonen for framtiden. Så også i Oslo kommune.

I kartleggingsundersøkelsen høsten 2016 ble virksomhetene i kommunen blant annet spurt om elektroniske fagsystemer. De ble spurt om de hadde slike systemer i bruk, og i så fall hvor mange. De ble også spurt om de fortsatt forvaltet systemer som ikke lenger var i daglig bruk, og i så fall hvor mange. Resultatet er oppsummert i tabellen nedenfor.

Virksomhet	Elektroniske fagsystem						
	I bruk			Ikke lenger i bruk			
	Ja	Antall	Nei	Ja	Antall	Nei	Vet ikke
Bydeler	6	1-5	7	2	4-5	3	9
Etater	25	1-5	6	4	1-4	19	7

Det er altså to bydeler og fire etater som oppgir at de har flere fagsystemer som ikke lenger er i bruk. Bydelene oppgir 4 til 5 systemer, etatene 1-4 systemer. Siden disse systemene er oppgitt å ha gått ut av bruk, skulle de etter regelverket ha vært overført til Byarkivet. Selv om det er visse forbehold knyttet til datagrunnlaget som gjør tallene usikre, må vi med utgangspunkt i dette kunne si at et forsiktig anslag tilsier at manglende overføring av data fra i alle fall 20 fagsystemer må betraktes som en restanse.

For å få en oversikt ikke bare over restansene men også over forventet vekst i antall uttrekk i Byarkivets forvaltning i årene fremover, ble virksomhetene også spurt om antall fagsystemer som fortsatt er i daglig bruk. Seks bydeler oppga høsten 2016 at de benyttet 1-5 ulike fagsystemer. Bydelene har i utgangspunktet svært like oppgaver og benytter erfaringsmessig samme fagsystem, som gjerne driftes av en fagetat. Det er rimelig å anta at bydelene til sammen benytter minst fem slike systemer. 25 av kommunens etater meldte at de i dag benytter 1-5 ulike fagsystemer. Dersom hver av de 25 etatene benytter 3 ulike fagsystemer, tilsier det at 75 fagsystemer er i bruk i fagetatene i dag. Legger vi til de fem fra bydelssektoren, kan vi anta at 80 fagsystemer benyttes i Oslo kommune i dag.

I Byarkivets kartlegging fra 2008⁷, oppgav virksomhetene da at 80-90 ulike fagsystemer var i bruk. Trondheim kommune gjennomførte i 2013 en vurdering hvor det oppgis at kommunen har ca. 70 fagsystemer med arkivverdige innhold. Ettersom Oslo kommune som fylkeskommune har noen flere oppgaver enn Trondheim, er det rimelig å anta at tallet for Oslo kommune noe er høyere. Alt i alt synes det rimelig å anta at det er 80-90 fagsystemer med arkivverdige innhold i drift i Oslo kommune i dag.

Fagsystemene som fortsatt er i bruk er ikke nødvendigvis restanser, men de kan likevel inneholde data som skulle vært overført til Byarkivet. Restansen på 20 fagsystemer som allerede burde vært overført må derfor regnes som et absolutt minimum.

⁷ Arkivsituasjonen i Oslo kommune (2009)

7.2 Noark-baserte systemer

I Byarkivets kartleggingsundersøkelse fra 2016 spørres det også om uttrekk fra Noark-systemer (sak-arkivsystemer).

Informasjon om virksomhetens Noark-system

Virksomhet	Bydeler	Etater	Totalt
Antall system som er i bruk	14	29	43
Tidsperiode systemene omfatter	1996-	1990-	
Systemene har avsluttede perioder som ikke er avlevert til Byarkivet			
JA	11	21	32
NEI	2	7	9
VET IKKE	1	1	2
Uttrekk av avsluttede perioder er planlagt deponert til Byarkivet	12	23	35
JA	-	4	4
NEI	2	2	4
VET IKKE			

12 bydeler og 23 virksomheter oppgir at de planlegger å deponere uttrekk til Byarkivet. Tilsvarende oppgir 11 bydeler og 21 etater at de har avsluttede perioder som ikke er avlevert til Byarkivet. Byarkivet kan med andre ord forvente at det vil komme inn uttrekk fra minimum 35 virksomheter, og siden det er snakk om avsluttede journalperioder, skulle de allerede vært overført og er altså å anse som restanse. Basert på erfaring med slike journaldatabaser, kan vi anslå at det i snitt er tre uttrekk fra hver virksomhet, noe som utgjør en restanse på anslagsvis 105 uttrekk fra sak/arkiv-systemer.

Manglende overføring av data fra sak/arkiv-systemer og fagsystemer kan altså anslås til en restanse på til sammen 125 uttrekk.

7.3 Mottak av uttrekk – oppgaver og ressursbruk

Overføring til og håndtering av arkivuttrekk i Byarkivet innebærer mange ulike oppgaver. Byarkivet veileder virksomhetene ved bestilling av uttrekk fra ulike systemer. Dette innebærer ofte flere møter før uttrekk kan bestilles. I tillegg skal mottak av arkivuttrekk avtales, uttrekkene overleveres til avtalt kontaktperson i Byarkivet, og avsender skal ha en kvittering.

Arkivuttrekkene skal deretter legges i karantene i noen uker, for å unngå eventuelle datavirus som er så nye at virusskanningen ikke avslører dem. Deretter skal uttrekkene skannes for virus, sjekksummer skal kontrolleres, innhold og metadata i uttrekkene skal kontrolleres og uttrekkene skal testes. Avvik skal meldes tilbake til arkivskaper, som bestiller nye, komplette uttrekk fra leverandør. Når nye uttrekk bestilles, gjentas prosessen med mottak, karantene og kontroller før testing.

Riksarkivet anslår at hvert uttrekk må utføres gjennomsnittlig 2,4 ganger før det oppnår kvalitet egnet for langtidslagring. Basert på dette, på egen erfaring og på kunnskapen om det arbeidet som utføres for å sikre integritet og autenticitet på materialet, anslår Byarkivet at håndtering av hvert enkelt arkivuttrekk krever gjennomsnittlig ett månedsverk.

7.4 Prosjekt for innsats med etterslep

Oslo kommune har nylig inngått samkjøpsavtale med Acos om sak/arkiv-systemet Websak. Det er forventninger om at de fleste av kommunens virksomheter tar i bruk Websak. Overgang i løpet av de nærmeste årene vil utløse overføring av data fra systemer under utfasing. Et pågående generasjonsskifte i kommunens fagsystemer vil gi tilsvarende effekt, og bydelene har i forbindelse med tilsyn fra Riksarkivaren fått pålegg om å overføre avsluttede journalperioder til Byarkivet.

For å sikre forsvarlig forvaltning og bevaring av data i systemer under utfasing og for å kunne arbeide systematisk med restansen, foreslår vi ekstraordinær innsats med innhenting av bevaringsverdige arkivdata i en periode på tre år. Byarkivet anslår at det vil kreve en ressursinnsats i Byarkivet på gjennomsnittlig ett månedsverk pr. uttrekk. Med utgangspunkt i dette og beregningen av restansene over, anslås det at det totalt vil kreve ca. 11 årsverk å avvikle restansene, og det er derfor lagt opp til en innsats på til sammen 11 årsverk i perioden 2019-2021. Med en kostnad på 700' NOK pr årsverk, vil innsatsen jevnt fordelt på tre år beløpe seg til NOK. 2,25 millioner i innsatsperioden, eller til NOK 7,7 millioner for hele prosjektperioden. Riksarkivet har nylig beregnet behovet for ressursinnsats for å få avviklet tilsvarende i en samlet kommunal sektor til over 300 millioner kr. I lys av dette må Byarkivets forslag til innsats i etterslepprosjektet for Oslo kommune vurderes som forsiktig.

Behovet for langtidsbevaring av digitale arkiver krever et annet sett med verktøy, metodikk og kompetanse enn når dokumentene er på papir. Begge deler handler om å bevare selve dokumentene, men også informasjon om dem, slik at de kan settes i rett kontekst. Dette betyr at det også er behov for å ta vare på informasjon om struktur og metadata fra systemet som håndterte dokumentene på en ensartet og forutsigbar måte, samt metadata som skapes i det dokumentene kommer til depot og videre i forvaltningen av dem.

Erfaringer med kvalitetssikring av mottatte uttrekk viser at det er svært varierende kvalitet på det som kommer fra systemene i virksomhetene. Selv to uttrekk fra forskjellige versjoner av samme system kan ha vidt forskjellig karakter. Disse avvikene er gjerne systemgenerert eller leverandørgenerert i forbindelse med uttrekk. Systemgenererte avvik som er dokumentert av virksomhet må man regne med, all den tid kvaliteten på registrering vil variere noe over tid. Leverandørgenererte avvik derimot, kan blant annet skyldes uklar bestilling av uttrekk. Dette er ikke spesielt for Oslo kommune, men et kjent fenomen i både stat og kommune. Uttreksfunksjonalitet har sjelden vært innebygget i journal-/arkivsystemene, og virksomhetene benytter som oftest konsulentbistand fra systemleverandøren for å få tatt uttrekket ved periodisering. Kvaliteten på slike uttrekk er altså i stor grad avhengig av kompetansen til den aktuelle konsulent hos leverandøren, noe som også er kostbart. Eksemplene er mange på at virksomhetene derfor gir dette arbeidet lav prioritet.

I første fasen (2018) har prosjektet som mål å videreutvikle rutiner og prosesser slik at Byarkivet kan bistå kommunens andre virksomheter i å få gjennomført gode nok uttrekk på en mer rasjonell måte, til en lavest mulig kostnad. Vi vil også utvikle våre egne rutiner for mottak i depotet. I denne fasen må det også undersøkes nærmere med enkelte etater/bydeler for å få et bedre grunnlag for

prioritering og rekkefølge på arbeidet. I dette ligger det også å undersøke hvilke av dagens aktive fagsystemer som inneholder data som skulle vært overført til Byarkivet. Parallelt med prosjektet fortsetter Byarkivets daglige drift med mottak, testing og oppbevaring av uttrekk.

Når første fasen er gjennomført kan prosjektet hente inn uttrekkene fra og med 2019. Det understrekes at det ikke er forsvarlig å trappe opp aktiviteter med innhenting av uttrekk i nevneverdig grad før en mer forsvarlig infrastruktur er på plass.

8 – Forslag til videre arbeid og samlet kostnadsestimat

Denne utredningen har beskrevet nåsituasjon, behov og overordnede krav og skisserer til ulike mulige konsepter for et digitalt arkivdepot, samt oppsummert behovet for innsats med etterslep. Utredningen er å anse som forarbeid til en anskaffelsesprosess for å få på plass infrastruktur som sikrer forsvarlig forvaltning og forhindrer tap av uerstattelige data om kommunens virksomhet og innbyggernes rettigheter.

Med utgangspunkt i krav i lov og forskrift, hensiktsmessig og rasjonell ressursbruk og arbeidsdeling i kommunen som helhet samt risiko og sårbarhetsmomenter i nåsituasjonen, anbefaler vi at det investeres i nødvendig infrastruktur for å få etablert et forsvarlig digitalt arkivdepot i kommunen. Vi anbefaler videre at det settes av ressurser til å utvikle etterslepet på overføring av historiske data til et slikt arkivdepot. Vi foreslår følgende tidsplan/faser i en helhetlig prosess, både anskaffelse og implementering av infrastruktur og arbeid med avvikling av etterslep:

Periode	Aktivitet
H2017	Utredningsarbeid ledes av ekstern konsulent: <ul style="list-style-type: none">• Videreutvikling og vurdering av ulike alternative konsepter for depot.• Utredning av kostnader knyttet til ulike alternativer.• Grunnlag for etablering av kontraktstrategi.
H2017	Kartlegging av etterslep. Bevaringsvurderinger og prioritering (Aktivitene i 2017 dekkes av økt bevilgning fra og med 2016 til arkivformål, deriblant utredning)
V2018	Anskaffelsesprosjekt ledet av ekstern konsulent. Utvikling av detaljert kravspesifikasjon, utlysning og gjennomføring av konkurranse, kontraktsinngåelse.
V2018	Planlegging av innsats med etterslep. Bevaringsvurderinger og prioritering
H2018	Mottak og implementering av løsningen. Ledet av ekstern konsulent.
H2018	Oppstart av arbeidet med etterslep. Bevaringsvurderinger og prioritering.
V2019- H2021	Intensiv innsats med etterslepet
2019- (løpende)	Infrastruktur i drift
2019- (løpende)	Forvaltning av økendemengde digitalt materiale i depot

Overnevnte plan med ulike aktiviteter forutsetter imidlertid økte bevilgninger i planperioden. Utredningen har gjort anslag på hva det vil koste å etablere et digitalt depot samt komme ajour med etterslepet av arkivuttrekk. Det er behov for ressurser i en kortere periode til å gjennomføre anskaffelsen og innføringen, i tillegg til midler til selve investeringen. UKE og Program for elektroniske tjenester bør delta i arbeidet med anskaffelsen.

Det er nødvendig å etablere en teknisk infrastruktur som Byarkivet kan benytte til forvaltningen av kommunens bevaringsverdige digitale arkivmateriale over tid. Infrastrukturen består i første omgang av en lagringsløsning og et forvaltningsverktøy. Vi understreker at det ikke er forsvarlig å arbeide med overføring av etterslepet eller overføring av materiale fra systemer under utfasing ved utskiftning av sak/arkiv-system eller fagsystem før infrastrukturen er oppgradert.

Videre er det behov for ressurser til drift av depotet, samt til vedlikehold og betjening av en stadig voksende mengde digitalt materiale i Byarkivets forvaltning. Det må utvikles rutiner, materialet i depotet må overvåkes med tanke på behov for vedlikehold, det må migreres til nye formater og bærere med jevne mellomrom, og det må gis tilgang til materialet, i første omgang etter forespørsel.

Planlagt innsats med etterslep på overføring av digitalt materiale fra virksomhetene, samt overgang til nytt sak/arkiv-system og utskiftninger i kommunens fagsystemportefølje vil dramatisk øke mengden av digitalt materiale i Byarkivets forvaltning. Riksarkivarens tilsyn med virksomhetene har samme effekt.

Et samlet kostnadsoverslag for økonomiplanperioden 2018-2021 ser slik ut:

(Tall i hele mill).

	2018	2019	2020	2021
1. Konsulent, anskaffelsesprosjekt	1,5			
2. Konsulent, implementering av anskaffelse	1,5			
3. Stillingsressurs, planlegging av avvikling av etterslep og arbeid med anskaffelse	1,0			
4. Investering og etableringskostnad, lagringsløsning og forvaltningsverktøy	2,0			
5. Driftskostnader, lagringsløsning og forvaltningsverktøy		4,0	4,0	4,0
6. Stillingsressurser, prosjekt for innsats med etterslep		2,25	2,25	2,25
7. Stillingsressurser, forvaltning av voksende mengder digitalt materiale i depot			2,25	2,25
Sum	6,0	6,25	8,5	8,5

Det er kun kostnadene i punkt 5 og 7 som er løpende utover økonomiplanperioden. Det varige, økte kostnadsnivået er altså estimert til 6,25 mill. Det er usikkerheter knyttet kostnaden i punkt 4, siden den påvirkes vesentlig av valg av lagringsløsning og delvis også forvaltningsverktøy, som blir klart først gjennom anskaffelsesprosessen. Dette vil imidlertid være klarere i løpet av første halvår 2018. Det er videre vanskelig å estimere omfanget på data som skal plasseres i lagringsløsningen, og dermed hvordan datamengden påvirker kostnaden. Lagringsløsningen vil imidlertid måtte være skalerbar, slik at kapasiteten i løsningen kan økes etter hvert som data kommer inn til den.

Det må også legges til at det ikke er tatt høyde for anskaffelse/utvikling av løsning for direkte tilgang til arkivmateriale i depot, heller ikke for nyere metoder for overføring av digitalt arkivmateriale fra virksomhet til depot. Her er både standarder og teknologi mer umodne, og dette må komme som senere utviklingsfaser. Men særlig mer direkte tilgang for virksomhetene til arkivmateriale i depot er en forutsetning for optimal gevinstrealisering – fordi det er først da virksomhetene helt kan avvikle vedlikehold og drift av løsninger hos seg som ikke lenger er i bruk til saksbehandling.

9 – Referanser

- Byarkivet. (2009). *Arkivsituasjonen i Oslo kommune*
- Byarkivet. (2016) *Arkivsituasjonen i Oslo kommune 2016 – Kartleggingsrapport.*
- The Consultative Committee for Space Data Systems (CCSDS). *Reference model for an open archival Information System (OAIS)*. Hentet fra <https://public.ccsds.org/Pubs/650x0m2.pdf>
- Meld. St. 3:13 (2009-2010). (2010). *Riksrevisjonens undersøkelse av arbeidet med å sikre og tilgjengeliggjøre arkivene i kommunal sektor*. Hentet fra <https://www.stortinget.no/no/Saker-og-publikasjoner/Publikasjoner/Dokumentserien/2009-2010/dok3-200910/dok3-200910-013/>
- Riksarkivet. (2012). *Vel bevart! Rapport fra samarbeidsprosjektet DIAS – digital arkivpakkestruktur*. Hentet fra http://www.kdrs.no/prosjekt/Dokumenter/Krav/DIAS_Sluttrapport-Del1.pdf
- Riksarkivet. (2017). *DIAS sluttrapport del II*
- Samdok. (2014). *Digitalt skapt materiale i kommunal sektor 1985 til 2010 – kartlegging*. Hentet fra https://samdokdotcom.files.wordpress.com/2015/01/rapport-samdok-2014_2-kartlegging-av-digitalt-skapt-materiale-i-kommunal-sektor.pdf