

Trine Nesland

ARK2000 Digitalt depot

Mappeeksamen vår 2014, innlevert 2. juni via Fronter.

Antall sider inkl. forside: 26

Bachelor i arkiv og dokumentbehandling.

Høgskolen i Oslo og Akershus, Institutt for arkiv-, bibliotek- og informasjonsfag.

Innhold

Sammendrag	3
Del 1: Forskjellen på et SIARD-uttrekk og et Noark-uttrekk	4
Del 2: Gjenfinning i depot.....	11
Del 3: Autentisitet, OAIS	17
Del 3-1: Autentisitet i et depotperspektiv	17
Del 3-2: Realisering (implementering) av OAIS i depot.....	20
Kilder	24

Sammendrag

Besvarelsen består av tre hoveddeler, hvorav tredje del har to underdeler.

Første del handler om SIARD, og hva som er forskjellen på et SIARD-uttrekk og et uttrekk fra databaser som er laget etter spesifikasjonene i Noark (Noark-4 og Noark 5). Her har jeg diskutert enkelte av disse forskjellene, og samtidig forsøkt å plassere uttrekk generelt, og Noark-/SIARD-uttrekk spesielt, i en større bevaringskontekst. Min konklusjon er at hver enkeltstående forskjell er påvirket av den overordnede strategiske forskjellen på Noark og et SIARD – det er ganske enkelt to meget ulike verktøy.

I andre del har jeg drøftet påstanden om at gjenfinning er viktig, men også utfordrende, i depotsammenheng. Jeg har forsøkt å se påstanden i lys av den generelle teknologiske utviklingen, samtidig som jeg har forsøkt å sette drøftelsen inn i en relevant, fortrinnsvis digital, sammenheng. Fremfinning, gjenfinning og søk er aktiviteter som i en bevaringsinstitusjon eller i et depot, påvirkes av virksomhetens overordnede brukerorientering, den generelle og spesielle brukeradferden, teknologiske muligheter og standardisering blant annet av metadata – problemstillinger som både sammen og hver for seg utelukkende er svært omfattende og også gjenstand for utstrakt debatt i fagmiljøet.

Tredje del har to underdeler. Her har jeg først drøftet autentisitet i et depotperspektiv, og deretter diskutert OAIS' rolle i depot, herunder at modellen kan realiseres på ulike måter. Dokumentasjon er en forutsetning for å lykkes med å bevare ethvert arkivmateriales troverdighet. Dette er nødvendigvis også sentralt i OAIS, som jeg har argumentert for at kan implementeres i ulik grad og på ulike nivåer i virksomheten, alt etter hva virksomheten måtte ha behov for.

Del 1: Forskjellen på et SIARD-uttrekk og et Noark-uttrekk

Begrepet «uttrekk» refererer til samlingen av digitale data som trekkes ut av opprinnelsessystemet/-systemene og lagres på formater som egner seg for bevaring ut over den tiden man kan forvente at teknologien som opprinnelsessystemet/-systemene bygger på vil være tilgjengelig. Uttrekk innebærer med andre ord å migrere dataene fra ett format til et annet for slik best å legge til rette for bevaring i et langtidsperspektiv, kanskje til og med i et evighetsperspektiv. I den følgende teksten vil jeg drøfte hva som skiller et Noark-uttrekk fra et SIARD-uttrekk, og samtidig forsøke å sette bruken av Noark og SIARD inn i en større arkivbevaringssammenheng.

Mange faktorer spiller inn når arkiv skal bevares. Avstand i tid til arkivdanningen er et eksempel på en slik faktor. Jo lenger i tid man beveger seg bort fra arkivdanningen, jo vanskeligere vil det bli å forstå dataene arkivmaterialet består av, så vel som arkivmaterialets indre og ytre proveniens, særlig hvis arkivdanningen har vært tilfeldig og lite regelstyrt. Dette gjelder både for papirarkiv og for digitale arkiv, men spesielt for digitalt arkivmateriale, fordi digitalt arkivmateriale lett lar seg korrumpere og også lett går tapt. Tidsvinduet for bevaring er med andre ord betydelig mindre for digitale data enn for data lagret på papir, ettersom teknologiutviklingen går så raskt at både maskin- og programvare foreldes i løpet av få år. (Consultative Committee for Space Data Systems [CCSDS] 2012, s. 2-1).

Det er utviklet en rekke logiske tilnærminger til og strategier for langtidsbevaring av digitalt arkivmateriale. Logisk bevaring skal sikre at dataene kan vises, leses og forstås i fremtiden, og kan skje ved hjelp av migrering, «den museale metode» eller emulering. Ved migrering (også kalt konvertering) bevares kun dataene, som løpende overføres til tidsriktige formater og medier, slik at samtidens maskin- og programvare er i stand til å aksessere og vise dataene. Den museale metoden innebærer å sikre tilgang til arkivmaterialet ved å bevare både maskin- og programvare og samtidig kontinuerlig holde denne vedlike. Ved emulering gjenskapes den originale maskin- og programvaren med bruk av nåtidig programvare. Emulering regnes i arkivkyndige kretser som en mulig bevaringsstrategi for noen typer arkivmateriale, dog ikke for alle. (Egeland 2013, s. 110). Kanskje er emulering særlig egnet ved bevaring av personarkiver, der det tekniske «miljøet» som arkivskaper opererte i og forholdt seg til, har betydning for å forstå

vedkommendes liv og virke. Emulering har f.eks. blitt benyttet som en del av bevaringsstrategien for Salman Rushdies personlige arkiv, som man finner i Emory Universitys Manuscript, Archives, and Rare Book Library (Carroll, Farr, Hornsby, & Ranker 2011, s. 61).

I Norge er det arkivloven med forskrifter som regulerer arkivbevaringsarbeidet i offentlig sektor. Lovverket gjelder tilsvarende for statlige som for fylkeskommunale og kommunale organ, mens det for fylkeskommunal og kommunal sektor også er gitt føringer i «Normalinstruks for avlevering av arkivmateriale til kommunale og fylkeskommunale arkivdepot», utarbeidet av Riksarkivaren i 2002. Bevaring av digitalt arkivmateriale er basert på migrering som strategi, dvs. at dataene skal gjøres uavhengige av systemene de er generert og lagret i. Normalinstruksen skisserer en generell bevaringsmetode for databaser og elektroniske register der data trekkes ut av opprinnelsessystemet (databasen) som definerte tabelluttrekk i sekvensielle tekstfiler, mens det er egne avleveringsspesifikasjoner for systemer som er bygget på Noark. (ABM-skrift 43, s. 10).

Hensikten med migrering som bevaringsstrategi er å legge best mulig til rette for at dataene ikke går tapt som følge av teknologiforgjengelighet. Digitalt arkivmateriale lagres gjerne tett integrert med den program-/maskinvaren som brukes for å skape og lese informasjonen, og dermed vil ikke informasjonen være tilgjengelig hvis ikke nettopp denne program-/maskinvaren samtidig er tilgjengelig. Når slik program-/maskinvare i tillegg er proprietær og ikke basert på åpen kildekode, alternativt (i det minste) åpne standarder, vil tilgang til informasjonen i et langtidsperspektiv utelukkende bero på at leverandørene sørger for tilgjengelighet. I en bevaringsinstitusjon vil man typisk operere med et evighetsperspektiv i bevaringssammenheng, hvilket normalt ikke er forenelig med perspektivene til kommersielle aktører. Hvis bevaringsinstitusjoner eller andre virksomheter med tilsvarende intensjoner og mandater ikke griper inn før program-/maskinvaren blir utilgjengelig, vil dataene, og informasjonen som kan utledes fra dataene, gå tapt. (eArkivsamarbeidet 2007, s. 5).

Det er imidlertid ikke slik at teknologiforgjengeligheten opphører etter at en bevaringsinstitusjon har tatt hånd om arkivmaterialet. Tvert imot kan samlinger av store og uensartede mengder data lagret på ulike formater bli vanskelige å håndtere hvis man

ikke har etablert en helhetlig strategi og metode for bevaring, forvaltning og formidling av arkivmaterialet. (ABM-skrift 43, s. 12). Dette er også blant problemstillingene som adresseres i «Open archival information system» (OAIS). Bevaringsstrategien som bevaringsinstitusjonen velger vil i alle tilfeller utgjøre en sentral del av bevaringsinstitusjonen overordnede styringsstrategi.

I Norge migreres informasjon fra databaser ved at utvalgte data trekkes ut og lagres i teknologiavhengige formater, dvs. at dataene gjøres uavhengige av opprinnelsesteknologien, dennes logiske oppbygging og de omgivelsene dataene er generert i hos arkivskaper. De negative konsekvensene av endringene arkivmaterialet gjennomgår når migreringen gjøres, for eksempel at dataenes autentisitet utfordres og at bruk av dataene blir vanskeligere, forsøkes redusert med beskrivelser av hvilke endringer som er gjort, hvordan endringene er gjort, hva som er resultatet av endringene, hvem som har forstått endringene, med hvilket mandat endringene er gjort osv. (ABM-skrift 43, s. 13). Hvilke metadata som produseres og bevares for å beskrive uttrekket bestemmes av hva slags system dataene trekkes ut av, dataenes anvendelse hos arkivskaper, bevaringsinstitusjonens behov, og en rekke andre faktorer. Generelt kan man si at det sjelden er ønskelig å avlevere og bevare alle data i en database, ettersom dette vil kreve en enorm dokumentasjonsinnsats, og alle dataene ikke nødvendigvis er relevante for forståelsen av innholdet i senere brukssammenheng (ABM-skrift 43, s. 24). Dette vil bevaringsinstitusjonen måtte vurdere i hvert enkelt tilfelle, i tråd med sin bevaringsstrategi.

Krav til arkivdanningen skal sikre både hva som bevares og hvordan, og arkivdanning bør derfor inngå som en del av bevaringsinstitusjonens bevaringsplanleggingsfunksjon, jf. OAIS (ABM-skrift 43, s. 20). Noark kan inngå som en del av en slik tilnærming til bevaring, der arkivdanningen forutsettes å skje i systemer som er bygget opp etter bestemte spesifikasjoner. Noark-4 og Noark 5 er svært ulike, men begge bygger på et prinsipp om at dataene skal trekkes ut av systemene med jevne mellomrom, med langtidsbevaring for øyet. Noark-4 er en kravspesifikasjon for elektroniske arkivsystemer i offentlig forvaltning. Spesifikasjonen angir krav til systemenes informasjonsinnhold (hvilke opplysninger som skal kunne registreres og gjenfinnes), datastruktur (utformingen av de enkelte dataelementer og forholdet mellom dem) og funksjonalitet (hvilke funksjoner systemene skal ivareta). Unntaksvis stilles det også krav til brukergrensesnitt

(hvordan systemene skal kommunisere med brukerne). (Riksarkivet 1999, s. 11). Noark 5 stiller krav til arkivstruktur, metadata og funksjonalitet, men ikke krav til hvordan dette faktisk skal løses i systemutviklingen. Noark 5 definerer derfor ikke et konkret system, men legger til rette for ulike modeller. For deponering, avlevering og migrering er kravene derimot sterkere: obligatoriske metadata skal inngå i uttrekket, og uttrekket skal ha en definert struktur. (Riksarkivet 2013, s. 11). Noark selv vil fungere som en overordnet beskrivelse av uttrekkets struktur og komponenter, og kan dermed, i tillegg til å lette senere bruk av dataene, bidra til å øke uttrekket og uttrekkets innholds autentisitet etter overføring til bevaringsinstitusjon.

Ofte er det imidlertid slik at bevaringsinstitusjonen først blir kjent med systemet (og kanskje til og med først blir kjent med systemets *eksistens*) når det blir snakk om at dataene som er lagret i systemets database muligens må bevares. Nettopp dette er noe av utfordringen med bevaring av data fra de mange fagsystemene som er i bruk, for eksempel, og kanskje særlig, i kommunal sektor i Norge. Ettersom variasjonene i systemene selv og også bruken av dem vil være store, er det vanskelig, kanskje umulig, og samtidig neppe hensiktsmessig, å bestemme én standardisert bevaringsmetode eller -tilnærming. Flertallet av systemene det her er snakk om er imidlertid ulike typer relasjonsdatabaser, og fra disse vil man kunne foreta tabelluttrekk til godkjente formater så lenge nødvendig dokumentasjon er tilgjengelig (ABM-skrift 43, s. 23). Nettopp her kan verktøy som SIARD spille en nøkkelrolle. «SIARD» står for «Software Independent Archiving of Relational Databases», og er, som navnet tilsier, et verktøy som muliggjør programvareuavhengig arkivering av relasjonsdatabaser. SIARD er både et format og en programvare. SIARD-formatet integrerer data, datalogikk, tekniske metadata og annen beskrivende informasjon i én ZIP-fil som utelukkende består av ikke-proprietære filformater egnet for langtidsbevaring av data (Heuscher, Järmann, Keller-Marxer, & Möhle 2004, s. 1). SIARD-programvaren består av tre hovedkomponenter: SiardFromDb, SiardEdit og SiardToDb, der SiardFromDb er migreringsverktøyet som migrerer metadata, data og relasjoner fra den opprinnelige databasen til SIARD-formatet, SiardEdit gjør det mulig å navigere i databasestrukturen og dataene, samt å dokumentere og oppdatere metadataene, mens SiardToDb lar brukere laste opp SIARD-filer til en av databasetypene som SIARD støtter, slik at man kan studere og bruke dataene i et spesifikt databasemiljø (Swiss Federal Archives 2010). Det er likevel ikke slik at man *må* bruke SIARD-programvaren for å aksessere innholdet i en SIARD-fil: nettopp fordi SIARD-

formatet er ikke-proprietært og utelukkende basert på åpne, internasjonalt anerkjente standarder som XML, Unicode og SQL:1999, øker mulighetene for å kunne gjøre bruk av dataene også i fremtiden, uavhengig av hva slags programvare man da måtte ha tilgjengelig.

Den største forskjellen på Noark og SIARD ligger kanskje på det overordnede, strategiske planet. Noark spesifiserer hvordan uttrekk skal bygges opp for å legge til rette for langtidsbevaring av arkivmaterialet, og gir samtidig en strategisk og helhetlig tilnærming til arkivmaterialets livsløp gjennom også å påvirke og styre arkivdanningen. SIARD kan med fordel inngå som en del av en overordnet bevaringsstrategi, som valgt format for lagring av data fra relasjonsdatabaser, men gir selv ingen føringer som påvirker eller er ment å påvirke arkivdanningen. Hvordan bevaringsinstitusjonen, eller arkivskaperen for den del, skal forholde seg til arkivdanningen før denne finner sted, og også underveis, hvordan systemet databasen er en del av dokumenteres ved utvikling og mens det er i bruk, sier SIARD med andre ord ingenting om, i motsetning til Noark.

Mange relasjonsdatabaser i dag er ekstremt komplekse, og består gjerne av hundrevis, kanskje tusenvis, av koblede tabeller der koblingene er avgjørende for å forstå dataene. Hvis disse koblingene forsvinner eller brytes, vil dataene i praksis gå tapt. (Heuscher et al. 2004, s. 2). SIARD forsøker å håndtere nettopp dette ved å trekke ut dataene samtidig som koblingene mellom tabellene opprettholdes. Dette skjer uavhengig av hvordan koblingene er etablert, til forskjell fra Noark, som selv, med henblikk på uttrekket, (konseptuelt) definerer disse koblingene. SIARD tillater samtidig at man legger til ytterligere beskrivelser av databasen og dens komponenter, hvilket i begrenset grad er mulig i et Noark-uttrekk. En slik utvidet beskrivelsesmulighet er avgjørende for å bevare dataenes autentisitet og integritet i et langtidsperspektiv, særlig når dataenes tilblivelse ikke har blitt styrt av en standard med en tilsvarende helhetlig tilnærming til bevaringsproblematikken som Noark har (dvs. allerede fra før arkivdanningen har funnet sted), og også ettersom SIARD selv ikke inneholder sikkerhetsmekanismer som har til hensikt å beskytte arkivmaterialets integritet. Der man for eksempel i Noark-uttrekksstrukturen til et sakarkivsystem lagrer hvert enkelt dokument (hver enkelt PDF-fil) som et objekt tilknyttet en sjekksum, vil man i et SIARD-uttrekk lagre dokumentene som BLOB-er eller CLOB-er uten tilsvarende attestasjon av filenes uforandrethet. En stor utfordring i arbeidet med SIARD-uttrekket blir dermed å sikre referanseintegriteten

mellom metadata og dokument (Groven 2014, forelesning om SIARD ved Høgskolen i Oslo og Akershus våren 2014).

Å produsere et Noark-uttrekk, dvs. å migrere dataene fra databasen til strukturene og formatene som Noark spesifiserer at uttrekket skal bestå av, innebærer i alle tilfeller å gjennomføre en temmelig omfattende endring av arkivmaterialet, og ofte også gjøre materialet nærmest ugjenkjennelig. Det er rimelig å anta at slike endringer vil påvirke forståelsen av arkivmaterialet, og selv om noe slik påvirkning er naturlig (særlig ettersom brukskonteksten endrer seg ved overføring fra arkivskaper til bevaringsinstitusjon), må man også forvente at senere brukere av materialet vil stille spørsmål om hvorvidt representasjonen av arkivmaterialet de får tilgang til gjennom Noark-uttrekket, er fullstendig og korrekt. Som en ekstra sikkerhetsmekanisme kan man derfor velge å bruke et verktøy som SIARD også på databaser i Noark-systemer, før uttrekket produseres. Særlig aktuelt er dette i tilfeller der uttrekket er vanskelig å gjennomføre fordi spesifikasjonene i Noark ikke er fulgt i utviklingen av systemet, eller arkivdanningen ikke har holdt tilstrekkelig kvalitet, slik at man er nødt til å gjøre endringer i databasen for å klare å trekke ut dataene i henhold til Noarks spesifikasjoner. Hvis man i slike tilfeller allerede har trukket ut dataene i SIARD-formatet, og deretter lar SIARD-filen inngå i arkivinformasjonspakken i bevaringsinstitusjonen, bevarer man også muligheten til å kontrollere uttrekkets korrekthet i en senere brukssituasjon. Dette vil styrke både dataene og uttrekkets troverdighet.

Som antydnet ovenfor, er det vanskelig å sammenlikne Noark og SIARD på grunn av de grunnleggende forskjellene. Noark (også Noark 5, selv om Noark 5s nedslagsfelt potensielt er større enn tidligere versjoners nedslagsfelt), er primært utviklet for offentlig sektor i Norge, dvs. arkivlovunderlagt virksomhet. SIARD har ikke en slik begrensning i seg, og henvender seg generelt til alle som bruker og kan tenkes å ville bevare innholdet i relasjonsdatabaser generelt. Det betyr at SIARD også kan brukes av private virksomheter som for sin egen del, for eksempel som følge av rettslige krav eller i bevis-/dokumentasjonssammenheng ellers, trenger å ta vare på data som ikke lenger brukes aktivt, og som befinner seg i databaser basert på proprietær eller på annen måte forgjengelig teknologi (og der man ikke i forkant av arkivdanningen forsto at man ville trenge å sørge for langtidsbevaring). SIARDs tilnærming til bevaring kan i slike tilfeller representere en klar fordel. Det vil være praktisk umulig å produsere et Noark-uttrekk fra

en database som ikke er bygget opp etter spesifikasjonene i Noark, men SIARD legger ingen slike føringer, og har således potensielt et betydelig mer omfattende nedslagsfelt enn Noark. SIARD er et verktøy man bruker i etterkant av arkivdanningen, når bevaringsbehovet melder seg, som muliggjør bevaring av arkivmateriale hvis data (og informasjon) ellers etter all sannsynlighet ville gått tapt. Noark-uttrekket fremstår dermed i større grad som et villet og konstruert resultat av en styrt prosess, mens SIARD muliggjør bevaring av data som man ikke tenkte på at man ville trenge å bevare da databasen ble anskaffet og tatt i bruk.

Verken Noark eller SIARD egner seg imidlertid til å løse bevaringsinstitusjonens mer overordnede oppgaver, slik OAIS legger opp til. For å oppnå komformitet med kravene i OAIS, kan bevaringsinstitusjonen for eksempel benytte seg av verktøy som Archivematica, som er laget for å sørge for tilgang til samlinger av digitale objekter i et langtidsperspektiv, og som også er basert på åpne, internasjonalt anerkjente standarder (Artefactual systems inc. 2014). Og nettopp her finnes det som synes å være fellesnevneren i alt arbeid med langtidsbevaring, noe også OAIS anerkjenner: langtidsbevaring av arkivmateriale er ikke mulig på en god måte uten bruk av åpne standarder og teknologi som med vilje legger til rette for interoperabilitet i tid og rom. Nettopp dette har Noark og SIARD til felles.

Del 2: Gjenfinning i depot

I det følgende vil jeg drøfte påstanden om at gjenfinning er elementært viktig, men også utfordrende, i depotsammenheng. Jeg vil forsøke å se påstanden i lys av den generelle teknologiske utviklingen, og samtidig sette drøftelsen inn i en relevant, fortrinnsvis digital, kontekst.

Den britiske arkivaren og arkivteoretikeren Charles Hilary Jenkinson ga sin begrunnelse for arkivdepotfunksjonen i 1922 gjennom å beskrive depotarkivarens plikter, som han mente var av to typer; en primær plikt og en sekundær plikt. Ifølge Jenkinson bør arkivarens virke først og fremst vies å iverksette alle mulige sikkerhetstiltak for å ivareta arkivet, det vil si å beskytte arkivmaterialets grunnleggende kvaliteter. Jenkinson satt hensynet til forskningen og forskeren til side til fordel for selve arkivmaterialet – et prinsipp han hentet fra de tre nederlenderne Muller, Feith og Fruin, som i sin manual fra 1898 ga uttrykk for at ved et arkivs ordning kommer hensynet til eventuelle historiske undersøkelser i annen rekke. Beskyttelsen av arkivmaterialet må med andre ord alltid ha førsteprioritet, og først når denne beskyttelsen er tilstrekkelig ivaretatt, kan man vie oppmerksomhet til den sekundære oppgaven, som er å betjene de som ønsker å bruke materialet. (Marthinsen 2012, s. 263-264). Flere ulike brukergrupper kan ha behov for at arkivmateriale blir bevart i kortere eller lengre tid. Hovedgrunnen til at arkiv i det hele tatt blir skapt, finnes hos arkivskaperene selv og deres umiddelbare etterfølgere, og i arkivskapers fremtidige virksomhet og behov. Øvrige interessenter må dermed langt på vei ta til takke med de rester som måtte være igjen etter at virksomheten er avsluttet. (Marthinsen 2012, s. 278). Ser man bort fra behovet for å styre arkivdanningen på en slik måte at man muliggjør langtidsbevaring av digitalt skapt informasjonen, slik man for eksempel gjør med Noark, og samtidig følger Jenkinsons argumentasjon, kan man si at en bevaringsinstitusjons primære oppgave er å ta vare på det som er igjen etter arkivskaper på måter som verner om og sikrer arkivmaterialets grunnleggende kvaliteter, for eksempel materialets indre og ytre proveniens. Først når man har lyktes med dette, kan institusjonen vie oppmerksomhet til den sekundære oppgaven, nemlig å betjene de som ønsker å bruke det bevarte materialet. Mange vil nok være uenige i Jenkinsons tilnærming, både i synet på arkivaren som (passiv) vokter av arkivmaterialet og den manglende brukerorienteringen, og heller hevde at hele poenget med å bevare arkiv er at det skal bli brukt – at hensikten med et arkivdepot er å gjøre dets materiale allment

tilgjengelig (Marthinsen 2012, s. 28) – og at man derfor i stor grad bør prioritere å (kunne) betjene (potensielle) brukere.

Arkivverket i Norge har bevart digitalt skapt arkivmateriale siden 1980-tallet, og kan vel sies, i alle fall praktisk sett, å ha fulgt Jenkinsons tilnærming i sitt arbeid, i det at man, i alle fall til nå, har konsentrert seg om trygg bevaring fremfor brukertilrettelegging. Mye av grunnen til dette har vært at materialet er lagret i rådataformat, og ikke er umiddelbart aksesserbart for alminnelige brukere, ettersom brukerterskelen er betydelig høyere for digitalt skapt arkivmateriale (i rådataformat) enn for materiale på papir. (Strand 2013, s. 20). Generelt kan man si at et forskningsarbeid basert på digitalt skapt arkivmateriale vil skje i to trinn: først må de relevante dataene finnes frem fra ulike kilder, settes sammen og struktureres på en slik måte at det er mulig å analysere dem, deretter må dataene analyseres. I sistnevnte trinn er det mulig å bruke standardiserte metoder, mens det i første trinn ikke finnes generelle prosedyrer. (Strand 2013, s. 74). Normalt vil man starte med å identifisere arkivskaper, deretter arkivskapers systemer, før man beveger seg videre nedover, til tabellene og dataene i uttrekket. I digitale arkiver tilsvarer de tekniske metadataene papirarkivenes kataloginformasjon, og de tekniske metadataene er dermed avgjørende for analysegrunnlagets fullstendighet og analyseprosessens suksess. (Strand 2013, s. 67-68).

Den normale, forventede brukeradferden er imidlertid mer en stereotyp enn noe annet. Bevaringsinstitusjoner antar gjerne at brukere av primærkilder er mer tilbøyelige til å nærme seg materialet gjennom materialets ytre proveniens (arkivskaper/objekt – virksomhet, person osv.), fremfor via tema eller (samfunns)funksjon. Denne antakelsen er imidlertid i liten grad støttet av forskning, snarere tvert i mot, og man må kunne forvente at etter hvert som bevaringsinstitusjonenes brukergrupper utvikler og endrer seg, vil også brukeradferden endres. Ettersom man gjerne ser en generell overvekt av nøkkelordsøk og andre tematiske tilnærminger til informasjonsinnhenting utenfor bevaringsinstitusjonene, må også bevaringsinstitusjonenes digitale fremfinningsystemer og underliggende surrogater, som metadatasett i MARC- og EAD-format, muliggjøre tilnærminger til materialet utover den tradisjonelle, proveniensorienterte innfallsvinkelen. (Gracy 2014).

Fremfinningsverktøy (eng. *finding aids*), som kataloger og beskrivelser for øvrig, skal ikke kun bidra med identifiserende opplysninger om et bestemt arkivmateriale (for

eksempel informasjon om arkivskaper, datoer, materialtyper og emner i arkivdanningen), men skal også beskrive materialets kontekst for slik å lette forståelsen av materialet selv og rekkevidden av det (Gracy 2014). I digital sammenheng skal ikke fremfinningsverktøyene som brukes kun presentere de ettersøkte arkivaliene, men også relasjonene mellom arkivaliene som er funnet og andre arkivalier, så vel som mellom arkivaliene og deres kontekst (Eito-Brun 2014). Arkivbeskrivelse har i løpet av de siste tiårene i stor grad blitt påvirket av den teknologiske utviklingen, og man har beveget seg fra 1980-tallets første maskinlesbare kataloginnførsler i Machine Readable Catalogue-format (MARC-format), via distribusjon av fremfinningsverktøy på Internett med HTML og Encoded Archival Description (EAD) på 1990-tallet, til dagens forsøk på utnyttelse av mer semantisk rettet teknologi, f.eks. ved bruk av Resource Description Framework (RDF) og koblede data (eng. *linked data*). Hvis arkivbeskrivelser konverteres til eller lages i RDF-format, vil man åpne for å kunne «linke» direkte fra én arkivdatabase til en annen, ekstern database, hvilket gir brukerne mulighet til å søke seg videre på tvers av beholdninger og på tvers av institusjoner. På samme måte kan koblede data brukes sammen med datagraving/dataminering (eng. *data mining*) og informasjonsvisualiseringsteknikker for å lette arbeidet med å analysere og presentere materiale med ulik proveniens og tilhørighet. (Gracy 2014). Dette kan gi helt nye muligheter for bruk og formidling av arkivmaterialet.

Mens man både med tradisjonelle fremfinningsteknikker og ved tilrettelegging for søk på nøkkelord eller tema, antar at brukeren vet hva han leter etter og hvordan han skal gå frem for å finne det, kan man med ny teknologi også skape en rekke andre muligheter:

«Providing catalogues of records [...] is not enough. Embracing a more user-centred view of data requires a rethink of database and interface design for collections data and a new emphasis on facilitating knowledge rather than just delivering information. Our collections need not exist only as inert data obtained through elusive search mechanisms. We can now empower users to explore, connect and contribute to the information that we hold and to develop and share the knowledge those activities create.» (Peacock, Ellis, & Doolan 2004).

Arkivinstitusjoner eller arkivdepoter som legger opp til en mer omfattende form for brukerorientering i sitt virke, omtales gjerne som deltakerarkiver (eng. *participatory*

archives), der brukere forventes å *delta i* beskrivingen og viderebruken av arkivmaterialet. Selv om de færreste bevaringsinstitusjoner går så langt som til å invitere brukeren til deltakelse, retter de fleste oppmerksomhet mot brukeren på en eller annen måte. Brukerorientering i form av identifikasjon av bevaringsinstitusjonens brukergruppe(r) er også en sentral del av «Open Archival Information System» (OAIS), som mange bevaringsinstitusjoner og arkivdepoter bruker i sitt arbeid. Hensikten med (videre)utvikling av fremfinningsverktøy er naturligvis å gjøre bruk av arkivmaterialet mulig (og enklere) – som nevnt ovenfor i økende grad bruk på tvers av bevaringsinstitusjoner, samfunns- eller brukergrupper og landegrenser. Formålet med EAD uttrykkes slik: «EAD makes it possible to provide union access to detailed archival descriptions and resources in repositories distributed throughout the world.» (Cornish 2004). Det er ikke lenger tilstrekkelig at gjenfinning kun skjer i den enkelte bevaringsinstitusjon, og man arbeider snarere for å legge til rette for at tilgang skal kunne skje fra hvor som helst, når som helst og av hvem som helst.

Søk på tvers av heterogene databaser og andre uensartede surrogatsystemer er uansett generelt sett svært krevende. Dette gjelder ikke kun for søk og frem-/gjenfinning i eller på tvers av bevaringsinstitusjoner, men også for søk i virksomheter:

«Many characteristics of enterprise search represent a significant challenge for IR system designers [...]. Information in the enterprise may be structured or unstructured. Documents are produced by a variety of sources, perhaps in many different languages, and generally without formatting standards. Metadata may be created according to a number of different schemes, or may not be added at all.» (Hawking 2011, s. 646).

Dette stiller seg naturlig nok noe annerledes i bevaringsinstitusjoner enn i virksomheter, ettersom bevaringsinstitusjonens beholdning ikke endrer seg etter mottak – det legges normalt ikke til nytt innhold, og dette skjer i alle fall ikke utenfor arkivarenes umiddelbare kontroll. Likevel er utfordringene mange. Standardisering og annen verktøy-/bestepsisutvikling av bevaringsinstitusjonsvirksomheten generelt (OAIS, TRAC), av metadata (MARC, Dublin Core, Noark m.fl.) og av representasjon, distribusjon og kommunikasjon av metadata og innhold (RDF, FRBRoo og andre semantisk orienterte

teknologier) er tatt i bruk for å imøtekomme utfordringene. En del av standardene gir imidlertid i seg selv nye utfordringer:

«[...] the inadequacies of the MARC and EAD standards – which requires archivists to place critical contextual information in unstructured fields and areas of the record – remain the most vexing problems that must be addressed. The continued use of free text strings in archival descriptive standards actually militates against data sharing, exchange, and interoperability, thus perpetuating the isolation of archival data and preventing its integration with data from other sources.» (Gracy 2014).

Mangelen på semantisk definisjon av de ulike enhetene i de ulike standardene er også problematisk, og begrenser gevinstene som bruk av standarder generelt gir, men selv om ideen om for eksempel et universelt metadataformat er tiltrekkende, er det mange forhold som gjør en slik tilnærming urealistisk (Aalberg 1998, s. 14). En unison enighet om design og implementering av standarder og systemer kan rett og slett vanskelig oppnås så lenge ulike brukergrupper og fagmiljøer har ulike behov, ulik forståelse av arkivmaterialet og dets kontekst og ulike ønsker om detaljnivå, tilgjengeliggjøring og innfallsvinkler til materialet. Likevel, hvis det er slik at man vil fortsette å bruke eksisterende verktøy for å finne, aksessere og forstå ulike arkivbeholdninger og arkivalier, bør verktøyene oppdateres og utvikles slik at man muliggjør full utnyttelse av eksisterende og ny teknologi. (Gracy 2014). Metadataheterogenitet kan være et resultat av at de samme dataene uttrykkes forskjellig, selv om dataene i seg selv ikke representerer forskjellig informasjon. Bruken av standardiserte formater kan bidra til å løse heterogenitetsproblemer, og dermed legge til rette for interoperabilitet. Standardisering er imidlertid vanligvis ikke et virkemiddel for å løse eksisterende heterogenitet, og brukes heller for å motvirke at heterogenitet oppstår eller får utvikle seg. (Aalberg 1998, s. 28-29, 35). For eksisterende arkivmateriale vil man derfor uansett måtte fortsette å arbeide med indeksering, katalogisering, beskrivelse, tilgjengeliggjøring og formidling innenfor rammene av nåværende fremfinningssystemer, samtidig som man forsøker å nyttiggjøre seg mulighetene ny teknologi til enhver tid gir.

Det er liten tvil om at fremfinning, gjenfinning og søk er viktig for bevaringsinstitusjoner og arkivdepoter. Samtidig er det utfordrende. Slik vil det nødvendigvis alltid være når

store informasjonsmengder skal håndteres og navigeres i. Søkeprosessen kan være en målrettet virksomhet fra et behov oppstår til behovet er tilfredsstillt, men den kan også være arbitrær og uten definerte mål (Aalberg 1998, s. 37). Samtidig går den teknologiske utviklingen svært raskt, brukernes forventninger stiger, og arkivarene må, i tillegg til å være tjenesteytere, ivareta arkivmaterialet på arkivfaglig og arkiv*etisk* forsvarlige måter. En del av dette innebærer å sørge for å beskytte materialets autenticitet og integritet, for selv om frem- og gjenfinning er viktig (kanskje også helt grunnleggende for en bevaringsinstitusjon eller et arkivdepot), har gode frem- og gjenfinningsmuligheter begrenset verdi hvis brukerne presenteres for materiale som det ikke er mulig å feste lit til. Den tillit man kan ha til materialet avhenger av at bevaringsinstitusjonen/arkivdepotet har vært tilstrekkelig oppmerksomme på og nøye med å sørge for at håndteringen av arkivmaterialet *arkivfaglig sett* har holdt høy kvalitet.

Del 3: Autentisitet, OAIS

Denne delen av besvarelsen består av to underdeler, der den første handler om tillit og den andre handler om hvordan OAIS kan realiseres (implementeres) i depot.

Del 3-1: Autentisitet i et depotperspektiv

I den følgende teksten vil jeg drøfte påstanden om at autentisitet er viktig i et depotperspektiv. Spørsmålet om dokumenters autentisitet hører til selve kjernen i arkivfaget. Stammer et dokument fra en kjent eller identifisert opphavsperson og dokumentet er innholdsmessig uendret i den tiden som har gått siden opprettelsen, er det, arkivfaglig sett, autentisk (Marthinsen 2012, s. 57-59). Tilsynelatende enkelt, men likevel meget komplekst, særlig i en digital sammenheng.

Begrepet «autentisitet» er nært knyttet til begrepet «tillit». Tillit er «[...] confidence in or reliance on some quality or attribute of a person or thing, or the truth of a statement.» (MacNeil 2000, s. xi). For at man skal kunne ha tillit til et arkivdokument, må arkivdokumentet være troverdig. Et troverdig arkivdokument er en sannferdig attestasjon av fakta og samtidig en genuin manifestasjon av disse faktaene. Arkivdokumenters troverdighet gis gjerne to kvalitative dimensjoner: pålitelighet og autentisitet, der pålitelighet innebærer at arkivdokumentet kan stå for faktaene det forteller om, for eksempel vurdert på bakgrunn av avsenders stilling og kompetanse, mens autentisitet betyr at arkivdokumentet er det det gir seg ut for å være. (MacNeil 2000, s. xi). Autentisitet viser til troverdigheten arkivdokumentet har *som arkivdokument*, og er under angrep hver gang arkivdokumentet overføres i tid og rom (InterPARES 2 Project, Creator Guidelines).

Autentisitet alene er likevel ingen garanti for at *innholdet* i arkivdokumentet er pålitelig (Strand 2013, s. 105). Det som står i et arkivdokument kan vel så gjerne være det reneste oppspinn som noe annet, og evnen til å utøve (historisk) kildekritikk er derfor helt fundamental i depotsammenheng, om ikke i mye større grad, så i alle fall på en ganske annet måte enn i arkivdanningssammenheng. Ansvar for kildekritikken ligger først og fremst hos brukeren av arkivmaterialet, og bevaringsinstitusjonens oppgave er vanligvis begrenset til å opprettholde og dokumentere arkivdokumentenes autentisitet, fremfor å forsøke å avgjøre hva som rent faktisk er sant eller korrekt (Marthinsen 2012, s. 66).

Dette innebærer nødvendigvis at bevaringsinstitusjonen må kunne klare å skille mellom arkivdokumentenes autentisitet og innholdets pålitelighet; fra et arkivfaglig ståsted spiller det liten rolle om historien som gjengis i arkivdokumentet er sann eller ikke, og bevaringen av arkivdokumentet *som arkivdokument* må skje uavhengig av innholdets sannferdighet. Viktige oppgaver for bevaringsinstitusjonen i denne sammenhengen er å verne om arkivdokumentets integritet (uforanderlighet og uforandrethet), og samtidig sørge for (trygg bevaring av) metadata om arkivdokumentenes identitet (identifikasjon, kontekst osv.).

Den enkleste, og samtidig mest korrekte måten å behandle enkeltarkiver på, er å beholde dem slik de ble til. Proveniensprinsippet er utformet som redskap for å holde orden på arkivmaterialet og for å bevare det med autentisiteten i behold. Prinsippet består av to deler. Den første delen dreier seg om å holde ethvert enkeltarkiv for seg selv og ikke blande arkivmateriale fra flere enkeltarkiver sammen. På norsk kalles dette *ytre proveniens*. Det andre leddet omhandler den indre ordningen av materialet i enkeltarkivet. Her skal man beholde den opprinnelige orden som materialet i sin tid fikk under arkivdanningen. Dette betegnes som *indre proveniens*. (Marthinsen 2012, s. 37). Teknologiforgjengeligheten er imidlertid blant faktorene som gjør det umulig å bevare digitale arkiver nøyaktig slik de ble til, i alle fall over lengre tid. Ettersom inngripen i og endring av arkivmaterialet utfordrer materialets troverdighet, er langtidsbevaring av digitalt arkivmateriale særlig krevende for bevaringsinstitusjonene. Det er ingen enkel oppgave å bekrefte at digitalt arkivmateriale er bevart med opprettholdt autentisitet (Riksarkivet 2010, s. 11). Til forskjell fra arkivdokumenter på papir, kan man ikke se hen til lagringsmediet for å utlede hvorvidt et digitalt dokument er autentisk eller ikke – koplingen mellom de digitale dataene og lagringsmediet er forbigående og generelt skjør (Strand 2013, s. 105). Tilliten brukere og andre har til hvert enkelt digitale arkivdokument kan sies å være uløselig knyttet til den tilliten man kan ha til bevaringsinstitusjonen som helhet. For å lykkes med sitt virke må bevaringsinstitusjonen sørge for å ha systemer som klart viser at den er i stand til å bevare arkivmaterialet på måter som sikrer dets troverdighet, *i det minste* fra og med avleverings-/mottakstidspunktet, ettersom de fleste bevaringsinstitusjoner sjelden er i posisjon til å påvirke arkivdanningen hos arkivskaper, og dermed vanskelig kan sørge for at arkivdokumentene blir skapt på måter som legger til rette for bevaring av dokumentenes autentisitet i et langtidsperspektiv.

Konformitet med etablerte og anerkjente standarder regnes gjerne som et kvalitetsstempel, og kan gjøre mye for å gi en bevaringsinstitusjon et godt omdømme, herunder bidra til å gjøre institusjonen verdig den tilliten som publikum generelt og brukere spesielt har til institusjonen. «Open archival information system» (OAIS) gir en helhetlig tilnærming til depotvirksomhet, og kan fungere som et rammeverk for virksomheten i bevaringsinstitusjonen. Denne virksomheten inkluderer ikke kun bevaringsinstitusjonens håndtering av arkivmaterialet, men også mer omkringliggende faktorer. Kompetansen og det eventuelle kompetansebehovet til de ansatte, samt institusjonens økonomiske stilling, både på kort og lang sikt, er faktorer som spiller inn. Det må for eksempel kunne forventes at bevaringsinstitusjonens ansatte har kompetanse til å sikre arkivdokumentenes uforanderlighet for eksempel med sjekksummer, at de forstår hva som skjer med digitale signaturer når arkivdokumenter migreres fra en teknologi til en annen og at de er i stand til å treffe nødvendige tiltak i den forbindelse. Samtidig må institusjonen ha en sunn økonomi, og man må ha sørget for fortsatt bevaring av beholdningen hvis institusjonen av ulike grunner skulle bli tvunget til å avvikle virksomheten. Hvis bevaringsinstitusjonen skal klare å verne om arkivdokumentenes autenticitet, må institusjonen med andre ord nærme seg problematikken på en helhetlig måte; en måte som omfatter hele virksomheten. Dette gjelder ikke bare, men likevel kanskje særlig, når institusjonen håndterer digitalt arkivmateriale.

Påstander om troverdighet er lette å fremme, men den reelle påliteligheten er vanskelig å bevise. Dokumentasjon og gjennomsiktighet står i alle tilfeller sentralt, og er en forutsetning i tilfeller der bevaringsinstitusjonen ønsker å la seg sertifisere av en uavhengig tredjepart. Sertifisering er i seg selv et solid virkemiddel i arbeidet med å bygge opp bevaringsinstitusjonens (dokumenterte) troverdighet, men også å tilgjengeliggjøre sertifiseringsresultater og sertifiseringsprosessdokumentasjon for allmennheten, vil støtte dette arbeidet. (Dale & Ambacher 2007, s. 3-4). For de langt fleste brukere av en bevaringsinstitusjons tjenester er det overdrevet å forvente en stillingtaken til potensielt kompleks virksomhets troverdighet for trygt å kunne bruke informasjonen som virksomheten tilgjengeliggjør, og her kan sertifisering være særlig nyttig.

Arkiv utgjør både en forutsetning og en garanti for rettssikkerheten. Dette gjelder åpenbart i arkivdanningsfasen og ellers så lenge arkivmaterialet er i arkivskapens

varetekt, men også etter at arkivmaterialet er overført til bevaringsinstitusjonen. For det første fungerer arkiv som et forebyggende tiltak mot korrupsjon, overgrep og andre forhold som utfordrer rettssikkerheten. For det andre muliggjør arkiv reparerende tiltak etter tapte rettigheter eller overgrep. Det viktigste elementet i denne sammenhengen er naturlig nok at arkivmaterialet i det hele tatt er bevart, dernest at det er bevart med sin autenticitet – og dermed sin opprinnelige utsagnskraft – i behold. (Marthinsen 2012, s. 326). Samfunnet som helhet, så vel som hvert enkelt individ, må kunne stole på at dokumentasjonen som skapes og bevares virkelig er det den utgir seg for å være (Marthinsen 2012, s. 59). Og det er nettopp dette som *bekrefter* den innledende påstanden: autenticitet er viktig i et depotperspektiv.

Del 3-2: Realisering (implementering) av OAIS i depot

I den følgende teksten vil jeg drøfte påstanden om at OAIS spiller en viktig rolle i depot, men kan realiseres på mange ulike måter.

«Open Archival Information System» (OAIS) er en referansemodell for arbeid med bevaring av digitale objekter, og har stor internasjonal utredelse og anerkjennelse. OAIS tar for seg alle prosesser i arbeidet med bevaring av digitalt arkivmateriale, både de praktiske sidene, som drift og organisering av virksomheten, arbeidet mot arkivskaper og bevaringsvurderinger, og de mer tekniske sidene, som lagring, testing og kvalitetssikring. OAIS gir et funksjonelt og konseptuelt rammeverk for organisering av depottjenesten og et grunnlag for implementering av dataverktøy for forvaltning av digitalt arkivmateriale. I funksjonsmodellen skisseres de mer generelle oppgavene som bevaringsinstitusjonen må utføre, mens det i den konseptuelle modellen beskrives mer spesifikt hvordan informasjonsobjektene skal behandles og håndteres i bevaringsinstitusjonen. Beskrivelsene handler blant annet om hvilke typer metadata som må inngå i hver enkelt informasjonspakke og knyttes til hvert enkelt informasjonsobjekt i en avlevering. (ABM-skrift 43, s. 13).

OAIS' informasjonsmodell er kanskje OAIS' viktigste bidrag i arbeidet med å forstå og behandle digitale objekter (ABM-skrift 43, s. 13). Terminologien i OAIS, for eksempel begrepene *Submission Information Package* (SIP), *Archival Information Package* (AIP) og *Dissemination Information Package* (DIP), er etter hvert svært godt innarbeidet i

fagmiljøet, både i Norge (med sine norske oversettelser, hhv. leveranseinformasjonspakke, arkivinformasjonspakke og distribusjonsinformasjonspakke) og i utlandet. Dette har en rekke fordeler, spesielt i nasjonale og internasjonale samarbeidssituasjoner (ABM-skrift 43, s. 14).

Ordet *depot* kan bety lagerbygning eller lagerlokale, men kan også vise til selve materialet som oppbevares der. Nordisk arkivleksikon (1979 nr. 13) definerer arkivdepot som en institusjon eller et lokale for ett eller flere enkeltarkiver. (Marthinsen 2012, s. 262). OAIS definerer *en OAIS* som «[...] an Archive, consisting of an organization, which may be part of a larger organization, of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community.» (Consultative Committee for Space Data Systems [CCSDS] 2012, s. 1-1). Det er ikke umiddelbart samsvar mellom de to forståelsene, og teknologit utviklingen må også antas å påvirke definisjonen av hva et depot er og også hva en depottjeneste gjør. I alle tilfeller forstår man at det dreier seg om lagring av arkivmateriale, og at denne lagringen skjer i en eller annen form for virksomhetssammenheng. Hensikten er uansett bevaring. For OAIS' del er det *digitalt* arkivmateriale som står i sentrum, og selv om konseptene og funksjonene i mange tilfeller vil være anvendbare også i en papirbasert virkelighet, uttrykkes det i OAIS at bevaring av informasjon i digital form er mye vanskeligere enn bevaring av informasjon på medier som papir eller mikrofilm (CCSDS 2012, s. 1-3).

Ettersom OAIS er en referansemodell, er den gitt en viss skalerbarhet. Det er ikke slik at OAIS legger opp til at man må anvende alle deler for å kunne ha nytte av standarden, eller at man må gjennomføre implementeringen på en bestemt måte. Tvert i mot sier OAIS at «[t]he OAIS Reference Model does not define or require any particular method of implementation of [the OAIS'] concepts.», og det antas at virksomheter som vil gjøre bruk av OAIS' konsepter, også kan bruke OAIS som en guide i utviklingen av egne systemer tilpasset egne forhold (CCSDS 2012, s. 1-3).

I norske arkivmyndigheters beskrivelser av OAIS, understrekes det at OAIS brukes som et teoretisk rammeverk, og ikke som en implementeringsmodell. Det pekes f.eks. på at OAIS har visse mangler, særlig i håndteringen av arkivmateriale fra *records management*-systemer (Arkivverket 2011). For å lykkes med å implementere OAIS'

konsepter, må man derfor i tillegg benytte seg av andre standarder, rammeverk, modeller og strategier, for eksempel standarder som METS og PREMIS. OAIS definerer metadata**typer**, men spesifiserer ikke innholdet i disse nærmere, og PREMIS (PREservation Metadata: Implementation Strategies) tar for eksempel OAIS' bevaringsbeskrivende informasjon-konsept videre ved å definere typetilhørende metadatafelt i et digitalt metadatabibliotek med tilhørende XML-skjema (Donaldson & Yakel 2012, s. 60). PREMIS som informasjonsmodell har på sin side ingen direkte verdi for bevaringspraksis hvis modellen ikke implementeres i et større system, for eksempel i en virksomhet som etterlever kravene i OAIS. Det legges dermed opp til en slags gjensidig avhengighet, så vel som et slags gjensidig utbytte, mellom de ulike standardene og rammeverkene som er tilgjengelige.

Enkelte bevaringsinstitusjoner vil kanskje kun bruke enkeltelementer fra OAIS i sitt arbeid, for eksempel kun terminologi eller utvalgte konsepter. Det er det ingenting i veien for, og kan nok også for mange passe bedre enn å legge opp til en mer fullstendig OAIS-implemtasjon. Det er også meget mulig at private virksomheter, f.eks. bedrifter som ønsker å ta vare på sin historiske informasjon på en betryggende måte, kan ha god nytte av å anvende elementer fra OAIS i sitt arbeid.

Målet med OAIS er å stille opp et konseptuelt rammeverk for et digitalt arkiv, og som redegjort for ovenfor, kan realiseringen av OAIS skje i ulik grad og eventuelt etappevis, avhengig av den enkelte bevaringsinstitusjonens behov. Det vil for eksempel ikke være hensiktsmessig å inkludere alle OAIS-elementene umiddelbart hvis målet for virksomheten i første omgang ikke er å etablere brukertjenester, men kun å legge til rette for *ingest* (mottak av arkivmateriale) og videre oppbevaring i et langtidsperspektiv. Dette vil kunne være tilfellet for bedrifter som ønsker å etablere en avdeling for historisk dokumentasjon, men som ikke kan eller vil legge til rette for bruk av informasjonen, i alle fall ikke i første omgang. Annerledes vil det naturligvis stille seg for fullverdige bevaringsinstitusjoner, som Riksarkivet eller en kommunearkivinstitusjon, som kanskje også vil ønske å motta sertifisering for å bevise sin etterlevelse av anerkjente standarder. På samme måte vil elementer som ikke er en del av OAIS-modellen kunne inkluderes i virksomheten, uten at det forringer OAIS' øvrige implementering.

OAIS kan, og kanskje bør, spille en sentral rolle i alle bevaringsinstitusjoner som følge av den generelle aksepten for OAIS' terminologi, konsepter og modell – samtidig muliggjør OAIS' fleksibilitet ulike implementeringer avhengig av behov.

Kilder

- ABM-skrift 43. (2007). *Minnehåndtering. Metode for digital langtidslagring i kommunal sektor*. Tilgjengelig 6. oktober 2013 via http://kulturradet.no/vis-publikasjon/-/asset_publisher/N4dG/content/publikasjon-minnehandtering
- Artefactual systems inc. (2014). *What is Archivemtica?* Tilgjengelig 26. april 2014 via https://www.archivemtica.org/wiki/Main_Page
- Carroll, L., Farr, E., Hornsby, P., & Ranker, B. (2007). A Comprehensive Approach to Born-Digital Archives. *Archivaria*(72), 61-92.
- Consultative Committee for Space Data Systems [CCSDS]. (2012). *Reference model for an open archival information system (OAIS)*. Tilgjengelig 28. september 2013 via <http://public.ccsds.org/publications/RefModel.aspx>
- Cornish, A. (2004). Using a Native XML Database for Encoded Archival Description Search and Retrieval. *Information technology and libraries | December 2004*. Tilgjengelig 24. mai 2014 via <http://research.wsulibs.wsu.edu/xmlui/handle/2376/1076>
- Dale, R. L. & Ambacher, B. (2007). *Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC). Version 1.0*. Tilgjengelig 4. mai 2014 via <http://www.crl.edu/>
- Donaldson, D. R., & Yakel, E. (2012). Secondary adoption of technology standards: The case of PREMIS. *Archival Science* (2013)(13), 55-83.
- eArkivsamarbeidet. (2007). *Metodebeskrivelse for digital langtidslagring i kommunal sektor*. Tilgjengelig 25. april 2014 via <http://elarkiv.nxc.no/elarkiv/content/download/2145/9901/file/Metodebeskrivelse%20eArkiv%20utkas>.

- Egeland, A.-K. (2013). Kan vi forstå de digitale arkivalier om 100 år? I Lange, H., & Spring, U. (Red.), *Til arkivet. Formidling for framtida – tanker om arkiv*. Oslo: ABM-media AS.
- Eito-Brun, R. (2014). Context-based aggregation of archival data: the role of authority records in the semantic landscape. *Archival Science*. doi:10.1007/s10502-014-9215-3
- Gracy, K. F. (2014). Archival description and linked data: a preliminary study of opportunities and implementation challenges. *Archival Science*. doi:10.1007/s10502-014-9216-2
- Hawking, D. (2011). Enterprise Search. I Baeza-Yates, R., & Ribeiro-Neto, B. (Red.), *Modern Information Retrieval. The concepts and technology behind search. Second edition*. Tilgjengelig 25. mai 2014 via <http://www.mir2ed.org/>
- Heuscher, S., Järman, S., Keller-Marxer, P., & Möhle, F. (2004). *Providing Authentic Long-term Archival Access to Complex Relational Data*. Tilgjengelig 24. april 2014 via <http://arxiv.org/abs/cs/0408054>
- InterPARES 2 Project. (udatert). *Creator Guidelines*. Tilgjengelig 4. mai 2014 via http://interpares.org/ip2/ip2_products.cfm
- MacNeil, H. (2000). *Trusting Records. Legal, Historical, and Diplomatic Perspectives*. Dordrecht: Kluwer Academic Publishers.
- Marthinsen, J. H. (2012). *Arkivteori – en innføring*. Oslo: Riksarkivarens skriftserie 36.
- Peacock, D., Ellis, D., & Doolan, J. (2004). *Searching For Meaning: Not Just Records*. Tilgjengelig 24. mai 2014 via <http://www.archimuse.com/mw2004/papers/peacock/peacock.html>

Riksarkivet. (1999). *Noark-4. Noark arkivsystem. Versjon 4.1. Del 1: Funksjonsrettet beskrivelse og kravspesifikasjon. (Oppdatert nettversjon 17.08.2005)*. Tilgjengelig 25. april 2014 via www.arkivverket.no

Riksarkivet. (2010). *Digitalt og autentisk. Planlegging av ny depotløsning for Arkivverkets digitalt skapte arkivmateriale. Prosjektrapport*. Tilgjengelig 4. mai 2014 via <http://www.arkivverket.no/arkivverket/Arkivbevaring/Elektronisk-arkivmateriale/Langtidslagring/Nytt-digitalt-depot>

Riksarkivet. (2011, 31. august). *OAIS*. Tilgjengelig 5. mai 2014 via <http://www.arkivverket.no/arkivverket/Arkivbevaring/Elektronisk-arkivmateriale/Standarder/OAIS>

Riksarkivet. (2013). *Noark 5 Standard for elektronisk arkiv. Versjon 3.1*. Tilgjengelig 25. april 2014 via www.arkivverket.no

Strand, B. (2013). *Digital archives in historical research*. Oslo: Riksarkivarens skriftserie 38.

Swiss Federal Archives. (2010). *Save your Databases. The SIARD Relational Database Archiving Solution, fact sheet*. Tilgjengelig 2. april 2014 via <http://www.bar.admin.ch/dienstleistungen/00823/00825/index.html?lang=en>

Aalberg, T. (1998). *Integrert gjenfinning i heterogene dokumentbaser*. (Hovedfagsoppgave). Trondheim: NTNU. Tilgjengelig 25. mai 2014 via www.idi.ntnu.no

Andre referanser:

Groven, A-K. (2014). Forelesning om SIARD i emnet Digitalt depot ved Høgskolen i Oslo og Akershus våren 2014.