

FAGFELLEVDERT ARTIKKEL

Olav Sata-slått

Noarkstandardens betydning for vår fremtidige evne til bevaring av elektroniske arkiv¹

SAMMENDRAG

Vår evne til å langtidsbevare elektroniske arkiv er basert på hvorvidt visse normative tekniske og arkivfaglige krav er oppfylt eller ikke. Artikkelen beskriver og analyserer langtidsbevaringen av elektroniske arkiv innenfor rammene av Noark- og fagsystemer. Den tar også for seg av juridiske krav til autentisitet, og forsøker å synliggjøre forholdet mellom kvalitet i arkivdanningen og vår evne til å bringe innholdet fra de elektroniske arkivene ut til de som etterspør dem. Den forsøker å vise at testprosessen knyttet til digitale arkivuttrekker ikke kun en del av en revisjonsmyndighet, men også en vesentlig faktor for i det hele tatt å kunne langtidsbevare de elektroniske arkivene, samt å sikre den fremtidige tilgjengeliggjøringen av dem.

Innledning og avgrensning

”Alt henger sammen med alt”. Gro Harlem Brundtland²

Denne artikkelen har som utgangspunkt fire vesentlige forhold knyttet til bevaring av elektroniske arkiv:

1. Dagens Noark-standard bygger på spesifikasjoner for elektroniske arkiv- og saksbehandlingssystemer innført så tidlig som i 1984 og utviklet gjennom en generasjon.
2. Forskriftsverket sier at elektroniske arkiv- og saksbehandlingssystemer som innføres i offentlig forvaltning skal oppfylle kravene i Noark.
3. Uttrekk skal produseres etter krav i regelverk, og innleveres til statlig eller kommunal depotinstitusjon.

¹ Synspunktene presentert i artikkelen representerer ikke Riksarkivets holdninger i de berørte problemstillingene.

² Wikiquotes.

3. Uttrekkene kan avvises av Riksarkivet hvis de ikke oppfyller forskriftsmessige krav
4. Det har vist seg tidkrevende og kostbart å få produsert uttrekk som oppfyller de forskriftsmessige kravene.

Artikkelen begrenser seg til å beskrive og analysere langtidsbevaringen av elektroniske arkiv innenfor rammene av Noark- og fagsystemer. Den er systemorientert, og tar bevisst ikke stilling til løse strukturer av dokumentasjon som vi vet finnes som filstrukturer på den enkeltes hjemmeområde eller organisasjonens fellesområde, e-post, websider osv. Årsakene til at det er en økende tendens til at de befinner seg der fremfor i Noark-systemet, ligger utenfor denne artikkelens rekkevidde. Den vil heller ikke forsøke å gi en analyse av hvorvidt en standardisering av arkivdanningen slik Noark innebærer for Norge har vært et fortinn eller en ulempe for selve arkivdanningen, spesielt innenfor perspektivet dokumentfangst og brukervennlighet.

Fra et teknisk perspektiv er det ikke spesielt utfordrende å bevare kategoriene e-post og dokumenter i løse filstrukturer, men det fordrer en arkivbeskrivelse og en eller annen form for pre-definert metadata-tagging som binder de ulike objektene til et kontekstuell rammeverk. Alle samlinger av digitale objekter i bibliotek eller arkiv er orientert rundt en akse av metadatabeskrivelser, uavhengig av om man står i bibliotek-verdenen eller arkivforvaltningen.

Et e-postarkiv eller et fillager på en PC er en samling digitale objekter som fordrer en merking med metadata for å gjøre oss i stand til å kunne beskrive innholdet og relasjonene mellom dem, så sant de ikke skal forbli et hav av data. Både bibliotek og arkiv tar utgangspunkt i en indeksering og strukturering av materialet. I et bevarings- og tilgjengeliggjøringsperspektiv vil et bortfall av denne tilnæringsmåten fordre et Googleperspektiv, der alt er søkbart og gjenfinnbart ut fra den samme tankegangen som ligger bak semantisk web³. Dette er på ingen måte en urealistisk tilnærming til arkiv, men i forhold til f.eks. Noark vil den fordre at vi gjør dataene og dokumentene «svangre» med metadata som setter oss i stand til å utføre tilfredsstillende søk og sikre fremtidig bevaring og tilgjengelighet. Den løfter begrepet arkiv ut av den tradisjonelle arkivfaglige eller samlingsorienterte sfæren, der alle dokumenter og all data må være sentrert rundt en journal eller en indeks. Semantisk web overført til arkiv eller samlinger av data og dokumenter fordrer egentlig bare at vi utstyrrer dokumentene og/eller dataene med litt ekstra metadata i dannelsesøyeblikket. Ved å gjøre dette, går vi bort fra den tradisjonelle journal eller indeks-tankegangen som verktøy for gjenfinning, og sier at all informasjon kan være søkbar, gjenfinnbar og meningsbærende på tvers av systemgrenser i det øyeblikket det skapes.

Betydningen av bevaring og tilgjengeliggjøring av elektroniske arkiv henger til dels sammen med kjensgjerningen at offentlig forvaltning i stadig økende grad blir full-elektronisk. Snarere enn å la oss imponere av en rivende teknologisk utvikling, bør vi ta

³ Semantisk web (oversettelse fra engelske *Semantic web*) er et begrep skapt av World Wide Web Consortium for å beskrive teknologier og metoder for å få maskiner til å forstå meningen - eller semantikken - bak informasjonen som finnes på internett.

inn over oss at dokumentasjon knyttet til oss som innbyggere, pasienter og rettighetshavere fra et gitt tidspunkt utelukkende eksisterer i elektronisk form. Majoriteten av den norske befolkningen vil kanskje hevde at dette er helt uproblematisk, men faktum er at forvaltningen står overfor enorme utfordringer knyttet til bevaringen av elektroniske arkiv. Tilstanden er beskrevet på overordnet nivå i Stortingsmelding om arkiv og i Riksrevisjonens undersøkelse av arbeidet med å sikre og tilgjengeliggjøre arkivene i kommunal sektor⁴. Elektroniske arkiv -i motsetning til papirbaserte arkiv, er ikke selvbevarende. Vel fordrer også bevaring av papirarkiv korrekte oppbevaringsforhold og i visse tilfelle også konservering, men differansen i kompleksitet mellom langtidsbevaring av papirarkiv og elektroniske arkiv er allikevel betydelig.

Uttrekk fra elektroniske arkiv- og saksbehandlingssystem i overordnet perspektiv

På mange vis er prosessene knyttet til bevaring av elektroniske arkiv sammenlignbar med et museum som beslutter å flytte, reise, bevare og tilgjengeliggjøre et historisk bygg. Når det gjelder elektronisk arkiv, er systemet enten spesifisert basert på kravene i Noark⁵ eller kravene til fagsystemer, definert i Riksarkivarens forskrift kapittel VIII, IX. Uttrekk fra fagsystemer følger den tekniske metadatastandard for arkivbeskrivelse, ADDML⁶. Prosessen som ligger i at et museum velger å flytte en bygård fra sine opprinnelige omgivelser for å gjenreise det på sitt museumsområde, kan beskrives på samme måte som migreringen av elektronisk arkivmateriale fra sin opprinnelige setting ute i en offentlig virksomhet og inn i en depotinstitusjon og senere ut til allmenheten.

Nå er det ikke slik at å skape et såkalt uttrekk fra databasen som ligger under et arkivsystem kan sammenliknes med det å dekonstruere et bygg. En del av prosessene er likevel de samme på overordnet nivå. For å kunne gjenreise bygget må hver eneste vinduskarm, planke og gavl – ja, hver eneste del av bygget, merkes og settes inn i en beskrivelse. På denne måten vet vi hvordan de ulike elementene er relatert til hverandre. På liknende vis må alle elementer, poster, attributter og relasjoner i uttrekket fra arkivsystemets database følge en predefinert plan eller struktur når det trekkes ut og pakkes for innlevering til depot, i tillegg til at en lang rekke beskrivelser må følge med. Dette gjøres ikke bare for å sette oss i stand til å gjenreise innholdet fra det elektroniske arkivet i fremtiden, men også for å kontrollere at de ulike elementene i uttrekket er i samsvar med de kravene vi har stilt for arkivdanning og uttrekksproduksjon. Kjernespørsmålet blir da om disse kravene – pri-

⁴ Riksrevisjonens undersøkelse av arbeidet med å sikre og tilgjengeliggjøre arkivene i kommunal sektor - Dokument 3:13 (2009–2010). Meld. St. 7 (2012–2013) - Melding til Stortinget om Arkiv

⁵ Noark er definert som spesifisering fra 1984 til 2008, og som standard fra 2008 frem til i dag.

⁶ ADDML: Archival Data Description Mark-up Language er en standard for teknisk beskrivelse av digital informasjon utviklet i samarbeid mellom de norske og svenske riksarkivene i 1998. ADDML er i dag en sentral komponent ved uttrekk fra Noark 5 og fagsystemer i Norge. I 2014 ble Finland innlemmet i ADDML-samarbeidet.

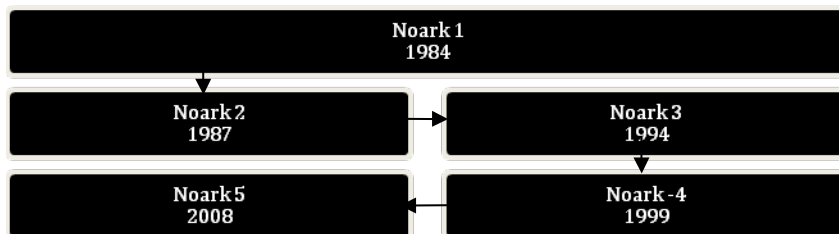
mært satt for å modellere selve arkivdanningen, bærer med seg implikasjoner for langtidsbevaringen.

Vår evne til å langtidsbevare elektroniske arkiv er basert på hvorvidt visse normative tekniske og arkivfaglige krav er oppfylt eller ikke. I Noarksammenheng bærer selve uttrekksproduksjonen og testingen av uttrekkene preg av å være en regnskapsrevisjon, selv om dette aspektet nok ikke er vektlagt i betydelig grad innen arkivfaget eller i forskriftsverket. I forhold til fagsystemer, som vi frem til Noark 5 kunne definere som arkivsystemer som ikke var underlagt kravene i Noark men allikevel skulle følge kravene i Riksarkivarens forskrift kapittel VIII og IX, er situasjonen tilsvarende. Selve uttrekksproduksjonen følger allikevel et noe annet oppsett enn for Noark, og en litt annen form for ”revisjonsvirksomhet” i Riksarkivet.

La oss dvele litt lenger ved allegorien fra museumsverdenen. Et bygg som ikke inneholder samtlige bygningselementer vil i fremtiden ikke kunne anses å være en fullstendig eller identisk gjenskapning av den opprinnelige bygningen. Museet må i verste fall erstatte opprinnelige deler som gikk tapt ved flyttingen med nye deler konservatorene utformer i dag. I forhold til elektronisk arkiv er situasjonen langt mer dramatisk. Hvis elementer i det opprinnelige uttrekket mangler, er feil strukturert eller innlevert i formater som ikke er egnet for langtidsbevaring, vil vår evne til bevaring og tilgjengeliggjøring av materialet bli betydelig redusert eller bortfalle. Hvis forholdet avdekkes umiddelbart ved innlevering til depot, kan depotinstitusjonen kreve at et nytt, korrigert uttrekk fremstilles. Hvis feil, mangler og inkonsistens oppdages ti eller tjue år etter at systemet er gått ut av bruk hos den offentlige institusjonen, er situasjonen i svært mange tilfelle umulig å utbedre. For å konkretisere forholdet kan vi gi et eksempel, hvor en etat en gang på 70-tallet innleverte et uttrekk fra et register for perioden 1967-1972. Uttrekket ble av forskjellige grunner ikke testet i forhold til struktur, elementer og relasjonskonsistens, verken da det ble innlevert til Riksarkivet eller i senere tid. Riksarkivet står da i dag i verste fall overfor betydelige utfordringer hvis et eller flere elementer mangler, eller hvis for eksempel tegnsnittet ikke lar seg konvertere til et format som er i tråd med dagens regelverk. Konsekvensen av dette blir at for å gjøre uttrekket konsistent (i ytterste konsekvens mulig å tilgjengeliggjøre) må nye elementer tilføres, og omfattende justeringer foretas. Hvorvidt dette rokker ved autenticiteten til arkivmaterialet vil drøftes senere. Eksempelet er ikke fiktivt, og belyser konsekvensen av å operere med et stort tidsspenn fra arkivdanningen til uttrekksproduksjon og testing.

Generasjonsskifter i Noark

Da Noark 1 ble introdusert som en spesifisering for elektroniske arkivsystemer i 1984, hadde man neppe den fremtidige tilgjengeliggjøringen av journalinnholdet som hovedfokus. Situasjonen var vel snarere den at fremfor å foreta et uttrekk ved periodisering eller overgang til ny versjon av Noark, valgte man rett og slett å konvertere eller migrere basen over til den nye løsningen.



Riksarkivet har mottatt uttrekk fra Noark 1 og 2-systemer basert på avleveringsformatet Riksarkivet spesifiserte i 1989. Noark 3 ble innført og tatt i bruk i offentlig forvaltning i 1994. En normal deponeringsperiode inntreffer etter fem år. Vi vil dermed kunne anta at det har foregått en kontinuerlig produksjon fra database til database fra Noark 1 til Noark 4 i tidsrommet 1986⁷ til la oss si 1999 i de uttrekkene Riksarkivet har mottatt. Likeledes viser erfaringene fra testing av innleverte uttrekk i Riksarkivet at de fleste Noark-4-uttrekk inneholder segmenter fra elektronisk arkivdanning som strekker seg langt tilbake i tid. Det er med andre ord slik at man har arkivdanning i f eks Noark-4-systemer som bærer med seg eldre segmenter fra f eks Noark 3. En interessant problemstilling er derfor om en kan spore semantiske avvik i innholdet fra Noark 1 til 4.

Om man foretar en komparativ analyse av strukturen i de ulike versjonene av standarden med selve uttrekkene, vil man antageligvis kunne si noe om hvor og når vesentlige endringer fant sted som følge av versjonering og så videre. I den grad dette er en interessant problemstilling avhenger helt og holdent av om uttrekksproduksjonen som følger av lov- og forskriftsverket i Norge byr på problemer eller ei. Det er et ufravikelig faktum at både uttrekksproduksjonen og testingen har vist seg svært problematisk, tidkrevende og til dels svært komplisert.

Embedsverkets intensjon har hele tiden vært at Noark-standardene skal være bakoverkompatibel. En logisk avledning av dette er at hvis arkivdanningen som skjer gjennom kravene i Noark er bakoverkompatibel, må vel også denne bakoverkompatibiliteten avspeiles gjennom syntaks i ethvert uttrekk som foretas fra et Noark-systems database. Vil ikke bakoverkompatibilitet måtte innebære at strukturen i uttrekk er retrospektivt koherent, uavhengig av hvor mange generasjoner av Noark som befinner seg i det? Antagelsen svikter av flere årsaker. For det første er arkivdanning i Noark-systemene i det ganske land ikke en homogen aktivitet. På det systemadministrative planet vet vi at individuelle rutiner, systemoppsett og tilpasninger i parametersettingene i sin ytterste konsekvens medfører avvik fra kravene i standarden. Dette avdekkes i testfasen. Noark-standardene er såpass streng at den for eksempel ikke tillater implementering av feltverdier i systemene som er lengre enn hva standarden spesifiserer. Fordi testverktøyene er programmert til kun å akseptere de standardiserte verdiene, må nødvendigvis testingen avdekke inkonsistens. Testingen er i realiteten en en-til-en verifisering av konsistens, hvor innholdet i uttrekket valideres mot oppsettet i testverktøyet, som igjen er definert ut fra regelverket. Testverktøyet er utformet på bakgrunn av hvordan vi har definert kravene i Noark og i

⁷ Systemversjoner av Noark 1 kom først i bruk i 1986.

regelverket.⁸ Igjen er den avgjørende problemstillingen om avvik fra Noark-oppsettet overhodet har noen betydning for fremtidig bevaring, og derigjennom tilgjengeliggjøring.

Versjonering og versjonshybrider

Noark-systemene som er blitt innført i offentlig forvaltning gjennom årenes løp har kommet i endrede versjoner fra leverandørenes side etter at de ble første gang godkjent. Årsaken er til dels den fortløpende implementeringen av ny og bedre teknologi, men kanskje i like stor grad endringer basert på arkivskapers innmelding av ønsker og spesifikke behov. Antar vi at et Noark-system er et finmasket nett av relasjoner, kan slike endringer få betydelige konsekvenser i forhold til et uttrekks konsistens.

Gjennom ikrafttredelsen av arkivloven med forskrifter i 1999, skal ethvert Noark-system som innføres i Norge først godkjennes av Riksarkivaren. Omfatter systemet sak-arkiv, skal det følge kravene i Noark. Det er ikke dermed sagt at lovgiverne påla leverandørene å presentere endrede versjoner for ny godkjenning hver gang slike endrede versjoner ble utviklet. Om dette hadde vært gjort, ville trolig en rekke av utfordringene skapt ved deponering og avlevering av uttrekk blitt betydelig redusert. For å forstå denne problemstillingen, er det nødvendig å analysere hvilken forskjell det er mellom systemet slik det ble opprinnelig godkjent, og hvordan uttrekket fra det fremstår i deponeringsfasen etter at dataene er blitt endret som resultat av stadig nye versjoner. Det er altså forskjellen mellom hva som i utgangspunktet ble godkjent da leverandøren presenterte sin løsning, og hva som etter mer eller mindre omfattende endringer der ute sendes inn for testing og godkjenning i Riksarkivet som er av avgjørende betydning. Den observante leser vil trolig se sammenhengen mellom de fire punktene i innledningen og behovet for en kontinuerlig re-sertifisering hver gang nye versjoner presenteres av Noark-leverandørene.

Fra et overordnet perspektiv innebærer dermed Noark-standarden at ingen endringer i systemet kunne forekomme uten at det får konsekvens i uttrekksøyeblikket. Sagt på en annen måte vil databasen bære preg av strukturering som ikke er i samsvar med Noark-kravene, og spesielt kapittel VIII i Riksarkivarens forskrift. Riksarkivet innehar slik dette fungerer en revisorrolle, selv om det aldri er stadfestet, hvor det gjennom testing av arkiv-uttrekket utfører revisjon av kravoppfyllelse i forhold til Noark-standarden og kapittel VIII i Riksarkivarens forskrift.

Det kan være interessant å presentere et annet eksempel, hvor Direktorat S gjennom en periode på et par år har forsøkt å generere uttrekk fra et av sine Noark-systemer. Systemet er definert som Noark-4. Direktoratet bestiller uttrekksproduksjon fra leverandøren av systemet, som ikke makter å skape et uttrekk basert på spesifikasjonene i Noark-4. Direktoratet leier inn et av landets fremste IT-konsulentbyråer, som også kommer til kort. Da direktoratet sender inn uttrekket til Riksarkivet med anmodning om bistand, fremgår det at uttrekket i utgangspunktet er fra et Noark 3-system, men at systemleverandøren har åpnet opp for digitalisering av innkomne dokumenter. Dokumentfiler for innkomne, digi-

⁸ Riksarkivarens forskrift Kapittel VIII.

taliserede dokumenter er knyttet opp til journalpostene på en måte som ikke oppfyller kravene i verken Noark 3 eller Noark-4. Dette er foretatt i det tidsmessige skjæringspunktet mellom Noark 3 og Noark-4, og hva en analyse avdekker er at systemleverandøren har levert en løsning som gir seg ut for å være et Noark-4-system. Dette er med andre ord et hybridssystem, hvor de eldre segmentene, det vil si hele lag av basen, ikke samsvarer med den elektroniske arkivdanningen som ble utført i la oss si en periode på fem år i ettertid. Når uttrekket generes som Noark 3 eller Noark-4 er ikke relasjonene mellom journalpostene og de elektroniske dokumentene konsistente. Hvor man skal begynne å nøste i en slik situasjon er en svært kompleks problemstilling. Eksempelet understreker viktigheten av å få etablert en ny ordning på nasjonalt nivå, hvor systemleverandørene faktisk må søke godkjenning for hver nye versjon som i vesentlig grad endrer datamodell og prosesser. Dette er, som vi skal se senere, en løsning som er innført med stort hell i Finland.

Videre om Noark-standarden som virkemiddel for autentisitet

Det er nok aldri tidligere blitt publisert en analyse av hvilke elementer i Noark-standarden som sikrer oppfyllelse av juridiske krav til autentisitet, og denne artikkelen har heller ikke som siktemål å berøre problemstillingen annet enn på et overordnet nivå. Generelt kan vi si at informasjonen i elektroniske arkiv inneholder viktig forvaltnings- og samfunnsdokumentasjon. I tillegg avdekker den i hvilken grad det offentlige har oppfylt sine forpliktelser overfor landets innbyggere -og vice versa. Alt dette er forsøkt sikret gjennom lovgivningen -nedfelt i forvaltningsloven, offentlighetsloven, arkivloven og tilhørende forskrifter og standarder. Uttalt i arkivlovens formålsparagraf kommer i tillegg naturligvis de forskningsmessige aspektene, samt den historiske verdien de elektroniske arkivene har.

Autentisitetsproblematikken på generelt nivå vil bli analysert i et senere avsnitt i denne artikkelen. Vi skal allikevel dvele litt ved funksjoner og parametersetting i Noark-systemene, og analysere om det finnes elementer som forhindrer manipulering, sletting og uautorisert endring av metadata og dokumenter. Autentisitet som underliggende krav er ikke eksplisitt uttrykt i Noark-versjonene før ved introduksjonen av Noark 5 i 2008. Heller ikke denne versjonen av standarden tar oss dypt inn i problemstillingene knyttet til autentisitetsspørsmålet i relasjon til elektroniske arkiv. Den internasjonale standarden ISO-15489-1, Information and documentation – Records Management⁹ er trolig på det nåværende tidspunkt mer relevant som en standard-basert tilnærming til området. I forhold til en sak¹⁰ danner den et bindeledd mellom autentisitet, pålitelighet, integritet og anvendbarhet. Oppsiktsvekkende nok er det punktet Anvendbarhet¹¹ som i sin essens bærer i seg det sentrale momentet som i realiteten er den norske testmetodikkens modus operandi.

Direkte oversatt sier standarden at den kontekstuelle forbindelsen mellom dokumenter bør bære i seg informasjonen som er nødvendig for å forstå transaksjonene som skapte og

⁹ ISO-15489-1, Information and documentation – Records Management – Part 1:General

¹⁰ Definert som Record, et begrep som har en annen betydning enn vår «sak» eller «mappe».

¹¹ s 7, 7.2.5 Useability

benyttet dem. Den sier videre at det skal være mulig å identifisere et dokument innenfor konteksten av bredere forretningsaktiviteter og funksjoner. Bindeleddene mellom dokumenter som dokumenterer en sekvens av aktiviteter bør ivaretas. Dette er selve kjernen i den norske uttrekksmetodikkens formål, nemlig å sørge for at data og dokumenter migreres ut av Noark-basene på en måte som gjør at den kontekstuelle forbindelsen mellom dokumentene i ettertid kan verifiseres. Gjennom dette kan aktivitetene organet gjennomførte dokumenteres for ettertiden.

Forvaltningsloven og offentlighetsloven er begge rettskilder som sikrer en gjennomføring av det offentliges plikter, krav og oppgaver i tråd med hva Stortinget har pålagt. Arkivloven med forskrifter, de tidligste Noark-spesifikasjonene og de nyere versjoner av Noark-standarden er i tillegg verktøy som skal sikre at saksbehandlingen lar seg dokumentere, verifisere, inspisere og på ulike vis autentisere. Arkivlovgivningen, Noark-standarden og Noark-systemene skal sikre at dokumentasjonen er registrert og lagret på en måte som ikke bare gjør fremtidig fremfinning og avlesning mulig. Vi skal i tillegg være sikre på at hva som fremfinnes og tilgjengeliggjøres er identisk med hva forvaltningen mottok av dokumenter eller informasjon, behandlet internt og til sist ekspederte eller unnløst å ekspedere. Årsaken til at unnløst fant sted skal som en digresjon bemerkes også la seg dokumentere. Det er på det rene at disse funksjonene muliggjør en rettslig definisjon av et elektronisk dokument lagret i et Noarksystem som et valid originaldokument¹². Man vil i tillegg kunne hevde at et sakskompleks dannet i og bevart fra et Noark-system bærer i seg samme grunnlag for verifisering av autentisitet, slik en konstellasjon av sakskompleks gjør det.

Et utgangspunkt for videre forskning vil kunne være komparasjonen mellom Noark og land som ikke har en arkivstandard, og analysere hvorvidt kravsetting har gitt et bedre grunnlag for datakvalitet og integritet. Samtidig kan en analyse avdekke om Noark gir et bedre grunnlag for verifisering av innholdet, noe som i seg selv må antas å være en forutsetning for materialets fremtidige anvendelse og verdi som dokumentasjon.

Forholdet mellom standarder for elektronisk arkiv, sertifisering og tilsynsmyndighet

Hvilke konsekvenser har manglende oppfyllelse av kravene i Noark for offentlige myndigheter, og hvilke konsekvenser følger for systemleverandør ved manglende oppfyllelse av de samme krav? Hva som er åpenbart er at kravstilling, kravoppnåelse og retting av feil og mangler har en kostnadmessig konsekvens. Dette gjelder ikke bare i form av ressursbruk knyttet til produksjon av nye uttrekk ved feil og avvik, men også i form av selve testprosessen, hvor analysearbeidet, kommunikasjonen og feilrettingen ute hos arkivskaper har strukket seg over måneder og år i statlig sektor. Testprosessen er ikke kun en del av en revisjonsmyndighet, men er også en vesentlig faktor for i det hele tatt å

¹² Justis- og bredskapsdepartementets prinsipputtalelse av 05.05.1999; Sivilprosess generelt - Adgangen til å treffe avtaler elektronisk og beviskraften av elektroniske dokumenter.

kunne langtidsbevare de elektroniske arkivene, samt å sikre den fremtidige tilgjengeliggjøringen av dem.

Det vil være av avgjørende betydning ikke bare for vår evne til å langtidsbevare og tilgjengeliggjøre elektroniske arkiv, å synliggjøre forholdet mellom kvalitet i arkivdanningen og vår evne til å bringe innholdet fra de elektroniske arkivene ut til de som etterspør dem. Riksrevisjonens rapport peker på manglende tilsyn med den elektroniske arkivdanningen i kommunal sektor, men det er interessant å forsøke å analysere forholdet mellom kravsetting og tilsyn, for om mulig å komme frem til hvordan og hvorfor tilsyn og kontinuerlig oppfølging av den elektroniske arkivdanningen er av avgjørende betydning for sikring av kvalitet – en kvalitet som er nødvendig for den fremtidige anvendelsen av materialet.

Sertifisering av Noark-systemer og flytting av testansvaret – en skissering av de svenske og finske løsningene

Det finske riksarkivet implementerte siste versjon av sine Sahke2 sertifiseringskriterier for elektroniske arkiv- og saksbehandlingssystemer i mars 2012. Arkivstandarden er inndelt i tre deler: 1. Retningslinjer for elektronisk arkiv

2. Metadatamodell

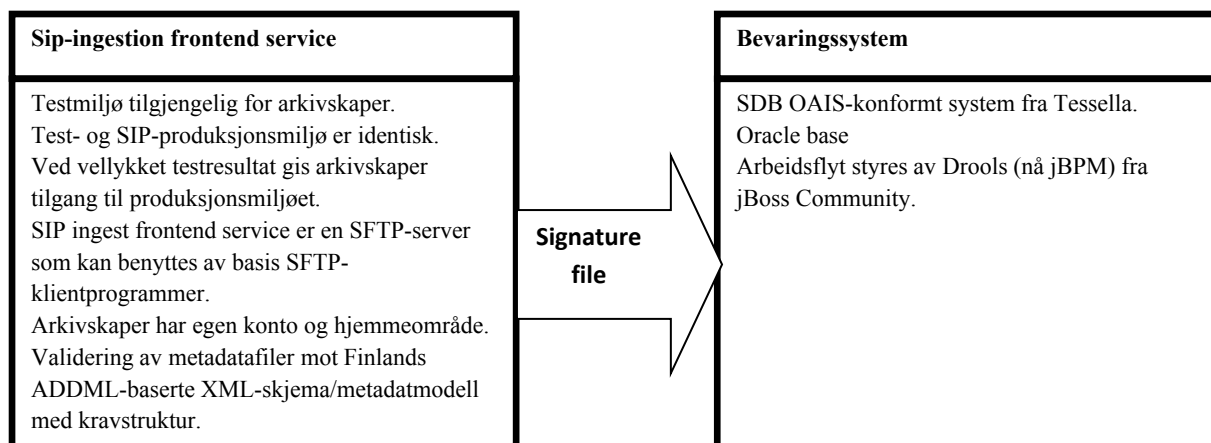
3. Tekniske retningslinjer for overføring av digitale arkivpakker til Riksarkivets bevaringssystemer.

Leverandørene av elektroniske arkiv- og saksbehandlingssystemer har mulighet til å få sine systemer sertifisert. Denne godkjenningsordningen er outsourcet til et internasjonalt anerkjent akkrediteringsselskap. Et slikt sertifikat/en slik godkjenning er gyldig i tre år. Etter dette kan leverandøren søke om ny godkjenning for produktet. Leverandørene må levere konsekvensanalyse til akkrediteringsselskapet hvor det beskrives eventuelle nye funksjoner som blir implementert og som påvirker egenskapene i systemet som allerede er blitt sertifisert. Hvis det er slik at det kun er brukergrensesnittet som er endret, vil ikke dette fordre ny godkjenning. Hvis det på den annen side er skapt vesentlige endringer i prosessstyring eller i metadatamodell, må dette søkes om godkjenning for påny – i realiteten er re-sertifisering av produktet. De to prosessene for godkjenning er identiske, men enklere og raskere i sistnevnte tilfelle.

Den finske modellen er interessant i forhold til områdene generering av digitale arkivpakker og testing, ettersom den har etablert en ordning hvor arkivskaper er gitt ansvaret for komparasjon av utrekket opp mot kravmodellen, samt teknisk test før innleveringspakken kvitteres ut for godkjenning til Riksarkivet¹³.

¹³ SIP (Submission Information Package) er et OAIS-basert begrep. OAIS står for *Open Archival Information System*. OAIS er en modell for å innlemme, administrere og bruke bevart arkivmateriale i et arkivdepot. Den beskriver et depots funksjoner, prosesser og informasjonsflyt med fokus på integritetssikring og definerer opplegg for vedlikehold innenfor et kontrollert miljø. OAIS-modellen beskriver bevaringsobjekter med vekt både på deres konseptuelle og tekniske aspekter. OAIS-baserte digitale sikringsmagasin implementeres i disse dager i Norge gjennom DIAS-prosjektet. DIAS står for *digital arkivpakkestruktur* og er gjennomført som et formalisert samarbeidsprosjekt mellom Riksarkivaren og fire kommunale aktører – Oslo byarkiv, Bergen byarkiv, IKA Trøndelag og IKA Møre og Romsdal – for å etablere et felles norsk rammeverk for digitale arkivdepoter.

Den finske modellen for digital arkivpakke-generering og test kan illustreres på følgende vis:



Det svenske riksarkivet har utviklet verktøyet Ralf (Riksarkivets LeveransFörberedelse-verktyg) for leveransekontroll av arkivuttrekk. Ralf er likhet med tilhørende skjemaer tilgjengelige fra det svenske riksarkivets nettsted. Ralf utfører grunnleggende kontroller

og danner en innleveringspakke (SIP). En grunnleggende forskjell i den norske og svenske modellen er at i Sverige må arkivskaper selv utføre testingen. Testingen kontrollerer kravoppfyllelse som er tilnærmet identisk de krav vi stiller i Riksarkivarens forskrift kapittel VIII. En sammenlikning av den svenske og norske modellen for uttrekksproduksjon er interessant ut fra ønsket om å optimalisere løsninger, redusere ressursbruk hos arkivskaper og arkivdepot og derigjennom effektivisere forvaltningen knyttet til langtidsbevaring av elektroniske arkiv.

Av avgjørende betydning for vår fremtidige håndtering av arkivpakker fra statlig og kommunal sektor er en reduksjon i tids- og ressursbruk knyttet til såkalt testing av innholdet i disse pakkene, inkludert dialog med arkivskaper for tilpasning av uttrekket henimot endelig godkjenning. Det svenske riksarkivet definerer denne testingen som mapping. Innenfor en semantisk begrepsterminologi er det konsistensen mellom arkivstandarden og dataleveransen vi kontrollerer. Det svenske riksarkivet har bakt inn i Ralf de kravelementene vi stiller i kapittel VIII av Riksarkivarens forskrift. Det vesentligste for oss her i Norge er at det svenske riksarkivet har flyttet tyngdepunktet for testing fra riksarkivet og ut til forvaltningen. Hva som gjenstår å teste i riksarkivet er bare hvorvidt det er en velformet arkivpakke, samt kontroll av autentisitet (sjekksum).

Tvinger Noark-standarden oss inn i en mer kompleks struktur- og innholdsverifisering enn hva som er nødvendig i de øvrige nordiske land? Betinget ikke dette en komparasjon av deres uttrekks- og godkjenningsprosedyrer med vår testing av tekniske og Noark-relaterte krav? Hvor ligger den semantiske forskjellen mellom kravene i Noark, kapittel VIII i Riksarkivarens forskrift og strukturen i innholdet? Det virker åpenbart at en slik forskjell

ikke skulle være nødvendig. Uttreksmetodikken er enten ikke tunet inn mot modellen i Noark eller Dias-skjemaene, ellers så har ikke SIP-genereringen ute hos arkivskaper luket vekk inkonsistens og avvik i databasen i forhold til kravene i standarden og Riksarkivarens forskrift.

Problemstillinger knyttet til autentisitet

Interpares-prosjektet definerer et fundament for autentisitet i spørsmålet om elektroniske arkiv i bevaringsperspektiv. Dette fundamentet utgjøres av at man kan demonstrere at ingen vesentlige endringer i den elektroniske mappen eller saken har funnet sted forut for innsendelsen av materialet¹⁴. En slik attest kan kun utstedes hvis det er etablert loggfunksjoner som bekrefter at enhver potensiell endring er identifisert og dokumentert. En slik loggfunksjon er etablert i Noark 5. Artikkelforfatteren har diskutert problemstillingen i et møte¹⁵ med Dr David Giaretta¹⁶ og Dr Victoria Lemieux¹⁷ i desember 2013. Begge er av den oppfatning at autentisitet ene og alene hviler på arkivskapers utstedelse av en bekreftelse. I følge Interpares kan dette kun skje ved at man vet hvilke elementer som utgjør det elektroniske arkivet når det ble fattet beslutning om at sakene eller mappene skulle velges ut for langtidsbevaring. Etter dette oppstår utfordringen som ligger i å demonstrere at ingen av endringene som kan og må inntreffe i ettertid påvirker vår evne til å demonstrere sakene eller mappenes identitet og integritet¹⁸. Spørsmålet knyttet til etterprøvbarehet av det enkelte objekts autentisitet før og etter det ble registrert i et arkivsystem er en helt annen problemstilling. Denne type komparasjon er av teknisk karakter og kan løses ved bruk av PKI¹⁹.

Med elektroniske arkiv gjør utfordringene knyttet til bevaring det nødvendig for bevaringsinstitusjonen å kunne produsere og vedlikeholde dokumentasjon som beskriver reproduksjonen av arkivene, slik at dette underbygger institusjonens bekreftelse på at den elektroniske saken eller mappen er autentisk²⁰. Det følger av dette at en kopi av en autentisk elektronisk sak eller mappe er autentisk hvis dette bekreftes av bevaringsinstitusjonen, og hvis en slik bekreftelse er understøttet av bevaringsinstitusjonens evne til å demonstrere at alle krav for fremstilling av autentiske kopier er ivaretatt. I kraft av denne bekreftelsen er kopien ansett å være identisk med saken eller mappen den reproducerer inntil bevis på det motsatte eventuelt kan fremlegges. Dette er sammenfallende med algoritmekomparasjon knyttet til sjekksummer²¹. I norsk sammenheng er det svært interessant å analysere om kravstrukturen i Noark danner premissene for et slikt rammeverk som Interpares definerer.

¹⁴ Interpares Book D part 1, side 22 og 23.

¹⁵ Møtet fant sted i Den Haag 06.12.2013 forbindelse med det UNESCO/International Council on Archives-initierte Digital Roadmap-prosjektet.

¹⁶ Leder for APA/APARSEN (den EU-finansierte Alliance for Permanent Access) og leder av arbeidet med utviklingen av OAIS-modellen.

¹⁷ Assistant Professor and Acting Director for Media and Graphics Interdisciplinary Centre [MAGIC] ved Universitet i British Columbia.

Problemstillinger knyttet til databasedump

Databasedump er ikke er en tilfredsstillende metode for fremtidig verifisering av autentisitet, ettersom en karakteristikk eller beskrivelse av de ulike elementene må ha funnet sted forut for dannelsen av den digitale arkivpakken. Det følger videre at dagens ordning med uttrekk produsert uten bevaringsinstitusjonens tilstedeværelse og bekreftelse på autentisitet i uttrekksproduksjonen fjerner vårt grunnlag for fremtidig bekreftelse på det elektroniske arkivets autentisitet. Det fremtidige grunnlaget for autentisitetsbekreftelse i en bevaringsinstitusjon hviler på dens evne til å loggføre endringer, og å demonstrere at disse endringene ikke i vesentlig grad har rokket ved det elektroniske materialets autentisitet²².

DIAS-skjemaene er deskriptive, dvs at det gis en arkivbeskrivelse ved genereringen av en digital innleveringspakke (OAIS-modellens Submission Information Package, eller SIP). En databasedump vil mangle denne beskrivelsen, som må foreligge før mottak i depot (Ingest²³) og generering av en depotbasert arkivpakke (AIP), hvis ikke bryter vi fullstendig med den stringente og koherente linjen som både OAIS og DIAS har lagt opp til (koherent semantikk hele veien fra SIP til AIP til DIP²⁴). Vi vil innenfor et slikt scenario importere ukurante SIPer som igjen vil gi ukurante AIPer som ikke har ønsket beskrivelse, hvilket igjen vil vanskeliggjøre dannelsen av arkivpakker for tilgjengeliggjøring, eller DIPer. Det skal ikke være akseptabelt å innlemme en ukurant SIP i et Oais-basert digitalt sikringsmagasin. Konseptet ukurant SIP bryter ned autentisitetsprinsippet. Den åpner i tillegg opp for en videreføring av omfattende manuelle tilpasninger internt i en depotinstitusjon som vi ut fra ressursmessige hensyn må søke å eliminere. Den store gevinsten DIAS-skjemaene representerer, er i tillegg til standardiseringen og forutsigbarheten, muligheten for automatikk i forvaltningen av de digitale arkivpakkene.

¹⁸ Se forøvrig David Giaretta: *Advanced Digital Preservation*, 2011 – Background to Authenticity, s 204

¹⁹ PKI betyr Public Key Infrastructure, som er et rammeverk for utstedelse, administrasjon og bruk av digitale sertifikater over datanettverk.

²⁰ Dette er et tema som analyseres i detalj i f eks *Authenticity in a Digital Environment*, the Council on Library and Information Resources, 2000.

²¹ Problemstillingen drøftes i en depotsetting, definert under termen 'Persistent identifiers' i *A DRIVER's Guide to European Repositories*, edited by Kasja Weenink, Leo Waaijers and Karen van Godtsenhoven, Amsterdam University Press 2008, s. 164

²² AIP= Archival Information Package, igjen et OAIS-begrep. En AIP dannes av en SIP (se fotnote 6).

²³ Ingest er selve aksesjonen eller mottaket av det digitalt skapte materialet – det å bringe materialet inn til forvaltningssystemet for det digitale sikringsmagasinet.

²⁴ DIP=Dissemination Information Package, en OAIS-betegnelse. DIP er en digital arkivpakke satt sammen for tilgjengeliggjøring.

Uttrekk fra fagsystemer

Fra et teknisk og arkivfaglig perspektiv er fagsystemer samtlige systemer som ikke følger kravmodellen i Noark. Databasestrukturen er dermed i utgangspunktet uforutsigbar i det øyeblikket et uttrekk skal genereres, noe som medfører en annen fremgangsmåte for fagsystemer enn for Noark-systemer. Arkivdanningen i fagsystemene har ikke fulgt en predefinert struktur som i Noark, og i en uttrekkssituasjon må en utføre en manuell kartlegging av de ulike elementene i basen man ønsker å trekke ut og bevare. Dette avhenger helt av systemets beskaffenhet, og det er helt klart at måten dataene og eventuelt dokumentene er strukturert på i Infotrygd-systemet i NAV²⁵ fra 1967 og samme etats Pensjons-system, Pesys, implementert i 2008, er grunnleggende forskjellig.

I uttrekksfasen starter en manuell prosess som beskriver hvert enkelt element i basen²⁶. I de fleste tilfelle har teknisk personell fra de kommunale eller statlige depotinstitusjonene måttet sitte sammen med systemkonsulenter fra de ulike etatene, og sammen identifisert og kartlagt de ulike elementene som skal bevares, og hvordan det hele er strukturert i systemet. Man påfører en teknisk arkivbeskrivelse på de ulike metadataelementene og relasjonen mellom dem, i tillegg til at man manuelt må definere hvordan relasjonene mellom metadataelementene og eventuelle filer (digitale objekter i form av f.eks. tekstdokumenter, kart, lydfiler osv.²⁷) skal struktureres ved uttrekk. Fra et overordnet perspektiv er vi ute etter det samme resultatet uansett om det dreier seg om et uttrekk fra et fagsystem eller et Noark-system. Filene på en frittstående Mac fra en kjent kulturpersonlighet som Arne Nordheim vil kunne trekkes ut, bevares for fremtiden og tilgjengeliggjøres som et fagsystem. Tilnærmingen er lik den som skisseres i innledningen i tilknytning til e-post og løse filstrukturer.

Utfordringer knyttet til tidsspennet mellom den elektroniske arkivdanningen og overføringen til bevaringsinstitusjonen

Arkivforskriften § 5-2 presiserer forskjellen mellom avlevering og deponering av elektronisk arkiv. Regelverket er utformet lenge før forvaltningen fikk erfaringsgrunnlag med utfordringene som ligger i å ha et stort tidsspenn mellom den elektroniske arkivdanningen og langtidsbevaringen. Det var ikke Noark-spesifikasjonene eller Noark-standardens kompleksitet som lå til grunn for å skyve testingen langt frem i tid. Det var tvert imot den arkivfaglige tilnærmingen til inndeling i arkivperioder som var definerende, i tillegg til frykten for altfor store datamengder. Begge disse prinsippene ble nedfelt i Noark allerede i 1984²⁸.

²⁵ Trygdeetatens system fra 1967 som fortsatt er i drift i NAV.

²⁶ Ved hjelp av ADDML.

²⁷ Tillate filformater er definert i Riksarkivarens forskrift kap IX, §§ 17-20.

²⁸ Noark 1, side 21.

I Riksarkivets budsjettforslag for 2015 lå søknad om midler til etablering av et Signalprosjekt²⁹. I dette prosjektet ligger planene om å etablere en ordning for det man kaller synkron overføring av data og elektroniske dokumenter fra forvaltningen til de respektive kommunale depotinstitusjoner og Riksarkivet. Som bakgrunn for forslaget fra Riksarkivet, som for øvrig har bred støtte også fra KS, ligger utfordringene knyttet nettopp til langtidsbevaringen og tilgjengeliggjøringen av digitalt skapt materiale. Forslaget adresserer, uten å uttrykke det eksplisitt, nødvendigheten av å etablere nye metoder og ny infrastruktur for overføring av digitale arkivpakker til bevaringsinstitusjonene, både innenfor kommunal og statlig sektor.

Synkron overføring av digitalt skapt materiale utfordrer den tradisjonelle arkivfaglige oppfatningen av at bevaring er noe som ligger langt fremme i tid. Den flytter bevaringsperspektivet fremover til selve arkivdanningen, både funksjonelt, teknisk og arkivfaglig. Samdok-prosjektet³⁰ har i tillegg brakt opp tilgjengeliggjøringen av dataene som element i selve arkivdanningen, ved å utrede bruk av RDF³¹ som en komponent i Noark 5³². Denne metodikken støper vår fremtidige evne til tilgjengeliggjøring av arkivmaterialet inn i selve arkivdanningen. RDF som format for langtidsbevaring fordrer kun en analyse av hvilke muligheter som ligger i XML transformert over i RDF³³.

Konklusjon

Dagens testregime har frem til nå medført et alt for stort tidsspenn fra det øyeblikket uttrekksproduksjonen initieres ute hos arkivskapers og frem til det er endelig godkjent i Riksarkivet, og dette må endres. Det er i den sammenheng påkrevet å understreke følgende forhold:

- Kvaliteten på uttrekkene har frem til nå vært for lav, og det gjelder for fagsystemuttrekk produsert av arkivskaper i like stor grad som for Noark-uttrekk.
- Selve den tekniske prosessen med gjennomkjøring av uttrekk i testverktøyet er ikke spesielt tidkrevende.
- Dialogen med uttrekksprodusent eller arkivskaper for å få til et nytt, forskriftsmessig strukturert uttrekk etter at det første er avvist, tar for lang tid.

²⁹ http://samdokdotcom.files.wordpress.com/2014/01/signalprosjekt_versjon-0-88_10-01-2014.pdf

³⁰ <http://samdok.com/>

³¹ RDF (Resource Description Framework) er en gruppe av W3C-spesifikasjoner som beskriver en modell for å representere metadata, data om data. Spesifikasjonene brukes i mange ulike sammenhenger for å beskrive informasjon og er et viktig element i visjonen om semantiske web.

³² http://samdokdotcom.files.wordpress.com/2014/06/sturtzel_to-rdf-or-not-to-rdf-17-06-2014.pdf

³³ RDF 1.1 XML Syntax: <http://www.w3.org/TR/rdf-syntax-grammar/>

- Standardiseringen av analysearbeidet har blitt forpurret ved at feilene har vært for uensartede, og at terskelen for hva som kan aksepteres av feil ikke egentlig har vært nøyaktig definert frem til nå.
- Gjennomgangen av feil og avvik, og analysen av disse, har frem til nå vært manuelle og svært tidkrevende operasjoner.
- Det har versert en oppfatning om at testløpene er umulig å standardisere. Det har i tillegg vært antatt at et testverktøy ikke kunne være ferdigutviklet uten behov for å gjøre spesielle tilpasninger for hvert enkelt uttrekk. Årsaken er at avvikene i uttrekkene pr i dag ikke har vært ensartede.
- En har gått ut fra at de ulike generasjonslagene i basene der ute (fra N3 til N5) har vanskeliggjort et endelig og definitivt oppsett i testverktøyet.

Langtidsbevaringen av elektroniske arkiv i betydningen Noark beveger seg langs en akse av kvalitetskrav som spenner fra selve arkivdanningen via uttrekksproduksjonen, testingen, bevaringen og til syvende og sist tilgjengeliggjøringen. Uttreksproduksjonen, innbefattet metodikk og verktøy, har i betydelig omfang vært prisgitt leverandører og fritts-tående konsulenter. Kostnadsrammen for arkivskaperene knyttet til uttrekk varierer fra 50 000 til 1,5 millioner kroner³⁴. Samspillet mellom uttrekksproduksjonen og testregimet i arkivverket har i for liten grad vært belyst og satt på dagsordenen frem til de siste par-tre årene. Satt på spissen kan vi hevde at elektronisk arkivdanning, bevaring og tilgjengeliggjøring er tre sider av samme sak.

I den digitale verden gir det samtidig ingen mening i å operere med et stort tidsspenn mellom arkivdanningen og bevaringen.

Sammenliknet med andre land som ikke har noen standard for elektronisk arkiv, står Norge i en privilegert særstilling ved at vi ikke bare kan definere metadataelementer og relasjonene mellom dem og de digitale objektene skal forholde seg til hverandre. Fordi vi gjennom utarbeidelsen av denne standarden har definert hvordan arkivdanningen skal modelleres, setter vi i realiteten arkivskaper i stand til selv å utføre testingen. De vet hva som skal være med i uttrekket, og de er takket være den pre-definerte modellen Noark er, bevisst hvilke elementer som befinner seg i det elektroniske arkivet. Alt dette muliggjør en effektivisering, ved at et standardisert og optimalisert testverktøy kan utvikles for arkivskaperene. For å komme dit må det etableres en standardisert kompetanseheving på dette området for arkivlederne, IT-arkivarene og systemadministratorene i offentlig sektor. Gevinsten ved en slik utvikling er åpenbar. Den vil redusere testløpene rent ressursmessig og tidsmessig til et absolutt minimum.

Det er svært få depotinstitusjoner internasjonalt som vier spesielt stor oppmerksomhet til testing av arkivuttrekk. Testingen i depotinstitusjonen består som regel av virustest,

³⁴ Tall fremsatt av representanter fra arkivskaperene ved bl. a Justis- og politidepartementets fagsamling 23. september 2014, Norsk Arkivråds kurs Periodisering og deponering av elektronisk arkiv, Trondheim 16. og 17. september 2014 og i kommunikasjon mellom arkivskaperene og medarbeidere i seksjon for digitalt depot i Riksarkivet i 2014.

sjekksumkontroll og versjonskontroll. Både i Sverige, Finland og Estland skal arkivskaper ha foretatt testingen og dokumentert feilfritt resultat før innlevering til depot. Det virker naturlig å flytte arbeidet knyttet til verifisering, validering og autentisering fra bevaringsinstitusjonen og frem til arkivskaper, noe som også er i harmoni med både Interpres og OAIS-modellen. Slik denne artikkelen forsøker å vise, vil det i tillegg være hensiktsmessig å se på mulighetene for å etablere en sertifiseringsordning der Noark-leverandørene må søke fortløpende godkjenning ved enhver versjonsendring i Noark-systemene. Det er nærliggende å tro at dette vil utelukke versjonshybrider og avvik fra standarden som har brakt så store problemer i testfasen her til lands. I tillegg ser vi konturene av en helt ny tilnærming til overføringen av data fra forvaltningen til depot, både rent metodisk og teknisk-funksjonelt. Samlet sett nødvendiggjør disse nye perspektivene en fullstendig revidering av den tradisjonelle tilnærmingen til bevaring og tilgjengeliggjøring av data som adskilte størrelser som tidsmessig er plassert langt frem i forhold til selve arkivdanningen.

Referanser

Sherry L. Xie: Building Foundations for Digital Records Forensics: A Comparative Study of the Concept of Reproduction in Digital Records Management and Digital Forensics, *American Archivist*, Volume 74, Number 2 / Fall-Winter 201.

Audrey Novak: Fixity Checks: Checksums, Message Digests and Digital Signatures, Digital Preservation Committee, 2006: <http://www.docstoc.com/docs/41459532/Fixity-checks-Checksums-Message-Digests-and-Digital-Signatures---PDF>

Wakeman, Gudes, Damsgaard Jensen, Crampton: Trust Management V: 5th IFIP WG 11.11 International Conference, IFIPTM 2011, Copenhagen, Denmark, June 29 - July 1, 2011, Proceedings, Springer 2011.

Mads Nygård, Torbjørn Skramstad et al: Development and Validation of Requirements for Evidential Value for Assessing Trustworthiness of Digital Records over Time, *Journal of Information*, volume 16.

Luciana Duranti, Corinne Rogers: Trust in digital records: An increasingly cloudy legal area, *Computer Law & Security Review*, Volume 28, Issue 5, October 2012

Mads Nygård, Guttorm Sindre : Empirical Study on Quality Requirements of Migration Metadata, i *Informatics Engineering and Information Science, Part III: International Conference, ICIEIS 2011, Kuala Lumpur, Malaysia, November 12-14, 2011 –Proceedings*, Springer forlag 2011.

Peter B. Hirtle: Archival Authenticity in a Digital Age, i *Authenticity in a Digital Environment*, the Council on Library and Information Resources, 2000.

David Giaretta: *Advanced Digital Preservation*, Springer forlag 2011.

Alistair Tough, Michael Moss: *Record Keeping in a Hybrid Environment: Managing the Creation, Use, Preservation and Disposal of Unpublished Information Objects in Context*, Tough & Moss 2006.

Sangchul Song, Joseph JaJa: Techniques to audit and certify the long-term integrity of digital archives, *Int. J. on Digital Libraries* 10(2-3), 2009.

ISO 13008:2012 - Information and documentation - Digital records conversion and migration process:

http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=52326

A DRIVER's Guide to European Repositories, edited by Kasja Weenink, Leo Waaijers and Karen van Godtsenhoven, Amsterdam University Press, 2008.

Riksarkivarens testrapporter for Noark 3 og Noark-4-systemer i perioden 2008 - 2014. Testrapporter for prøveuttrekk ifm midlertidig og endelig godkjenning av Noark 5-systemer 2012-2013.

Dias-skjemaene: <http://www.arkivverket.no/arkivverket/Arkivbevaring/Elektronisk-arkivmateriale/DIAS-prosjektet>

Funksjonsbeskrivelse ESSarch: <http://www.arkivverket.no/arkivverket/Arkivbevaring/Elektronisk-arkivmateriale/DIAS-prosjektet>

Kravspesifikasjon for Noark-4-testverktøyet ArkN4

ADDML - Archival Data Description Markup Language: <http://www.arkivverket.no/arkivverket/Arkivbevaring/Elektronisk-arkivmateriale/Standarder/ADDML>

Brev til fra Riksarkivet til Noark-leverandørene 13.09.2010.

Noark-4 del 1 og 2: <https://www.arkivverket.no/arkivverket/Offentleg-forvaltning/Noark/Noark-4>

Noark 5 v 3.1 med vedlegg: <https://www.arkivverket.no/arkivverket/Offentleg-forvaltning/Noark/Noark-5>
Kapittel VIII og IX i Riksarkivarens forskrift

Interpares Book D part 1, side 22 og 23: <http://www.interpares.org/ip2/book.cfm>

REFERENCE MODEL FOR AN OPEN ARCHIVAL INFORMATION SYSTEM (OAIS)- RECOMMENDED PRACTICE CCSDS 650.0-M-2 – Magenta Book June 2012.

ISO 16363:2012 - Audit and certification of trustworthy digital repositories - CCSDS 652.0-M-1 - September 2011

Rapporten FGS Paketstruktur Förvaltningsgemensam Specifikation (FGS) för paketstruktur för e-arkiv under Prosjekt e-arkiv og e-diarium (eARD)versjon 2.0 2013-01-14 fra det svenske riksarkivet.

Det svenske Riksarkivet, IT-avdelningen; Användarhandledning för RALF av 07.10.2013
“Long-term preservation permission procedure for governmental agencies”, National Archives of Finland, presentasjon for de nordiske riksarkivene ved El-arkivseminar 2011 Selfoss, Island 26.-27. mai 2011

Møte mellom Dr David Giaretta, Dr Victoria Lemiux og Olav Sataslåtten i forbindelse med UNESCO/ICA's Digital Roadmap Conference, Den Haag, 06.12.2013.

E-postkorrespondanse Sataslåtten - Merenmies i det finske riksarkivet 07-18.01.2014.

Det finske riksarkivet: Sähke2 certification criteria: Records managemens schedule system, Operative ERD-MS system, Preservation system (not long-term), EDRMS requiremens, Metadatamodel, XML schema, SÄHKE2 CERTIFICATION CRITERIA ARCHIVING SYSTEM v. 1.0 (15.3.2012) –foreløpig og uoffisiell oversettelse mottatt fra det finske riksarkivet 15. 01. 2014.

Graham Moore: Noark5 in RDF – presentasjon på SAMDOK2014.

Resource Description Framework (RDF): <http://www.w3.org/RDF/>