

Prosjektrapport PEANUTS

Private Elektroniske Arkiver i Norge – Uttrekk, Testing, Sikring

Arbeiderbevegelsens arkiv og bibliotek

2016

Innhold

Del 1 – Innledning.....	s. 3
- <i>Prosjektets formål.....</i>	<i>s. 3</i>
- <i>Bakgrunn for prosjektet.....</i>	<i>s. 3</i>
- <i>Fil- og mappearkiver – organisasjon og teknologi.....</i>	<i>s. 4</i>
- <i>Attac Norge – et typisk case.....</i>	<i>s. 5</i>
- <i>Archivematica/AtoM.....</i>	<i>s. 5</i>
Del 2 – Deponering og ordning.....	s. 7
- <i>Deponering av Attacs elektroniske arkiv.....</i>	<i>s. 7</i>
- <i>Ordning og bevaringsvurdering i Windows.....</i>	<i>s. 9</i>
Del 3 – Behandling av materialet ved hjelp av Archivematica.....	s. 13
- <i>Arbeidet med Archivematica og AtoM.....</i>	<i>s. 13</i>
- <i>AIC-funksjonen i Archivematica.....</i>	<i>s. 17</i>
- <i>Riksarkivarens forskrift.....</i>	<i>s. 18</i>
Del 4 – Vurdering av AtoM.....	s. 20
- <i>Brukervennlighet – en historikers inntrykk av ArchivesCanada.....</i>	<i>s. 20</i>
- <i>Prosjektgruppens oppfatning av søkemulighetene i AtoM.....</i>	<i>s. 21</i>
- <i>Funksjoner – en interessant alternativ søkeinngang.....</i>	<i>s. 22</i>
- <i>Brukertilgang – muligheter for crowdsourcing?.....</i>	<i>s. 24</i>
Del 5 – Archivematica og AtoM-workshop.....	s. 27
Del 6 – Andre programmer.....	s. 28
Konklusjon.....	s. 30
Kildeliste.....	s. 33

Del 1 – Innledning

Prosjektets formål

Arbeiderbevegelsens arkiv og bibliotek (Arbark) søkte våren 2015 Arkivverket om prosjektmidler (utviklingsmidler for privatarkiv) for testing av depotprogrammet Archivemata og tilgjengeliggjøringsprogrammet Access to Memory (AtoM). Det ble innvilget til sammen 400.000 til prosjektet i 2015.¹

Formålet med prosjektet var å teste ut programmenes evne til å bevare og tilgjengeliggjøre et fil- og mappearkiv, samt å teste deres brukervennlighet. Forhåndshypotesen var at AtoM og Archivemata kunne fylle et metodologisk tomrom når det kommer til bevaring av fil- og mappearkiv, siden veldig mye av bevaringsarbeidet til nå har tatt utgangspunkt i databaser.

Prosjektet, som fikk navnet PEANUTS, har også nær tilknytning til SAMDOK (Samla samfunnsdokumentasjon) ved at en av Arbarks medarbeidere deltok (og fortsatt deltar) i arbeidsgruppen *Metodikk for høsting og langtidslagring av privatarkiv*.

Prosjektet hadde oppstart i september 2015, og ble sluttført i april 2016. Prosjektansvarlig har vært Frank Meyer. Prosjektgruppen har bestått av Martin Ellingsrud og Dag Juvkam. I tillegg var Lise Rødland en periode tilknyttet prosjektet for å ordne arkivet.

Bakgrunn for prosjektet

Bakgrunnen for prosjektet er den store mengden av private fil- og mappearkiver, med andre ord ustrukturert data, som bygger seg opp hos arkivskaperne.

For Arbarks del er det nødvendig å ta store grep for å sikre viktig samfunnsdokumentasjon. Fagbevegelsens organisasjoner, Arbarks kjerneområde, har mange fil- og mappearkiver på regionalt (fagforeninger) og lokalt (bedriftsklubber) nivå. På sikt er det overhengende fare for at regionale og lokale arkiver som Arbark har bevaringsansvar for, ikke vil bli bevart.² Vi mener at prosjektet som sådan har overføringsverdi både for IKAer og andre kommunale

¹ Kr. 100.000 gikk til Attac Norge, Arbarks samarbeidspartner, som formelt sto som arbeidsgiver for Lise Rødland, prosjektmedarbeideren som ordnet arkivet. Beløpet dekket lønn og andre utgifter vedrørende Rødlands ansettelsesforhold.

² Arbark har i tillegg til et nasjonalt bevaringsansvar også ansvar for fylkene Oslo, Akershus og Buskerud.

arkivinstitusjoner og statsarkiver, som vil møte regionale og lokale private arkivskapere som kan tenkes å ha slike arkiver.

Fil- og mappearkiv – organisasjon og teknologi

Fil- og mappearkiv er ikke den type arkiver arkiv-Norge primært ser på som elektronisk arkiv. Oppmerksomheten har først og fremst vært rettet mot databaseteknologi, med andre ord teknologi for å lage strukturert data, som lenge har blitt brukt i sak-/arkivsystemer (og fagsystemer) for å lage en bestemt struktur. Fil- og mappearkiver er derimot såkalt ustrukturert data, og kan lettere endres på enn mer rigide databasetabeller.

Her er det sammenheng mellom teknologi, arkiv og organisasjon. De store aktørene har behov for og anskaffer ofte avanserte systemer med en viss struktur (les: ofte databasesystemer) for lagring av informasjon. De små har ofte ikke behov for (eller ser ikke behov for) avanserte systemer og klarer seg med enkle fil- og mappearkiver. De har som regel heller ikke råd til noe mer avansert. Store aktører har også behov for mer stabile rutiner enn små, på grunn av datamengden som behandles. De små kan ha mer *ad-hoc*-pregede ordninger i sine arkiver.

Det vil alltid være flytende grenser mellom hva som er en stor arkivskaper og hva som er en liten. Vi finner fil- og mappearkiv i store organisasjoner og bedrifter også, men disse kommer da i tillegg til de arkivene som dannes i de avanserte (fag)systemene³, siden det er de avanserte systemene som oftest dokumenterer kjernevirksomheten eller kjernefunksjonene til en større virksomhet.

Privatarkiver kan være fulle av ulike produksjonsformater som ikke egner seg for langtidslagring. I den grad det skal være mulig å håndtere slike arkiver på en god måte, bør depotprogrammet klare å konvertere flest mulig formater (eventuelt kan produksjonsformatet også følge med i en avlevering). I all hovedsak er det ikke nødvendig å tenke alt for mye på filformater. Vi tror at det ofte vil dreie seg om for eksempel tekstfiler, regneark⁴, bilde- og lydfiler og filmer. Dette vil variere fra arkivskaper til arkivskaper. En grafisk designers personarkiv kan være fult av spesielle filformater, mens et arkiv fra en fagforening i stor grad vil bestå av tekstformater.

³ Et eksempel her kan være løpende prosjektarkiv for grupper eller mer eller mindre viktige dokumenter som den enkelte medarbeider lagrer på sitt område.

⁴ Regneark i seg selv er en utfordring. PDF/A er her ikke alltid et godt egnet langtidsformat hvis man vil ha med formler og/eller regnearket går utover en A-4-side.

Attac Norge – et typisk case

For å studere bruken av fil- og mappearkiver, valgte Ar bark organisasjonen Attac Norge, som kan betraktes som typisk for mønstre man vil finne hos mange små organisasjoner.

Attac er en verdensomspennende bevegelse, som oppstod i Frankrike i 1998, med det hovedmål å presse frem en mer demokratisk kontroll over verdens finansmarkeder, spesifikt blant annet gjennom innføring av såkalt Tobin-skatt. I Norge ble organisasjonen etablert i mai 2001. Attac Norge kan sies å tilhøre den ”nye” typen sosiale bevegelser. Slike bevegelser kjennetegnes av at de er ideelle og ekspressive, og i mindre grad enn tidligere fremmer materielle mål.⁵

I dag har Attac Norge om lag 2 500 registrerte medlemmer. Organisasjonen er medlemsbasert og partipolitisk uavhengig. Det er leder av organisasjonen og to ansatte som driver hovedkontoret, så organisasjonen må karakteriseres som liten. Virksomheten er ofte kampanjedrevet – for eksempel drives det informasjonsvirksomhet om internasjonale handelsavtaler som er på forhandlingsstadiet (eksempelvis TTIP og TISA). For å få oppmerksomhet om sine saker må Attac være synlig både på sosiale medier og i mediebildet generelt. De ansatte ”lever for nuet” og har i mindre grad tid og ressurser til å etablere og revidere arkivrutiner.

Archivematica/AtoM

Programmene som skulle testes i PEANUTS-prosjektet er Archivematica og Access to Memory. Archivematica er et digitalt depotprogram utviklet av Artefactual Systems i Vancouver, Canada.⁶ I OAIS⁷-terminologi håndterer Archivematica primært SIP (mottak) og AIP (lagring), men også konstruksjon av DIPer (tilgjengeliggjøring).⁸

Archivematica i seg selv er bygd opp rundt mikrotjenester. Mikrotjenestene eller –oppgavene er konkrete bevaringsoppgaver, for eksempel utregning av sjekksum⁹ eller konvertering til bevaringsformater. Mikrotjenestene kjøres gjennom en kombinasjon av Archivematica Python

⁵ Kjeldstadli 2010, s. 19.

⁶ Se nettside: <https://www.artefactual.com/>.

⁷ CCSDS 2012. OAIS – Reference Model for an Open Archival Information System – er et rammeverk for lagring av digitalt materiale. Standarden ligger til grunn for en rekke andre standarder og digitale depotprogrammer som håndterer slikt materiale.

⁸ Riksarkivet 2011.

⁹ Sjekksummer brukes for å undersøke om digitale objekter har blitt utsatt for endringer.

scripter og ett eller flere verktøy med åpne kildekoder som befinner seg inne i Archivematica.

¹⁰ Slik sett kan Archivematica best sammenlignes med en overbygning, som har innebygd mange små deler eller programmer.

AtoM er programmet for tilgjengeliggjøring (DIP). I praksis kan AtoM sies å være en parallell til Arkivportalen. Man registrerer et hierarki av arkiv (*fonds*), serier, mapper og *items* (dokumenter). Her refereres det spesifikt til den internasjonale arkivbeskrivelsesstandarden ISAD(G), siden det var denne vi brukte i AtoM, i tillegg til ISAAR(CPF) (standarden for beskrivelse av arkivskapere). Det er også mulig å bruke andre standarder, for eksempel Dublin Core. Et interessant aspekt her er at ISDF, standarden for funksjonsbeskrivelse, er implementert i systemet.¹¹

Kort fortalt går prosessen ut på at man først ferdigbehandler arkivet i Archivematica for så å laste opp en DIP i AtoM, der arkivbeskrivelsen og de konverterte digitale filene kan ses på og aksesseres.

¹⁰ Artefactual Systems. "Micro-services". I *Archivematica wiki*.

¹¹ Artefactual Systems. "Descriptive standards". I *AtoM user manual (2.3)*.

Del 2 - Deponering og ordning

I det følgende tar vi for oss avleveringen av Attacs elektroniske arkiv. Her beskriver vi hva vi gjorde underveis i prosessen, samt hvordan vi jobbet med å forberede arkivet til prosessering i Archivematica. Vi kommer også inn på bevaringsvurdering av materialet og problemer knyttet til denne.

Deponering av Attacs elektroniske arkiv

I september 2015 kopierte to av Arbarks medarbeidere Attacs elektroniske arkiv. Kopieringen foregikk i Attacs lokaler. Konkret kopierte Arbarks medarbeidere fil- og mappearkivet fra Attacs maskin over på en ekstern, mobil harddisk. Perodeskilledato ble satt til 31.12.2014¹², ved hjelp av kopiprogrammet Beyond Compare.¹³ Dette programmet kan definere regler som går ut fra ”sist endret”-dato. Fanestien for regelsetting er *Rules – Folder Compare – Other Filters – Date*. Her kan man ta med ”filer eldre enn” eller ”filer nyere enn”. Det går an å sette begge regler samtidig, noe som gjør at man tar med uttrekk for eksempel fra 2002-2004, selv om arkivet strekker seg fra 2001-2015. På denne måten avleverte Attac arkivet sitt fra årene 2001-2014.

Attac beholdt etter eget ønske det ubehandlede arkivet hos seg. Siden dette var et prøveprosjekt og siden Arbark ikke har sanksjonsmuligheter, ble dette akseptert. Dette vil uansett neppe føre til store problemer senere, så lenge vi kan bruke Beyond Compare eller tilsvarende programmer til å hente ut det materialet vi ønsker.

Ved kopiering får nye mapper samme opprettelsesdato som da kopieringen fant sted. Ettersom disken er kopiert, fikk vi dermed ikke med informasjon om når mapper er opprettet. På dokumentnivå kunne vi imidlertid se opprettelsestidspunkt. Uttrekket hadde på kopieringstidspunktet nærmere 30 000 dokumenter. Filene bestod av forskjellige formater, men i hovedsak var det doc/docx, PDF, power point, excel og JPEG-filer.

¹² I ettertid ser vi at vi heller kunne satt perodeskilledato til 1.1.2015 for å være sikre på å fange opp alt. Å sette perodeskilledato til 31.12.2014 hadde ingen praktiske konsekvenser denne gangen, men det er verdt å huske på ved senere avleveringer.

¹³ Scooter Software.

På øverste nivå har arkivet denne strukturen¹⁴:

 100 Grunndokumenter	21.10.2015 15:59	Filmappe
 200 Administrasjon	21.10.2015 16:00	Filmappe
 300 Intern organisasjon	21.10.2015 16:01	Filmappe
 400 NORAD, UD, FrittORD	21.10.2015 16:01	Filmappe
 500 Politikk	21.10.2015 16:03	Filmappe
 600 Arrangementer	21.10.2015 16:04	Filmappe
 700 Utgående brev	21.10.2015 16:04	Filmappe
 800 Informasjonsarbeid	21.10.2015 16:14	Filmappe
 C-disk	23.10.2015 11:31	Filmappe

Arkivet hadde en hierarkisk forgreining nedover, med flere mappenivåer ned mot dokumentene. Strukturen nedover mot dokumentene var ved mottak noe tilfeldig. Noen dokumenter lå direkte under øverste mappe, men ofte måtte man gjennom flere mappenivåer for å komme frem til dokumentene. Noen av undermappene hadde ved mottak årstall i mappenavnet, men dette var heller unntaket enn regelen. Noen av klassene var langt mer omfattende enn andre. Særlig 500 Politikk og 800 Informasjonsarbeid, de to største klassene, var uoversiktlige og delvis uten arkivkode. Attac har i løpet av sin relativt korte virkeperiode utvidet sitt saksområde. Nye emner har løpende blitt lagt til uten en revidering av arkivnøkkelen, slik at nøkkelen gradvis har mistet betydning for gjenfinning og struktur.

Oppsummert kan vi si at arkivene til mindre organisasjoner med små ressurser ofte vil bære preg av å være noe kaotiske. Klassifisering bør, slik Arbarks prosjektgruppe ser det, være mest mulig generell for organisasjoner som Attac på grunn av den relativt hurtige utskiftningen av ansatte. Jo mer generell, desto færre koder og desto mindre sjanse for skjønnsutøvelse og feilarkivering. Ikke minst vil det ta kortere tid å lære opp nyansatte. Ulempen er at klassifiseringen kan bli for grovkornet, slik at skillet mellom ulike funksjoner og aktiviteter blir for uskarpt. Dermed kan klassifiseringssystemet få mindre verdi for gjenfinning osv. Men jo mer avansert system (les: flere koder), desto større er sjansen for at det brytes og kanskje etter hvert kolliderer. En slik risiko er åpenbart større hos små organisasjoner som Attac enn hos organisasjoner med egne arkivansatte.

Rådgivning overfor offentlige, særlig statlige arkivskapere er en aktivitet som arkivdepotene i Norge lenge har påtatt seg. Imidlertid, uten ekstra ressurser eller støtteordninger til å

¹⁴ C-disk-mappen inneholdt eldre arkivmateriale, det vil si materiale fra ca. 2001. Opprinnelig var arkivmaterialet på en separat disk, men vi betraktet det som en serie i Attacs arkiv og førte det sammen med resten av arkivet.

gjennomføre slik rådgivning overfor private arkivskapere, vil depotinstitusjoner stå i fare for å ende opp med en nokså tilfeldig og ensidig bevaring. Samfunnshukommelsen vil bli dominert av de rike og mektige organisasjonene og institusjonene, fordi andre aktører ikke makter å foreta avleveringer, eller fordi arkivene som avleveres viser seg å være lite tilgjengelig for utenforstående (det vil si arkivbrukere).

Arbark fikk i liten grad gitt råd til Attac angående arkivdanning, slik Arbarks søknad opprinnelig skisserte.¹⁵ Dersom Arbark skulle fulgt opp dette, måtte vi ha satt i gang et relativt omfattende og langsiktig arkivdanningsprosjekt, med analyse av funksjoner med mer, noe Arbark ikke hadde ressurser til. Vi valgte derfor å prioritere testingen av Archivemata og AtoM, ikke Attacs arkivdanning.

Ordning og bevaringsvurdering i Windows

En ordningsarbeider var i en periode knyttet til prosjektet for å ordne arkivet. Hun hadde bakgrunn fra Attac og dermed inngående kjennskap til både materialet og organisasjonen. Målet var å få mer struktur i materialet samt å foreta arkivbegrensning og bevaringsvurdering. Likevel, siden deler av arkivet i utgangspunktet var inkonsekvent organisert og fordi arkivet i liten grad har vært gjenstand for arkivbegrensning, forble det noe uoversiktlig også etter ordning.

Arkivet ble ordnet i Windows etter virustest og produksjon av sikkerhetskopi, slik Riksarkivet anbefaler.¹⁶ Ordningsarbeideren arbeidet med en arbeidskopi som Arbark hadde laget ved hjelp av Beyond Compare. Under ordningen ble det lagt til ytterår i mappenavn.

Bevaringsvurdering ble foretatt ut fra avtale med Attac, og gikk hovedsakelig ut på å fjerne alt materiale som hadde å gjøre med personalsaker og personaladministrasjon. Attac hadde ikke ønske om å overføre disse sakene til Arbark (til tross for klausuleringsmuligheten). I tillegg ble åpenbare arkivverdige filer, for eksempel testfilmer for opptak, fjernet.¹⁷

Umiddelbart etter virustest og produksjon av sikkerhetskopi anbefales bruk av DROID eller et tilsvarende program for å få en oversikt over filer og mapper. DROID er gratis og produsert av The National Archives i Storbritannia. DROID har for eksempel mulighet til å oppdage

¹⁵ Arbark 2015, s. 3.

¹⁶ Riksarkivet 2005, s. 6.

¹⁷ For nærmere drøftelse av arkivverdige filer, se for eksempel Riksarkivet 2005, s. 6.

kopier, identifisere (antatt) uleselige filer og lage forskjellige rapporter med mer som kan brukes i analysearbeidet av arkivet.¹⁸ Ettersom DROID var ukjent for Arbarks prosjektgruppe frem til april 2016, fikk vi ikke brukt programmet så mye som vi i etterkant kunne ønsket. Det at ordningsmedarbeideren hadde god kjennskap til Attac gjorde uansett at mye av arbeidet kunne gjøres relativt raskt manuelt. Likevel, DROID er et program som går inn i vår ”verktøykasse” for fil- og mappearkiver, og vi vil, etter å ha bygget opp mer kompetanse på programmet, sannsynligvis anbefale det til arkivskapere.

Arkivet hadde mange medlemslister (i Excel) som var trekt ut fra et medlemssystem. Det var Attac selv som genererte listene. Når det gjelder medlemsopplysninger, er det etter personopplysningsloven slettepåbud for personopplysninger straks formålet med behandlingen er gjennomført. Et unntak kan være hvis opplysningene antas å ha viktig historisk verdi.¹⁹ I Attacs tilfelle vil det imidlertid være nærliggende å tenke seg at listene bare blir brukt til umiddelbare, administrative formål, ikke forskningsmessige.

Datatilsynet ga nylig klare og strenge signaler om sletting av utgåtte medlemskap hos politiske partier, og partimedlemskap faller under kategorien sensitive personopplysninger.²⁰ Gitt likhetsprinsippet, er det grunn til å tro at Datatilsynets slettepåbud gjelder alle sensitive personopplysninger (med mindre annen lovgivning tilsier noe annet). Siden Datatilsynet har strenge slettingsregler for medlemsorganisasjoner som forvalter sensitive personopplysninger, har Arbark restriktiv politikk på mottak av elektroniske medlemsregistre. Ettersom et nytt EU-regelverk snarlig vil innføres, kan det bli aktuelt å revurdere denne politikken.²¹

Det er et noe åpent spørsmål om Attac er en organisasjon som faktisk forvalter sensitive personopplysninger i juridisk forstand, gitt at den ikke er et politisk parti. Vi bestemte oss her likevel for å se hva resultatet av en omfattende fjerning av medlemslister, med andre ord en streng tolkning av personopplysningsloven, ville gjøre med arkivet. Det ville gi oss en god mulighet til å vurdere slettepåbudet generelt sett, men det skal påpekes at erfaringen blir noe begrenset. Det er kun Attac det dreier seg om.

I bevaringsvurderingen viste det seg at del medlemslister i Excel hadde svært liten verdi for å dokumentere medlemsmassen. Noen medlemslister hadde for eksempel kun

¹⁸ The National Archives 2011. Arbark takker Pål Mjørund (Fylkesarkivet i Oppland/Opplandsarkivet Maihaugen) for tipset om programmet.

¹⁹ Personopplysningsloven § 28.

²⁰ Datatilsynet 2012 og 2015a; personopplysningsloven § 2, punkt 8.

²¹ Datatilsynet 2015b.

medlemsnummer. Andre igjen var åpenbart produsert for umiddelbare administrative behov, for eksempel for purring på medlemskontingent. Vi har vanskelig for å se at medlemslister med kun noen medlemmer (de som ikke betalte i en gitt måned i et gitt år) har en bevaringsverdi, siden det kun gir oversikt over noen medlemmer. Noen medlemslister kunne vi nok bevart, siden de dokumenterte deltakere på ulike arrangementer.

Det må understrekes at det her ikke er snakk om sletting i absolutt forstand.

Ordningsmedarbeideren førte nemlig alle dokumenter som ikke var bevaringsverdig til et kassasjonsarkiv (en kopi av mapper i arkivet, opprettet for at filene uten bevaringsverdi kan spores tilbake til opprinnelig plassering), slik at prosessen kunne kontrolleres og vurderes i etterkant. Siden dette var et prøveprosjekt, anså vi dette som nødvendig. Ved en permanent avlevering måtte vi ha vurdert nærmere kassasjonsarkivet opp mot lovverket – hvis vi da hadde kommet til at materialet faller inn under lovverkets definisjon av personsensitiv informasjon, måtte vi ha foretatt endelig sletting.

Uansett, vår vurdering er at det hadde lønnet seg med mer komplette uttrekk av medlemsmassen fra selve medlemssystemet, i det minste de med verv (som ikke er omfattet av slettepåbudet), og/eller statistiske opplysninger om medlemsmassen over tid.

19 prosent av materialet ble ikke bevart. Prosentandelen i seg selv er ikke av betydning, det er det kvalitative aspektet som må fremheves. Et arkiv fjernet for ”støy” blir presumptivt mer brukervennlig. Den store ”gevinsten” i så måte er arkivbegrensning, for eksempel fjerning av tomme filer. Det viste seg å være svært vanskelig å finne flere håndfaste kriterier for å sile ut materiale uten bevaringsverdi, ettersom strukturen i arkivet var såpass tilfeldig. For vårt vedkommende ble det da vanskelig å være sikker på hva man egentlig kasserte på overordnet nivå. Ordningsmedarbeideren måtte gå dypt ned i mappehierarkiet for å få nok informasjon til å vite hva som ikke var bevaringsverdig og hva som skulle bevares. Et eksempel kan være at klasse 500 Politikk og klasse 600 Arrangementer overlappet mye. Arrangementer var det også mange av i klasse 500.

Ordningsmedarbeideren i Windows må betraktes som det stadiet der arkivskaper selv rydder opp i arkivet (som sannsynligvis vil skje nettopp i Windows, ikke i et depotprogram), i tillegg til deler av det arbeidet som Riksarkivets rapport beskriver som en del av depotets arbeid.²²

²² Riksarkivet 2005, s. 5-6. Rapporten omhandler utfordringer knyttet til blant annet bevaring, ordning og katalogisering. Målgruppen for rapporten var depotinstitusjoner som mottar elektronisk materiale fra private arkivskapere.

Ordningsarbeideren hadde som nevnt bakgrunn fra Attac, og hun fokuserte dermed på at arkivet skulle være mest mulig til hjelp for organisasjonen. Det innebar at hun ikke fjernet like mye som vi ville ha gjort, for eksempel maler (dokumenter uten reelt innhold).

Ordningsmedarbeideren lagde også en ordningsrapport. Dette ble gjort på grunn av dårlige automatiske loggingsmuligheter i Windows og for å dokumentere hvorfor noe ble gjort. Rapporten i seg selv beskriver rent overordnet hvilke grep som ble gjort og hvilke problemer som ble støtt på.²³

²³ Et av de vanskeligere problemene var datomal, det vil si at konvertering av dokumentet førte til at datoen inne i selve dokumentet ble endret til dagens dato ved konvertering og/eller åpning av fil. Dette er en stor autentisitet utfordring Arbark ikke klarte å finne et godt svar på. Dette gjaldt noen av de eldre filene i Word, og de var relativt få. Vi har dermed foreløpig ikke noen grunn til å anta at slike datomaler vil være et stort problem (spesielt ikke i nyere arkivmateriale).

Del 3 - Behandling av materialet ved hjelp av Archivemata

I denne delen gir vi en beskrivelse av hva som ble gjort under prosesseringen i Archivemata. Fokuset vil være på problemer som oppsto og hva som ble gjort for å imøtegå disse problemene. Videre tar vi for oss AIC-funksjonen i Archivemata, før vi ser nærmere på om hvorvidt lovverket kan bidra til å veilede i deponeringsprosesser av denne typen.

Arbeidet med Archivemata og AtoM

Installering av Archivemata/AtoM lokalt krever Linux/Ubuntu-plattform. Arbark hadde innledningsvis problemer med denne installasjonen. Vi kontaktet derfor utvikleren Artefactual Systems Inc., som anbefalte oss å bruke en skyløsning i testperioden, slik at vi kunne benytte Windows. Løsningen innebar at vi leide lagringsplass, både for software og opplastet materiale, på en server lokalisert i Frankfurt. Vi lastet materiale opp til denne serveren via programmet Filezilla – en såkalt FTP-klient.²⁴ Vi leide 100 GB plass. Teknisk støtte fra Artefactual var inkludert. Kostnadene kom på kr. 65 000. Testperioden var fra 7. desember 2015 til 8. april 2016.²⁵ Etter 8. april hadde vi ikke lenger tilgang til serveren.

Skyløsninger er den siste tiden blitt diskutert i arkivmiljøet, spesielt på grunn av utførselsforbudet i arkivloven, som da gjelder offentlige arkiv, ikke private arkiver som ikke er særskilt verneverdige etter arkivlovens bestemmelser. Vi så det som hensiktsmessig å benytte oss av en skyløsning i en begrenset periode, særlig for å klare å gjennomføre prosjektet innen avtalt tid.²⁶ Ikke minst sparte vi IT-ressurser (både arbeidstid og hardware), som allerede var presset.

Etter et opplæringsmøte via nettet begynte vi å prøve ut programmet. Planen var å utføre tester for å finne ut hva som eventuelt måtte gjøres av endringer i programmet (på serveren) før vi begynte prosesseringen av Attacs arkiv. Konkret går materialet under prosesseringen gjennom en *pipeline*, hvor det først lastes opp til Archivemata og behandles i en

²⁴ En FTP-protokoll er en operativsystemuavhengig protokoll som brukes for å overføre filer mellom en FTP-klient og en server.

²⁵ Opprinnelig hadde vi tilgang til serveren til 7. mars 2016, men ettersom vi ønsket å holde en workshop for institusjoner i Oslo-regionen forlenget Artefactual tilgangsperioden til 8. april, uten ekstra kostnader. Vi fikk også økt lagringsplassen på serveren til 200 GB.

²⁶ Det at serverne var i Tyskland, gjorde at tilsvarende lovgivning angående personvern ble gjeldende, siden norsk og tysk personopplysningslovgivning bygger på samme EU-direktiv.

overføringsfase (Transfer). Deretter sendes det videre til Ingest, før det ender opp som ferdigbehandlede AIPer i Archival storage. Herfra kan pakkene sendes til AtoM som DIPer.

Testmappene hentet vi opp til Archivematica via Filezilla. På en av de første testene støtte vi på problemer da konverteringen av dokumentene feilet. Problemet viste seg å være en norsk bokstav – ø – i tittelen på en av dokumentfilene. Etter å ha undersøkt dette nærmere, fikk Artefactual løst problemet ved å tillate ”ø” på vår server.

Vi fikk også problemer med bokstaven ”å” i transfer-navnet. Artefactual hadde dette å si om problemet: ”The issue is that Archivematica currently handles multi-byte unicode sequences on a case by case basis, and we had not added å to that list. We are reviewing how we are supporting unicode to improve this.”²⁷ Den aktuelle mappen – *årsavslutning* – altså med liten å, kom aldri gjennom.

Problemet med spesielle norske bokstaver gjaldt både for transfer-, fil- og mappenavn. Bokstavproblematikken henger sammen med tegnstandardene som brukes i AtoM og Archivematica. Både i AtoM og Archivematica er UTF-8 anbefalt. De første 128 tegnene (US-ASCII) trenger én byte for å kode, mens de neste 1920 tegnene trenger to bytes.²⁸ Våre spesielle norske tegn inngår i den delen av standarden som trenger to bytes. Denne delen er, slik vi forstår det, ikke implementert i standardversjonen av Archivematica, noe som betyr at programmet ikke håndterer norske bokstaver. Ting blir ytterligere komplisert av at Archivematica støtter seg på en rekke forskjellige *open source*-verktøy.

En annen problematisk sak er at Archivematicas standardversjon ikke konverterer til PDF/A-formatet. Artefactual er klar over problemet og har jobbet med det, men har foreløpig ikke funnet gode løsninger. Foreløpig, inntil tilgjengelige verktøy forbedres eller eventuelt nye kommer på banen, mener Artefactual at bevaring og tilgjengeliggjøring i opprinnelig Microsoft-format er å foretrekke (i standardversjonen).

Artefactual vil ideelt sett ha formater i standardversjonen som oppfyller fire krav:

- Formatet, eller nærmere bestemt formatspesifikasjonen, må være fritt tilgjengelig
- Formatet skal ikke være lisensiert eller patentert
- Etablerte depoter må bruke eller anbefale bruk av formatet

²⁷ E-post fra Artefactual Systems til Arbark 16.03.2016.

²⁸ ”UTF-8”. I *Wikipedia*.

- Det skal være en variasjon av konverteringsverktøy for formatet.²⁹

Dette får konsekvenser for valg av konverteringsverktøy, siden Artefactual vil ha *open source*-verktøy til å foreta konvertering. Når det gjelder PDF/A, er det da få verktøy som oppfyller en slik spesifisering for batch-konvertering, og de som gjør det (som OpenOffice) er avhengig av ”reverse engineering of closed specifications”³⁰. Ikke minst kan det bli forandringer i formateringen hvis man for eksempel bruker OpenOffice, som igjen gjør at PDF/A-dokumentet ikke nødvendigvis er helt likt eller så likt som det ideelt sett bør være originalen for at konverteringen skal bli oppfattet som autentisk.³¹

Vi har her ikke til hensikt å gå inn på fordeler og ulemper med proprietær og åpen kildekode. Faktum er at Artefactual Systems vil benytte seg av *open source*. I den grad proprietære verktøy skal benyttes, må det gjøres enkeltvis, og det er et åpent spørsmål om Artefactual Systems, med sin kompetanse primært rettet mot *open source*-verktøy, kan bidra med spesielt mye for å få verktøyene til å samarbeide med Archivemata.

For oss var det uansett viktig å få konvertert Attacs tekst-dokumenter til PDF/A, både for bevaring og tilgjengeliggjøring. Artefactual løste dette ved å koble til LibreOffice i konverteringsprosessen i Archivemata på Arbarks installasjon.

Også excel-filer ble konvertert til PDF/A. Det visuelle resultatet av denne konverteringen var imidlertid lite holdbar. Det beste er trolig å konvertere excel-filer (i det minste de mindre avanserte) til PDF/A, men også ta vare på excel-filene i opprinnelig format for slik å få med formler og en bedre presentasjonsform – i hvert fall for de excel-filene som ikke passer inn i et A-4-format.

Når det gjelder bildefiler, blir JPG og PNG konvertert til TIFF i Archivematas standard-versjon. Denne konverteringen gjelder imidlertid bare for bevaring. Vi ønsket å konvertere bildefiler til TIFF også for tilgjengeliggjøring. Denne modifikasjonen endret vi selv inne i Archivemata.

²⁹ McLellan 2010, s. 31.

³⁰ McLellan 2010, s. 32. *Reverse engineering* er ”(«omvendt konstruksjon» eller «omvendt ingeniørkunst») er en arbeidsteknikk hvor man tar utgangspunkt i et ferdig produkt, plukker dette fra hverandre og analyserer det, og ut fra dette utarbeider detaljerte tegninger og spesifikasjoner på hvordan produktet fungerer.” ”Reverse engineering”. I *Wikipedia*.

³¹ McLellan 2010, s. 32.

Vi ønsket også å konvertere MSG-filer (det vil si Outlook) til PDF/A. Dette var ikke mulig nå, men Artefactual jobber med epost-bevaring for en annen klient, og hevder at det kommer en ny versjon med bedre epost-funksjonalitet i framtiden.³²

Vi etterspurte også nærmere informasjon om hva utviklerne/programmererne gjorde ”back end” i programmet. De ga oss da tilgang til deres *ticket tracking system*, altså deres system for innmelding fra brukere og utviklernes rettinger av feil. Et skjermbilde av dette ligger som vedlegg til denne rapporten.

Når det gjelder kommunikasjonen mellom AtoM og Archivemata, er den på den versjonen vi testet ut (1.4) utilfredsstillende. Det går ikke an automatisk å sende hele arkiver (eller *fonds*) på en enkelt måte. Vi måtte først opprette hele mappe-/arkivhierarkiet i AtoM, og deretter laste én og én mappe (altså mappene nederst i hierarkiet som hadde dokumenter i seg) inn i AtoM etter behandling i Archivemata. Denne løsningen ble til slutt valgt etter en lengre dialog med Artefactual. Det finnes andre løsninger, for eksempel å ”flytte” i AtoM (kan sammenlignes med ’flytt’-funksjonen i Asta). Det skal sies at alle disse løsningene er svært arbeidskrevende.

I versjon 1.5 er kommunikasjonen mellom AtoM og Archivemata forbedret. I denne versjonen kan man selv merke mapper med en betegnelse, for eksempel serie eller underserie inne i Archivemata, som da gjør at de faller automatisk på plass inne i et arkiv i AtoM uten at noe veldig stort forarbeid må gjøres i AtoM først. Man må her kun opprette arkivnivå i AtoM i forkant.³³

Ettersom dette var et testprosjekt valgte vi å ikke prosessere hele Attacs arkiv. Det ble for tidkrevende. Poenget var å få en forståelse av hvordan Archivemata fungerer. Ved slutten av prosjektet lastet vi ned flere arkivpakker fra Archivemata for å kunne ta vare på noe av arbeidet vi har gjort med Attacs arkiv. Samtidig fikk vi tilsendt en database-dump³⁴ av hele AtoM-siden, slik at vi har innholdet derfra for framtidig bruk. For å få med mappestruktur eksporterte vi i tillegg EAD-filer fra serie-nivåene i AtoM. To av seriene var for store til at vi kunne eksportere selv, så Artefactual måtte gjøre det for oss gjennom kommandolinjen.

³² E-post fra Artefactual Systems til Arbak 26.01.2016.

³³ Artefactual Systems. ”Arrange a SIP for AtoM”. I *Archivemata user manual (1.5)*.

³⁴ Database-dumper inneholder struktur/informasjon fra en base. Dumper brukes ofte for å kunne gjenopprette informasjon og dermed forhindre potensielt informasjonstap.

AIC-funksjonen i Archivematica

Ifølge den internasjonale OAIS-standarden er en AIP en informasjonspakke bestående av innholdsinformasjon og tilhørende bevaringsbeskrivende metadata.³⁵ Hver mappe vi lastet opp til Archivematica tilsvarer en AIP-arkivpakke. Med et så omfattende arkiv som Attacs måtte det nødvendigvis bli en stor mengde opplastede mapper i Archivematica, noe som innebærer at vi ble sittende med veldig mange AIPer. Problemet er da at man trenger overbyggende arkiv/metadainfomasjon for å kunne dokumentere arkivets struktur og sammenhenger. Det er her AICen kommer inn.

AIC står for *Archival Information Collection*. I Archivematica består en AIC av relaterte AIPer og en METS-fil som linker elementene i strukturen til innholdsfiler og metadata med fortegnelse over alle relaterte AIPer. I Archival storage angir en såkalt pointer.xml-fil lagringsinformasjon for hver AIC-METS-fil og hver AIP.³⁶ Målet med en AIC ville for vår del være å samle alle våre AIPer innenfor én og samme AIC.³⁷ Etersom både selve AICen og de tilhørende AIPene ville ha egne tilhørende metadata, ville vi dermed få med metadata på alle nivåer i hierarkiet.

Det er imidlertid problemer knyttet til denne funksjonen. I Archivematica må en AIC lages samtidig som man prosesserer i *ingest*-taben (altså når man oppretter SIP, AIP og DIP). En AIP skapes samtidig som en AIC skapes, og AIPen som skapes i *Ingest*-prosessen er da den første AIPen som tilknyttes AICen. Deretter kan eldre AIPer tilknyttes AICen. Problemet oppstår når en ny AIP skal knyttes til en AIC. Hvis man skal legge til flere AIPer til en allerede eksisterende AIC, må den originale AICen slettes.³⁸ Etter at nye AIPer er lagt til, må så en ny AIC skapes.

Det går med andre ord ikke an å oppdatere en gammel AIC for å lenke den til yngre AIPer. Etersom det er naturlig med tilvekster i mange arkiver (i hvert fall for oss på Arbark), og at man dermed risikerer å stadig måtte oppdatere prosesserte arkiver, blir dette etter vår mening en uflexibel og tungvint løsning.³⁹ Dette er hovedårsaken til at vi ikke prioriterte å opprette en AIC for Attacs arkiv, selv om vi nok måtte ha gjort det dersom vi skulle prosessert hele

³⁵ CCSDS 2012, s. 1-9.

³⁶ Artefactual Systems. "AIC". I *Archivematica wiki*.

³⁷ I Archivematica implementeres AICer som del av "dataset preservation"-workflowen.

³⁸ Artefactual Systems. "AIC". I *Archivematica wiki*.

³⁹ DIAS-sluttrapporten spesifiserer at AIC skal kunne oppdateres. *Vel bevart! Rapport fra samarbeidsprosjektet DIAS– Digital arkivpakkestruktur. Del 1*. Riksarkivaren 2012, s. 18.

arkivet. Det understrekes at vi testet ut versjon 1.4, og at AIC-funksjonen kan komme til å endre seg i nyere versjoner.

Riksarkivarens forskrift

Vi er av den oppfatning at lovverket for offentlige arkiver ofte med hell også kan brukes som veiledning for privat arkivarbeid. Regelverket er tross alt bygd på arkivfaglige betraktninger og begrunnelser. Vi vil i det følgende se på noen paragrafer fra forskriften og undersøke i hvilken grad Archivemata lever opp til eller samsvarer med forskriften.⁴⁰ Det kan også være relevant å se på hvordan Riksarkivarens forskrift kan tolkes i en ny sammenheng (fil- og mappearxiv) og se om den er fornuftig å bruke i denne sammenhengen.

Dokumentasjon av den opprinnelige orden som Archivemata lager, kan sies å tilsvare INNHold.TXT i en vanlig avlevering.⁴¹ Det er imidlertid ikke noe i veien for at det legges ved filer som tilsvare INNHold.TXT og INFO.TXT i selve avleveringen/deponeringen før Archivemata kjøres.⁴² Et sentralt poeng i forskriften er jo nettopp at arkivskaper skaper metadata som siden kontrolleres av depotet. Her kan man for eksempel bruke DROID til å lage en nøyaktig oversikt av den opprinnelige fil- og mappestrukturen som kan følge med avleveringen (jf. INNHold.TXT). DROID kan også brukes for å telle antall filer og mapper.⁴³ I tillegg kan DROID lage sjekksummer, også ved bruk av SHA-256.⁴⁴

I Archivemata legges det opp til at filer som INFO.TXT og INNHold.TXT (eller tilsvarende overordnede metadatafiler) kan inngå i deponeringen/avleveringen. Man kan legge ved såkalt *submission documentation*, noe INNHold.TXT og INFO.TXT vil fungere som. Det er også mulig å legge ved sjekksummer og få disse kontrollert av AtoM. Men den funksjonen synes vi var utilfredstillende beskrevet i bruksanvisningen.⁴⁵ Det var ikke åpenbart for oss hvordan tekstfilen med sjekksummer egentlig skulle være strukturert. Funksjonen i seg

⁴⁰ Arbark 2015, s. 4. Siden Riksarkivarens forskrift i første rekke tar for seg (database)tabelluttrekk, er det mye av regelverket som blir uaktuelt for ustrukturert data.

⁴¹ Riksarkivarens forskrift § 8-32, tredje ledd.

⁴² Jf. Riksarkivarens forskrift § 8-32, andre og tredje ledd.

⁴³ Strengt tatt står det i forskriften ”antall poster” og ”annen tallfesting av omfang” (Riksarkivarens forskrift § 8-30, bokstav e). Arbark velger her å ”oversette” dette til antall filer og mapper.

⁴⁴ SHA står for Secure Hash Algorithm. Slike algoritmer regner ut sjekksummer, som brukes til integritetssjekk av digitale objekter.

⁴⁵ Artefactual Systems. ”Create a transfer with submission documentation” og ”Create a transfer with existing checksums”. I *Archivemata user manual (1.5)*.

selv har vi ikke testet ut, men vi vurderer den som positiv forutsatt at det gjennom veiledning er mulig å forstå den.

Oppsummert mener Arbarks prosjektgruppe at Archivematica er godt egnet til å oppfylle Riksarkivarens forskrift på de punktene vi har sett på. Imidlertid, siden avlevering av fil- og mappearkiv ikke er det samme som databaseuttrekk, må det et tolkningsarbeid til. Slik vi ser det, er det viktig at det utarbeides klare bestemmelser eller retningslinjer rundt avleveringer/deponeringer av slike arkiver.⁴⁶ Det eksisterer muligens noen rutiner rundt omkring i landet, men det ser ikke ut som slike rutiner, i den grad de finnes, har sivet ut til resten av arkivsektoren. Riksarkivets rapport om private elektroniske arkiver foreligger, men en enkel veiledning (steg for steg) for mottak av elektroniske fil- og mappearkiver tror vi likevel vil være et udiskutabelt gode. En slik veiledning bør omhandle fasen før depotprogrammet gjør sitt arbeid, for eksempel hva arkivskaper bør gjøre, og hvordan arkivskaper bør sende inn materialet (tillatt hardware osv.). Riksarkivets rapport kan brukes som et grunnlag, men en veiledning bør være skrevet i et språk som henvender seg mer til arkivskapere uten spesiell arkivkompetanse enn arkiv(depot)sektoren. Den kan også være fri for de mer analyserende betraktningene vi finner i Riksarkivets rapport.

Arbeidet med retningslinjer har delvis begynt i regi av SAMDOK-prosjektet. Blant annet pågår det et arbeid om en veiledning om elektronisk arkivdanning i privat sektor og ikke minst høsting av slike arkiver. Derfor er det prosjektgruppens håp at det også er mulig å få til en veiledning (eller inkorporere i en annen veiledning) om avleverings- eller deponeringsfasen i regi av SAMDOK-prosjektet.

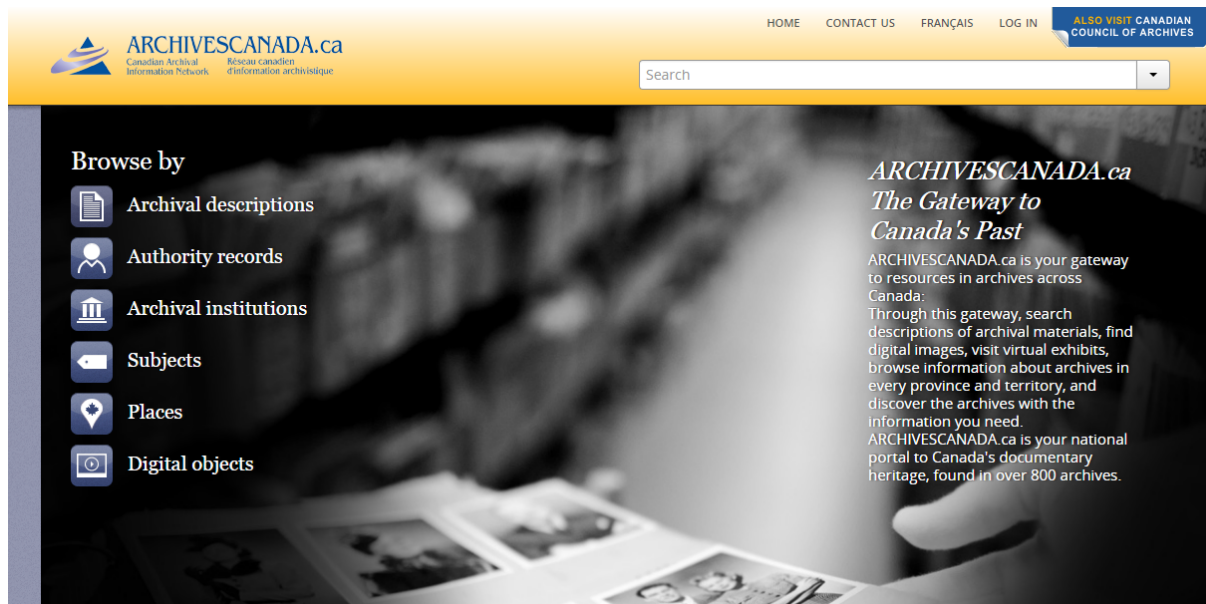
⁴⁶ Et unntak her er filformatbestemmelser, siden Riksarkivarens forskrift er ganske klar når det kommer til arkivformater.

Del 4 – Vurdering av AtoM

I det følgende tar vi nærmere for oss AtoM. Først ser vi på brukervennlighet og søkemuligheter, før vi belyser en interessant og for oss ny søkeinnngang – funksjonssøket. Deretter ser vi nærmere på mulighetene AtoM gir for brukertilgang.

Brukervennlighet - en historikers inntrykk av ArchivesCanada

For å få noe kunnskap om brukervennligheten til AtoM, fikk Arbark en norsk historiker til å prøve ut nettsiden *ArchivesCanada*.⁴⁷ Dette er en canadisk tilgjengeliggjøringsportal som egentlig er en AtoM-installasjon flere arkivinstitusjoner i Canada benytter seg av, litt som vår hjemlige Arkivportalen. Det understrekes at forskerens undersøkelse var kort – omtrent 10 minutter, og han hadde heller ikke nevneverdig kjennskap til arkivene han fant på *ArchivesCanada*. Denne portalen ble valgt fordi vi ikke lenger hadde tilgang til AtoM, men også fordi *ArchivesCanada* gir en mer realistisk brukeropplevelse enn Arbarks AtoM-installasjon, siden den canadiske portalen har mange flere arkivbeskrivelser (med eller uten digitale filer).



Fra nettsiden archivescanada.accesstomemory.ca. Merk de likeverdige søkeinnngangene.

Forskeren var først skeptisk. Han var vant med Arkivportalen og syntes de nye sidene var rotete. Han trakk selv særlig frem det uvante brukergrensesnittet. Men etter hvert som han

⁴⁷ *ArchivesCanada*.

manøvrerte seg gjennom sidene forstod han relativt raskt hvordan ting fungerte. Med andre ord kan sidene sies å være intuitive, selv om man først kan oppfatte dem som rotete. Forskeren sier selv om den norske Arkivportalen at søkere bør ha god kjennskap til materialet de er ute etter. Det er ikke nødvendigvis tilfelle med den canadiske portalen. Gjennom de canadiske sidene søkte han på sted og etter emne, noe som gjorde fremfinningen av det han søkte etter enkel (det ble søkt etter skoler i British Columbia).

Det at man kan åpne digitale dokumenter direkte fra portalen var særlig positivt. Sammenlignet med Arkivportalen virket de digitaliserte dokumentene mye lettere tilgjengelige. Er det digitale dokumenter i arkivet, vil de vises på alle nivåer i arkivet (fra arkivnivå til dokumentnivå), og det er også en *scrolling*-funksjon som gjør at man kan skifte mellom ulike digitale dokumenter. Her er det selvfølgelig en fordel å gå lengre ned i arkivhierarkiet (for eksempel mappe) for å redusere antall dokumenter som vises, siden alle digitale dokumenter i arkivet vil vises i rekkefølge på arkivnivå. Det samme gjelder for underarkiver, serier og mapper.

Den ukjente, engelskspråklige terminologien trakk imidlertid ned forskerens brukeropplevelse. Begreper som 'fonds' kan nok virke fremmed både for norske arkivbrukere og for folk som jobber med arkiv, som er vant med norsk begrepsbruk. Her tar vi det for gitt at eventuelle nyttige deler av AtoM (eller hele AtoM, hvis arkiv-Norge vil det) vil komme i norsk språkdrakt.

Oppsummert kan vi si at historikeren fikk et positivt inntrykk av den canadiske portalen. Det understrekes at dette bare er én enkelt forskers oppfatning.

Prosjektgruppens oppfatning av søkemulighetene i AtoM

Vår egen oppfatning er at emnetilgangen via emneord og geografi er spesielt godt integrert i portalen. Begge disse er på portalen like fremtredende som provenienstilgang (altså søk i arkivbeskrivelser/-kataloger). I den grad arkiv-Norge vil legge til rette for nybegynnere, er det grunn til å merke seg at LEADERS-prosjektet konkluderte med at nybegynnere kan være fornøyde med emnetilgang.⁴⁸ Det bemerkes her at forskeren vi forhørte oss med, som må sies å være en erfaren arkivbruker, umiddelbart fant det attraktivt å bruke emnetilgang fordi han

⁴⁸ Yeo 2010, s. 103.

var ukjent med canadisk arkivmateriale og dermed en ”nybegynner” i en canadisk arkivsetting.

Når man skal tenke søkemuligheter, er det viktig å huske på at ulike brukergrupper arbeider ulikt med arkiv.⁴⁹ Det at AtoM legger til rette for en enkel emnetilgang, samtidig som de profesjonelle har provenienstilgang, øker brukervennligheten. Arkiv-Norge trenger ikke nødvendigvis å bruke AtoM, men stiftelsen Asta kan i sitt videre arbeid med Arkivportalen trolig lære av å se nærmere på hvordan det er lagt til rette for emne- og geografitilgang i AtoM, selv om for eksempel stedssøk eksisterer i Arkivportalen i dag.

Det må understrekes at også digitale objekter (dokumenter)-tilgang er like fremtredende som de andre tilgangene. Det er mulig å trykke på ikonet *digital objects* og kun søke etter digitale eller digitaliserte dokumenter gjennom fritekstsøk.

Når det gjelder PDF og PDF/A ser vi det som positivt at tekstgjenkjenning er mulig. Det vil si at hvis PDFen er utstyrt med tekstgjenkjenning, for eksempel gjennom OCR, skal det være mulig å få treff på ord i teksten i PDF-filen gjennom AtoMs generelle søk.⁵⁰

Det er alt i alt mange forskjellige søkemuligheter, også boolske søk. Nybegynnere vil muligens finne de mer avanserte søkemulighetene lite relevante og vanskelige. Men dette er ikke noe stort problem så lenge vanlige fritekstsøk også tilbys og er et godt synlig alternativ.

Likevel, alt er ikke like bra. For eksempel, å gå ned på serienivå fra arkivnivå vil gjøre at de andre seriene i arkivene forsvinner slik at de hierarkisk underordnede elementene i serien du ser på (for eksempel mapper) får plass på skjermen. Da kan en bruker fort miste oversikten over hvor hun eller han er i arkivet.

Funksjoner – en interessant alternativ søkeinnngang

I AtoM er det mulig å finne fram til materiale ved å søke etter funksjoner. Etter vår mening er dette et interessant alternativ til tradisjonelle søkeinnnganger som emne og navn. I den internasjonale standarden for funksjonsbeskrivelse, ISDF, defineres funksjon som

⁴⁹ Yeo 2010, s. 103-104.

⁵⁰ Artefactual Systems. ”Upload PDF”. I *AtoM user manual* (2.3).

[a]ny high level purpose, responsibility or task assigned to the accountability agenda of a corporate body by legislation, policy or mandate. Functions may be decomposed into sets of co-ordinated operations such as subfunctions, business processes, activities, tasks or transactions.⁵¹

Funksjoner er ifølge standarden mer stabile enn administrative strukturer, siden sistnevnte over tid kan endres, for eksempel ved omstrukturering. Funksjoner kan fungere som grunnlag både for organisering, klassifisering og beskrivelse av dokumenter, som grunnlag for bevaringsvurdering, og som verktøy for gjenfinning og analyse.⁵²

ISDF er implementert i AtoM i sin helhet, noe som gjør det mulig å koble funksjonsbeskrivelser til arkivbeskrivelser.⁵³ La oss bruke Attac som eksempel på hvordan brukere kan benytte seg av denne søkefunksjonen.

Hvilken type dokumenter oppretter Attac og for hvilket formål? I Attacs vedtekter heter det om organisasjonens formål:

Attac Norge er en selvstendig, antirasistisk organisasjon som, uavhengig av religion og partipolitikk, har som formål å samle og mobilisere en bred, folkelig og demokratisk bevegelse mot den makten finanskapitalen og multinasjonale selskaper utøver på bekostning av demokratiet, og de negative konsekvensene dette har politisk.⁵⁴

Hvilke konkrete aktiviteter utfører så Attac for å oppnå sine mål? Det beste eksemplet er informasjonsvirksomhet, ettersom spredning av informasjon er en viktig måte for Attac å stimulere til engasjement. Dersom brukeren søker etter ”spredning av informasjon” eller ”informasjonsvirksomhet” via funksjonssøk i AtoM, vil han da få opp Attacs arkiv som treff, eventuelt kun den aktuelle serien dersom funksjonsbeskrivelsen kobles til serienivået, ikke arkivnivået. Det kan tenkes at søk på funksjon vil gjøre studiet av historiske dokumenter interessant for nye brukergrupper. Medie- eller samfunnsvitere kan for eksempel være interessert i et funksjonsbasert søk på ”informasjonsvirksomhet” for å finne velegnede caser.

ISDF-implementeringen i Atom er fleksibel – bortsett fra de helt essensielle skrivefeltene er mange felter i valgfrie. I *Identity area* er for eksempel to av fem felter obligatoriske. Vi mener at funksjonssøket er et interessant søkealternativ, og det er verdt å merke seg at Arkivportalen enda ikke har denne søkemuligheten. Det å få flere søkeinnganger til materialet, ikke kun det tradisjonelle innholdssøket, kan være positivt. Men før prosjektgruppen konkluderer med at funksjoner er nyttige nok for brukeren, vil vi understreke behovet for brukerstudier av funksjonsbeskrivelser.

⁵¹ ICA 2007, s. 10.

⁵² ICA 2007, s. 7.

⁵³ Artefactual Systems. ”ISDF”. I *AtoM user manual* (2.3).

⁵⁴ Attac 2015.

Det er foreløpig hovedsakelig arkivarer som (i økende grad) er vant med en funksjonell tankegang. Brukere har i hvert fall i norsk sammenheng så langt blitt lite utsatt for funksjoner i arkivbeskrivelsessystemer. En studie fra Skottland understreker at det var profesjonelle arkivarer og arkivstudenter som positivt kommenterte funksjonsbeskrivelser, mens vanlige brukere ikke kommenterte dem i det hele tatt. Enten fant vanlige brukere funksjonssøket såpass lett å manøvrere at de ikke fant det bryet verdt å kommentere det, eller så overså de funksjonssøket helt.⁵⁵

Brukertilgang – muligheter for crowdsourcing⁵⁶?

AtoM har mange forskjellige roller for tilgang: forsker, bidragsyter, redaktør, oversetter og administrator. Dette er utgangspunktet til AtoM. Det er mulig å lage egne brukergrupper også, og gi en bruker flere tilgangsprivilegier og muligheter ved å gjøre ham del av ulike brukergrupper. Kort sagt, han kan for eksempel være administrator og oversetter.⁵⁷

Forskere (eller vanlige arkivbrukere) har kun lesetilgang til publiserte arkivbeskrivelser (også publiserte digitale objekter/dokumenter). Det understrekes at forskerne ikke trenger å være registrert i AtoMs system.

Bidragsyterne kan blant annet søke, lage og oppdatere arkivbeskrivelser. De har også tilgang til det digitale objektet, men de kan ikke publisere arkivbeskrivelsene. Redaktører kan på sin side skape, publisere, oppdatere og slette arkivbeskrivelser.

Oversettere kan oversette for eksempel arkivinnholdsbeskrivelser.⁵⁸ Akkurat det er et interessant element når det for eksempel kommer til ulike minoritetsarkiv. Å ha tilgjengelig en arkivbeskrivelse på både majoritets- og minoritetsspråk, men kun ett om gangen (du velger hvilket språk som skal vises når oversettelsen er gjort), gjør arkivene mer tilgjengelige for både majoritets- og minoritetsbefolkningen. Man kan også tenke seg oversettelser fra nynorsk til bokmål og vice versa. Det at du velger hvilket språk som vises, gjør at skjermen ikke blir ”overlasset” med informasjon.

⁵⁵ Glasgow University Archive Service, 2006, s. 42-43.

⁵⁶ *Crowdsourcing* innebærer at mange går sammen om å anskaffe tjenester eller å gjennomføre noe som man har felles interesse av.

⁵⁷ Artefactual Systems. ”User roles”. I *AtoM documentation (2.3)*.

⁵⁸ Artefactual Systems. ”Multilingual design principles”. I *AtoM user manual (2.3)*.

Administratorrollen har alle tilganger og administrerer de ulike brukerne og deres tillatelser. Administrator har muligheter til å skreddersy brukerprofiler, slik at for eksempel en oversetter har noen ekstra privilegier, men ikke like mange som en redaktør. Likevel, fordelene med standardiserte brukerroller er at de gjør det enklere å sette opp nye brukerkontoer, og samtidig lett å holde oversikt over hvilke privilegier ulike brukere har.⁵⁹

Vi mener det er positivt med ulike roller. De ulike rollene gjør det mulig å skreddersy tilgang etter behov. Alle trenger ikke alle funksjoner. Vanlige ansatte trenger kanskje bare redaktør- eller bidragsytertilgang for å gjøre jobben sin, mens noen få ansatte trenger all tilgang for systemadministrasjon. Dette er også bra hvis arkivskaper vil ha tilgang til klausulerte filer kun arkivskaperen skal ha tilgang til. Da kan det opprettes en brukerprofil som gir innsyn i filer, men ikke gir mulighet for å endre selve arkivbeskrivelsen.

I den grad man vil benytte seg av *crowdsourcing* er det også et positivt element at det går an å skreddersy tilgang.⁶⁰ Arkivbrukere kan potensielt bidra til arkivbeskrivelser uten at man behøver å frykte at de sletter noe. De kan gjøre dette i sin egen stue ved at de får tilgang til AtoM via web. Skreddersydd tilgang er også ett av kriteriene til Huvilas såkalte *participatory archive*.⁶¹

Men selv om AtoM kan legge opp til *crowdsourcing*, har organiseringen mye å si for om slike prosjekter lykkes eller ikke.⁶² Bidragsytere må også få noe opplæring, men er de allerede viderekomne arkivbrukere og kjent med arkivbegreper som serier og mapper, bør de klare å orientere seg rundt omkring i arkivbeskrivelsene. De kan også kanskje bidra med emneord for å forbedre tilgangen til arkivmaterialet.

Likevel, AtoM gir tilgang til hele basen, ikke ett og ett arkiv, for bidragsytere. Dette kan være negativt i den forstand at brukere kan se digitale dokumenter som egentlig er klausulert. Foreløpig mangler det altså noe før AtoM fullt ut kan sies å være et verktøy for *crowdsourcing*. Samme problem melder seg når arkivskaper skal ha tilgang til klausulerte filer – da kan han få tilgang til filer i andre arkiver også.

En mulig løsning kan være å vente med å legge til arkiver med klausulerte dokumenter i AtoM-basen, men da kan tiden fort bli en dimensjon, siden eventuelt beskrivelsesarbeid må

⁵⁹ Dette kan gjøres gjennom *list user*-siden i AtoM, der administratoren kan se de ulike rollene til hver enkelt bruker i stedet for å klikke seg inn på hver enkelt.

⁶⁰ For en god evaluering av *crowdsourcing*-prosjekter, se Spindler 2014.

⁶¹ Spindler 2014, s. 19.

⁶² Spindler 2014, s. 12 & 17.

vente. En annen løsning er å utvikle en mer finfordelt tilgangsstyring i systemet enn det som finnes i dag (eller anskaffes ved hjelp av andre programmer/verktøy). Her kan man tenke seg at utgangspunktet for tilgang til bidragsytere er alt som er lagret i AtoMs base, men mer finjustert tilgang kan gis etter behov. Så skal vi heller ikke overdrive effekten av *crowdsourcing*, det er et åpent spørsmål om hvor mange som faktisk vil bidra når systemene legger til rette for det.

Del 5 – Archivemata og AtoM-workshop

I slutten av mars, mens prosjektgruppen fortsatt hadde tilgang til programmene, avholdt Arbark en workshop hvor vi viste frem programmene for arkiv- og museumsinstitusjoner i Oslo-regionen. Workshopen varte i omtrent to timer. I tillegg til å introdusere programmene for de andre institusjonene, var det for vår del også interessant å få et inntrykk av hvor lette de er å håndtere for helt nye brukere. Deltakende institusjoner var Oslo byarkiv, Vestfoldarkivet, Norsk teknisk museum, Norsk hjemmefrontmuseum og Nasjonalmuseet.

Ingen av deltakerne hadde erfaring med bruk av digitale depotprogrammer, men det var bred enighet om at det er et stort behov for slike. Flere ga uttrykk for at behovet er særlig stort når det kommer til fil- og mappearkiver. Etter at vi holdt en kort introduksjon fikk deltakerne selv prøve ut programmene på PCer vi hadde rigget opp. Tendensen var omtrent som vi hadde sett for oss. Det var flere som synes det var vanskelig å henge med på vår introduksjon, noe som er naturlig ettersom de hadde liten grad av forhåndskunnskap. Men når de selv fikk teste ut programmene (med veiledning fra oss), virket det som flere syntes at dette absolutt var overkommelig. Testingen gikk ut på å sende små mapper gjennom Archivemata, for deretter å laste dem opp i Atom – med andre ord, det samme som vi gjorde under prosesseringen av Attacs arkiv.

Med tanke på workshopens relativt korte varighet, samt at institusjonene ikke hadde noe kunnskap om programmene på forhånd, synes vi workshopen var konstruktiv.

Prosjektgruppen sitter igjen med to hovedinntrykk: For det første, for institusjonene i Oslo-området er det et veldig stort og hittil nesten udekket behov for digitale depotløsninger. For det andre, deltakerne synes at dette var interessant og at både Archivemata og AtoM var overkommelig når det gjelder vanskelighetsgrad. Det understrekes at det må forventes at en arkivar som jobber med slike depotprogrammer setter seg inn i hva mikroprosessene faktisk gjør. Selv om mange av prosessene er automatiserte, må man ha en forståelse av hva som foregår.

Del 6 – Andre programmer

I vår opprinnelige søknad skrev vi at vi skulle sammenligne Archivematica/AtoM med andre systemer, i den grad det ville bli mulig. Her var det meningen at vi skulle lese oss opp på andre programmer, for å bli i stand til å gjøre gode sammenligner. Siden den gang har det blitt konkretisert at Archivematica/AtoM primært skal sammenlignes med Preservica og EssArch i regi av DIAS II-prosjektet. For ikke å gjøre en jobb som likevel skal gjøres grundigere enn det vi kan få til innenfor PEANUTS-prosjektets ramme, har vi her valgt bort sammenligningen for heller å fokusere på hvordan Archivematica fungerer sammen med andre systemer. De konkrete eksemplene vil komme fra Zuse Institute Berlin.⁶³ Siden Archivematica kan sammenkobles med mange programmer, er det bedre å fokusere på dette enn å foreta en ren sammenligning, siden ulempene med Archivematica kan veies opp for ved å ta i bruk andre programmer og vice versa.

En av Archivematicas ulemper er at det ikke har en ”speilingsfunksjon”, altså at det ikke lager en kopi til lagring et annet sted på en enkel måte. Det vil kreve at arkivet behandles helt på nytt igjen. Zuse-instituttet bruker her iRODS for sitt *data management*-system. iRODS er hos Zuse-instituttet ansvarlig for å lage replikater av alle AIPer.⁶⁴ Speilingsfunksjonen er dermed altså på plass.⁶⁵

For tilgang er det Islandora/Fedora som gjelder. Islandora er en såkalt Fedora-integrasjon. I en omfattende undersøkelse gjort av South Carolina Departement of Archives and History ble det bemerket at Islandora var veldig brukervennlig. Den hadde imidlertid andre store ulemper, for eksempel ingen viruskanning.⁶⁶ Gitt at Zuse-instituttet kun bruker Islandora til tilgang, kan slike mangler overses. Hos Zuse blir DIPen sendt til en Fedora Object Store (”container/lager”), mens selve tilgangen for brukeren skjer igjennom Islandora. Det interessante her er at det er to forskjellige ”visninger” av DIPen, én for administrativt arbeid, for eksempel for rapportering av lagrede filformater, og en ”vanlig” for å se innhold.⁶⁷

Dette viser bare en av mange mulige konfigurasjoner Archivematica kan settes opp sammen med. Slik sett blir Archivematicas mangler mindre presserende. Likevel, jo flere programmer man benytter seg av, desto mer systemadministrasjon kreves for å få systemene til ”å snakke

⁶³ Zuse Institute Berlin er et forskningsinstitutt for informatikk og anvendt matematikk, som betjener en regional regnesentral med behov for lagring av enorme mengder data.

⁶⁴ Klindt og Amrhein 2015, s. 6.

⁶⁵ For detaljer rundt speilingsfunksjonen, se *iRods Technical Overview*, 2016, s. 4-7.

⁶⁶ Thomas 2015, s. 13-14.

⁶⁷ Klindt og Amrhein 2015, s. 7.

sammen” best mulig. I tillegg vil det kreve at depotet har kompetanse på flere programmer. Vi ser det helt klart som en fordel at *ett* program gjør mest mulig for ikke å belaste IT-administrasjonsressursene mer enn nødvendig. Men hvis flere depotinstitusjoner går sammen, kan Arbark være interessert i å se på en bevaringsløsning der flere programmer har ulike funksjoner.⁶⁸

⁶⁸ iRODS er her interessant, siden programmet muliggjør at flere institusjoner kan gå sammen, men likevel ha sine soner for sitt materiale. Klindt og Amrheim 2015, s. 6; *iRods Technical Overview*, 2016, s. 12.

Konklusjon

Det er flere positive sider ved Archivemata og AtoM som kan trekkes frem⁶⁹:

- Programmene gjør depotinstitusjoner ikke avhengig av proprietære systemer og leverandører.
- Selv om installering koster penger, er programmene gratis å bruke. Dette er selvfølgelig et viktig poeng for arkivinstitusjoner med begrensede budsjetter.
- Arbarks prosjektgruppe fant programmene lette og intuitive å bruke. Det virket som institusjonene på workshopen delte denne oppfatningen.
- Programmene har mange brukere, noe som gjør at flere kan snakke sammen om felles utfordringer. Institusjonene kan altså selv være med å påvirke utviklingen.
- Programmene er i konstant utvikling, med stadig nye versjoner, noe som gjør at programmene forbedres kontinuerlig. Det er stor åpenhet fra Artefactual Systems om hva som kommer.
- Artefactual Systems er svært løsningsorienterte og gode å samarbeide med.
- Programmene følger OAIS-rammeverket.
- Hele ISAD(G)-standarden er implementert i AtoM, noe som legger til rette for rikere arkivbeskrivelser enn det som er vanlig i norsk kontekst.
- Flere andre standarder er implementert i AtoM, også biblioteksstandarder. Vi anser dette som et gode, siden det kan medføre at ulike typer materiale kan presenteres for brukeren i samme system (Arbark er også et bibliotek).
- Det er mulig å fange den opprinnelige orden til et fil- og mappearkiv, for så å bevaringsvurdere og ordne arkivet senere.
- Det er mulig å utvikle egen bevaringspolitikk (det vil si handlinger i programmet, for eksempel formatkonvertering). Altså er Archivemata høyst konfigurerbart.
- Det skal være mulig å innarbeide *fixity check*.⁷⁰

Men det er også ulemper:

- Gratisaspektet kan tones noe ned. Det var nødvendig for Arbark å kjøpe støttetjenester. Selvinstallering er mulig, men vanskelig.

⁶⁹ Listen er basert på Arbarks egne og andres opplevelser. Se Mitcham 2015; Thomas 2015, s. 14-15.

⁷⁰ Artefactual Systems. "Fixity checking and reporting". I *Archivemata wiki*. *Fixity check* refererer til en metode (utregning av sjekksum) for å verifisere at et digitalt objekt ikke er blitt endret.

- Dette henger sammen med at programmene kun er kjørbare på Ubuntu/Linux-plattform, noe som igjen kan utløse behov for kompetanse og ressurser som små depotinstitusjoner ikke har.
- Artefactual Systems utvikler ikke egne verktøy for konvertering. De er avhengig av at andre gjør den jobben for dem.
- Norske bokstaver/tegn inngår (foreløpig) ikke i tegnsettingsstandarden i standardversjonen i Archivemata, noe som skapte problemer for oss i vår testfase.
- Rapporteringsfunksjonene i Archivemata er dårlige. Fordi rapporter kan være svært praktiske i bevaringssammenheng (for eksempel hvilke pakker som har format X, som skal byttes ut til format Y), er dette en ulempe.
- Det er en viss usikkerhet rundt skalerbarhet/mengdehåndtering for både Archivemata og AtoM. For eksempel måtte vi ha hjelp fra Artefactual Systems for å laste ned en av EAD-filene fra AtoM.
- Brukerrettighetene i AtoM er gode, men likevel ikke ideelle.
- Vi merket oss at dokumentasjon og bruksanvisning for AtoM og Archivemata har svakheter og mangler.

Vi er litt usikre på om såkalte betalte prosjekter skal listes opp som positivt eller negativt. Ved å betale en viss sum, kan man få Artefactual Systems til å utvikle funksjoner i programmet. Det er dermed mulig å være ”gratispassasjer” i den forstand at andre som har samme behov som deg blir med og betaler for utvikling du senere drar nytte av. Men selvfølgelig, jo mer spesielt og særegent behovet er, desto større er sjansen for at du må betale for det (hvis du da ikke utvikler det selv). Dette kan for eksempel skje hvis Archivemata ikke oppfyller DIAS-krav, et forhold vi ikke har sett på i denne omgang. Det er også mulig å sende inn utviklingsønsker til Artefactuals ønskeliste (utviklingsønsker som utvikles gratis), men disse får lavere prioritet.

Dette prosjektet har helt klart gjort Arbark bedre i stand til å ta imot elektroniske arkiver. Vi har fått testet ut og økt kunnskapen ikke bare om Archivemata, men også andre og enklere programmer som kan benyttes i en avleveringsprosess, eksempelvis DROID og Beyond Compare. Archivemata og AtoM er programmer som absolutt er gode nok til å brukes av Arbark. Slik vi ser det, kan Arkivportalen med fordel bli likere AtoM, for eksempel gjennom implementering av likeverdige søkeinnganger for emne, sted og proveniens (og kanskje funksjonssøk). Å bytte ut Arkivportalen med AtoM er trolig verken nødvendig eller ønskelig, fordi man da må begynne en lang prosess med å flytte data til AtoM og bygge opp en helt ny

nasjonal plattform. Poenget er at for Arbark er det avgjørende å ha en tilgjengeliggjøringsportal med tilknytning til depotprogrammet som brukes, slik at brukerne kan søke i og nyttiggjøre seg av materialet. Arkivskapere som vil avlevere sine privatarkiver til oss vil også ha tilgang til dem så å si umiddelbart, et krav som er blitt enda tydeligere med elektroniske arkiver. Vi kan ikke kun pakke dem inn i et arkivdepot for oppbevaring.

En endelig avgjørelse om bruk av Archivematica tas imidlertid ikke nå av Arbark. Endelig avgjørelse tas først ved fullførelsen av DIAS II-prosjektet for å få et bedre komparativt grunnlag.

Kildeliste

Brukermanualer og wiki for Archivemata og AtoM

Artefactual Systems. *Archivemata user manual (1.5)*:

- ”Arrange a SIP for AtoM”. Hentet 10.08.2016 fra <https://www.archivemata.org/en/docs/archivemata-1.5/user-manual/ingest/ingest/#hierarchical-dip>.
- ”Create a transfer with existing checksums”. Hentet 10.08.2016 fra <https://www.archivemata.org/en/docs/archivemata-1.5/user-manual/transfer/transfer/#create-a-transfer-with-existing-checksums>.
- ”Create a transfer with submission documentation”. Hentet 10.08.2016 fra <https://www.archivemata.org/en/docs/archivemata-1.5/user-manual/transfer/transfer/#create-submission>

Artefactual Systems. *Archivemata wiki*:

- ”AIC”. Hentet 09.06.2016 fra <https://wiki.archivemata.org/AIC>.
- ”Fixity checking and reporting”. Hentet 09.06.2016 fra https://wiki.archivemata.org/Fixity_checking_and_reporting.
- ”Micro-services”. Hentet 09.06.2016 fra <https://wiki.archivemata.org/Micro-services>.

Artefactual Systems. *AtoM user manual (2.3)*:

- ”Descriptive standards”. Hentet 10.08.2016 fra <https://www.accesstomemory.org/en/docs/2.3/user-manual/overview/descriptive-standards/>.
- ”ISDF”. Hentet 10.08.2016 fra <https://www.accesstomemory.org/en/docs/2.3/user-manual/data-templates/isdf-template/#isdf-template>.
- ”Multilingual design principles”. Hentet 09.09.2016 fra <https://www.accesstomemory.org/en/docs/2.3/user-manual/translate/multilingual-design-principles/#multilingual-design-principles>.
- ”Upload PDF”. Hentet 10.08.2016 fra <https://www.accesstomemory.org/en/docs/2.3/user-manual/import-export/upload-digital-object/#upload-pdf>.
- ”User roles”. Hentet 10.08.2016 fra <https://www.accesstomemory.org/en/docs/2.3/user-manual/getting-started/getting-started/#user-roles>.

Artikler og nettkilder med mer

ArchivesCanada. Hentet 07.06.2016 fra <http://archivescanada.accessmemory.ca/>.

Arbark. *Prosjektbeskrivelse PEANUTS*. Arbark, 2015.

Attac. *Attacs vedtekter*. Attac, 2016. Hentet 21.06.2016 fra <http://attac.no/organisasjon/vedtekter/>.

Datatilsynet. *Politiske partier må slette medlemsopplysninger*. Datatilsynet, 2012. Hentet 01.09.2016 fra <https://www.datatilsynet.no/Regelverk/Tilsynsrapporter/2012/Politiske-parti-ma-slette-medlemsopplysningar-/>.

Datatilsynet. *Medlemskap i politiske partier*. Datatilsynet, 2015a. Hentet 06.09.2016 fra <https://www.datatilsynet.no/Sektor/Foreninger-idrett/Medlemsskap-i-politisk-parti/>.

Datatilsynet. *Ny personvernlovgivning på vei fra EU*. Datatilsynet, 2015b. Hentet 09.08.2016 fra <https://www.datatilsynet.no/nyheter/2015/ny-personvernlovgivning-pa-vei-fra-eu/>.

Glasgow University Archive Service. *Developing Archival Context Standards for Functions: Final Report*. Gateway to Archives of Scottish Higher Education, 2006. Hentet 08.08.2016 fra http://www.gashe.ac.uk/news/final_report.pdf.

International Council on Archives [ICA]. *International Standard for Describing Functions*. ICA, 2007. Hentet 11.07.2016 fra http://www.ica.org/sites/default/files/CBPS_2007_Guidelines_ISDF_First-edition_EN.pdf.

iRODS Technical Overview, 2016. Hentet 06.07.2016 fra <http://irods.org/wp-content/uploads/2016/06/technical-overview-2016-web.pdf>.

Kjeldstadli, K. "Kollektive bevegelser". I Helle I., Kjeldstadli K. og Sørvoll J. (red.). *Historier om motsand. Kollektive bevegelser i det 20. århundret*. Abstrakt forlag AS, 2015.

Klindt M. og Amrhein K. *One Core Preservation System for All your Data. No Exceptions!*. School of Information and Library Science, University of North Carolina at Chapel Hill, 2015. Hentet 06.07.2016 fra <https://phaidra.univie.ac.at/view/o:429551>.

McLellan, E. P. "Selecting Formats for Digital Preservation: Lessons learned during the Archivemata project". I *Information Standards Quarterly*, volum 22, nr. 2, 2010. Hentet 21.06.2016 fra

http://www.niso.org/apps/group_public/download.php/4237/IP_McLellan_Selecting_Formats_isqv22no2.pdf.

Mitcham, J. "The first meeting of Archivmatica users (or explorers)". I *Digital Archiving at the University of York*. Bortwick Institute for Archives, University of York, 2015. Hentet 10.08.2016 fra <http://digital-archiving.blogspot.no/2015/01/the-first-meeting-of-archivematica-uk.html>.

"Reverse engineering". I *Wikipedia*. Hentet 16.09.2016 fra https://no.wikipedia.org/wiki/Reverse_engineering.

Riksarkivet. *Prosjekt elektroniske privatarkiver. Vedlegg 1A: Bevaring av frittstående elektroniske dokumenter*. Riksarkivet, 2005. Lastet 21.04.2016 fra <http://www.arkivverket.no/arkivverket/Privatarkiv/Elektroniske-privatarkiver/Prosjektrapport>.

Riksarkivet. *OAIS-modellen*. Riksarkivet, 2011. Hentet 27.04.2016 fra <https://www.arkivverket.no/arkivverket/Arkivbevaring/Elektronisk-arkivmateriale/Standarder/OAIS>.

Scooter Software. *Intelligent Comparison*. Hentet 10.08.2016 fra http://www.scootersoftware.com/features.php?zz=features_focused.

Spindler, R. P. *An Evaluation of Crowdsourcing and Participatory Archives Projects for Archival Description and Transcription*. Arizona State University Libraries, 2014. Hentet 05.07.2016 fra <https://repository.asu.edu/attachments/135630/content/Research%20Paper%20v3.pdf>.

The Consultative Committee for Space Data Systems [CCSDS]. *Reference model for an open archival information system (OAIS). Recommended practice*. CCSDS, 2012. Hentet 16.06.2016: <http://public.ccsds.org/publications/archive/650x0m2.pdf>

The National Archives. *DROID: How to use it and how to interpret the results*. The National Archives, 2011. Hentet 09.08.2016 fra <http://www.nationalarchives.gov.uk/documents/information-management/droid-how-to-use-it-and-interpret-results.pdf>.

Thomas, B. "Selecting an Electronic Records Repository Platform at the South Carolina Department of Archives and History". I *Journal of Contemporary Archival Studies*, volum 2,

artikkel 2, 2015. Hentet 06.07.2016 fra

<http://elischolar.library.yale.edu/cgi/viewcontent.cgi?article=1009&context=jcas>.

”UTF-8”. I *Wikipedia*. Hentet 06.06.2016 fra <https://no.wikipedia.org/wiki/UTF-8>.

Vel bevart! Rapport fra samarbeidsprosjektet DIAS – Digital arkivpakkestruktur. Del 1.

Riksarkivaren, 2012. Hentet 10.08.2016 fra

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

Yeo, G. ”Debates about Description”. I MacNeil H., og Eastwood T. (red.). *Currents of Archival Thinking*. Libraries Unlimited, 2010.

Lov og forskrift

Personopplysningsloven. Lov av 14.04.2000 nr. 31 om behandling av personopplysninger.

Hentet 07.09.2016 fra [https://lovdata.no/dokument/NL/lov/2000-04-14-](https://lovdata.no/dokument/NL/lov/2000-04-14-31?q=personopplysnings)

[31?q=personopplysnings](https://lovdata.no/dokument/NL/lov/2000-04-14-31?q=personopplysnings).

Riksarkivarens forskrift. Forskrift av 01.12.1999 nr. 1566 om utfyllende tekniske og arkivfaglige bestemmelser om behandling av offentlige arkiver. Hentet 09.08.2016 fra

<https://lovdata.no/dokument/SF/forskrift/1999-12-01-1566>.

Personlig kommunikasjon

E-post fra Artefactual Systems til Arbark 26.01.2016.

E-post fra Artefactual Systems til Arbark 16.03.2016.