

Digitalt Depot for Privatarkiv (DDP)

Prosjektrapport



5. oktober, 2020

Prosjekt: Digitalt Depot for Privatarkiv (DDP)

Prosjektsammendrag

<i>Prosjektets formål:</i>	Utvikle en modell for langtidslagring og tilgjengeliggjøring av digitalt skapt privatarkiv.
<i>Prosjektperiode:</i>	August 2018 – mars 2020
<i>Prosjekteier:</i>	Snorre D. Øverbø, leder Arkivavdelingen, Aust-Agder museum og arkiv, avdeling KUBEN
<i>Prosjektleder:</i>	Egil Blomsø, Aust-Agder museum og arkiv, avdeling KUBEN
<i>Prosjektmedarbeidere:</i>	Anne Gaaserud, Marit Gressli og Yngve Schulstad Kristensen, Aust-Agder museum og arkiv, avdeling KUBEN
<i>Samarbeidspartnere:</i>	• Lillesand seilforening, Lillesand (Avleverende virksomhet) • ArkivIT AB, Stockholm (Drift/support Archivematica og AtoM) • Piql AS, Drammen (Drift/support Archivematica og AtoM)
<i>Hovedleveranser fra prosjektet:</i>	• Gjennomført planlegging, mottak og behandling av digitalt skapte privatarkiv med Archivematica som hovedverktøy. Etablert Digitalt Depot • Tilgjengeliggjort arkivene på Internett: http://atom.aama.no

1. Innledning

Aust-Agder museum og arkiv, avd. KUBEN, har i oppgave å arbeide for at arkiv og andre informasjonsbærere av historisk og administrativ verdi fra Aust-Agder blir tatt vare på og gjort tilgjengelig for ettertiden. Dette inkluderer også arkiv skapt av ikke-offentlige enheter, som bedrifter, foreninger og privatpersoner. KUBEN oppbevarte pr. 1.1.2020 om lag 3200 hyllemeter med privatarkiv fordelt på over 1600 private arkivskapere. En bestandsanalyse fra 2015 viser imidlertid at det er få arkiver fra etter år 2000, og at ingen av privatarkivene er digitalt skapt.

I dag skaper og lagrer de fleste private aktører sine arkiver digitalt, men dette gjenspeiler seg foreløpig ikke på depotsiden. Det er nok flere grunner til dette. Depotarkivarer, i hvert fall når det gjelder privatarkiv, opererer i primært i et papirbasert paradigme og har nok ikke vært flinke nok til å omstille seg til nye, digitale medier. Mangel på IT-kompetanse har nok også gitt seg utslag i en viss berøringsangst i forhold til dette. Siden bevaring av private arkiver ikke er lovregulert er det heller ikke utviklet en "norsk standard", dvs. en type bransjestandard/«Best Practice» for mottak, bevaring og tilgjengeliggjøring av digitalt skapte privatarkiv i Norge.

Det ligger innenfor Aust-Agder museum og arkiv IKS (fork. AAmA) sitt ansvarsområde å samle inn og bevare også digitalt skapt privatarkiv. Det er imidlertid en rekke utfordringer knyttet til dette. Elektroniske dokumenter som er lagret i ulike filformater, og på ulike digitale medier, er nemlig svært sårbare sammenliknet med papirdokumenter.

Utfordringene ved langtidsbevaring av digitale arkiver er mange og sammensatte. Vi vil her trekke fram noen viktige momenter: Digitale dokumenter vil gå tapt i løpet av få tiår dersom det ikke blir gjort en *aktiv* innsats for å bevare arkivet. For det første er selve lagringsmediene og tilfredsstillende backup-systemer ressurskrevende. Videre så må en sørge for at datasettene til enhver tid lagres i bærekraftige lagringsformater. Og endelig må bevaringsinstitusjonen jevnlig påse, gjennom aktiv testing, at dataene faktisk kan hentes fram, leses og brukes.

Prosjektet Digitalt Depot for Privatarkiv (DDP) er et forsøk på å løse noen av disse utfordringene, samtidig som man utnytter mulighetene digitalt skapte arkiver gir.

2. Prosjektets målsetning

Målet med dette prosjektet har vært, i henhold til AAmA's samfunnsoppdrag, å etablere en modell for langtidslagring og tilgjengeliggjøring av digitalt skapte privatarkiv. I dette ligger etablering av nødvendige teknisk infrastruktur, kompetansebygging samt etablering av rutiner for mottak, ordning og lagring.

3. Bakgrunn for prosjektet

I 2015 fikk KUBEN arkivutviklingsmidler pålydende 200 000 kroner fra Arkivverket til prosjektet *eArkiv – fra idrettskrets til arkivdepot*. Planen med prosjektet var å foreta uttrekk fra Aust-Agder idrettskrets sak- og arkivsystemer, som var lagret på Norsk Idrettsforbunds sentrale servere, for så å sikre dette for langtidslagring og tilgjengeliggjøring. Prosjektet ble påbegynt, men da prosjektleder gikk ut i en toårig studiepermisjon høsten 2016 måtte det utsettes inntil videre. Den opprinnelige prosjektplanen forutsatte et aktivt teknisk og administrativt samarbeid med Idrettskretsen både lokalt og sentralt, noe som ble vanskelig å gjenoppta etter en toårig pause. I 2018, da prosjektet skulle gjenopptas, fant vi at det derfor hensiktsmessig å gå bort fra den opprinnelige planen og heller starte opp prosjektet på nytt med en ny case.

4. Endringer i prosjektet - ny case

Lillesand Seilforening ble etablert i 1885 og har i dag ca. 250 medlemmer. Seilforeningen har tidligere avlevert papirarkiv til KUBEN fra helt tilbake til 1920-tallet.¹ I 2016 foretok Lillesand Seilforening en ny avlevering, og denne omfattet også foreningens digitalt skapte materiale fra perioden 2011-2016. Materialet, som var kopiert over på en minnepenn, viste seg å inneholde et omfattende bildemateriale, samt en god del dokumenter som ikke fantes på

¹ For arkivbeskrivelse se: https://www.arkivportalen.no/entity/no-AAKS_arkiv000000211598

papir, som f.eks. foreningens styreprotokoll. Materialet var organisert i en mappestruktur ut fra emne og årstall. Slike fil-/mappearkiver er nok typiske for norske lag og foreninger og det egnet seg derfor godt som case til vårt prosjekt.

Da tilgjengeliggjøring var et sentralt element i prosjektet, ble det inngått en avtale med Lillesand Seilforening om publisering av avlevert arkiv. Avtalen ligger vedlagt. Avtale går i korthet ut på at AAmå får rett til å publisere fotografier og dokumenter på Internett, enten i regi av egen publiseringsløsning eller hvis en tredjepartsløsning blir aktuell en gang i framtiden, f.eks. digitalarkivet. AAmå skal etter eget skjønn vurdere dokumentene før de publiseres. Det ligger også begrensninger i avtalen, bla at styreprotokoller yngre enn 5 år ikke kan publiseres, ei heller medlemslister og fotografier fra private sammenhenger.

5. Resultater fra prosjektet



Arkivar Yngve Schulstad Kristensen mottar arkivmateriale fra Lillesand Seilforening i 2016. Bjarte Olsen fra seilforeningen til høyre.

Det digitalt skapte arkivet fra Lillesand Seilforening var på 464 MB og besto av 1617 filer fordelt på 196 mapper. En nærmere analyse i DROID² viste at arkivet totalt besto av 47 forskjellige filformater.³ Sortert på filtype fordelte det seg på denne måten:

Filtype	Antall filer	Størrelse
jpg/jpeg	858	288,5 Mb
pdf	300	84,1 Mb
doc/docx	101	80,7 Mb
html/htm	189	6,1 Mb
xls/xlsx	34	1,6 Mb
ppt/pptx	7	285 Kb
avi/wmv	2	21,9 Mb
annet	126	

² DROID er et verktøy utviklet av britiske National Archives og som blant annet kan foreta automatisk identifikasjon av alle filformater i en fil-/mappestruktur

³ 47 forskjellige PUIDer

DROID v6.4

File Edit Run Filter Report Tools Help

New Open Save Export Add Remove Start Pause Filter On Report

Untitled-1 x

Resource	Extension	Size	Last modified	Ids	Format
D:\Lillesand Seilforening			09.05.18 12:25		
Bilder			09.05.18 12:25		
Brettgruppa			04.07.16 13:24		
Jollegruppa			09.05.18 12:25		
2003			09.05.18 12:25		
2004			09.05.18 12:25		
2005			09.05.18 12:25		
2006			09.05.18 12:25		
2007			09.05.18 12:25		
2008			09.05.18 12:25		
2009			09.05.18 12:25		
2010			09.05.18 12:25		
2011			09.05.18 12:25		
2012			09.05.18 12:25		
2013			09.05.18 12:25		
2014			09.05.18 12:25		
2015			09.05.18 12:25		
2016			09.05.18 12:25		
Instruksjon			09.05.18 12:25		
Baatregister_joller.htm	htm	84,2 KB	09.08.12 21:10		Hypertext Markup Language
påmeldingsskjema.doc	doc	41 KB	09.08.12 21:15		Microsoft Word Document
Snipeinfo.htm	htm	8,4 KB	09.08.12 21:11		Vector Markup Language
Seilbasen på kokkenes			09.05.18 12:25		
Styredokumenter			09.05.18 12:25		
Tur og Hav			09.05.18 12:25		
.dropbox		34 bytes	15.01.16 13:46		
desktop.ini	ini	163 bytes	15.01.16 13:46		Text Configuration file
LSF - Hedersbevisning v...	docx	2 MB	06.11.13 21:55		Microsoft Word for Windows

Arkivstrukturen slik den ser ut DROID

5.1 Ordning og bevaringsvurdering

I kategorien "annet" skjulte det filtyper som ikke kunne regnes som dokumenter og filer som ikke hadde noe dokumentasjonsinnhold eller som ikke var lesbare. Det var dessuten var det en del dubletter (f.eks. dokumenter i både doc- og pdf-format), samt mapper uten innhold. I likhet med et vanlig, uordnet papirarkiv måtte altså arkivet gjennom en ordningsprosess. Ordning av et fil-/mappearkiv kan i prinsippet kan foregå på samme måte som ordning av et papirarkiv, altså ved at materialet gjennomgås, og innhold som ikke er bevaringsverdig kasseres.

Mappestrukturen ble i det store og hele beholdt i sin opprinnelige form, men tomme mapper ble slettet. Etter ordning var arkivet redusert til 417 MB, 1356 filer og 175 mapper.

5.2 Valg av metodikk og verktøy

For å gjennomføre prosjektets målsetting har sett nærmere på hva andre har gjort før oss og hvilke anbefalinger som foreligger.

Det endte opp med at vi valgte å bruke dataprogrammet *Archivematica*⁴ for digital bevaring og «AtoM» for tilgjengeliggjøring. Archivematica er en løsning som etter hvert har fått et solid fotfeste i mange bevaringsinstitusjoner rundt om i verden, særlig i engelskspråklige land, og har et aktivt «community» av fagfolk. I Norge var Arbeiderbevegelsens arkiv og bibliotek tidlig ute med å prøve ut Archivematica. Ellers har Norsk Helsearkiv på Tynset tatt i bruk programmet og bruker det aktivt i sin drift.

Archivematica er utviklet som en Open Source Programvare, som sikrer at programmet blir tilgjengelig for ettertiden. Tanken rundt langtidsbevaring som til og med overlever programutvikleren, gjør at det kjennes tryggere for en institusjon som AAMA å bruke systemet. Open Source muliggjør tanken på langtidslagring fri for proprietære begrensninger, også for digitalt skapte privatarkiver.

5.3 Installering, oppsett og drift av bevaringsløsning

Erfaringene fra det første prosjektet (2016 – 2017) tilsa at vi måtte hyre inn IT-kompetanse for på å installere og drifte Archivematica i en skyløsning. Valget falt på det svenske firmaet ArkivIT AB. De kunne også tilby support og opplæring på systemet.

For å høre hvilke erfaringer andre arkivinstitusjoner hadde med Archivematica/AtoM besøkte prosjektgruppa i januar 2019 arkiv Sørmland i Eskilstuna, som allerede hadde fått installert systemet gjennom ArkivIT. Deretter deltok på en kort workshop i regi av ArkivIT i Stockholm.

Både bevaringsløsningen Archivematica og tilgjengeliggjøringsløsninger ble gjort tilgjengelig fra leverandørens datasenter-leverandør i Stockholm:

- <http://earkiv.aama.no> (Archivematica-installasjonen for KUBEN og Digitalt Depot)

⁴ Eller retttere sagt en løsning som består av en gruppe av programvareverktøy. Se: <https://www.archivematica.org/en/>

- <http://atom.aama.no> (AtoM-installasjon, tilgjengeliggjøringsløsning som pt. er åpen for alle)

Fra februar 2020 overtok Piql AS i Drammen kontrakten for drift og support av systemene. Hovedårsaken til skiftet av leverandør var en betydelig lavere årskostnad, samt at det er enklere juridisk å forholde seg til en leverandør som er bundet av norsk lovverk.

5.4 Opplæring i og bruk av Archivematica

Det er ikke til å legge skjul på at Archivematica er relativt komplisert og krevende å bruke. I tillegg kommer at programmet bruker metodikk og fagtermer som ikke nødvendigvis er velkjente og en del av den norske «arkivtradisjonen». Dessuten er programmet på engelsk. Det finnes omfattende brukerdokumentasjon både fra et teknisk ståsted og arkivfaglig (bla videoinstruksjoner), men å omsette denne kunnskapen til praksis knyttet til bevaringsarbeid i Archivematica viste seg å være til tider krevende.

En stor fordel med elektroniske arkiver er søkbarheten. Dette krever en annen tilnærming til ordning og katalogisering, da en i AtoM har muligheten til å søke direkte i dokumenter. Selv om mye av strukturen i arkivet er tradisjonelt oppbygd, er ikke tradisjonelle detaljerte katalogen like nødvendig som ved tradisjonelt papirarkiv.

Ett av de mest arbeidskrevende handlingene med programmet var at arkivpakkene blir lagret i en flat struktur. Vi så dessverre at ikke all metadata fra Archivematica ble med over til AtoM. Dette betød at vi måtte legge inn informasjonen manuelt. Blant annet må en godkjenne hvert enkelt foto for publisering. Dette var et nitidig arbeid ettersom store deler av arkivet bestod av bildematerialet. Arkivet var også ganske nytt, så krav om personvern står også sterkt. Det er dog mulig å begrense tilgang til enkeltdokumenter og serier. Det er også mulig å enten å sperre av eller ha begrenset inngang til hele eller deler av arkivet, men dette må gjøres manuelt inne i AtoM noe som også tar tid.

Det viste seg derfor svært nyttig og inspirerende å delta på en «Archivematica Camp» i Geneve, Sveits 22 – 24. okt. 2019. Deltakerne kom hjem med en utvidet forståelse om digital langtidsbevaring og om Archivematica som system. Campen ble også en arena for å utveksle erfaringer med andre brukere av Archivematica og bygge nettverk. Kontaktene vil bli nyttig i det videre arbeidet.

6. Erfaringer

Avtaler – personvernregler mv

Samfunnsoppdraget for en bevaringsinstitusjon eller et arkivdepot, ligger ikke bare i å bevare arkiver eller annen dokumentasjon for ettertiden, men å tilgjengeliggjøre dem for publikum og forskere. Arkivarer formulerer det ofte slik: «-Vi bevarer for å tilgjengeliggjøre». Det å ikke kunne åpne et arkiv for bruk, oppleves som svært krevende for de fleste som arbeider med arkivbevaring.

Når det gjelder digitalt skapte privatarkiver et utfordringen knyttet til dette temaet to-delt:

1) Skape forståelse og aksept gjennom dialog med avleverer (giver av arkivet), at det er nødvendig at arkivet skal åpnes for bruk. Dette kan oppleves som særlig vanskelig ift. digitalt skapte arkiver, som ofte er av nyere dato og som gjelder personer/hendelser som fortsatt lever og er aktuelle.

2) Utarbeide gode avtaler mellom avgiver og bevaringsinstitusjoner. Avtaler som både holder vann juridisk, er forsvarlig faglig sett og forståelige for «vanlige folk». Det er mange relevante lover og forskrifter som må tas hensyn til.

Det ville vært sterkt ønskelig hvis Arkivverket kunne utarbeidet noen standardavtaler for bevaringsinstitusjonene. Avtaler som vedrører ulike former for tilgjengeliggjøring, databehandleravtaler og personvern.

Kompetansebehov

Vårt inntrykk er at det er få som jobber faglig med privatarkiv i Norge som arbeider med bevaringsprogramvare spesifikt utvikla for digitalt skapte arkiver, slik som Archivematica. Skal bruken av slike systemer som Archivematica bli mer utbredt, må det etableres en

brukerbase av noe omfang, og som kan være et utgangspunkt for kompetansespredning knyttet til systemet rundt om i Norge. Arkivverket – som den største bevaringsaktøren innen privatarkiv i Norge – har en viktig rolle her. Den løsningen de faller ned på (dvs. valg av arkivbevaringsløsning) vil bety mye for den videre utviklingen og adaptasjonen av systemet i Norge.

Installering og drift av arkivbevaringsløsninger og digitalt depot

Det er vår erfaring at eksterne leverandører er de beste til å sette opp og drifte Archivematica. Både fordi det krever teknisk kompetanse som få arkivinstitusjoner har, og fordi stabil og robust it-drift krever et profesjonelt driftsmiljø.

Utfordringer at dette koster penger. En slik tjenesteutsetting er en relativt kostbar affære. Vi snakker om priser fra ca 80.000 kr pr. år og oppover. Dette vil utgjøre nye årlige driftsutgifter inn i allerede pressede driftsbudsjetter. Å etablere digitale bevaringstjenester for privatarkiv med tilhørende inntektsstrøm er vanskelig. For bevaringsinstitusjoner vil det bety at det ikke er ikke økonomisk bærekraftig, slik situasjonen er i dag.

De samme utfordringene gjelder for etablering og drift av digitalt depot for privatarkiv. I alle fall hvis en tenker seg det som en skybasert løsning med tilstrekkelig kapasitet og sikkerhet. Her må en først og fremst langsiktighet, robusthet og høy sikkerhet.. Det kreves et profesjonelt driftsmiljø, og vil bety store årlige beløp på leie av maskin- og nettverkskapasitet. KDRS (Kommunearkivinstitusjonenes Digitale RessursSenter) har etablert et slik depot for sine medlemmer. Men det er utelukkende snakk om digitalt skapte arkiver fra kommunesektoren.

7. Anbefalinger – hva bør gjøres framover?

7.1 Nasjonale fellesløsninger

Vi tror at de fleste som arbeider profesjonelt med privatarkiv i dag, erkjenner at de bevarer for lite av digitalt skapte privatarkiv fra sitt område. Mange opplever at de rett og slett ikke makter å ivareta bevaringsoppdraget sitt.

Skal arbeidet med bevaring og tilgjengeliggjøring av digitale privatarkiver få et løft, må det etableres nasjonale fellesløsninger, slik som ASTA var det i sin tid.

Det grunnleggende går på å skaffe bevaringsinstitusjonene tilgang til moderne og relevante verktøy for digital bevaring, og tilgang et felles digitalt depot, drevet etter langsiktige avtaler og med god sikkerhet. Vi ser å det som naturlig at Arkivverket går i spissen for å etablere slike fellesløsninger, gjerne i partnerskap med Arkivforbundet. Arkivforbundet representerer en stor del av institusjonene som bevarer privatarkiv i Norge. Slike fellesløsninger må naturligvis prissettes slik at de aller fleste har muligheter får å ta dem i bruk.

7.2 Kompetanse

Like viktig som nasjonale fellesløsninger, er kompetanse i sektoren på behandling og bevaring av digitale privatarkiver. Det er behov for å etablere brede og robuste fagmiljøer for digital «konservering» av privatarkiver. De finnes så å si ikke i dag.

Det er behov for å lage guider og håndbøker i en norsk kontekst knyttet til tema, som kan sette standarder og informere om «best practice» i fagfeltet. Målet bør være at alle bevaringsinstitusjoner innen 5 år, og som har som mål å bevare digitale privatarkiv, selv har kompetanse, eller har tilgang til fagfolk med kompetanse, på bevaring, lagring og tilgjengeliggjøring av digitale privatarkiv. Dette på samme måte som det er en selvfølge å ha fagpersonell med kompetanse på «papirarkiv» i egen organisasjon.

Vedlegg

Avtale om publisering av avlevert arkiv, mellom AAMA og Lillesand Seilforening