



ARKIVVERKET
RIKSARKIVET

Samdok

samla samfunnsdokumentasjon

RAPPORT 2016



PRIORITERT OPPGAVE

**7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter
Prosjektrapport – sluttrapport med anbefalinger**

Utarbeidet av
Anette Skogseth Clausen m.fl.

Rapportdato
6. jan. 2017



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 1 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

OPPGAVE	<i>Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter</i>
Ansvarlig delprosjekt	<i>Kommunale arkiv</i>
Arbeidsgruppens leder	<i>Anette Skogseth Clausen (Arkivverket)</i>
Arbeidsgruppens medlemmer	<i>Espen Tønnesen (Arkivverket), Kristian Hunskaar (Arkivverket), Ingrid Jørgensen (Arbak), Aasta Karlsen (Arkiv i Nordland), Espen Sæterbø (Fylkesarkivet i Sogn og Fjordane), Øystein Eike (Oslo byarkiv), Ottar André Anderson (Ika Møre og Romsdal), Snorre Dag Øverbø (Aust-Agder Museum og Arkiv), Tove Wefald Pedersen (Norsk folkemuseum)</i>
Målformulering	<i>Målet for forprosjektet er en beskrivelse av hvordan nasjonal publiseringsplattform kan realiseres faglig, teknisk, driftsmessig, økonomisk og organisatorisk – både i en første versjon og på lengre sikt. Skal omfatte «back end» (for innholdsleverandører) og «front end» (for brukere). Løsningen skal kun gjelde for skannede arkivdokumenter.</i>
Sammenheng	<p><i>Forprosjektet har i løpet av 2016 sett på hva som må til for å tilby alle depotinstitusjoner en plattform for publisering/tilgjengeliggjøring av sitt arkivmateriale slik at publikum skal ha en inngang til digitalisert arkivmateriale. Det lå i arbeidsgruppens mandat å kun se på publisering av skannet arkivmateriale.</i></p> <p><i>Arbeidet ble strukturert etter delleveransene som skulle leveres i løpet av prosjektperioden (se kapittel 2. Prosjektets produkter). Arbeidsgruppen skisserte først de funksjonelle og tjenesterelaterte kravene til en slik publiseringsplattform (delrapport 1a). Noe av det sentrale i disse kravene er stabile, automatiserte prosesser som gjør innholdsleverandørene (depotinstitusjonene) selvbetjente og effektive i publiserings- og administrasjonsoppgavene for sitt materiale. Samtidig er det et krav at publikum skal ha gratis tilgang til alt arkivmateriale som kan være fritt tilgjengelig. Det skal være enkelt for publikum å finne fram til arkivmateriale og publikum skal være selvbetjente. I neste delrapport (delrapport 1b), som beskriver det tekniske og skal være et konkret løsningsforslag, har gruppen vært innom flere alternativer for en plattform. Det var likevel tydelig ganske tidlig i prosjektperioden at Digitalarkivet ville egne seg best som publiseringsplattform for arkivmateriale også tatt i betraktning de krav som lå til grunn (se kapittel 3c. om Digitalarkivet). Ettersom det allerede var etablerte planer som ville muliggjøre mange av kravene i første delrapport, forsterket dette ytterligere arbeidsgruppens syn om Digitalarkivet som det beste alternativet (se vedlegg 2). Det konkrete løsningsforslaget tok derfor utgangspunkt i Digitalarkivet og de planene som allerede foreligger. Arbeidsgruppen legger vekt på at videreutviklingen av Digitalarkivet må skje i tett dialog med innholdsleverandørene. Organiseringen av et hovedprosjekt for utviklingen av en slik publiseringsplattform er foreslått slik at innholdsleverandørene er sterkt involvert (delrapport 1c). Det vil være naturlig at Arkivverket som eier og drifter av Digitalarkivet, leder utviklingsarbeidet, men at prosjektorganisasjonen er bredt sammensatt med innholdsleverandører representert i styringsgruppe, arbeidsgruppe og referansegruppe. Prosjektperioden er anbefalt å gå over to år for å kunne ha en god utviklingsprosess hvor moduler i løsningen tas i bruk av innholdsleverandører fortløpende. Det er også anbefalt at Arkivverket tar de største kostnadene for dette, men at depotinstitusjonene dekker egne kostnader i forbindelse med prosjektet. Totalt for den to-årige prosjektperioden er det estimert 11,2 millioner i kostnader, inkludert estimerte</i></p>



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 2 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

	<p>årsverk og reiseutgifter. Her vil Arkivverket stå for 7,6 millioner. Både for hovedprosjektet og for så vidt også videre drift, er det viktig at dette er en tjeneste alle depotinstitusjoner kan ta i bruk gjennom litt egeninnsats og godt skanneutstyr. Arbeidsgruppen anbefaler derfor at tjenesten er gratis, eller i hvert fall svært rimelig å ta i bruk. I ROS-analysen (delrapport 1d) er det avdekket få alvorlige risikoer for hovedprosjektet, det mest alvorlige berører informasjonssikkerhet hvor ukyndig bruk eller sikkerhetshull kan gi brukere tilgang til informasjon de ikke skulle sett. Men samtidig er det ikke vurdert som stor sannsynlighet at dette skal inntreffe. Det er større sannsynlighet for forsinkelser i prosjektet, f.eks. gjennom manglende utviklingsressurser eller tap av kompetanse. For å redusere sannsynligheten for dette er det derfor anbefalt at prosjektorganisasjonen består av en egen utviklergruppe som arbeider utelukkende med dette prosjektet.</p> <p>Denne sluttrapporten er dels en oppsummering av forprosjektet og dels noen anbefalinger for det videre arbeidet i hovedprosjektet. Rapporten påpeker også hva som ikke blir løst i løpet av hovedprosjektet og som evt. må løses i nye prosjekter.</p>
Oppfølging 2017	<p>- Det anbefales å iverksette et 2-årig hovedprosjekt i 2017. Prosjektet bør ledes av Arkivverket, men med bred og god sammensetning av representanter fra depotinstitusjoner i prosjektorganisasjonen.</p>
Vedlegg	<p>1 skjematisk oversikt over krav og om Digitalarkivet svarer på kravene 1 redegjørelse om digitaliseringsanbefaling 4 delrapporter (som separate filer)</p>



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 3 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Innhold

Sluttrapport med anbefalinger	4
1. Arbeidsform og måloppnåelse for prosjektet	4
2. Prosjektets leveranser	4
3. Anbefalinger.....	4
4. Avgrensinger ved hovedprosjektet	9
Vedlegg 2:	11
Vedlegg 3 Minimumskrav til kvalitet i digitalisering av papirbasert materiale.....	26
Filformater og oppløsning.....	26
Metode og valg av utstyr i «volum-digitalisering».....	26
Metode og valg av utstyr i digitalisering av store format	28
Metode og valg av utstyr for digitalisering av skjøre originaler.....	30
Kvalitetskontroll.....	32
Vurdering av behovet for utarbeidelse av veiledning for «best practice» for digitalisering av papirbasert materiale.....	32



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 4 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Sluttrapport med anbefalinger

1. Arbeidsform og måloppnåelse for prosjektet

Arbeidsgruppen hadde sitt første møte 12. april og det siste 15. november. Det ble gjennomført fire fysiske møter i Riksarkivet og to skype-møter. Det har ikke vært fulltallig på møtene, men det har vært god oppslutning og gode diskusjoner.

Arbeidsgruppen har levert grundige leveranser - delrapporter - underveis i prosjektet, som har svart på bestillingen. Samtidig har det vært avklaringspunkter som gruppen har valgt å overlate til hovedprosjektet dels fordi det ikke har vært tid i løpet av forprosjektet til å gå i dybden på noen problemstillinger, og dels fordi det må en viss modenhet/erfaringer til i prosjektet før man kan gi en god anbefaling/konklusjon. Dette gjelder problemstillinger om samspill/interaksjon med andre relevante innholdstjenester (kapittel 3d), kontekstinformasjon for uvante brukere (se kapittel 3e), samt å se på drift av en slik løsning på lengre sikt.

2. Prosjektets leveranser

Produkt	Status	Ansvar videre
1a: Funksjons- og tjenestebeskrivelse ("back end" og "front end")	Levert 14.06.2016	Arkivverket, gjennom arbeidsgruppen for hovedprosjektet
1b: Teknisk beskrivelse og løsningsforslag	Levert 01.11.2016	Arkivverket, gjennom styringsgruppen og utviklergruppen for hovedprosjektet
1c: Forslag til eierskap, organisering, driftsorganisasjon, milepæler og kostnadsramme for hovedprosjekt (utvikling, etablering, og drift første år)	Levert 15.12.2016	Arkivverket, gjennom styringsgruppen for hovedprosjektet
1d: Risiko- og sårbarhetsvurderinger for hovedprosjektet	Levert 16.12.2016	Arkivverket, gjennom styringsgruppen og arbeidsgruppen for hovedprosjektet

3. Anbefalinger

Det er arbeidsgruppens anbefaling at det iverksettes et to-årig hovedprosjekt for å etablere en nasjonal publiseringsplattform for skannede arkivdokumenter. Som rapporten viser vil Digitalarkivet være den anbefalte løsningen å bygge videre på. Ettersom Arkivverket er ansvarlig for Digitalarkivet og har allerede pågående utvikling i gang for tjenesten, vil det være mest hensiktsmessig at Arkivverket er prosjekteier, men at det samtidig etableres en bred sammensatt prosjektgruppe med representanter fra depotinstitusjoner. I det følgende vil ytterligere anbefalinger være beskrevet. Det henvises også til delrapportene for detaljerte beskrivelser.

a. Utviklingsløp

Utviklingsløpet for en nasjonal publiseringsplattform er beskrevet godt i delrapport 1c. Utviklingsløpet er anbefalt å gå sammen med den allerede planlagte utviklingen for Digitalarkivet. Dette henger sammen med at arbeidsgruppen anbefaler Digitalarkivet som utgangspunkt for den nasjonale publiseringsplattformen (se kapittel 3c). Utviklingen vil foregå i flere faser og utvikle moduler underveis, hele tiden i tett dialog med arbeidsgruppen og referansegruppen for hovedprosjektet. I disse



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 5 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

gruppene vil det være stor representasjon av depotinstitusjoner som vil være innholdsleverandører til den nasjonale publiseringsplattformen og ta i bruk løsningene etterhvert som de ferdigstilles. På den måten får utviklergruppen gode tilbakemeldinger på hva som fungerer og hva som må bli bedre. Utbedringer gjøres underveis og nye versjoner lanseres for innholdsleverandørene. En slik iterativ utviklingsmetode skaper god brukermedvirkning.

Selve utviklingen av moduler som er nødvendige for den nasjonale plattformen er skissert til å ta ett år, men ettersom det også legges opp til justeringer i modulene underveis, vil det ventelig gå noe mer tid før alle moduler er på plass. Samtidig vil det være vesentlig å tilpasse nettsidene for Digitalarkivet til å bli en portal/plattform for alle innholdsleverandører, og dette vil også ta noe tid. Utviklingsløpet er derfor satt til to år.

Her er den foreslåtte utviklingsplanen for nasjonal publiseringsplattform (hentet fra delrapport 1c):

Leveranse	Beskrivelse	Iverksettes	Avsluttes	Avhengigheter
Serverpark	Servere dimensjonert for store mengder billedfiler	2017	2017	Innkjøp hos Arkivverket
Indekseringsmodul	Metadata-tilknytning til hvert bilde. Modulen vil erstatte dagens løsning som kun kan brukes bak brannmur hos Arkivverket. Indekseringsmodulen vil kunne brukes av alle depotinstitusjoner, og i tillegg kan det iverksettes crowdsourcing for indeksering av bilder hos publikum.	01.09.16	20.12.16	
Prosesseringsrigg	Et system for å kjøre bildene gjennom ulike automatiserte prosesser fra skanning til langtidslagret fil og publisert fil via filkonverteringsprosess, sladdeprosess (ved behov), OCR-prosess (ved behov) ol. Denne riggen vil også kunne utvides med ytterligere funksjonalitet/prosesser for behandling av andre type filer enn de produsert gjennom skanning, f.eks. lyd og film, store datasett fra digitalt skapte arkiver mv.	02.01.17	01.04.17	
Produksjonsadministrasjon	En administrasjonsløsning for å følge hele produksjonslinjen fra planlegging av skanning til ferdig publisert arkivdokument i Digitalarkivet. Her kan det planlegges hva som skal skannes på hvilke skannere og når, overordnet metadataarbeid for arkivdokumentet kan gjøres her og det som fins i ASTA/Arkivportalen kan hentes ut, mv.	01.04.17	01.10.17	
Nettsider	Etablere Digitalarkivet på domene og løsning uavhengig av Arkivverket. Nytt design og løsninger som ivaretar innholdsleverandørens behov for profilering.	14.11.17	01.04.18	

b. Faglige problemstillinger i forbindelse med etablering

I delrapport 1c er det belyst noen problemstillinger som hovedprosjektet må se på i forbindelse med etableringen av en nasjonal publiseringsplattform. Det gjelder:



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 6 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

- redaktøransvar og publiseringspolicy
Drøfte problemstillinger knyttet til publisering av dokumenter og hva som kan ligge fritt tilgjengelig og hva som bør skjermes, juridiske avklaringer og redaktørrollene. Evt. utarbeide felles retningslinjer for publisering.
- lisensiering av innhold
Drøfte problemstillinger knyttet til lisensiering av materialet - er alt CC eller skal det andre lisenser på materialet? Evt. utarbeide felles retningslinjer for lisensiering av innhold.
- samspill med andre tjenester
Ha dialog med Arkivportalen, DigitaltMuseum, Nasjonalbiblioteket mm. med tanke på interaksjon eller samspill i tjenestene
- driftsorganisasjon og finansiering
Beskrive driftsorganisasjonen og håndtering av daglig drift når prosjektfasen er avsluttet
- vurdere annet innhold enn skannet arkivmateriale
Se på muligheter for å utvide plattformen til også å inkludere andre formater, som data (digitalt skapt og transkribert fra analoge kilder), lyd, video, mm.

Avklaringer av disse problemstillingene er innbakt i en milepælsplan foreslått av arbeidsgruppen (hentet fra delrapport 1c):

Leveranse	Beskrivelse	Avhengigheter	Frist
Hovedprosjektorganisasjonen etableres	Mandat defineres, detaljert milepælsplan for 2017	vedtak hos Arkivverket om hovedprosjektet, prosjektleder	1.kvartal 2017
Skannede arkivdokumenter	Bilder av arkivdokumenter, skannet i henhold til minimumskrav, klargjort for publisering	beskrivelse av minimumskrav for skanning	1.kvartal 2017, løpende
Innspill til forbedringer	Teste og gi innspill til utviklede moduler for NAPSA		løpende
Delrapport 1: Anbefaling om redaktøransvar og publiseringspolicy	Drøfte problemstillinger knyttet til publisering av dokumenter og hva som kan ligge fritt tilgjengelig og hva som bør skjermes, juridiske avklaringer og redaktørrollene. Evt. utarbeide retningslinjer for publisering.		2. kvartal 2017
Delrapport 2: Anbefaling om lisensiering	Drøfte problemstillinger knyttet til lisensiering av materialet - er alt CC eller skal det andre lisenser på materialet? Evt. utarbeide retningslinjer for lisensiering.		2. kvartal 2017
Vurdere samspill med andre tjenester	Dialog med Arkivportalen, DigitaltMuseum, Nasjonalbiblioteket mm. mtp interaksjon og samspill i tjenestene		3. kvartal 2017
Delrapport 3: Anbefaling om finansiering	Dialog med KulturIT, Asta-Stiftelsen, ledelse i Arkivverket m.fl. om mulige løsninger for finansiering av videre drift.		4. kvartal 2017
Delrapport 4: Anbefaling om drift	Beskrive driftsorganisasjonen og operering av daglig drift		2. kvartal



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 7 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Leveranse	Beskrivelse	Avhengigheter	Frist
			2018
Vurdere annet innhold enn skannet arkivmateriale	Se på muligheter for å utvide plattformen til også å inkludere andre formater, som data, lyd, video		3. kvartal 2018
Sluttrapport	Oppsummerer prosjektet og foreslår videre løp		4. kvartal 2018



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 8 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

c. Vurdere hvorvidt digitalarkivet.no med tilhørende programvare kan benyttes/utvikles

Prosjektgruppa anbefaler at Digitalarkivet blir valgt som løsning og utgangspunkt for realisering av nasjonal publiseringsplattform. Underveis i forprosjektet har prosjektgruppa imidlertid vurdert ulike modeller og løsninger for å kunne realisere en slik plattform.

Vi vurderer det slik at det grovt sett er tre hovedalternativer – i tillegg til Digitalarkivet – når det gjelder utvikling, etablering og drift av NAPSA:

Alternativ 1: Kjøpe tilsvarende tjenester i markedet. («Hylleware») Det finnes en rekke tjenester på Internett for opplasting, tilgjengeliggjøring og deling av mediefiler av ulikt slag (tekst, lyd, bilde, video). Det finnes også tjenesteleverandører som spesialiserer seg inn mot kulturarvssektoren med å tilgjengeliggjøre arkivdokument, foto, manuskripter osv. I USA og Europa har flere nasjonale aktører innen museum, arkiv og bibliotek kjøpt slike tjenester. Velges en slik løsning, - uansett «hylleware» eller ikke – må det trolig gjøres relativt omfattende tilpasninger og utvidelser for å støtte kravspesifikasjonene og funksjonaliteten som ligger til grunn for den nasjonale publiseringsplattformen (se delrapport 1a). Det vil også være betydelig usikkerhet rundt de budsjett- og framdriftsmessige konsekvensene i et utviklingsløp. Utviklings- og driftskompetanse blir værende hos eksterne leverandører, og vil i liten grad tilflytte arkivsektoren ved Arkivverket. *Prosjektgruppa anbefaler at en ikke går videre med en slik løsning.*

Alternativ 2: Bygge Nasjonal publiseringsplattform fra bunnen av. Det er også mulig å bygge plattformen fra «bunnen av». Det kan gjøres ved å sette bort teknisk kravspesifikasjon, systemutvikling og drift til eksterne konsulenter/firma etter en anbudsprosess. Her er også økonomi og framdrift hovedankepunkt. Det vil være høye kostnader forbundet med å inngå sikre avtaler med eksterne konsulenter og tjenesteleverandører for å innfri kravspesifikasjonene, funksjonaliteten og framdrift i hovedprosjektet. I den sammenhengen vil Arkivverket som allerede har en tjeneste for publisering av skannet arkivmateriale, måtte vurdere om det er kostnadmessig forsvarlig for etaten å reinvestere i en ny plattform og legge ned den eksisterende, eller å bli værende i den eksisterende, med de konsekvenser det vil få for å etablere en felles plattform for alle depotinstitusjoner. Dessuten vil utviklings- og driftskompetanse bli værende hos eksterne leverandører, og i liten grad tilflytte arkivsektoren ved Arkivverket. *Prosjektgruppa anbefaler at en ikke går videre med en slik løsning.*

Alternativ 3: Kjøpe seg inn i / samarbeide med eksisterende innholdstjenester. Det finnes flere innholdstjenester innen kulturarvssektoren i Norge som det kunne vært aktuelt å samarbeide med. De mest relevante er trolig Arkivportalen og Nasjonalbiblioteket. (DigitaltMuseum er også en mye brukt innholdstjeneste, men konsentrerer seg mest om foto og levende bilder, og blir derfor ikke omtalt videre her). Gjennom en drifts- og utviklingsavtale med aktørene bak en av disse tjenestene, kan vedtatte løsningsbeskrivelsene for Nasjonal publiseringsplattform gjennomføres innenfor rammene av en eksisterende innholdstjeneste.

- *Spesielt om Arkivportalen.* Tjenesten er eid av Stiftelsen ASTA, og er i stor grad bygd opp rundt arkivkatalogen, slik vi kjenner den fra ASTA5. For å kunne være en nasjonal publiseringsplattform der hovedfokus er digitaliserte dokumenter, måtte denne løsningen blitt rigget om. En del av plattformens innholdsleverandører vil trolig aldri bli ASTA-kunder/-brukere, og som sådan være utelukket fra å bruke Arkivportalen, i alle fall slik avtalene ligger nå. Når vi videre har sett på de mulighetene som ligger i en utvidelse/modifisering Digitalarkivet, er prosjektgruppa overbevist om at det både er raskere, enklere og billigere å videreutvikle Digitalarkivet, som allerede er inne i et pågående utviklingsløp i regi av Arkivverket, og som uansett skal ta opp i seg mye av den funksjonaliteten som er planlagt innbakt i den nasjonale publiseringsplattformen.
- *Spesielt om Nasjonalbiblioteket (NB).* De har en omfattende innholdstjeneste, som hovedsakelig inneholder digitaliserte versjoner av trykte bøker, periodika og aviser. Men tjenesten inneholder også en del «upublisert» materiale, som manuskripter og brevsamlinger. Interessant er også at NB benytter seg OCR-prosessering av mediefilene der det er mulig, noe som også er viktig for den nasjonale publiseringsplattformen. NB har en omfattende digitaliseringskapasitet som dekker sitt ambisiøse nasjonale digitaliseringsprogram for trykte (publiserte) medier. Den nasjonale publiseringsplattformen legger opp til en åpen modell der



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 9 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

innholdsleverandørene (sammen med Arkivverket) står for all produksjon, og er likestilte seg imellom, med Arkivverket som systemeier i en mer fasiliterende og tilretteleggende tilleggsrolle. Det kan være utfordringer innenfor hovedprosjektperioden å få på plass gode avtaler som sikrer alle sider ved et slikt mulig samarbeid med NB. I tillegg må det gjøres en del endringer og utvidelser av NBs innholdstjeneste for å imøtekomme publiseringsplattformens funksjonskrav. Når prosjektgruppa har vurdert NBs løsning i forhold til de mulighetene som ligger i en utvidelse/modifisering Digitalarkivet, er prosjektgruppa overbevist om at det både er raskere, enklere og billigere å videreutvikle Digitalarkivet, som er inne i et pågående utviklingsløp i regi av Arkivverket, og som uansett skal ta opp i seg mye av den funksjonaliteten som er planlagt innbakt i den nasjonale publiseringsplattformen.

Selv om prosjektgruppa anbefaler å bygge på Digitalarkivet som løsning, ser vi det som viktig å etablere god dialog med Arkivportalen, Nasjonalbiblioteket og DigitaltMuseum underveis i hovedprosjektet. Dette er derfor foreslått i milepælsplanen. Målsettingen bør være størst mulig grad av integrasjon og datadeling/-utveksling til glede for fagmiljøer og sluttbruker

d. Samspill med andre relevante innholdstjeneste

Arbeidsgruppen har sendt ut brev til Nasjonalbiblioteket, Asta-stiftelsen, Kultur-IT (Digitalt museum), Kulturrådet (Norvegiana) og orientert om prosjektet og også lagt ved funksjons- og tjenestebeskrivelsen (delrapport 1a). Responsen har ikke vært overveldende (kun Asta og Kulturrådet) og samtidig har det ikke vært god nok tid i prosjektet til å gå i tett dialog med disse innholdstjenestene. Den lave responsen kan også være en reaksjon på at det er vanskelig å reagere/uttale seg om muligheter for samspill basert på en funksjons- og tjenestebeskrivelse. Arbeidsgruppen føler derfor at dette vil være en dialog det er lettere å få til når hovedprosjektet er i gang og det er konkrete resultater underveis det kan vises til. Det vil ikke være et forsinkende element at samspillet ikke er avklart før prosjektet går i gang.

e. Vurdere innlegging av kontekstinformasjon for uvante brukere

Arbeidsgruppen har i liten grad drøftet behovet for kontekstinformasjon. Ettersom det anbefalte alternativet er å videreutvikle Digitalarkivet til en nasjonal publiseringsplattform, er det relevant å se på planene som allerede ligger for Digitalarkivet. I disse planene satses det i stor grad på å gjenbruke metadata/arkivbeskrivelse fra Arkivportalen. Det vil være tungvint og uhensiktsmessig for depotinstitusjoner å vedlikeholde arkivinformasjon flere steder. Samtidig vil det være depotinstitusjoner som ikke benytter Arkivportalen og for disse må det finnes alternative løsninger for å kunne legge inn metadata om arkivmaterialet. Denne vurderingen om kontekstinformasjon bør derfor tas i forbindelse med vurderingene om samspill med andre relevante innholdstjenester (kapittel 3d) og underveis i videreutviklingen av modulene for den nasjonale plattformen. Denne avklaringen flyttes derfor til hovedprosjektet.

4. Avgrensinger ved hovedprosjektet

En realisering av hovedprosjektet vil gi publikum en inngang til det digitaliserte arkivmaterialet og det vil gi depotinstitusjonene et felles verktøy for publisering av sitt arkivmateriale. Samtidig vil ikke dette prosjektet løse alle behovene hos depotinstitusjonene.

Prosjektet vil arbeide med digitalisert papirmateriale, i stor grad protokoller, dokumenter, brev og lignende. Kilder på andre formater som digitalt skapte data, lyd og film, vil ikke være en del av prosjektet. Samtidig vil løsningen som utvikles for en stor del være generisk, slik at med forholdsvis enkle grep (særlig knyttet til presentasjon på nettsidene), vil det være mulig å utvide løsningen også til disse formatene.

Videre vil ikke løsningen som tilbys være en innsynsløsning som også er et stort behov blant depotinstitusjonene. Likevel, ettersom løsningen vil ha tilgangsstyring, kan institusjonene tildele enkeltpersoner innsyn i enkelt dokumenter, men det vil ikke tilsvare en fullgod og automatisert innsynsløsning.



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 10 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Selv om prosjektet tar for seg digitalisering av papirbaserte arkivdokumenter, vil ikke løsningen kunne tilby produksjon, dvs. skanning/digitalfotografering, av kildene. Løsningen vil likevel kunne tilby depotinstitusjonene et administrasjonsverktøy som gjør det enklere å planlegge produksjonen og styre køen av kilder ved det forskjellige skanningsutstyret. Det ligger også i prosjektets plan å sette opp minimumskrav for produksjonen, slik at publikum er sikret digitale reproduksjoner av arkivdokumentene i god nok kvalitet.



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 11 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Vedlegg 2:

Nedenfor er kravene fra funksjons- og tjenestebeskrivelsen satt opp skjematisk, med ei vurdering av i hvilken grad Digitalarkivet oppfyller kravene.

NAPSA=Nasjonal publiseringsplattform for skanna arkiver

+ = oppfyller krav

- = oppfyller ikke krav

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
NAPSA skal være en gratis, alternativt veldig rimelig, tjeneste for innholdsleverandørene.		N/A	Et strategisk valg som ikke angår Digitalarkivet <i>som løsning</i> .	Fillagringsløsningen kan måle brukt lagringskapasitet i forskjellige segmenter (buckets) av den totale bruken. Det skal utvikles telling av hvor mye hver kilde er i bruk, slik at det kan knyttes opp mot arkivinstitusjon for en eventuell avregning.	N/A
NAPSA skal være en skalerbar tjeneste mtp mengde og innholdstyper.	Tjenesten skal være bygd for store datamengder og massepublisering i industriell skala.	+	Digitalarkivet håndterer store datamengder. Det gjenstår å utvikle ulike administrative funksjoner som vil forenkle metadata- og publiseringsarbeid. Arkivverket har egeninteresse av å utvikle disse funksjonene, og arbeidet er i planleggingsfasen.	Alle systemer og løsninger er skalerbare med automatisk lastbalansering innen utgangen av august. Det betyr at ytelsen kan økes ved behov, og systemet vil ikke gå ned hvis en server skulle få en fatal feil.	+
	Tjenesten skal ikke være låst for	+	Digitalarkivet er ikke låst for videreutvikling.		+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 12 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	videreutvikling.				
	Nye innholdstyper skal kunne komme til.	-	Arkivverket har egeninteresse av å utvide Digitalarkivet med f.eks. lyd og levende bilder, men dette ligger foreløpig et stykke fram i tid.	Den planlagte prosesseringsriggen for mottak og konvertering av filer skal kunne utvides. Fillagringsløsningen håndterer alle typer filer, og vi har en prototyp på en løsning for lydfiler allerede, men det må mer til for å streame video. Men det vil ikke være noe teknisk hinder for å utvide med slik støtte på sikt.	+
NAPSA skal bygge på åpne standarder.		+	Digitalarkivet bygger på åpne standarder.		+
NAPSAs innholdsdata skal være tilgjengelige gjennom et åpent API. APIet kan forholde seg til tilgangsstyringa i NAPSA, og det er derfor ikke nødvendig å holde deler av innholdsdataene utenfor APIet.		+	Digitalarkivet har et API for Gammel grunnbok som inkluderer tilgangsstyring. Det er intensjonen å lage tilsvarende APIer for det øvrige innholdet i Digitalarkivet.		+
Den enkelte innholdsleverandøren skal ha fullt eierskap til sitt innhold.	Publisert innhold skal kunne trekkes tilbake, helt eller delvis.	+	I Digitalarkivet kan bilder enkelt avpubliseres.		+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 13 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	Innhold som er publisert fritt tilgjengelig, kan siden bli regulert med adgangskontroll.	+	I Digitalarkivet kan tilgangsstyringa for enkeltbilder og hele arkivenheter når som helst endres.		+
Den enkelte innholdsleverandøren er alene ansvarlig for at sitt fritt tilgjengelig publiserte innhold kan publiseres på Internett.	Taushetsbelagt innhold skal ikke publiseres fritt tilgjengelig.	+	Det er mulig å publisere med adgangskontroll i Digitalarkivet.		+
	Innholdsleverandører (organ) som er underlagt Offentleglova, skal ikke publisere opplysninger som nevnt i Forskrift til offentleglova § 7 (http://lovdata.no/forskrift/2008-10-17-1119/§7), fritt tilgjengelig.	N/A	Dette må innholdsleverandørene selv følge opp, uavhengig av løsning.		N/A
	Det må framgå tydelig hvem som er innholdsleverandør, m.a.o. hvem som er ansvarlig.	-	I Digitalarkivet vises det i dag tydelig hvor den skanna arkivenheten oppbevares (f.eks. Oslo byarkiv: https://media.digitalarkivet.no/db/browse?depot%5B%5D=OBA). Informasjon om hvem som er innholdsleverandør finnes delvis og til dels tilfeldig, og er under alle omstendigheter godt gjemt. Det er ikke mulig å markedsføre innholdsleverandøren tydelig. I Digitalarkivet er det p.t. underforstått at Arkivverket/Riksarkivaren er ansvarlig for alt innhold.	Nettsidene er planlagt oppgradert i løpet av 2018 nettopp for å kunne ha en bedre profilering for innholdsleverandører.	+
NAPSA skal være organisert med ett		+	Digitalarkivet oppfyller disse kravene.		+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 14 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
sentralt fillager, én sentral database med metadata på arkivenhetsnivå (metadata/kataloginformasjon som er felles for alle bilder knyttet til en arkivenhet) og metadata på filnivå (metadata som gjelder ett eller noen av bildene knyttet til en arkivenhet), og ett sentralt brukerregister.					
Innholdsleverandørene skal kunne utføre operasjoner mot fillager og database via et administrasjonsgrensesnitt, som de har tilgang til gjennom innlogging med brukernavn og passord.	<p>En innholdsleverandør kan ha flere brukerprofiler (brukernavn/passord), for eksempel én brukerprofil for hver medarbeider.</p> <p>Hver brukerprofil skal kunne gis spesifikke og avgrensede rettigheter etter behov. Typisk vil en brukerprofil ha rettigheter til å administrere filer og metadata som tilhører vedkommendes arbeidsgiver (innholdsleverandør).</p> <p>Et begrenset antall brukerprofiler skal ha admin-rettigheter for bl.a. å kunne tildele rettigheter til andre brukere.</p>	-	<p>Det er ikke mulig å utføre alle nødvendige operasjoner i Digitalarkivets nåværende administrasjonsgrensesnitt for databasen. Mange operasjoner kan kun utføres med SQL-kommandoer direkte mot databasen. Som nevnt lenger oppe, har Arkivverket egeninteresse av å utvikle administrasjonsgrensesnittet videre. Det finnes ikke noe administrasjonsgrensesnitt for fillageret. Det finnes et eget administrasjonsgrensesnitt for tilgangsstyring/adgangskontroll.</p> <p>Det er mulig å ha én brukerprofil for hver medarbeider, og admin-rettighetene er begrenset til enkelte brukerprofiler. Det er så langt ikke mulig å avgrense rettighetene til å gjelde innhold fra én innholdsleverandør.</p>	De nødvendige administrasjonsverktøyene er under planlegging, og det vil være naturlig å ta høyde for disse tingene under planlegging og utvikling.	+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 15 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
Det skal være mulig å registrere metadata på arkivenhetsnivå i databasen via administrasjonsgrensesnittet.	Dersom en arkivenhet er registrert i Arkivportalen, skal registreringen starte med å opprette ei kobling til oppføringa i Arkivportalen og høste relevante data derfra. Databasen skal jevnlig synkroniseres mot Arkivportalen, slik at endringer i arkivenhetens kataloginformasjon tilflyter databasen (forutsatt at arkivenhetens Asta-ID er uendret).	+	Ved å legge inn Asta-ID eller Asta-streng, kan man hente inn alle relevante data om arkivenheten (arkivet, serien, stykket, mappa eller dokumentet) fra Arkivportalen til Digitalarkivets database. Databasen kan synkroniseres mot Arkivportalen. Dette skjer p.t. ved manuell igangsetting og ikke automatisk/jevnlig (men det vil være mulig).	Vi har ønsker om å gå mer i denne retningen og være tettere mot arkivkatalogen uavhengig av om Digitalarkivet blir løsningen. De nødvendige administrasjonsverktøyene er under planlegging, og det vil være naturlig å ta høyde for disse tingene under planlegging og utvikling.	+
	Det skal være mulig å masseregistrere flere arkivenheter, f.eks. alle stykker i en serie, i én operasjon.	-	P.t. ikke mulig i Digitalarkivets administrasjonsgrensesnitt.	Dette er et behov og savn allerede. De nødvendige administrasjonsverktøyene er under planlegging, og det vil være naturlig å ta høyde for disse tingene under planlegging og utvikling.	+
	Dersom en arkivenhet ikke er registrert i Arkivportalen, skal det være mulig å registrere relevante data manuelt, eller man kan registrere den overordnede arkivenheten med merknad om at bare en del av denne er digitalisert.	+	Dette er mulig i Digitalarkivets administrasjonsgrensesnitt. Arkivenheten tildeles da et løpenummer som erstatning for den manglende Asta-IDen.		+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 16 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	Det skal være mulig å registrere ytterligere informasjon som ikke finnes i Arkivportalen, men som kan forbedre brukernes søkemuligheter, for eksempel emneord og geografiske tilknytninger.	+	Digitalarkivets database har felt for bl.a. emneknagger, 1947-kommuner og dagens kommuner, som alle kan fylles ut via administrasjonsgrensesnittet. Det er mulig å utvide databasemodellen med flere felt.		+
	Det skal være mulig å definere én eller flere "profiler" som legger den samme tilleggsinformasjonen på en serie av arkivenheter.	-	P.t. ikke mulig i Digitalarkivets administrasjonsgrensesnitt. Arkivenhetene må oppdateres én for én manuelt, eller med SQL-kommandoer direkte mot databasen.	De nødvendige administrasjonsverktøyene er under planlegging, og det vil være naturlig å ta høyde for disse tingene under planlegging og utvikling.	+
	Det skal være mulig å registrere at en arkivenhet skal være omfattet av tilgangsstyring og definere tilgangsgrupper som skal ha tilgang.	+	Dette er mulig i Digitalarkivets administrasjonsgrensesnitt.		+
Det skal være mulig å laste opp et filsett til fillageret via Internett.	Følgende filformater skal støttes: TIFF, JPEG2000, JPEG, PNG og PDF (for uthenting av bildefiler i PDF-filene).	-	Det er p.t. kun mulig å laste opp filsett med en intern programvare i Arkivverket, og pga. begrenset linjekapasitet sender statsarkivene sine skanna bilder på disk til Riksarkivet, hvor opplastinga foregår. Det er kun TIFF og JPEG som støttes.	De nødvendige verktøyene og løsningene er under planlegging, og det vil være naturlig å ta høyde for disse tingene under planlegging og utvikling. Arkivverket planlegger en funksjon for opplasting via Internett, da ordningen med å sende disk er ressurskrevende	+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 17 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
				og forsinkende.	
	Opplasting skal skje ved å velge en forhåndsregistrert arkivenhet fra databasen som filsettet skal kobles til.	-	P.t. ikke mulig, men inngår i Arkivverkets planer for ny opplastingsfunksjon.	De nødvendige administrasjonsverktøyene er under planlegging, og det vil være naturlig å ta høyde for disse tingene under planlegging og utvikling.	+
	<p>Ved opplasting skal filsettet valideres (virussekk etc.). Originalfilene konverteres til visningsfiler i JPG-format med tilhørende UUIDer, men de originale filnavnene lagres også.</p> <p>Av hensyn til hurtig nedlasting for brukerne, bør ikke visningsfiler være særlig større enn 1 MB. Der det er behov for større visningsfiler (detaljerte tegninger og kart etc.) må en tiles-funksjon være tilgjengelig.</p> <p>De originale filene slettes, slik at NAPSA kun lagrer visningsfilene.</p> <p>Innholdsleverandørene må selv ta ansvaret for å lagre de originale filene i de originale formatene, dersom de ønsker å</p>	-	P.t. foregår det ingen validering. Opplastede TIFF-filer konverteres til JPEG2000-filer (for lagring) og komprimerte JPEG-filer (for visning), mens TIFF-filene slettes. Opplastede JPEG-filer importeres direkte uten komprimering eller annen modifisering,	En ny konverteringsrigg for opplastede filer er under planlegging. Dette er i henhold til planene for hvilke muligheter som skal være tilgjengelig også for internt bruk i Arkivverket.	+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 18 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	<p>lagre dem for ettertida (anbefalt).</p> <p>Opsjon om at originalfilene kan lagres i en NOARK5-kjerne (vil medføre kostnader for innholdsleverandører som benytter opsjonen).</p> <p>Innholdsleverandørene skal kunne laste ned en "mapping" mellom originale filnavn og permanente UUIDer i NAPSA.</p>				
	En "default" visningsrekkefølge for filene etableres ved at filene sorteres alfabetisk-numerisk etter originale filnavn.	-	P.t. ikke mulig.	Inngår i Arkivverkets planer for ny opplastingsfunksjon.	+
Det skal være mulig (men ikke nødvendig) å registrere ytterligere metadata på filnivå (for enkeltfiler eller grupper av filer). Dette inkluderer:	Beskrivelse av hva hvert bilde (hver fil) viser, så som hva (f.eks. "Formannskapsmøte"), når (f.eks. "25.01.1933"), sidetall i kilden osv.	-	<p>Det er p.t. nødvendig å registrere ytterligere metadata på filnivå, men det er ikke nødvendig med omfattende metadata på filnivå, slik at det kan utføres forholdsvis raskt, avhengig av ambisjonsnivået.</p> <p>Metadataene på filnivå registreres i en hierarkisk struktur med inntil fire felt, samt felt for sidetall og kategorisering av filene.</p>	I forbindelse med ny løsning for indeksering, som skal være en del av media.digitalarkivet.no , ser vi på muligheten for å kunne tagge bilder med emenknagger.	+
	Endring av visningsrekkefølgen, herunder ekskludering av bilder fra visningsrekkefølgen.	+	I forbindelse med registrering av metadata på filnivå er det mulig å ekskludere bilder fra visningsrekkefølgen, men dersom man senere ønsker å ekskludere ytterligere bilder, eller dersom man likevel ønsker	I forbindelse med ny løsning for indeksering, som skal være en del av media.digitalarkivet.no , vil disse	+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 19 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
			<p>å inkludere et bilde som i første omgang er ekskludert, må dette gjøres med SQL-kommandoer direkte mot databasen.</p> <p>Mindre endringer av visningsrekkefølgen, sammenliknet med skanningsrekkefølgen, er også mulig i forbindelse med registrering av metadata på filnivå, men større endringer må i praksis gjøres med SQL-kommandoer direkte mot databasen.</p>	mulighetene være tilgjengelig.	
	Klausulering / tilgangsstyring av enkeltfiler.	+	Klausulering / tilgangsstyring av enkeltfiler er kun mulig med SQL-kommandoer direkte mot databasen.	I forbindelse med ny løsning for indeksering, som skal være en del av media.digitalarkivet.no , vil disse mulighetene være tilgjengelig.	+
	Automatisk OCR-gjenkjenning av maskinskrevet tekst skal kunne utnyttes til registrering av metadata på filnivå.	-	Digitalarkivet har ikke automatisk OCR-gjenkjenning.	En ny konverteringsrigg for opplastede filer er under planlegging. OCR vil inngå som funksjonalitet i denne riggen, og det kan lagres OCR-tolket tekstfil i ønsket format sammen med fila som ble tolket. Videre bruk av det tolkede data for metadata og søk via indeks må kravsettes.	+
	Registrering av ytterligere metadata skal tilrettelegges for crowdsourcing, slik at innholdsleverandørene kan velge å åpne	+	Crowdsourcing av metadata på filnivå er p.t. ikke mulig i Digitalarkivet, men en løsning for dette er på tegnebrettet og vil muligens kunne lanseres innen utgangen av 2016. Det vil da bli mulig å regulere hvilke	Spesifisering av en løsning for indeksering som innebærer å åpne opp	+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 20 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	for crowdsourcing på hele eller deler av sitt materiale (taushetsbelagt materiale kan f.eks. ikke crowdsources).		arkivenheter som skal være tilgjengelige for crowdsourcing.	for crowdsourcing pågår i disse dager.	
Det skal være mulig å laste opp nye filsett til en arkivenhet som allerede har filsett knyttet til seg, erstatte filer etc.		+	Dette er allerede mulig i Digitalarkivet, men det innebærer mye manuelt arbeid, inkludert SQL-kommandoer direkte mot databasen.	De nødvendige administrasjonsverktøyene er under planlegging, og det vil være naturlig å ta høyde for disse tingene under planlegging og utvikling. Ny indekseringsløsning vil tillate slik erstatning for enkeltfiler.	+
Publisering / avpublisering skal skje med et tastetrykk, enten for en enkelt arkivenhet eller for ei gruppe arkivenheter.		+	Én enkelt arkivenhet kan avpubliseres med et tastetrykk i administrasjonsgrensesnittet. Grupper av arkivenheter kan avpubliseres med en SQL-kommando direkte mot databasen.	De nødvendige administrasjonsverktøyene er under planlegging, og det vil være naturlig å ta høyde for disse tingene under planlegging og utvikling.	+
Statistikkverktøy skal gjøre det mulig å hente ut:	Antall besøk og sidevisninger totalt, for den enkelte innholdsleverandør, for enkeltkilder og for grupper av kilder (også på tvers av innholdsleverandører, for eksempel likeartede kilder). Antall publiserte bilder og kilder, fordelt	-	Digitalarkivet har foreløpig begrensede statistikkverktøy. Mye statistikk utarbeides på grunnlag av manuelle SQL-spøringer og excel-ark som vedlikeholdes med relativt stor arbeidsinnsats. Fra Google Analytics er det mulig å hente ut antall besøk og sidevisninger totalt, og med varierende behov for innsats kan tallene splittes opp på enkeltkilder og (makro)grupper av kilder, men i praksis	Vi høster allerede en del data. Vi har på tegnebrettet en løsning som i all hovedsak vil kunne dekke disse kravene med noen unntak.	+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 21 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	<p>på innholdsleverandør og/eller fordelt kronologisk (minimum måned).</p> <p>Antall opplastede bilder og kilder, fordelt på innholdsleverandør og/eller fordelt kronologisk (minimum måned). Dette inkluderer bilder og kilder som er lastet opp, men som ikke er publiserte.</p> <p>Antall registrerte kilder, fordelt på innholdsleverandør og/eller fordelt kronologisk (minimum måned). Dette inkluderer kilder som er registrert, men som ikke har fått bilder koblet til seg.</p>		ikke på innholdsleverandør.		
NAPSA og alt innhold skal være gratis å bruke for brukere (publikum).		N/A	Et strategisk valg som ikke angår Digitalarkivet <i>som løsning</i> .		N/A
Det skal være færrest mulig begrensninger på gjenbruk av innhold, men f.eks. åndsverksrettigheter må respekteres.	<p>Den enkelte innholdsleverandør er ansvarlig for å sette lisens på sitt materiale; fri lisens er standard og idealet.</p> <p>Den enkelte innholdsleverandør må ha oppmerksomhet mot bestemmelser om nettpublisering i forbindelse med avtaler</p>	N/A	Dette må innholdsleverandørene selv følge opp, uavhengig av løsning.		N/A



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 22 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	med arkivskapere.				
NAPSAs brukergrensesnitt skal være generøst, enkelt og moderne og tilgjengelig på alle digitale plattformer (PC, Mac, tablet, mobiltelefon).	Brukergrensesnittet skal være generøst, enkelt og moderne og tilgjengelig på alle digitale plattformer (PC, Mac, tablet, mobiltelefon).	+	I utviklinga av Digitalarkivets nye visningsprogrammer for skanna arkivmateriale (https://media.digitalarkivet.no/) er det lagt vekt på tilgjengelighet på alle plattformer. Visningsprogrammene vil i den forbindelse bli "pussa opp", trolig i løpet av høsten 2016.		+
	Søk etter publiserte arkivenheter skal være raske og effektive. Det skal være mulig å søke etter: <ul style="list-style-type: none"> • oppbevaringssted • arkiv • administrativ tilknytning / arkivskaper (ikke alltid identisk med arkiv) • emneord • geografisk tilknytning • tidsintervall • fritekst • OCR-lest tekst (innholdstekst) 	+	Jf. https://media.digitalarkivet.no/db/browse er det mulig å søke etter oppbevaringssted, arkiv, geografisk tilknytning, tidsintervall, emneord/emneknagger og fritekst. Søk etter administrativ tilknytning/arkivskaper og søk i OCR-lest tekst mangler.	Se punkter over	+
	Blading og fremvisning skal være raskt og effektivt, og det skal være god lesefunksjonalitet: <ul style="list-style-type: none"> • bilder skal kunne skaleres i 	+	Vi mener at Digitalarkivets nye visningsprogrammer oppfyller disse kravene. Bildene kan skaleres i størrelse (man kan se for seg videreutvikling med tiles), og bildene kan roteres.		+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 23 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	<p>størrelse</p> <ul style="list-style-type: none"> bilder skal kunne roteres 				
	Det skal være mulig å dele innhold på e-post og i sosiale medier.	+	Deling på e-post, Facebook og Twitter er mulig.		+
	<p>Det skal være mulig å kommunisere med NAPSA og/eller innholdsleverandørene via tilbakemeldingsfunksjoner. Funksjonene skal legge til rette for toveis-kommunikasjon:</p> <p>Tilbakemeldingsfunksjon for tekniske spørsmål og meldinger om driftsfeil betjenes av et sentralt ledd på vegne av NAPSA.</p> <p>Tilbakemeldingsfunksjon for spørsmål om arkivinnhold rutes til den enkelte innholdsleverandøren, basert på metadata om oppbevaringssted, for eksempel med e-postskjema eller ved at e-postadresse oppgis. Den enkelte innholdsleverandør er ansvarlig for å svare, men det er viktig at brukerne blir oppmerksomme på at ikke alle innholdsleverandører vil ha like stor</p>	+	Digitalarkivet har en tilbakemeldingsfunksjon for tekniske spørsmål og driftsfeil som brukerne i noen grad har tatt i bruk. Spørsmål om arkivinnhold kommer til Arkivverket i Digitalarkivets forum og på e-post. Mange brukere benytter forum og/eller e-post også for tekniske spørsmål og meldinger om driftsfeil.	I forbindelse med ny løsning for indeksering kommer vi til å innarbeide en mulighet for rapportering. Vi opplever dagens tilbakemeldingsfunksjon som leveres av en annen leverandør som noe ustabil. I den forbindelse vil vi kunne få muligheten til å rute tilbakemeldinger til den enkelte innholdsleverandør.	+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 24 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
	kapasitet til å svare, og at det kan ta tid før de får svar.				
En bruker skal ikke være nødt til å logge inn med brukerprofil for å bruke NAPSA, men det skal være mulig å logge inn for å få tilgang til særskilte brukervalg og arkivenheter med tilgangsstyring, samt personlig toveis-kommunikasjon med innholdsleverandørene.		+	Med unntak av leddet om personlig toveis-kommunikasjon med innholdsleverandørene, oppfyller Digitalarkivet dette kravet.		+
Publisert materiale skal vises med korrekt og detaljert kildereferanse, dvs. oppbevaringssted, arkivnavn og eventuelt ytterligere arkivbeskrivelse fra arkivkatalog.		+	Digitalarkivet oppfyller dette kravet.		+
Dersom det publiserte materialet er registrert i Arkivportalen, skal det etableres toveis lenke mellom NAPSA og Arkivportalen.		+	Digitalarkivet lenker automatisk til Arkivportalen. Det er lagt til rette for at Stiftelsen Asta kan innarbeide lenke fra Arkivportalen til Digitalarkivet, men Stiftelsen Asta har så langt ikke prioritert dette.		+
Innholdet i NAPSA skal kunne indekseres av markedsledende søketjenester som		+	Google indekserer skanna innhold i Digitalarkivet, men det er åpenbart mer å hente mht til å komme høyere på Googles trefflistene.		+



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 25 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Krav	Delkrav	*/-	Vurdering av Digitalarkivet pr. 8.7.2016	Planlagte utvidelser	+/- etter utvidelse
Google.					
Det skal være god integrasjon mot andre relevante innholdstjenester som:	<ul style="list-style-type: none">• Arkivportalen• Norvegiana• Digitalt Museum• bokhylla.no (nb.no)	-	Digitalarkivet har p.t. kun integrasjon mot Arkivportalen, jf. over.		-
Det skal være mulig å lenke mellom arkivenheter i NAPSA og relevant innhold i andre innholdstjenester.		+	Lenkene til de skanna bildene (brukslenker for sidevisning) i Digitalarkivet er permanente og kan benyttes i andre innholdstjenester. Lenker på arkivenhetsnivå er ikke garantert permanente.		+
Intelligente søk skal gi relevante treff i NAPSA og andre innholdstjenester samlet.		-	Krever formodentlig tilrettelegging.		-



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 26 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Vedlegg 3

Minimumskrav til kvalitet i digitalisering av papirbasert materiale

Filformater og oppløsning

Viser til delrapport 1a: Funksjons- og tjenestebeskrivelse, kapittel Innholdsleverandørsiden («back end») som beskriver filformater som skal støttes, opsjon om lagring i NOARK-5 kjerne og delrapport 1b: Teknisk beskrivelse og løsningsforslag, kapittel Lagring og bildebehandling beskrivelse av langtidslagring og brukerversjon til Digitalarkivet.

Filformat til langtidslagring bør være i PDF/A og filformat til visning bør være i format som støttes av tiles-funksjon, eksempel JPEG. Det er viktig at visningsfiler på detaljerte tegninger, kart etc. har god nok oppløsning slik at bruker kan gjennomføre dybde-zoom.

Alle originaler over A3 i fysisk størrelse bør bli digitalisert i format 1:1 i 300 dpi. Eksempel som kart i fysisk størrelse HxB 1000 x 1500 mm gir digitalisert versjon i samme HxB format i 300 dpi.

Originaler i A4, folio og A3 bør digitaliseres i format 1:1 i 300 dpi.

Originaler mindre enn A4 bør digitaliseres større enn 1:1 i 300 dpi for å gi god nok mulighet for dybde-zoom.

Metode og valg av utstyr i «volum-digitalisering»

Originaler som er i god stand kan digitaliseres ved hjelp av «bunskannere» som fører dokumentene over skannerfeltet for digitalisering. Dette medfører ei fysisk påkjenning som materialet må tåle uten at skade blir påført. Skanneropertøren må ha derfor ha god nok kunnskap til originalmaterialets fysiske stand, samt kjenne godt nok til utstyrets begrensninger og muligheter før ei digitalisering blir gjennomført.



Eksempel på digitalisering av flere originaler ved hjelp av «bunskanner».



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 27 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Originaler som ikke vil tåle den fysiske påkjenninga med å bli ført over skannerfeltet må digitaliseres ved hjelp av «flatbed» skanner – der den fysiske originalen ligger stille under digitaliseringa.



Eksempel på «flatbed» skanner der original ligger stille under digitaliseringa.



Det er viktig at binders, stifter og evt. andre gjenstander blir forsiktig fjernet før digitalisering. Dette for å unngå å skade digitaliseringsutstyret.

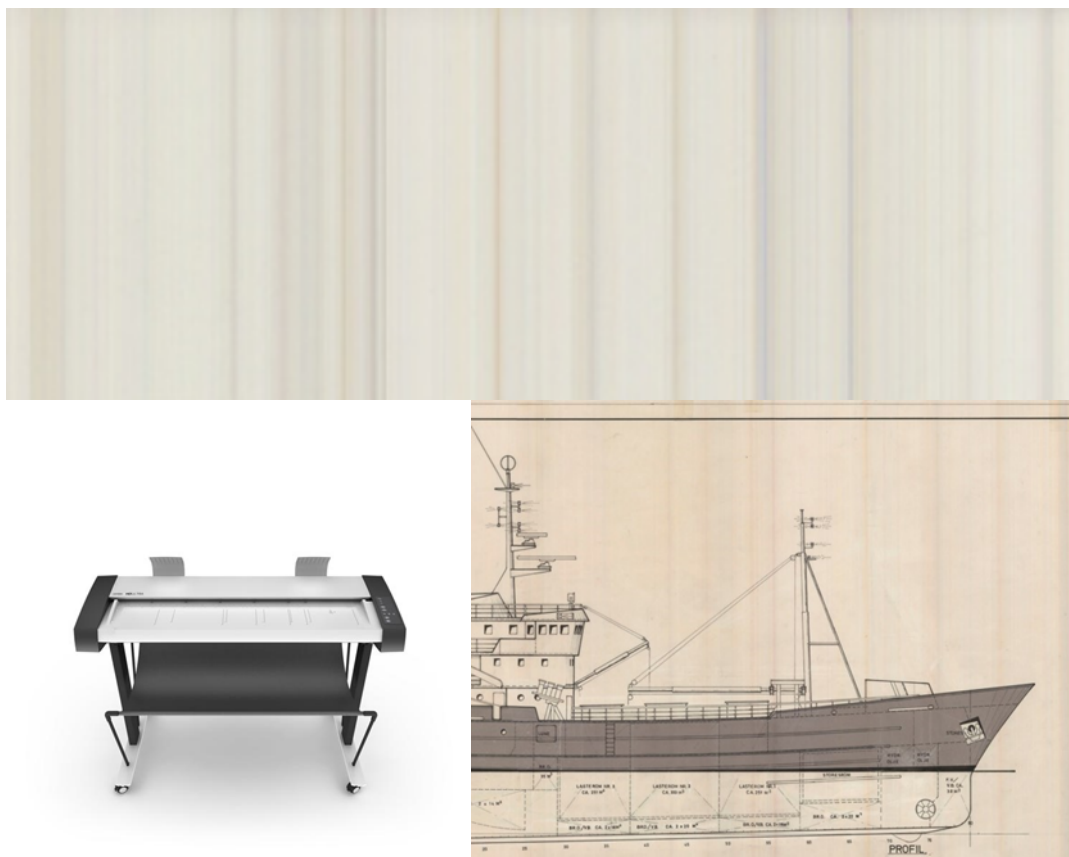


SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 28 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Metode og valg av utstyr i digitalisering av store format

En metode som har vært mye i bruk ved store format er bruk av «rulleskanner». Her blir originalene ført over skannerfeltet slik som ved en «bunskanner». Det kommer veldig fortløpende striper på de digitale filene grunna slitasje på glassplaten over skannerfeltet. Det blir striper i samme retning som originalmaterialet har blitt ført gjennom skanneren.

Det er å anbefale at store format blir digitalisert ved bruk av digitalt mellomformatkamera. Dette for å unngå stripene på de digitale filene, men også for å unngå slitasje på originalmaterialet. Et digitalt mellomformatkamera har større bildesensor enn et standard digitalt kamera og kan derfor registrere alle detaljer i ei stor original. Originalmaterialet må belyses jevnt fra begge sider ved hjelp av fastlys eller blitzlys.



Eksempel på striper på digital fil etter digitalisering ved bruk av «rulleskanner» for store format.



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 29 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			



Digitalisering ved bruk av digitalt mellomformatkamera for store format. Originalene festes ved hjelp av magneter på magnetisk tavle.



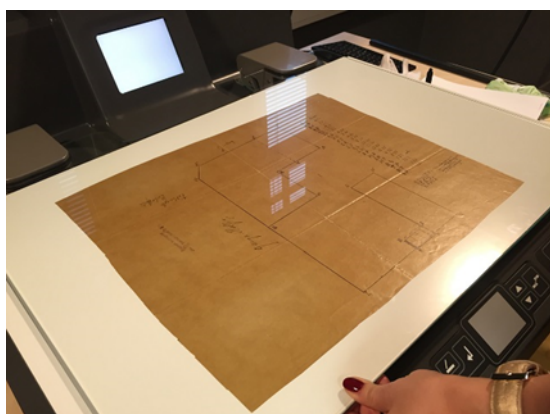


SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 30 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Digitalisering kan også gjennomføres ved at materialet ligg flatt på plate.

Metode og valg av utstyr for digitalisering av skjøre originaler.

Ved digitalisering av skjøre originaler er det å anbefale bruk av digitalt kamera. Det er en velprøvd metode som gir operatøren veldig god kontroll over resultatet uten at materialet blir utsatt for unødvendig fysisk påkjenning.



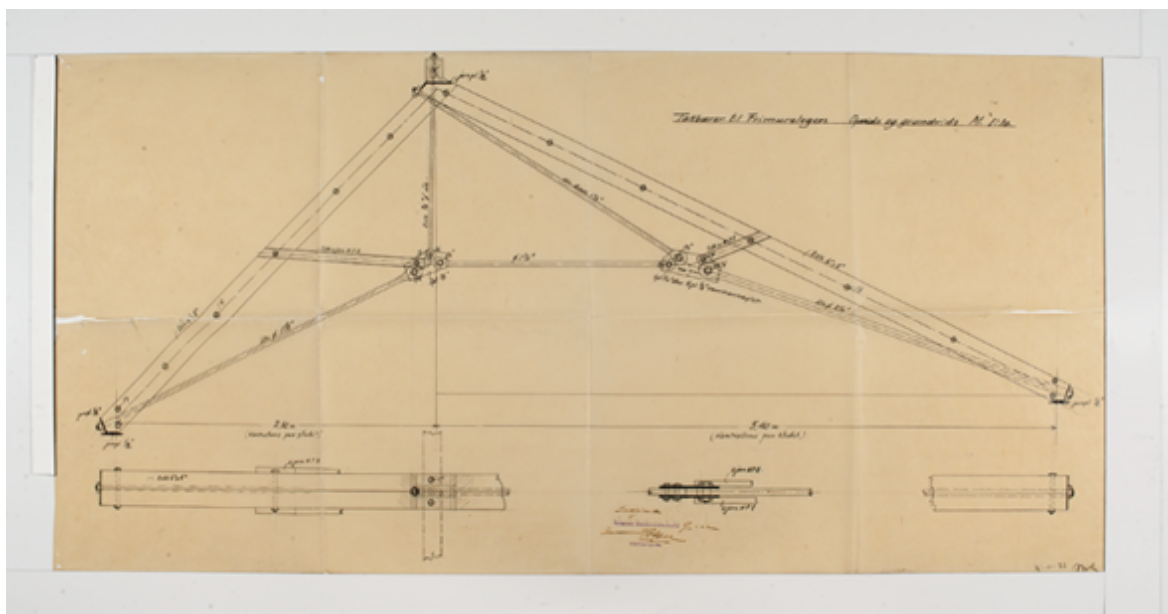
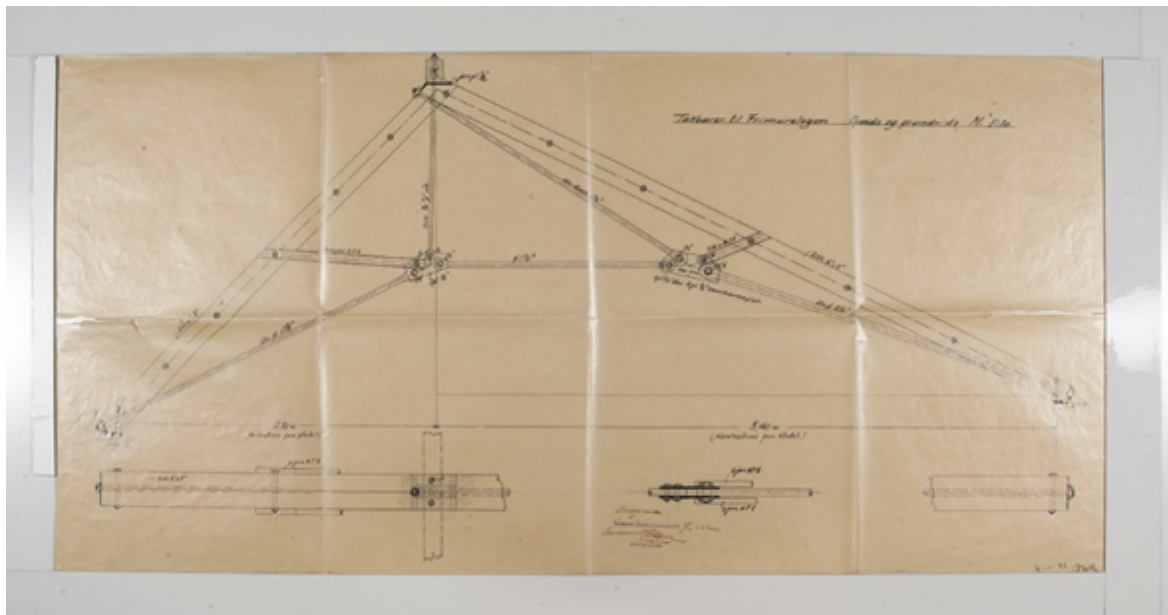
Eksempel på original med og uten press fra glassplate under digitalisering ved bruk av kamera.



Oppsett fra Riksarkivets fotoavdeling, digitalisering ved bruk av kamera og glassplate for press over original. Belysning ved bruk av blitzlys.



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 31 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			



Digitalisering ved bruk av kamera med «krysspolarisert» blitzlys for å jevne ut fysiske bretter på originalmaterialet. Første eksempel viser resultat uten «krysspolarisert» lys.

Operatør må ha god kjennskap til fototeknikk for å meste denne teknikken godt. Her er det å anbefale bruk av fagutdanna fotograf.



SAMDOK delprosjekt: Kommunale arkiv	Prioritert oppgave: 7 Forprosjekt: Nasjonal publiseringsplattform for skanna arkivdokumenter	SAK (ePhorte):	Dato: 06.01.2017	Side: 32 av 33
Forfatter: Anette Skogseth Clausen m.fl.	Tittel: Rapport 2016 – sluttrapport med anbefalinger			

Kvalitetskontroll

For at resultatene skal ha en jevn god kvalitet uavhengig av operatør er det helt nødvendig med faste rutiner på kvalitetskontroll. Det bør ved enhver digitalisering brukes farge- og kontrastreferanser under digitaliseringa for å ha mulighet til sammenligning av resultatet opp mot referansen.

Her finnes ulike typer, både enkle og mer avanserte referansekort/striper/paletter. En enkel linjal gir mulighet for å sjekke fysisk størrelsesforhold til original. Gode skjermer som er kalibrerte vil være en forutsetning for å kunne gjennomføre en slik kvalitetskontroll. Dagens teknologi gir skjermer som kalibrerer seg selv etter faste tidspunkt uten at operatør trenger å bruke tid og videre resurser på dette.



Eksempel på fargepalett, linjal og kontraststripe for kvalitetskontroll.

Vurdering av behovet for utarbeidelse av veiledning for «best practice» for digitalisering av papirbasert materiale

Det er arbeidsgruppens anbefaling at hovedprosjektet jobber med utarbeidelse av en slik veileder. Erfaringer medlemmer i arbeidsgruppen har gjort, gjennom flere besøk og samtaler rundt om kring både på regionale, men også nasjonale institusjoner tydeliggjør et behov. Siden ny teknologi har kommet til de siste par årene så har flere «gamle» praksiser og metoder blitt utrangert og bør byttes ut med ny praksis og metode.

Alle bilder er fra SEDAK og er fotografert av Ottar André Brevik Anderson uten at annet er oppgitt.