



ARKIVVERKET
RIKSARKIVET

Samdok

samla samfunnsdokumentasjon

RAPPORT 2014



PRIORITERT OPPGAVE

Arkiv i e-forvaltning

(3b) "Synkron" avlevering (STAT)

/Statens kartverk

Utarbeidet av
Tor Anton Gaarder og Erik Aaberg

Rapportdato
22.12.2014



SAMDOK delprosjekt: eForvaltning	Prioritert oppgave: Samdok 3 – Synkron overføring	SAK (ePhorte): 2014/15569	Dato: 22.12.2014	Side: 1 av 6
Forfattere: Tor Anton Gaarder, Erik Aaberg	Tittel: Rapport 2014 – Synkron overføring: Statens kartverk			

OPPGAVE	(3) "Synkron" avlevering (STAT) /Statens kartverk
Ansvarlig strategigruppe	<i>Arkiv i e-forvaltning</i>
Arbeidsgruppens leder	Tor Anton Gaarder
Gruppens medlemmer	Erik Aaberg, Arne-Kristian Groven
Målformulering	Ha på plass metode og infrastruktur for stegvis overføring av digitalisert pantebok fra Statens kartverk til Riksarkivet
Sammendrag	<p>Prosjektet har en lang forhistorie der det har vært mange tekniske skjær i sjøen. På begynnelsen av år 2014 så man seg om etter alternativ metodikk for å hente ut data fra panteboken, også med tanke på det etterslep som hadde bygd seg opp.</p> <p>I løpet av høsten er det etablert rutiner for å hente data over linje fra Kartverket til Riksarkivet. Data kan hentes i større eller mindre kvanta. De to første uttrekkene er hentet, og data er kvalitetssikret.</p> <p>Det gjenstår imidlertid fortsatt en del arbeid, spesielt for å automatisere prosesser som nå styres manuelt. Rutiner for tilbakemeldinger ved feilsituasjoner er heller ikke utviklet.</p>
Vedlegg	Grafisk skisse - Inkrementell deponering/avlevering av pantebok fra Kartverket Teknisk skisse - Inkrementell deponering/avlevering av pantebok fra Kartverket
Oppfølging 2015	De viktigste punkter: <ul style="list-style-type: none">• Etablere metodikk for tilbakemeldinger til Kartverket etter overføring• Automatisere flest mulig av de steg som nå er styrt manuelt.• Utvikle rutiner for ombrekking fra overføringsformat til endelige arkivformat• Hente inn etterslepet. Til sammen skal ca. 14.5 Tb overføres.
Status	Ferdig



SAMDOK delprosjekt: eForvaltning	Prioritert oppgave: Samdok 3 – Synkron overføring	SAK (ePhorte): 2014/15569	Dato: 22.12.2014	Side: 2 av 6
Forfattere: Tor Anton Gaarder, Erik Aaberg	Tittel: Rapport 2014 – Synkron overføring: Statens kartverk			

Innledning

Riksarkivet og Statens kartverk har over en periode på flere år jobbet med å få på plass en avtale om deponering og avlevering av digitalisert pantebok. Den seneste versjonen av denne avtalen er under utarbeidelse i disse dager og kan finnes i sak 2014/20939 i ePhorte.

En grunnleggende utfordring i de senere årene har vært å produsere et arkivuttrekk fra den digitale panteboken som tilfredsstillende Riksarkivets krav til slike uttrekk. En del av utfordringen er knyttet til det faktum at Kartverket ønsker å gå over til fullelektronisk pantebok. For å kunne gjøre trenger de godkjenning fra Riksarkivaren. Grunnlaget for en slik godkjenning må bygges på et godkjent arkivuttrekk fra Regina, systemet som brukes i forvaltningen av panteboken. Overføringsformatet er utviklet i samarbeid mellom Statens kartverk og Riksarkivet. Utfordringen kan deles opp i tre deler:

1. Metode for produksjon av uttrekk i henhold til spesifikasjonen
2. Testverktøy i Riksarkivet
3. Volumet på dataene som utgjør det uttrekket som skal deponeres i Riksarkivet.

Statens kartverk har hatt innvendinger mot det arbeidet som er blitt gjort med produksjon av arkivuttrekk fra panteboksystemet og med testingen i Riksarkivet. Produksjon av uttrekket er blitt utført av et konsulentselskap. Tjenesten har vært kostbar og ikke av tilfredsstillende kvalitet. Testingen av uttrekket i Riksarkivet har tatt svært lang tid. Årsaker har vært varierende tilgang til kompetanse, ikke tilgang til tilfredsstillende testverktøy og materialets volum, (ca. 5 terrabyte i første uttrekk). Volumet har gitt betydelige utfordringer både for eksisterende testverktøy og for maskinvare.

I et forsøk på å rette opp alle disse forholdene, er det nedsatt en arbeidsgruppe som skal utvikle en ny metode og ny infrastruktur for overføring av pantebokdata fra Kartverket til Arkivverket. Denne innsatsen ble innlemmet i Samdok-programmets oppgave 3: Synkron overføring.

Sammensetning av arbeidsgruppen

Arbeidsgruppen er sammensatt av representanter fra Riksarkivet og Statens kartverk. Fra Riksarkivet er følgende mest involvert:

- Arne-Kristian Groven
- Erik Aaberg
 - o Ole Myhre-Hansen
 - o Jon Atle Haugen
 - o Tor Anton Gaarder
 - o Torunn Gaasemyr har også vært noe involvert



SAMDOK delprosjekt: eForvaltning	Prioritert oppgave: Samdok 3 – Synkron overføring	SAK (ePhorte): 2014/15569	Dato: 22.12.2014	Side: 3 av 6
Forfattere: Tor Anton Gaarder, Erik Aaberg	Tittel: Rapport 2014 – Synkron overføring: Statens kartverk			

Hos Statens kartverk har følgende personer vært mest involvert:

- Robin Paulsen
- Stig Asle Rogne
- Einar Gautun

- Laila Vågen
- Geir Vidar Mørner

Beskrivelse av det konkrete arbeidet

Data og bevaring

Det er oppnådd enighet om hvilke data som skal overføres i henhold til hva som er bevaringsverdig informasjon.

Det arbeides foreløpig ut fra den antagelse at skjermede data også skal overføres. Data som blir overført til Riksarkivet vil **ikke** være synlige for andre enn de som arbeider i prosjektet. Unntak kan kun gjøres ved enighet mellom partene. Ref. også den kontrakt som utarbeides.

LINJE

Man ønsket seg en sikker fast linje for overføring av data fra Kartverket til Riksarkivet. Denne linjen ble ca. 2 måneder forsinket, mye skyldtes leverandør av linje. Prosjektet led noe under det, men ble ikke like mye forsinket som linja, da mye kunne testes før linje var på plass.

Databaser og tabeller hos Kartverket

Kartverket har etablert rutiner og databaser/tabeller for å legge opp data for overføring, og for å ta imot kvitteringer. De har også etablert bruker og rettigheter (kun for lesing) slik at de relevante tabeller kan nås fra Riksarkivet via riktig konfigurering av nettverk og brannvegger.

Server hos Riksarkivet

Riksarkivet startet tester fra en vanlig ansatt-PC. I oktober kom imidlertid en egen server på plass, kun tilegnet prosjektet. Etter noe diskusjon ble Windows valgt som plattform, da dette stiller oss noe friere i forhold til permanent fremtidig plassering av serveren. Denne server har en fast IP adresse, og av sikkerhetsgrunner kan ikke Kartverkets database nås fra noe annet sted i Riksarkivet.

Programvare hos Riksarkivet

Riksarkivet bruker p.t. **SIARD** (http://coptr.digipres.org/SIARD_Suite) for å hente data. I prinsippet kan hvilken som helst databaseklient brukes.

Riksarkivet har utviklet egen programvare for å lese SIARD-filer, og så teste data både i forhold til komplettitet og kvalitet. Første versjon av denne programvare er ferdig utviklet, og det første uttrekket kan dermed godkjennes.



SAMDOK delprosjekt: eForvaltning	Prioritert oppgave: Samdok 3 – Synkron overføring	SAK (ePhorte): 2014/15569	Dato: 22.12.2014	Side: 4 av 6
Forfattere: Tor Anton Gaarder, Erik Aberg	Tittel: Rapport 2014 – Synkron overføring: Statens kartverk			

Status i arbeidet ved årets slutt

- Linje er testet og virker
- Ett uttrekk er overført (2000 poster) - og ingen feil er funnet. Formell godkjenning er foreløpig ikke sendt. Ytterligere et uttrekk på 10 000 poster er overført, men ikke testet.
- Programvare for kvalitetstest av data er utviklet
- Samarbeidet mellom etatene fungerer godt

Det videre arbeidet i 2015

Data og bevaring

Det må tas en endelig beslutning angående skjermede data. Det vil ha vidtrekkende tekniske konsekvenser dersom skjermede data skal unntas, primært fordi det medfører vedlikehold av allerede overførte data (ny skjerming/avskjerming).

Linje

Linje var opprinnelig på 10 Mbit/s. Selv med full belastning av linja ville det tatt for lang tid å hente alle data. Linja ble så oppgradert til 100 Mbit/s, men foreløpig har vi ikke klart å "fylle" mer enn 20-30Mbit/s. Dette må det finnes ut av. Med 25 Mbit/s tok 10 000 dokumenter ca. 1 time. Det vil medføre overføringstid på minst 1500 timer for det datavolum vi snakker om. I tillegg kommer all overhead for administrasjon.

Databaser og tabeller hos Kartverket

Kartverket må ferdigstille rutiner for mottak av kvitteringer fra Riksarkivet. Dette innebærer blant annet å markere det overførte som "ferdig", for deretter å etablere en ny datamengde for uttrekk.

Programvare hos Riksarkivet

Riksarkivet må videreutvikle program for testing av data. Det kan optimaliseres en del, og rutiner for kvitteringsfunksjon tilbake til Kartverket er ikke implementert ennå. Det er foreløpig ikke bestemt om dette skal skje via filoverføringer eller "fjernoppdatering" av databaser hos Kartverket.

Riksarkivet må også utvikle rutiner for å brette om data til arkivformat i forhold til ADDML standarden (<http://arkivverket.no/arkivverket/Arkivbevaring/Elektronisk-arkivmateriale/Standarder/ADDML>).

Til slutt må det etableres rutiner for permanent lagring i henhold til de prinsipper som gjelder (skjerming, lagring i 3 kopier, lagring på 2 forskjellige media på 2 forskjellige steder).

Automatisering



SAMDOK delprosjekt: eForvaltning	Prioritert oppgave: Samdok 3 – Synkron overføring	SAK (ePhorte): 2014/15569	Dato: 22.12.2014	Side: 5 av 6
Forfattere: Tor Anton Gaarder, Erik Aaberg	Tittel: Rapport 2014 – Synkron overføring: Statens kartverk			

Foreløpig må alle steg i prosessen startes manuelt, og resultater sjekkes manuelt (etter lesing av logger fra de maskinelle testrutiner). Det største arbeidet i 2015 vil antagelig bli å automatisere dette i størst mulig grad, og underveis ta hensyn til **alle** de feilsituasjoner som kan (og vil - -) oppstå.



SAMDOK delprosjekt: eForvaltning	Prioritert oppgave: Samdok 3 – Synkron overføring	SAK (ePhorte): 2014/15569	Dato: 22.12.2014	Side: 6 av 6
Forfattere: Tor Anton Gaarder, Erik Aaberg	Tittel: Rapport 2014 – Synkron overføring: Statens kartverk			

Innhenting av etterslep

I løpet av 2015 er ambisjonen å hente inn etterslepet på mellom 14 og 15 Tb. En fast tid for slutføring av dette kan ikke gis på det nå værende tidspunkt, men det antas å vare til over sommeren 2015.

Prosjektets konsekvenser utover sine egne oppgaver

Dette prosjektet peker på en helt ny metodikk for overføringer og uttrekk fra arkivskapere til Arkivverket.

Det vil forenkle prosesser ved at de automatiseres, og dermed medføre mindre arbeid og kostnader både for arkivskapere og Arkivverket. I tillegg vil det sikre kvalitet på data, og i den grad feil oppstår vil det oppdages mye raskere. Man vil da forhåpentligvis slippe problemet med feil som kan akkumuleres over år før det oppdages.

Ved å bruke automatiserte, inkrementelle overføringer vil man også løse problemet med til tider lang responstid fra Arkivverket når det gjelder test av avleveringer.

Detaljer i **koden og programvare** som utvikles vil være forskjellig fra arkivskaper til arkivskaper, men mange av **prinsippene** kan gjenbrukes. På sikt må det være et mål å generalisere koden slik at man kun må oppdatere en konfigurasjonsfil hver gang en ny type uttrekk skal hentes.