

DIAS-prosjektet

Status mht. gjennomføring, resultater og produkter

LLP-konferanse, Trondheim 24.04.2012

Trond Sirevåg



DIAS - Digital arkivpakkestruktur

- Samarbeidsprosjekt april 2010 - mai 2012 for å etablere et felles rammeverk for digitale arkivdepoter i stat og kommune
- Gjennomført - med økonomisk støtte fra ABM-utvikling (nå Norsk kulturråd) - av følgende prosjektpartnere:
 - Riksarkivaren, Bergen byarkiv, Oslo byarkiv, IKA Møre og Romsdal, IKA Trøndelag
 - KDRS (fra august 2011)
- Opprinnelige hovedmål:
 - Spesifisere OAIS-struktur for arkivpakker for alle typer digitale avleveringer til statlige og kommunale depoter
 - utvikle hjelpemidler for praktisk bruk av strukturen
 - etablere et bredt, permanent samarbeid om digitale arkivdepoter mellom Riksarkivaren og de sentrale aktørene fra kommunal sektor

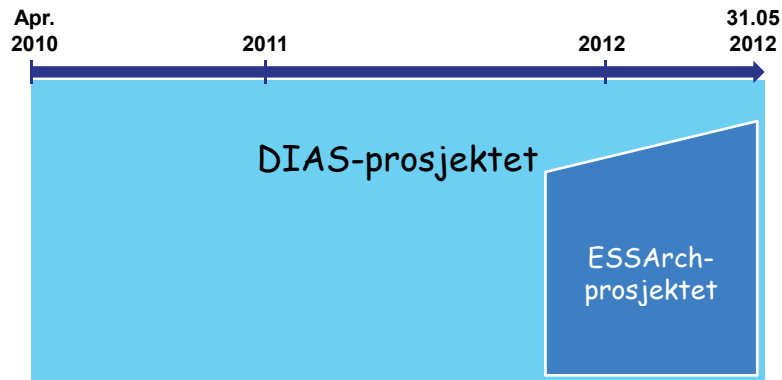
Bakgrunn – og forløpere

- Det kommunale eArkivsamarbeidet 2004 - 2006
 - jf. rapporten "Minnehåndtering - Metode for digital langtidslagring i kommunal sektor", ABM Skrift # 43, 2007
- LLP-prosjektet OpenARMS - initiert av eArkivsamarbeidet for å utvikle et OAIS-basert depotstyringssystem (2005-2006)
- IKA Trøndelags prosjekt for å teste tilgjengelige styrings-systemer for digitalt depot
- Riksarkivets prosjekt Elmag-2 for å planlegge nytt digitalt depot for Arkivverket, jf. rapporten "Digitalt og autentisk" (2010)
 - Riksarkivarens søknad til ABMU om prosjektmidler til DIAS skjøt ut av dette prosjektet
 - Etter en mer beskjeden målsetting innledningsvis har DIAS samlet i seg alle målene fra de tidligere kommunale prosjektene

Planutvikling

- ABMU bevilget 1 mill. kr. til ett-årig prosjekt (fra april 2010)
- Høsten 2010 ble prosjektrammen utvidet, og slutføringen utsatt til utgangen av 2011
 - Formål: å utvikle et DIAS-tilpasset forvaltningssystem basert på et eksisterende "repository"-system
 - I 2011 ble prosjektet tilført 1/2 mill. kr. i ekstra ressurser ved overføring av ABMU-prosjektmidler fra LLP og IKA-Trøndelag
- Implementeringen av DIAS-forvaltningssystemet (ESSArch) ble organisert som et eget prosjekt under DIAS-styringsgruppen
 - Sluttdato for ESSArch-prosjektet ble i planen satt til 01.04.2012, men senere utskutt til 31.05.2012
- DIAS-styringsgruppen utsatte også slutføringen av DIAS-prosjektet som helhet til 31.05.2012

Forholdet DIAS - ESSArch



DIAS styringsgruppe

Arnt Ola Fidjestøl, *IKA Møre og Romsdal*
 Karin Gjelsten, *Bergen byarkiv*
 Kari Remseth, *IKA-Trøndelag*
 Tore Somdal-Åmodt, *Oslo kommune/Byrådsavdeling for finans*
 Ole Gausdal, *Riksarkivet*, leder for styringsgruppen
 Anne Mette Dørum, *Riksarkivet*
 Hans-Herman Fischer, *Riksarkivet*
 Olav Hagen Sataslåtten, *Riksarkivet*
 Tor Eivind Johansen, *KDRS - Kommunearkivinstitusjonenes digitale ressurscenter*
 Terje Pettersen-Dahl, *Riksarkivet*, prosjektleder/ESSArch
 Trond Sirevåg, *Riksarkivet*, prosjektleder/DIAS

DIAS prosjektgruppe

Jan Tore Helle, *Bergen byarkiv*
 Frode Kirkholt, *oppr. Oslo byarkiv, fra aug. 2011: Riksarkivet*
 Petter Pedryc, *IKA-Trøndelag*
 Torbjorn Aasen, *IKA Møre og Romsdal*
 Hans Fredrik Berg, *Riksarkivet*
 Terje Pettersen-Dahl, *Riksarkivet*
 Terje Furuseth, *Statsarkivet i Hamar*
 Trond Sirevåg, *Riksarkivet, prosjektleder*

DIAS referansegruppe

Synne Stavheim, *Norsk kulturråd*
 Endre Grøtnes, *Difi - Direktoratet for forvaltning og IKT*
 Thomas Sødning, *Høgskolen i Oslo*
 Kjetil Reithaug, *KS IKT-Forum*
 Torleif Lind, *LLP - Landslaget for lokal- og privatarkiv*
 Lars Gaustad, *Nasjonalbiblioteket*
 Øyvind Rekdal, *Norsk arkivråd*
 Sigmund Evjen, *Oslo kommune/Byrådsavdeling for finans og næring*
 Tommy Maurud, *Oslo byarkiv*

Fra DIAS-styringsgruppen og -prosjektgruppen:
 Ole Gausdal, Terje Pettersen-Dahl, Hans Fredrik Berg og Trond Sirevåg, *Riksarkivet*

Egen ESSArch-prosjektgruppe (h. 2011-v. 2012)

Terje Pettersen-Dahl, *Riksarkivet/DD-seksjonen (prosjektleder)*
 Jan Tore Helle, *Bergen Byarkiv*
 Tor Eivind Johansen, *KDRS*
 Hallstein Bakken, *Riksarkivet/DD-seksjonen*
 Arne-Kristian Groven, *Riksarkivet/Elark-seksjonen*
 Helge Holte, *Riksarkivet/DD-seksjonen*
 Anthony Lærdal, *Riksarkivet/DD-seksjonen*
 Petter Svendsen, *Riksarkivet/DD-seksjonen*
 Stian Skindlo, *Riksarkivet/IT-avdelingen*
 Jørgen Skjånes, *Riksarkivet/IT-avdelingen*

ESSArch har i tillegg en referansegruppe med medlemmer fra de kommunale arkivmiljøene, Riksarkivet og statsarkivene

Prosjektmål og status for måloppfyllelse

<i>Fire mål i ABMU-søknad :</i>	<i>Status for måloppfyllelse:</i>
1) Utprøve og komplettere Elmag-arkivpakkemodellen	Modell definert som prosjektets arbeidsgrunnlag i juni 2010
2) Velge implementeringsstandarder for modellen	METS, PREMIS, EAD, EAC-CPF og ADDML valgt i 2010
3) Spesifisere xml-skjemaer for arkivpakkestrukturen	METS og PREMIS tilpasset for DIAS i egne xml-skjemaer i 2011
4) Utforme krav til programvare for å generere arkivpakker iht. xml-skjemaene	Fullstendig kravspesifikasjon til forvaltningssystem for DIAS-arkivpakker utarbeidet i 2010/11
<u>Utvidet mål høsten 2010:</u> 5) Utvikling av forvaltningssystem iht. DIAS-kravspesifikasjonen	Systemet ESSArch tilpasses for DIAS med sikte på driftsstart i Riksarkivet 01.06.2012

DIAS som en standard

- Definerer en standard for statlige og kommunale arkivdepoters implementering av OAIIS-modellen
 - Angir hvordan en arkivpakke skal organiseres ved bruk av eksisterende standarder (med tilhørende xml-skjemaer)

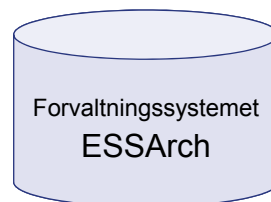
Dvs.: DIAS = en standard for bruk av standarder

DIAS-prosjektets produkter

• Spesifikasjoner

- 1) Arkivpakkemodell:
 - Basis arkivpakkestruktur
 - Prosessmodell for vedlikehold i digitalt depot
- 2) Tilpassede xml-skjemaer til støtte for modellen:
 - DIAS-METS og DIAS-PREMIS
 - Regler for mapping av (fullstendige) EAD- og EAC-CPF-skjemaer fra Asta
- 3) Kravspesifikasjon til forvaltningssystem

- Programvare for å anvende spesifikasjonene



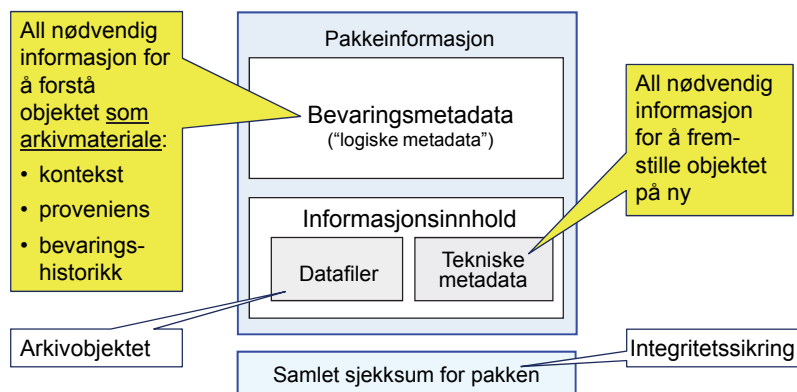
• Rutineopplegg

for håndtering av DIAS-arkivpakker i ESSArch

Arkivpakkemodell

Arkivpakke (AIP) iht. OAIS

En "container" hvor et arkivobjekt er lagret som en selvstendig enhet - integritetssikret, og fast forbundet med all tilhørende informasjon



Varianter: Leveringspakke (SIP), Bevaringspakke (AIP) og Brukspakke (DIP)

Implementeringsstandarder: DIAS xml-skjemaer

Bevaringsmetadata:

- **METS**
(DIAS-METS) Overordnet organisering av arkivpakker ("container"-beskrivelse) og oversikt over pakkeinnhold
- **PREMIS**
(DIAS-PREMIS) Bevaringsmetadata – i betydningen håndteringshistorikk (i arkivdepot)
- **EAD** Arkivbeskrivelse
- **EAC-CPF** Beskrivelse av arkivskapere o.a. aktører

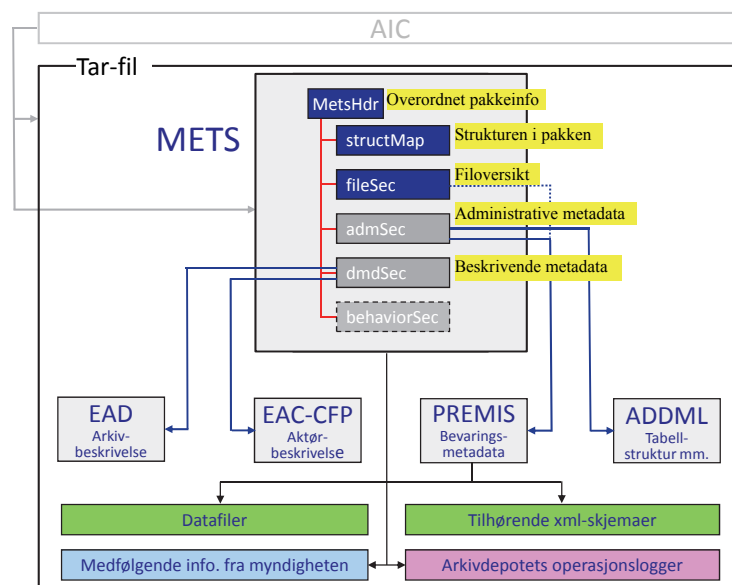
Tekniske metadata:

- **ADDML** Struktur- og postbeskrivelse av tabell-uttrekk fra databaser mm.

DIAS-METS og DIAS-PREMIS er spesifikke tilpasninger av METS og PREMIS for DIAS.

EAD/EAC brukes uten tilpasninger, men mappes mot Asta

Organisering av metadata i DIAS arkivpakke (AIP)



Arkivpakkemodellen som prosessmodell

Fokus på integritetssikting og oppfølging av kravene i
OAIS og TRAC

Utfordringene mht. integritet og autentisitet

- Digital informasjon kan enkelt endres og manipuleres
- Digitale arkiver kan ikke bevares statiske. De er i en prosess med stadig transformasjon gjennom hele livssyklusen
 - til fundamental forskjell fra papirarkiver
- Men hvordan kan folk stole på at stadig omskapt arkivmateriale fortsatt er autentisk?
- Hvordan garanterer vi oss når det settes søkelys på hva ulike aktører har hatt mulighet for å endre?
 - Hvordan lar det seg bekrefte at en aktør har nøyd seg med å foreta de endringer som sies å være utført?
 - Jf. O. J. Simpson-saken i USA i 1995, hvor retten underkjente politiets fingeravtrykk-bevis

TRAC-standarden

TRAC - Trustworthy Repository Audit and Certification - Criteria and Checklist (2007)

- Requirements to ensure the sustained integrity and authenticity of objects in repository custody

Trustworthy Repositories
Audit & Certification:
Criteria and Checklist



Contents:

Introduction
Establishing Audit and Certification Criteria
Towards an International Audit & Certification Process
Using this Checklist for Audit & Certification
Applicability of Criteria
Relevant Standards, Best Practices & Controls
Terminology
Audit and Certification Criteria
Organizational Infrastructure
Digital Object Management
Technologies, Technical Infrastructure & Security
Audit Checklist
Glossary
Appendices

Version 1.0
February 2007

<http://www.crl.edu/PDF/trac.pdf>

Krav i OAIS og TRAC

Endret virkelighet for digitale arkivdepoter

- Et godt renommé er ikke nok
- Bevisbyrden snus
 - Depotet må eliminere muligheter for tvil eller spekulasjon om materialets autentsitet, integritet og evt. konfidensialitet
 - Depotet må selv kunne verifisere med dokumentasjon
 - TRAC spesifiserer 90 kriterier som ledd i et opplegg for en sertifisering av arkivdepoter
- Arbeidsoppgave nr. 1 for et arkivdepot: Å kunne verifisere at informasjonsinnhold er bevart uendret fra og med mottak
 - Garanterer ikke at informasjonen er ekte og troverdig i seg selv - men da er vi over i et annet tema

Fysisk vs. logisk integritetskontroll

- **Integritetskontroll er enkelt på fysisk nivå (bit-nivå)**
 - Sjekksommer kan brukes for å bekrefte at digital informasjon er lagret fysisk uendret (v/kontroll av uendret bit-strøm)
- **Men digital bevaring uten fysiske endringer er umulig**
 - Bevaring av digital informasjon krever regelmessig vedlikehold med fysisk transformasjon av data for å opprettholde lesbarheten
- **Det er til å leve med at arkivinformasjon er fysisk transformert**
 - bare vi kan bekrefte et uendret logisk informasjonsinnhold
- **Men det eksisterer ingen enkel metode tilsvarende sjekksommekontroll for å bekrefte at logisk informasjonsinnhold er uendret**

Integritetskontroll på fysisk nivå

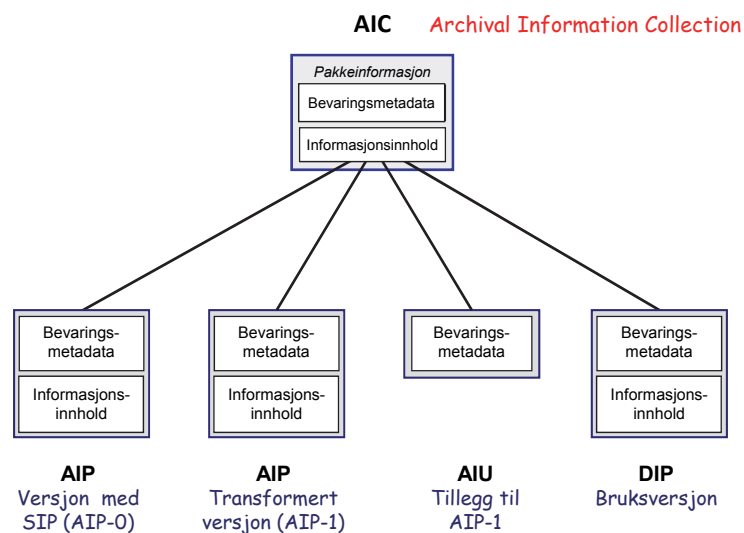
Uendret bit-strøm kontrolleres ved bruk av unik sjekksom

The slide displays two examples of bit streams for Word-2000 and Word-2007 files. The top example is for Word-2000, showing a bit stream starting with '00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F' and a corresponding hex dump. The bottom example is for Word-2007, showing a bit stream starting with '00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F' and a corresponding hex dump. The slide is part of a presentation by V-FI(0)-LongRec/Prezentasjoner/RA-TilretteDepoter_31102008.ppt.

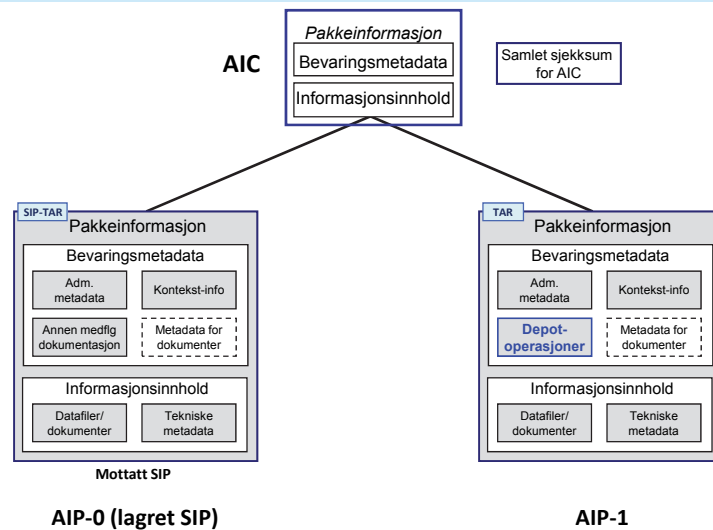
Modell for kontinuerlig integritetssikring i DIAS

- Integritetsproblematikken løses ved å bygge på resepten i TRAC: Alt bevares som tillegg - uten å endre eldre informasjon
- Sjekksommer kan da brukes på alle objekter/filer - og tjene som bestandige mekanismer for verifisering
 - En mottatt arkivversjon (SIP) bevares uendret og sjekksumsikret slik den ble mottatt - for alltid
 - Operasjoner i depot etter mottak bevares og dokumenteres som tillegg (med tilhørende sjekksommer) - utenfor SIP-en
 - Ved senere konverteringer/transformeringer bevares en ny generasjon av vedk. arkivpakke (AIP) i tillegg til den opprinnelige
 - *Alternativt kan enkle AIU-er brukes for å lagre tillegg, både fordi en eksisterende AIP aldri skal oppdateres, og fordi det er behov for å unngå at enhver depotoperasjon skal kreve en ny, transformert AIP*

Organisering i arkivsamlepakker (AIC)

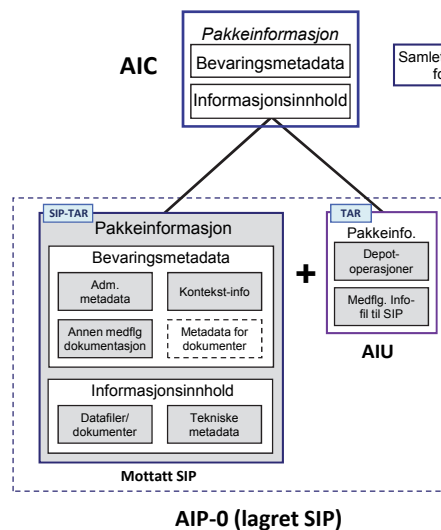


Prosess ved generering av arkivpakker (eks. 1)



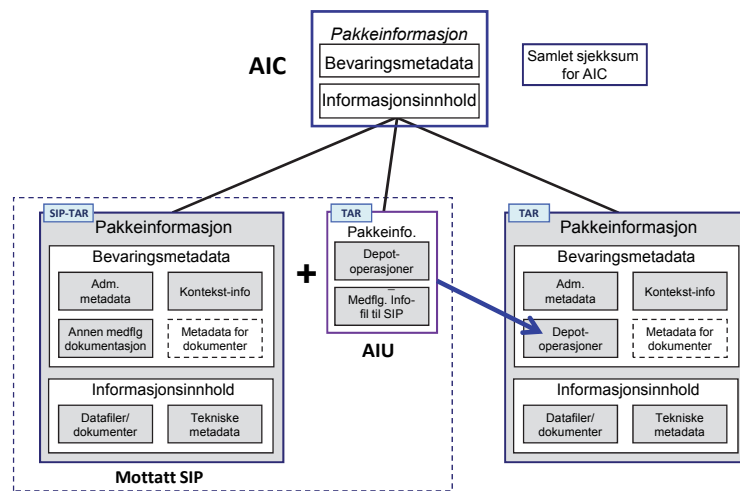
Når SIP lagres, genereres her også en transformert AIP (AIP-1). SIP-ens medfølgende info. + dokumentasjon av depotoperasjonene ved mottak innarbeides i AIP-1

Prosess ved generering av arkivpakker (eks. 2)



Dersom en SIP lagres uten at det samtidig blir generert en transformert AIP (1), skal det brukes en AIU i tillegg til SIP for medfølgende info. + depotoperasjoner

Prosess ved generering av arkivpakker (eks. 3)



AIP-0 (lagret SIP)

SIP ble her lagret uten at det samtidig ble generert en transformert AIP. Dette krevde en AIU i tillegg

AIP-1

Ved en senere generering av en transformert AIP (1) innarbeides AIU i denne

DIAS-prosjektets forvaltningssystem

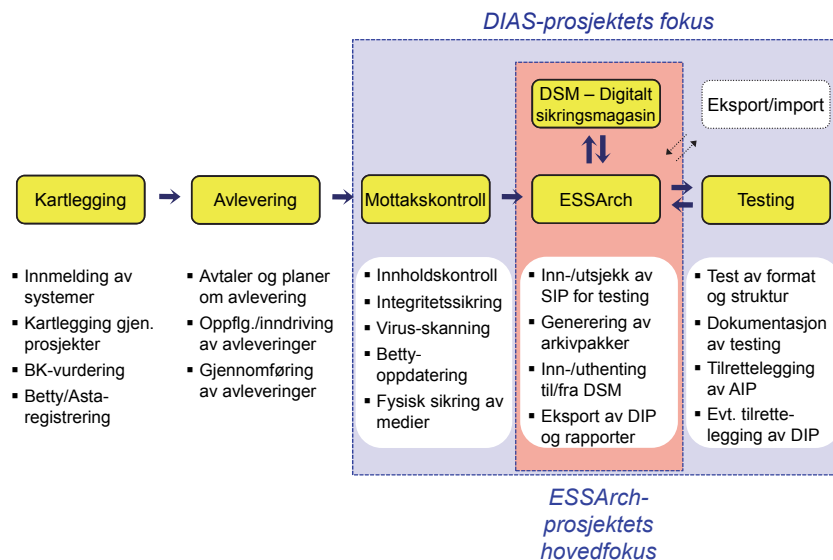
ESSArch: DIAS-prosjektets forvaltningssystem

- Svensk system som tilpasses for DIAS slik at det oppfyller samtlige punkter i prosjektets kravspesifikasjon
 - Avtaler undertegnet 21.11.2011 etter åpen tilbudsinnhenting
 - Leverandøren, ES Solutions AB, er også ansvarlig for tilpasning og videreutvikling
- Systemet er basert på åpen kildekode (GPLv3), og vil bli distribuert vederlagsfritt til "alt og alle"
- Riksarkivets vedlikeholdsavtale for ESSArch kan også fungere som en sentral vedlikeholdsavtale for kommunale depoter
 - Dvs. at all leverandørkontakt må kanaliseres via en sentral instans i Riksarkivet
 - Kommunale depoter har også adgang til å inngå egne avtaler om vedlikehold

ESSArch: Sentrale egenskaper

- 1) Genererer arkivpakker iht. DIAS-modellen og på grunnlag av DIAS xml-skjemaer
- 2) Styrer arbeidsprosessene fra innsjekking til lagring i depot, ivaretar integritetssikring i alle ledd, og utfører tilgangskontroll
- 3) Styrer vedlikehold av arkivpakker iht. DIAS-prosessmodellen
- 4) Styrer import/eksport av informasjon til/fra digitalt depot
- 5) Har egen database med nøkkelinformasjon om arkivpakkene i digitalt depot og arbeidsprosesser i tilknytning til dem
- 6) Ivaretar også *lagringsadministrasjon* (skriving til "multiple" media, kopiering/migrering mellom medier + verifisering)
- 7) Kan tilpasses for ulike lagringsløsninger/mediekonfigurasjoner

DIAS/ESSArch i prosessene tilknyttet digitalt depot

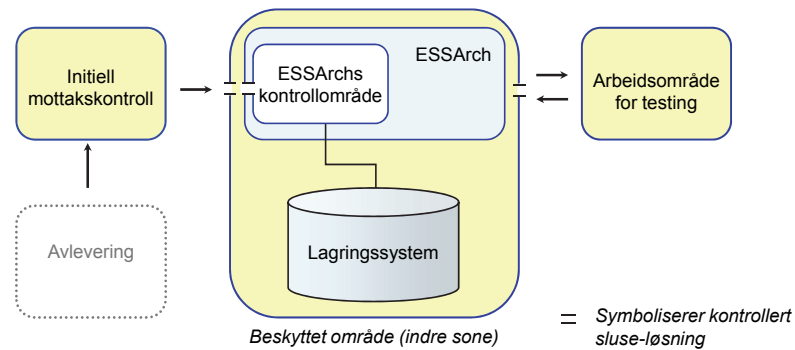


Utenfor ESSArchs styring: Initiell mottakskontroll

Mottakskontrollen omfatter:

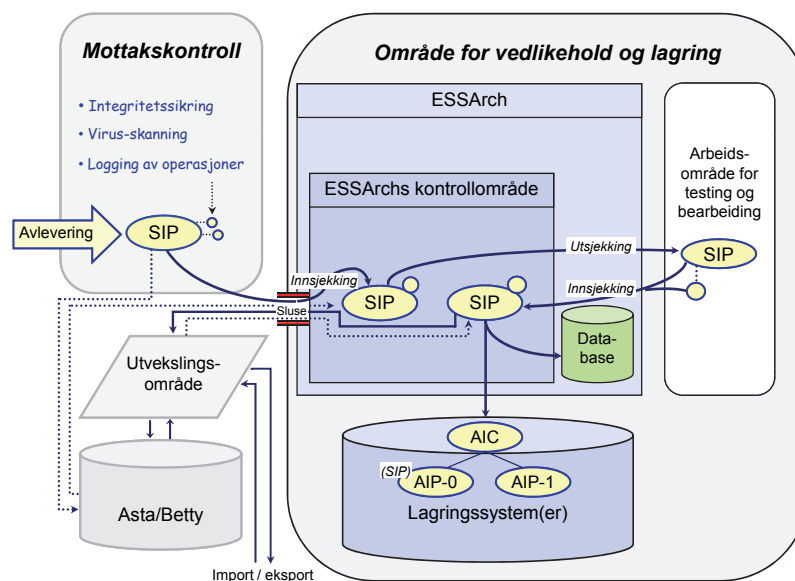
- Registrering (inkl. journalføring)
- Integritetssikring av mottatt SIP ved å generere samlet sjekksum + verifisering av evt. medfølgende sjekksum
- Enkel kontroll av at SIP-ens innhold er som fastsatt/avtalt
- Virus-skanning etter periode med karantene (min. 3 uker)
- Registrering/oppdatering av Asta (via Betty-systemet)
 - Herunder: fremstilling av arkivbeskrivelser (EAD/EAC) til ESSArch
- Tilrettelegging av SIP + medfølgende dokumentasjon + mottakskontrollens operasjonslogger for innsjekk i ESSArch

Inn- og utsjekking til/fra ESSArch



- Alle ESSArch-operasjoner skjer innenfor et kontrollert miljø, og all import/eksport til/fra lagringssystemet styres av ESSArch
- Dokumentasjon fra mottakskontroll og arbeidsområder for testing importeres samlet til ESSArch - og sjekksumsikres

Prosesser i et ESSArch-styrt digitalt depot



DIAS forvaltningsforum

- 1) Permanent samarbeidsforum for DIAS-oppfølgning f.o.m. prosjektslutt 31.05.2012
- 2) Eget prosjekt støttet av Kulturrådet i 2012 for å etablere forumet (*start 01.04.2012*)

Oppgaveramme for DIAS-forum

Ramme mht. utførende oppgaver:

- 1) Vedlikehold og videreutvikling av DIAS-standarder (xml-skjemaer)
- 2) Forvaltning av ESSArch og ESSArch-avtaler for et samlet brukermiljø
- 3) Kompetansesenter for fag- og driftsmiljøer

*Oppgaver for DIAS-forum***Oppgaver i tilknytning til DIAS-standarder**

- Vurdere behov for endringer eller tillegg i xml-skjemaer ut fra en vurdering av behov og praktisk-økonomiske implikasjoner
- Iverksette endringer i skjemaer eller styre iverksettelsen
- Vurdere implementering av endringer i relasjon til ESSArch
- Ivareta samordningsbehov med det svenske riksarkivet og andre

*Oppgaver for DIAS-forum***Vedlikehold og videreutvikling av ESSArch**Vedlikehold:

- 1) Overvåke vedlikeholdsavtale med ES Solutions (ESS)
- 2) Være sentral for mottak av ESSArch-feilmeldinger o. l. fra brukermiljøer
- 3) Vurdere om feilmeldinger skal oppfølges og videreformidles til ESS.
- 4) Om nødvendig simulere innmeldt feilsituasjon på et anlegg i Riksarkivets driftsmiljø for å skaffe grunnlag for en videre vurdering
- 5) Gjøre avtaler med ESS om implementering når det foreligger oppgraderte programmoduler fra ESS

Brukerstøtte:

- 6) Besvare henvendelser, evt. videreformidle henvendelser til ESS

Videreutvikling:

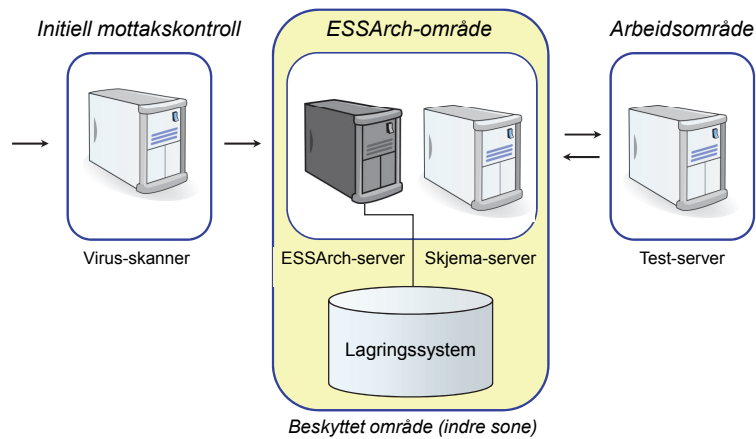
- 7) Fremme forslag om utviklingstiltak til styringsgruppen med utgangspunkt i hendelser og behov

Oppgaver for DIAS-forum
Opplæring og kurs

- Sørge for at det foreligger materiell og ressurser som grunnlag for opplæring.
- Arrangere kurs for fag- og driftspersonell.

Annet

Utstysrbehov ved ESSArch-innføring



Funksjoner i Betty

- 1) Understøtte ekstern kartlegging og bevaringsvurdering av arkivmateriale
- 2) Registrere eksternt arkivmateriale i interaksjon med Asta
- 3) Administrere avleveringer
- 4) Generere arkivbeskrivelser til ESSArch (EAD- og EAC-CPF-filer)
- 5) Utføre prosessstyring ved å vise status for behandlingen av avleveringer fra mottak til innlemmelsen i digitalt depot