

Orientering om E-ARK4ALL

Et pågående delprosjekt av
CEF eArchiving building block

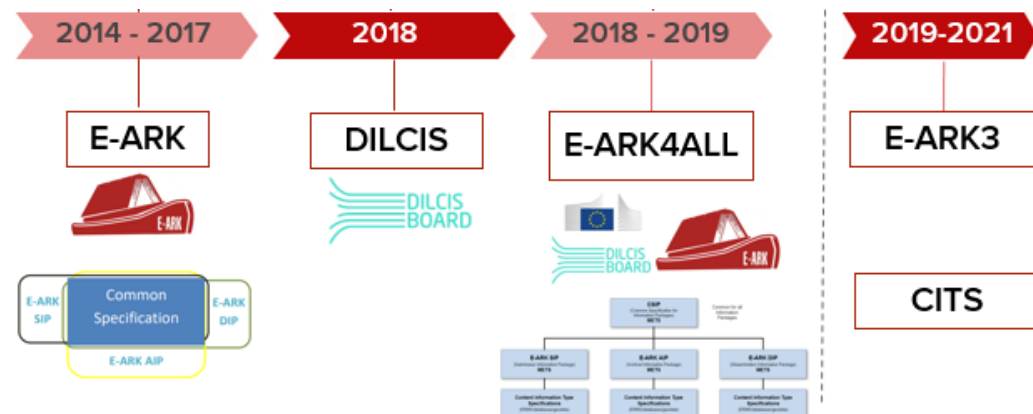


Agenda

- › Introduksjon og historie
- › Hvorfor er dette viktig for oss?
- › Hva består prosjektet av?
- › SIARD
- › Veien videre → EARK3?

Introduksjon og historie

- Prosjektene er en del av CEF eArchiving building block
- Formål:
Forenkling av langtidsbevaring, fellesløsninger som brukes av flest mulig
- Innhold:
Spesifikasjoner, verktøy og veiledning for digital arkivering, inkludert langtidslagring
- Meet eArchiving, the new CEF Building Blocks:
<https://youtu.be/VbR-zvr45so>



Hvorfor er dette viktig for oss?

- › Internasjonalt samarbeid
- › Innspill og beste praksis på ting vi jobber med
- › Felles verktøy
- › Direkte tilgang til ekspertise
- › Bidra til å lage standardløsninger som flere kan bruke
- › Påvirke den internasjonale utviklingen
- › Kan resultere i føringer fra EU



E-ARK (2014 – 2017)

I E-ARK utviklet man metodene og teknologien som bruke i eArchiving:

- › Verktøy for pre-ingest og generering av SIP
 - › RODA-in, ETP (ESSArch Tools for Producer), ETA (ESSArch Tools for Archives)
- › Verktøy for arkivering av relasjonsdatabaser
 - › Database Preservation Toolkit (DBPTK), dbptk-gui-backend, Database Visualization Toolkit
- › Verktøy for langtidsbevaring
 - › RODA, ESSArch Preservation Platform, E-ARK Web



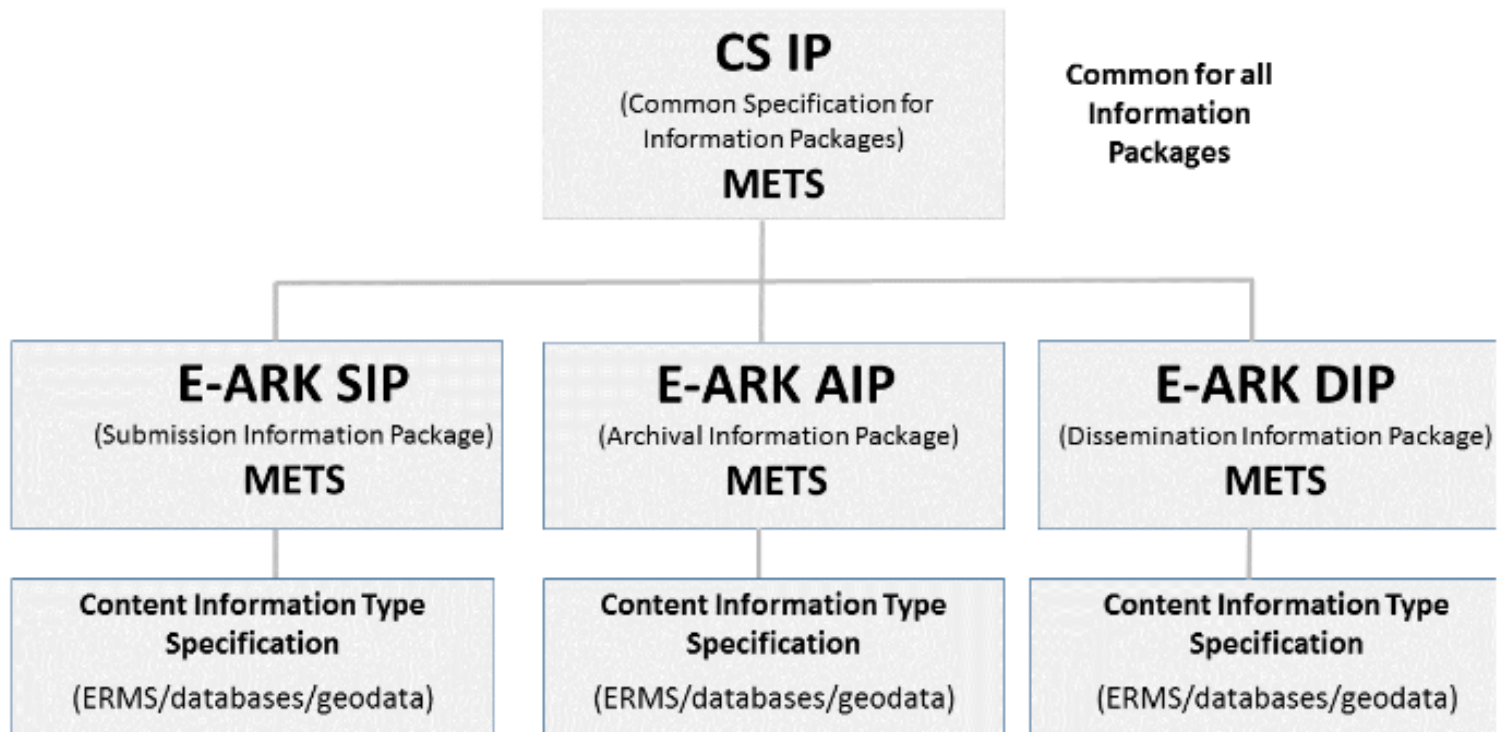
E-ARK (2014 – 2017)

- › Verktøy og komponenter for access
 - › Querying, ordering and viewing of E-ARK DIP
 - › Order Management Service (backend), E-ARK Platform UI (frontend)
- › Tekniske spesifikasjoner: E-ARK CS IP, E-ARK profiler (SIP,AIP,DIP), SIARD 2.0
- › Tjenester for interessenter: Maturity model, E-ARK knowledge centre, E-ARK general model

E-ARK4ALL (juni 2018 – juni 2019)

- › A1 Coordination
- › A2 Specification stabilisation
 - › Kjernen i prosjektet
- › A3 Sample software
- › A4 Training
- › A5 User engagement and outreach

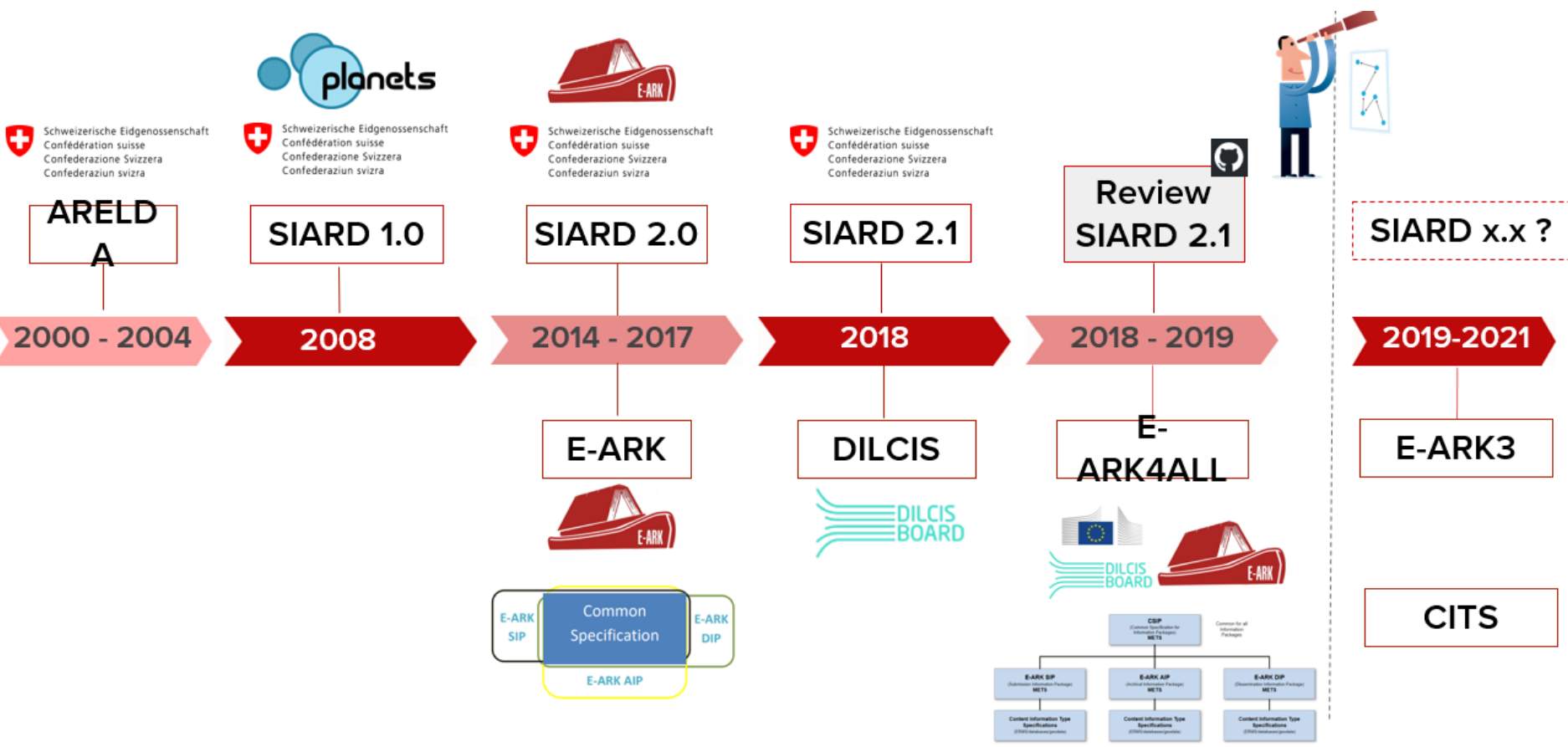
Kjernen av eArchiving



Standarder som brukes i CS IP

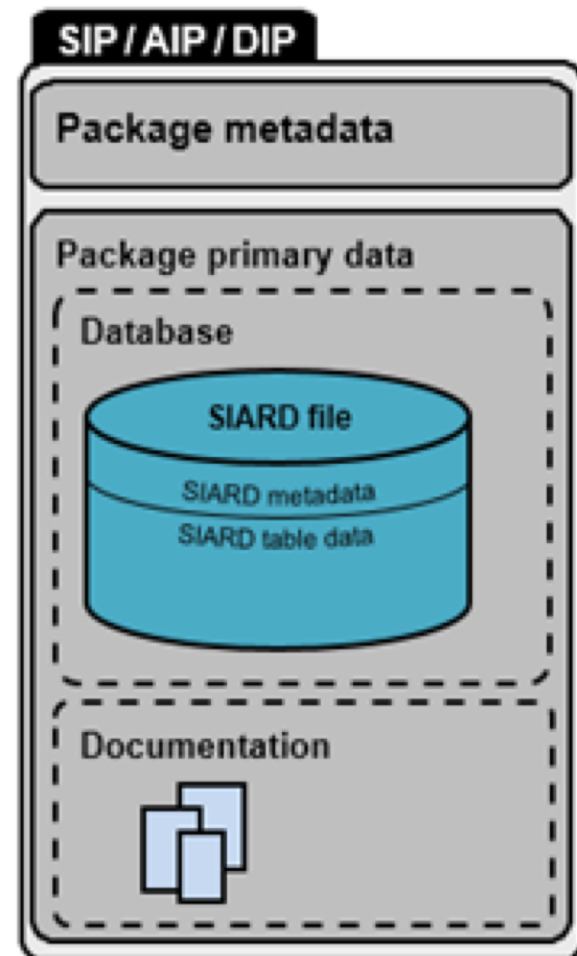
- › Reference model for an Open Archival Information System (OAIS) ([OAIS Reference model](#))
- › Metadata Encoding and Transmission Standard ([METS](#))
- › Preservation Metadata Implementation Strategies ([PREMIS](#))
- › General Data Protection Regulation 2016/679/EU (GDPR)
- › A Content Information Type specification
- › Data Entity Dictionary Specification Language (DEDSL) – XML/DTD: ISO 22643:2003 (CCSDS 647.3-B-1:2002)
- › Information technology – Document Schema Definition Languages (DSDL) – Part 3: Rule-based validation – Schematron: ISO/IEC 19757-3:2016 (ISO standard)

SIARD – Historie og overblikk

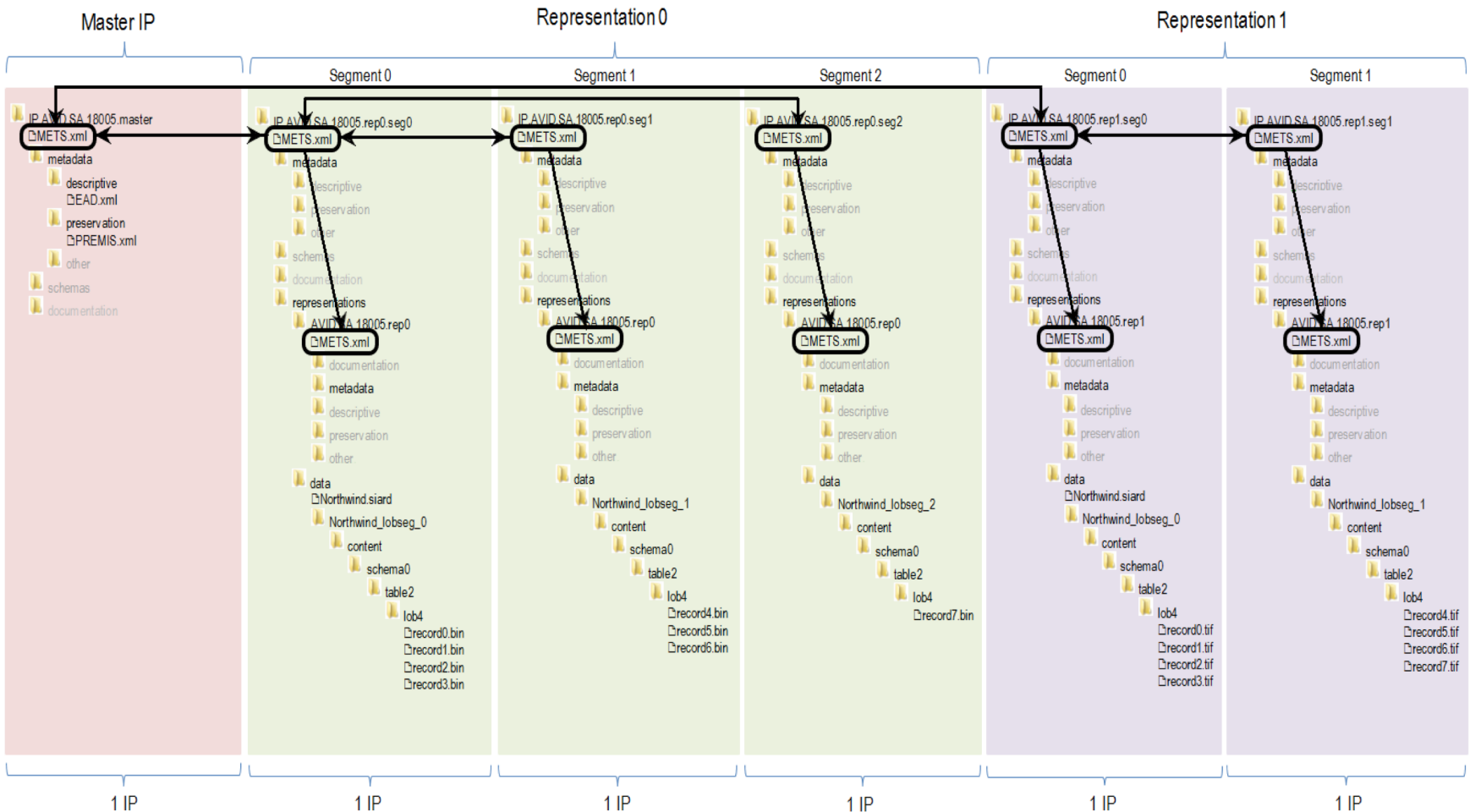


SIARD som del av en Information Package

- SIARD som langtidslagringsformat for relasjonsdatabaser er helt uavhengig av pakkestruktur
- Prosjektet ønsker å definere en CITS som kan brukes i pakkestrukturene som har blitt laget
- Det har blitt laget et utkast til hvordan en SIARD-database kan bli pakken i en IP



SIARD som del av en information package



1 IP

1 IP

1 IP

1 IP

SIARD og LOBer i E-ARK

- › LOBer utenfor SIARD-filen ble introdusert for SIARD 2.0
- › Databaser og håndtering av LOBer anses som utfordrende, uavhengig av om de behandles
 - › Internt – lagret i en celle i relasjonsdatabasen
 - › Eksternt – lagret utenfor relasjonsdatabasen
 - › Direkte referanser (sti- og filnavn)
 - › Indirekte referanser (fil-ID)

Project 620998: European Archival Records and Knowledge Preservation - E-ARK



RECOMMENDATION

for storing large objects outside the SIARD file

Project Acronym: E-ARK

Grant Agreement Number: 620998

Project Title: European Archival Records and Knowledge Preservation

AUTHOR(S)	
Name(s)	Organisation(s)
Anders Bo Nielsen	Danish National Archives (DNA)
Kathrine Hougaard Edsen Johansen	Danish National Archives (DNA)
Phillip Tømmerholt	Danish National Archives (DNA)

REVIEWER(S)	
Name(s)	Organisation(s)
Kuldar Aas	National Archives of Estonia (NAE)
Bruno Ferreira	KEEPS
Andreas Voss	Swiss Federal Archives

SIARD og metadata i E-ARK

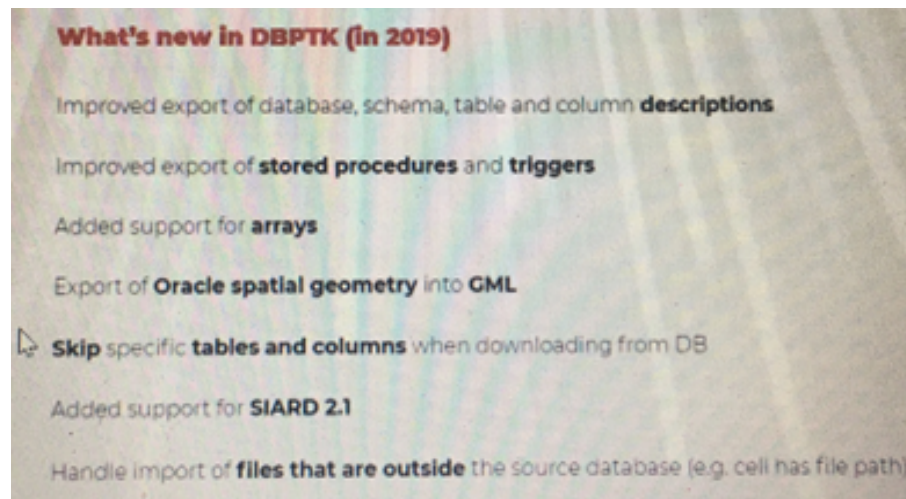
- › Hvordan tilføre metadata?
 - › Metadata inni SIARD-filen er begrenset til SQL-standarden
 - › Metadata utenfor SIARD-filen har ikke disse begrensningene
- › Dette vil få mer fokus i tiden fremover

SIARD verktøy

SIARD Suite

- › Pågående arbeid
- › Vil oppdatere ihht nye SIARD-versjoner (2.2 eller 3.0)

Database Preservation Toolkit



SIARD verktøy

Database Visualization Toolkit

What's new in DBVTK (in 2019)

Performance and stability improvement (Solr 7.X upgrade, one collection per db)

Support for **SIARD 2.1**

Configurable search stemming **language**

New **deployment** method (using docker)

What's next for DBVTK (in 2019)

Improve and **simplify design** and information organization (hide technical info)

Support for **materialized views**

Improve navigation through references/foreign keys (also try to reduce the need for it)

Log all searches and provide an interface to audit them

Automatically translate codes in associative table resolving **many-to-many relationships**

Veien videre for SIARD i prosjektet

- › Oppdatere anbefalingen for LOBer ihht SIARD 2.x / 3.0?
- › Lage anbefaling for SIARD og metadata
- › Utarbeide en felles validator for SIARD 2.x / 3.0

Veien videre generelt → EARK3

- › Det er søkt om midler til EARK3 (2019 – 2021)
 - › Videreutvikling av arbeidet fra E-ARK4ALL, lage flere CITSer,
- › Arkivverkets deltagelse
 - › SIARD CITS
 - › Er med i oppstarten rundt arbeid med referansearkitektur

- › Get involved!

anehov@arkivverket.no

arngro@arkivverket.no