



Riksarkivet

**Nordisk EI-Arkivseminar
Island 25-27 maj 2011**

**Svensk SIP och AIP
Mats Berggren, Riksarkivet**



Förstudie om E-Arkiv och E-Diarium

Förvaltningsgemensamma specifikationer

”Syftet med förvaltningsgemensamma specifikationer är att förenkla utveckling, upphandling och införande av enhetliga lösningar, att minska kostnaderna samt att skapa förutsättningar för att underlätta sökning och återanvändning av information i handlingar och ärenden. För att detta ska vara genomförbart behöver obligatoriska metadata identifieras. Dessa kan utgöras av en minimiuppsättning av nödvändiga basuppgifter som uppfyller grundläggande krav hos samtliga myndigheter.”

Samverkan mellan stat och kommun

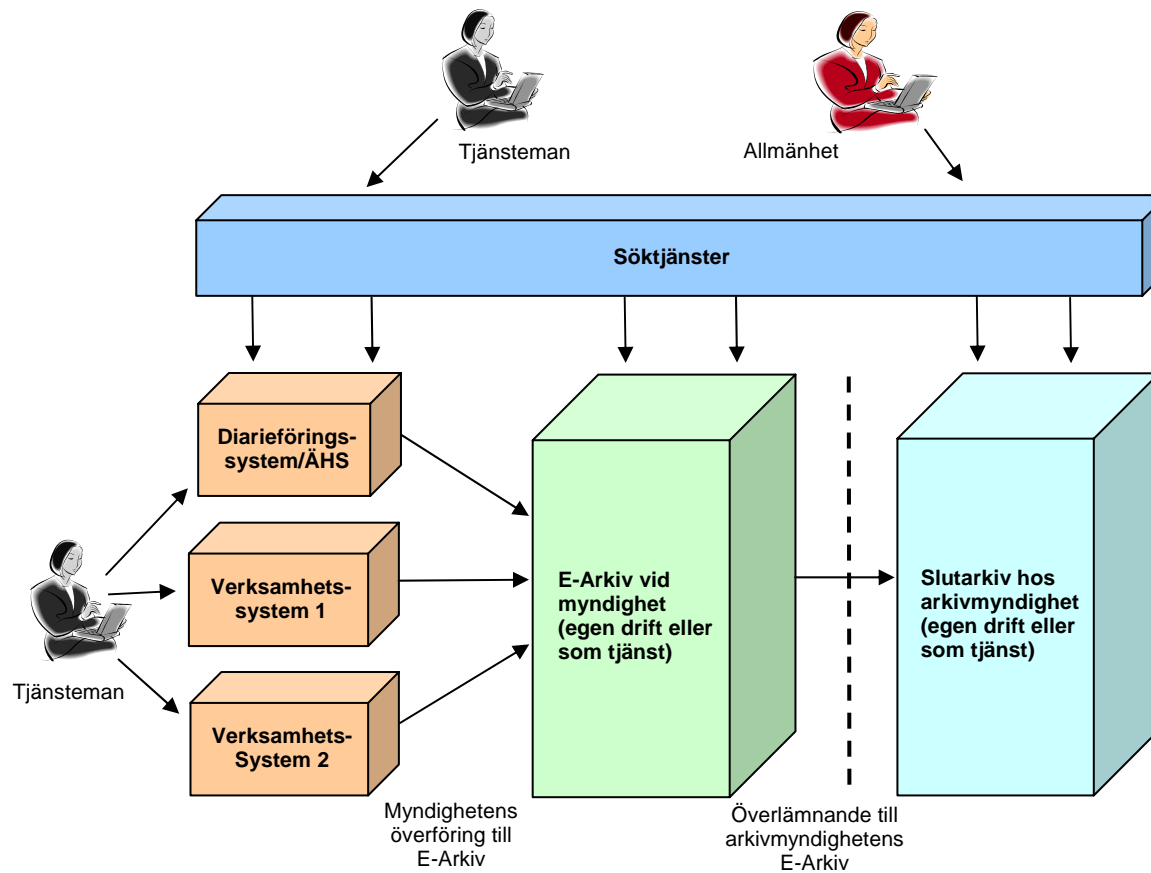
”Förstudien föreslår att förvaltningsgemensamma specifikationer för e-arkiv och e-diarium utarbetas i samverkan mellan statliga myndigheter, kommuner och landsting. Kompetens och erfarenheter från genomförda och nu pågående projekt hos myndigheter och leverantörer bör tas tillvara i detta arbete.”

Generisk modell för metadata (ur förstudiens bilagor)

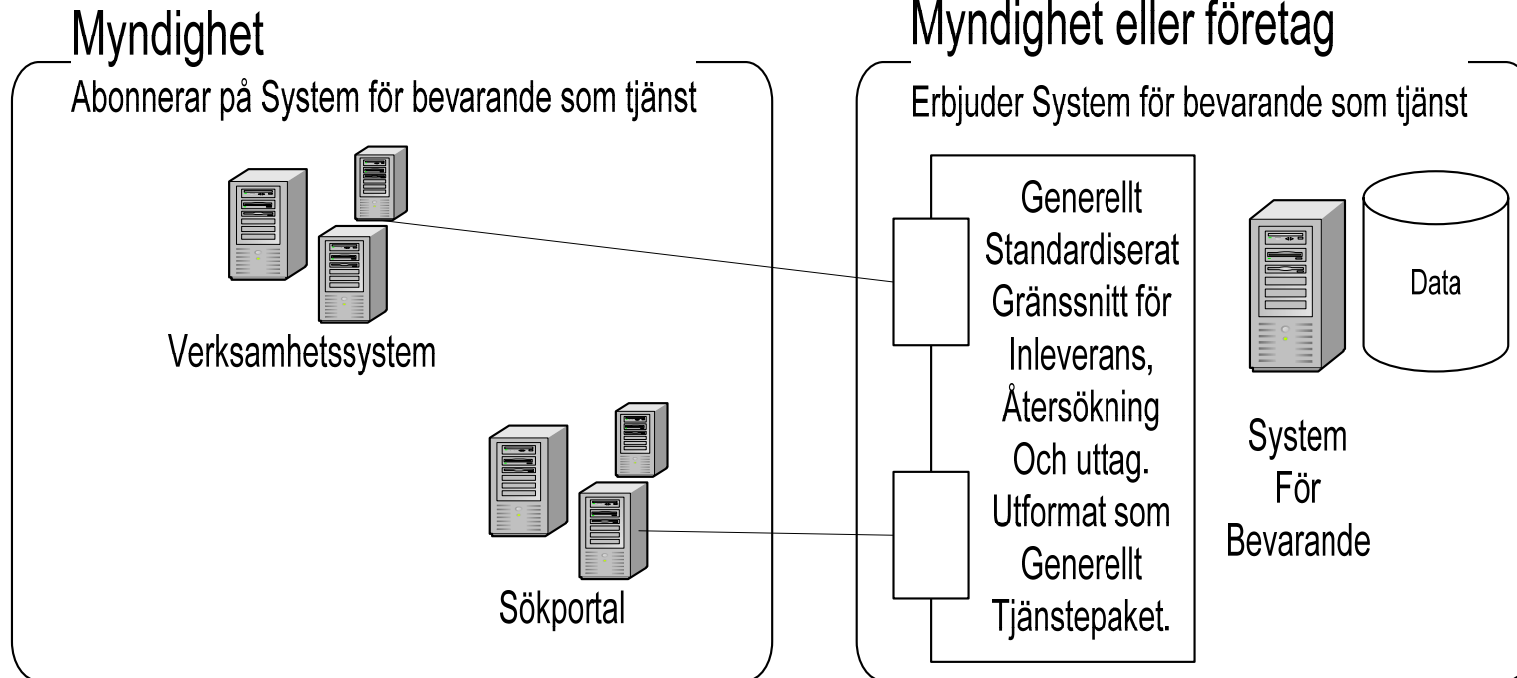
”På samma sätt (som det tjänstebaserade gränssnittet) ska även paketstrukturen och metadata vara generiskt uppbyggda. En viss basuppsättning metadata (som säkerställer identifiering och dataintegritet) ska vara obligatorisk, men det ska vara möjligt för myndigheter att också tillföra ytterligare myndighetsspecifik metadata. Obligatorisk metadata ska under alla förhållanden vara sådan att den på ett grundläggande sätt säkerställer långtidsbevarandet enligt Riksarkivets författningar.”



System för långsiktig informationsförsörjning



som tjänst



Uppdrag om E-Arkiv och E-Diarium

Uppdrag från E-delegationen

-Utarbeta förvaltningsgemensamma specifikationer för E-Arkiv och E-Diarium

Samverkan med myndigheter i Västernorrland

-Samarbete kring piloter för E-Arkiv och E-Diarium med myndigheter i Västernorrlands län
-Koordineras av landsarkivarien i Härnösand

Samverkan kring förvaltningsgemensamma specifikationer

-ENSAM-gruppen (Definition av grundläggande METS-profil 2010-05-12)
-Sambruk (47-kommuner i samverkan)
-KSL (Kommuner i Stockholms län)
-REDA (Region Skåne)
-R7 (Flera landsting i mellansverige)
-Intresserade myndigheter: Kungliga Biblioteket, Transportstyrelsen, Rikspolisstyrelsen mfl



Generell paketstruktur för E-Arkiv

Förvaltningsgemensam specifikation för stat, kommun och företag

- Myndigheter som bygger egna OAIS-baserade E-Arkiv
- Kommuner som bygger egna OAIS-baserade E-Arkiv
- Arkivmyndigheter
- Företag som erbjuder E-Arkiv som tjänst

Generisk specifikation för olika typer av system

- Generisk struktur, samma grundstruktur och samma basinformation oavsett systemtyp
- Systemspecifika definitioner på lägre nivå. Utarbetas av myndigheter i samverkan
- Regelverk för hur specifikationer utarbetas och underhålls
- Möjlighet till myndighetsspecifik metadata
- Tillämpa existerande standarder

Flexibel paketstruktur

- Stora paket, exempelvis när ett helt ärendehanteringssystem (exempelvis W3D3) exporteras och levereras som en enda SIP till arkivmyndigheten
- Små paket, exempelvis ett paket per ärende i de E-Arkiv som byggs av myndigheter (exempelvis RPS)
- Möjlighet att överföra många paket på en gång, behov av en "container"-struktur för paket



Standarder

- [OAIS](#) - Reference Model for an Open Archival Information System (CCSDS/ISO) **(2002)**
- [METS \(Metadata Encoding & Transmission Standard\)](#) - Structure for encoding descriptive, administrative, and structural metadata (DLF/LOC) **(2004)**
- [PREMIS \(Preservation Metadata\)](#) - A data dictionary and supporting XML schemas for core preservation metadata needed to support the long-term preservation of digital materials (OCLC/LOC) **(2005)**
- [MIX \(NISO Metadata for Images in XML\)](#) - XML schema for encoding technical data elements required to manage digital image collections (ANSI/NISO) **(2006)**
- [TAR \(Tape Archive\)](#) – Filformat för paketering av filer för effektiv lagring på tape och disk.

Övriga format

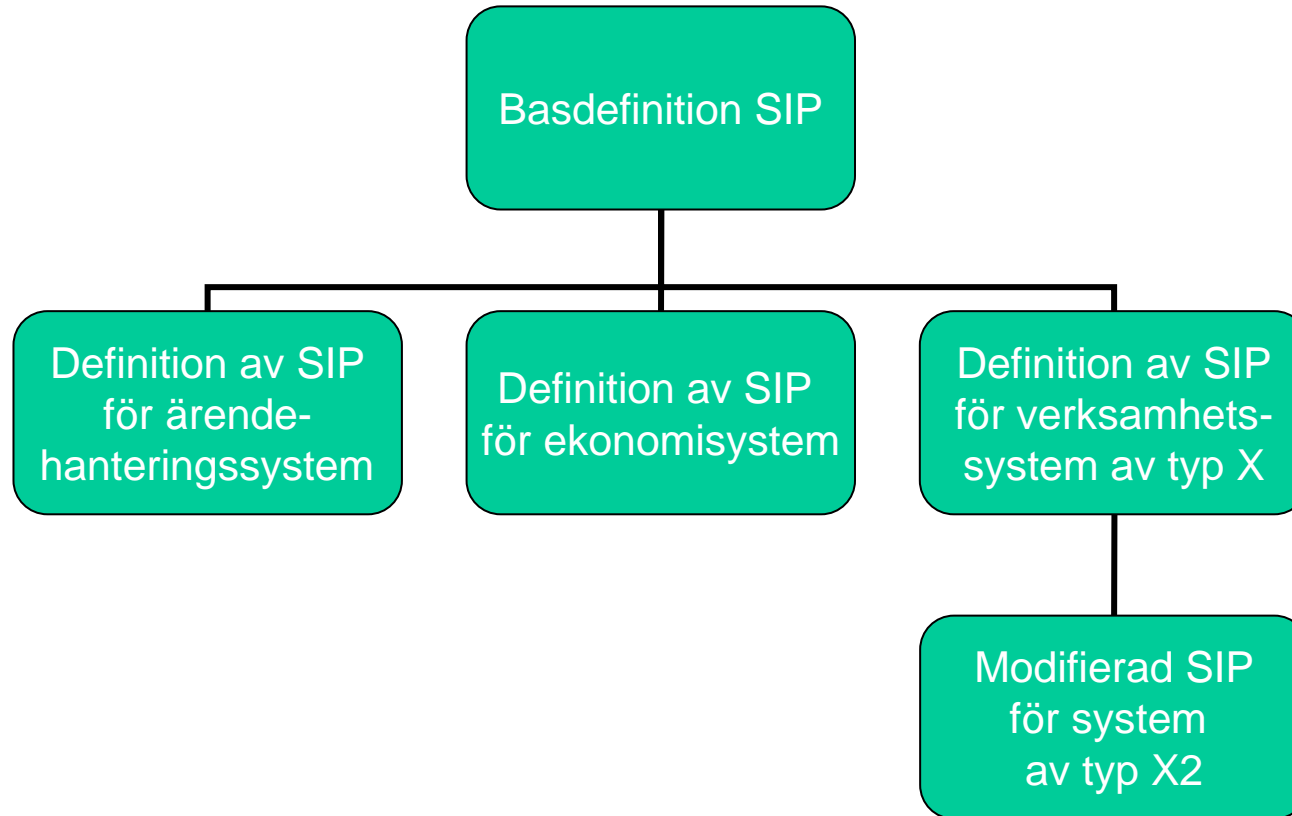
- [ADDML](#) (Archival Data Description Markup Language) – Norska riksarkivets, och nu även Svenska riksarkivets, XML-format för beskrivning av levererade filer, i synnerhet flata filer med postbeskrivningar **(2001, 2008)**

Under utredning

- [WARC](#) (Web ARChive, ISO 28500) – Format för paketering och arkivering av webbsidor **(2008)**
- [AudioMD](#), [VideoMD](#) och [AES X098-B](#) – Föreslagna scheman för teknisk metadata om audiofiler och videofiler.



paketstruktur för E-Arkiv



PKETSTRUKTUR I METS

METS Attribute/Element	Kommentar
LABEL	Beskrivning av paketet
OBJID (M)	Unikt ID för paketet
PROFILE (M)	METS-Profil
TYPE (M)	"AIC", "AIP", "AIU", "DIP" eller "SIP"
<metsHdr>	
<fileSec>	
<structMap>	

<metsHdr> Attribute/Element	Kommentar
CREATEDATE (M)	Tidpunkt när METS-filen skapats
<metsDocumentID> (M)	METS-filens namn
<altrecordID>	
<agent>	

<altrecordID> Attribute/Element	Kommentar
TYPE	"Systemtyp"
Elementinnehåll	Unik beteckning för systemtyp Används för att identifiera paket baserat på typ av system. Förvaltningsgemensamma beteckningar för olika systemtyper. Möjliggör en generisk modell för paketstruktur. Varje systemtyp skall vara definierad på ett sådant sätt att det möjliggör paketering och uppackning av motsvarande paket.

<altrecordID> Attribute/Element	Kommentar
TYPE	"Referens till arkivredovisning"
Elementinnehåll	Unik referens till arkivredovisning Används för att knyta paket till arkivredovisningen. Varje paket ska kunna ges en unik identifikation i en arkivredovisningsstruktur.



Metadata i baspaket för E-Arkiv

Det bör fastställas en struktur för ett baspaket som sedan kan modifieras beroende på systemtyp. Baspaketet ska innehålla den metadata som utgör "minsta gemensamma nämnare" för alla typer av paket. Grunden är METS, men sedan kan det diskuteras om även PREMIS ska vara obligatoriskt för alla typer av paket.

METS (obligatoriskt)

- Paketidentitet (OBJID)
- Pakettyp (TYPE)
- Systemtyp (altrecordID)
- Referens till arkivredovisning (altrecordID)
- Avsändare (agent)
- Mottagare (agent)
- Referenser till metadata (mdRef, mdWrap)
- Filreferenser (fileSec)
- Checksummor (fileSec)

PREMIS:Object (obligatoriskt?)

- Filformat och version
- Referens till filformatregister (PRONOM etc)

PREMIS:Rights (obligatoriskt?)

- Hänvisning till sekretesslag och personuppgiftslag

PREMIS:Agent (ej obligatoriskt?)

- Uppgift om person/mjukvara

PREMIS:Event (ej obligatoriskt?)

- Information om händelser i samband med leverans och mottagande



SIP-struktur för E-Arkiv

METS (content) Objekt-ID, filidentiteter, checksummor, information om avsändare och mottagare samt referenser till metadata i andra format (mdRef eller mdWrap). **Typfält (altrecordID) anger typ av paket beroende på systemtyp. Hänvisning till arkivredovisning.**

<PREMIS> Teknisk metadata om filer, metadata om bevarandeprocessen

<ADDML> Postbeskrivningar för flatfiler

<EAD>, <EAC-CPF> Eventuell bifogad arkivredovisning i XML-form

Andra metadataformat, beroende på systemtyp (se ovan), exempelvis MODS för Kungliga Biblioteket

→ <Datafil> Flatfil, XML-fil eller PDF-fil från myndigheter eller TIFF-fil från egen skanning

→ <Datafil>

→ <Datafil>

→ <XML-schema> SIP:en skall innehålla scheman för alla ingående XML-filer

→ <Dokumentationsfil> PDF-fil från myndigheten innehållande dokumentation av leveransen



AIP-struktur för E-Arkiv

METS (content) Objekt-ID, filidentiteter, checksummor, information om avsändare och mottagare samt referenser till metadata i andra format (mdRef eller mdWrap). **Typfält (altrecordID) anger typ av paket beroende på systemtyp. Hänvisning till arkivredovisning.**

<PREMIS> Teknisk metadata om filer, metadata om bevarandeprocessen

<ADDML> Postbeskrivningar för flatfiler

<EAD>, <EAC-CPF> Eventuell bifogad arkivredovisning i XML-form

Andra metadataformat, beroende på systemtyp (se ovan), exempelvis MODS för Kungliga Biblioteket

<Datafil> Flatfil, XML-fil eller PDF-fil från myndigheter eller TIFF-fil från egen skanning

<Datafil>

<Datafil>

<XML-schema> SIP:en skall innehålla scheman för alla ingående XML-filer

<Dokumentationsfil> PDF-fil från myndigheten innehållande dokumentation av leveransen

<Dokumentationsfil> PDF-fil upprättad av arkivmyndighet innehållande information om leveransen



SIP-container för E-Arkiv

METS (container) Identitet, avsändare, filidentiteter och checksummor för en eller flera SIP:ar. En SIP består av en content-METS samt refererade metadatafiler och datafiler. En SIP kan paketeras i TAR eller ZIP.

TAR TAR-fil (eller ZIP-fil) som innehåller metadatafiler och datafiler för en eller flera SIP:ar

METS (content) Objekt-ID, filidentiteter, checksummor, information om avsändare och mottagare samt referenser till metadata i andra format (mdRef eller mdWrap). **Typfält (altrecordID) anger typ av paket beroende på systemtyp.**

<PREMIS> Teknisk metadata, metadata om bevarandeprocessen

Andra metadataformat, beroende på systemtyp (se ovan), exempelvis MODS för Kungliga Biblioteket

<Datafil> Flatfil, XML-fil eller PDF-fil från myndigheter eller TIFF-fil från egen skanning

<Datafil>

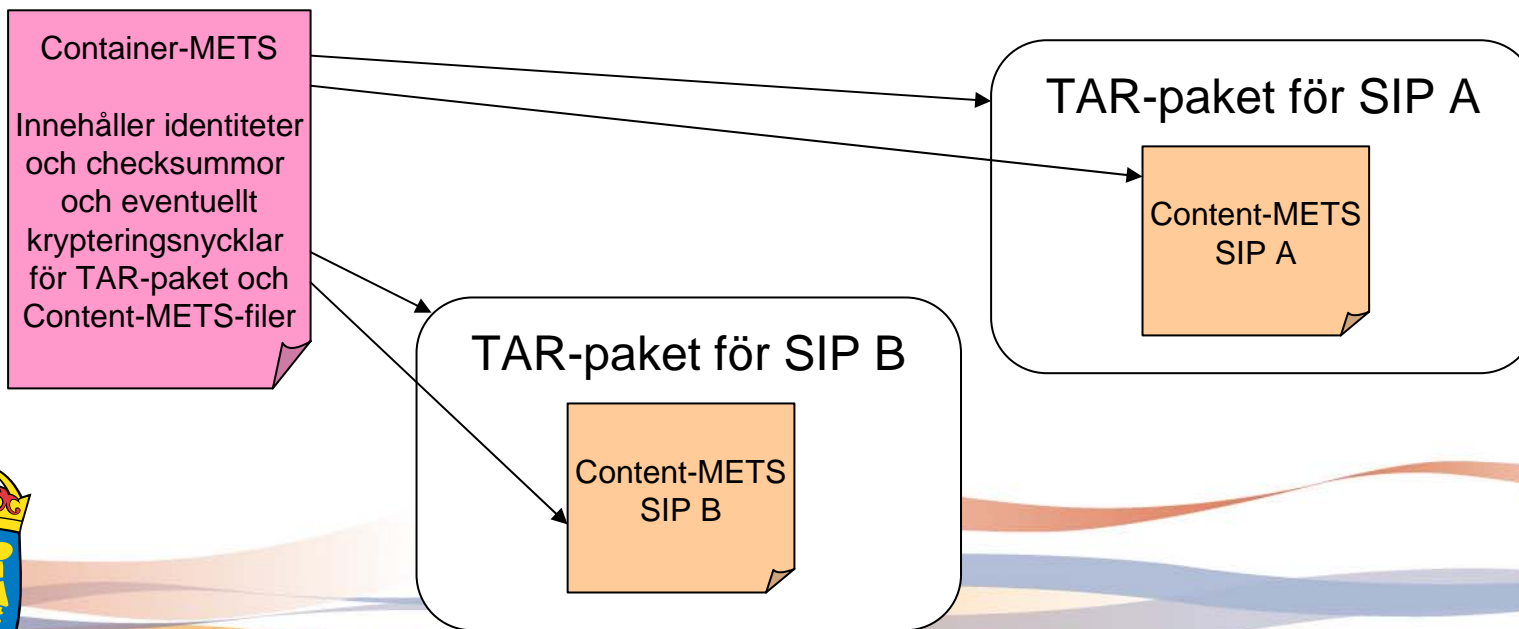
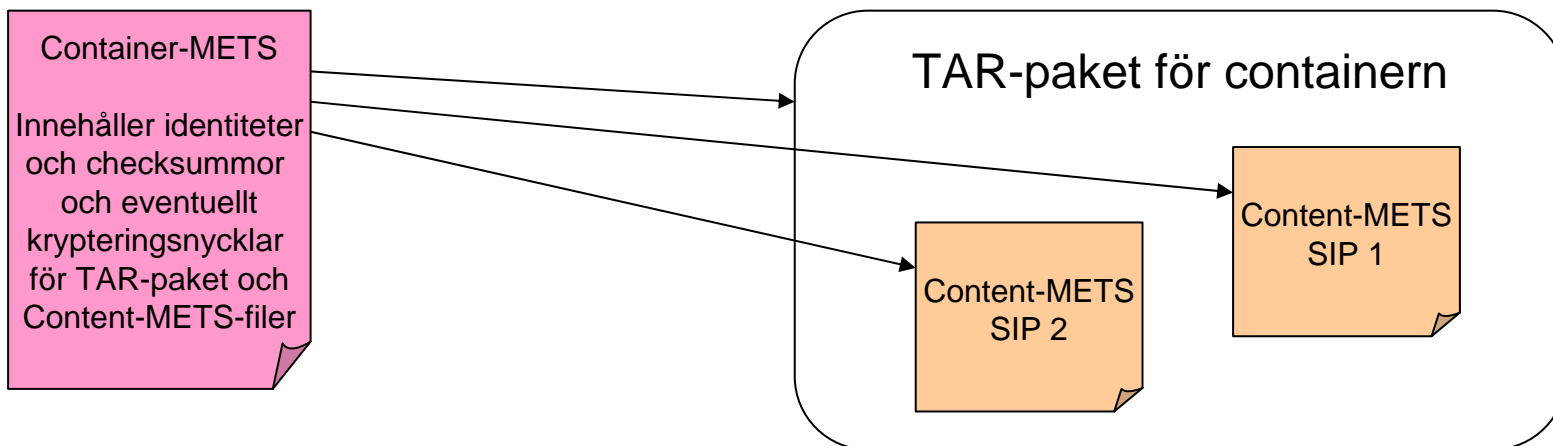
<Datafil>

<XML-schema> SIP:en skall innehålla scheman för alla ingående XML-filer

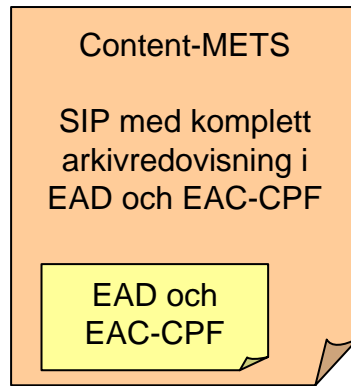
<Dokumentationsfil> PDF-fil från myndigheten innehållande dokumentation av leveransen



SIP-containers för E-Arkiv

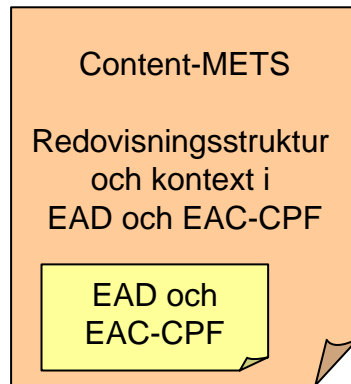
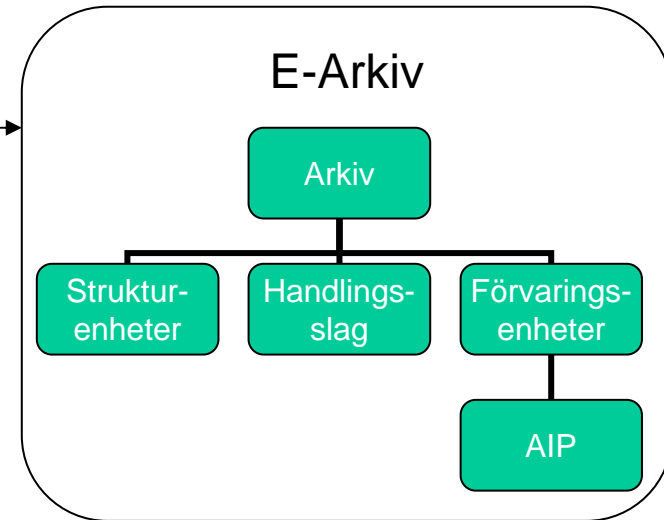


Arkivredovisning i E-Arkiv



En SIP kan innehålla en komplett arkivredovisning i EAD och EAC-CPF.

För E-Arkiv där varje SIP motsvarar ett enda ärende (sak) är det rimligare att överföra redovisningsstrukturen initialt. Därefter behöver varje SIP bara innehålla information om sig själv samt en referens till sin plats i redovisningshierarkin.

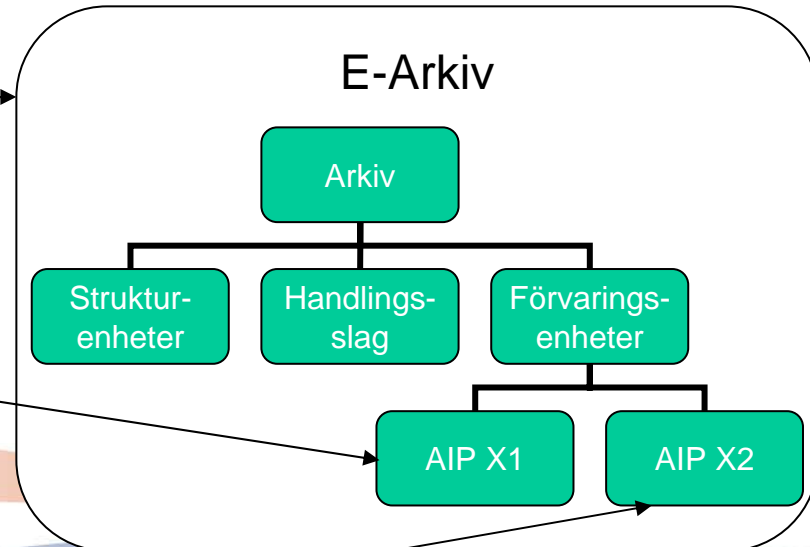


Content-METS SIP X1

Referens till redovisningen i <altrecordID>

Content-METS SIP X2

Referens till redovisningen i <altrecordID>



Riksarkivets digitala arkiv som E-tjänst (RADAR2)

