



KANSALLISARKISTO

Digitaliserat bildmaterial - skapa överföringspaket

Innehåll

1.	SYFTET MED ANVISNINGEN	1
2.	MATERIAL SOM SKA ÖVERFÖRAS OCH ÖVERFÖRINGSPOSTER: SKAPA ÖVERFÖRINGSPOSTER	2
3.	SKAPA ÖVERFÖRINGSPAKET	3
	KATALOGSTRUKTURER.....	3
	ÖVERFÖRINGSPAKET	6
1.	BESKRIVA INFORMATIONSMATERIAL.....	7



1. Syftet med anvisningen

I anvisningen beskrivs hur man skapar ett överföringspaket som ska överföras till Riksarkivet av material som digitaliserats av myndigheten själv. Anvisningen gäller material som digitaliserats i enlighet med Riksarkivets digitaliseringsspecifikationer KA/3357/07.01.02.04.02/2019 och KA/15780/07.01.02.04.02/2020 (se tabell 1). Anvisningen ska användas tillsammans med dokumenten som nämns i tabell 1.

I denna anvisning avses med Överlåtaren den organisation som är personuppgiftsansvarig/administratör för det informationsmaterial som ska överföras till Riksarkivet.

Tabell 1. Hänvisningar till andra anvisningar och dokument

Dokument	Innehåll
Anvisningen Vastaanotettavia tiedostoja koskevat vaatimukset om filer som tas emot och dess BILAGA 1 Luettelo Kansallisarkistoon vastaanotettavista tiedostomuodoista (förteckning över filformat som tas emot)	I anvisningen beskrivs de allmänna kraven på filer som tas emot av Riksarkivet, tillåtna teckenuppsättningar i filer i textformat och godkända filformat för överföring. I anvisningen specificeras i vilken överföringsstruktur varje filformat kan överföras. Anvisningen finns på Riksarkivets webbplats.
Riksarkivets krav på digitalisering som syftar till förstöring (2019) KA/3357/07.01.02.04.02/2019	Riksarkivets specifikationer för digitalisering som syftar till förstöring av det analoga uttrycket.
Riksarkivets specifikationer för digitalisering av specialmaterial som ska arkiveras (2021) KA/15780/07.01.02.04.02/2020	Riksarkivets specifikationer för digitalisering av specialmaterial som ska arkiveras.
Metadatablankett	Överlåtaren meddelar Tjänsten för digital arkivering metadata om överföringsposten på en metadatablankett. Överlåtaren kan använda metadatablanketten för att åskådliggöra materialpaketets struktur, metadata för arkivkontexten och användningsbegränsningarna som helhet. Blanketten finns på Riksarkivets webbplats.

2. Material som ska överföras och överföringsposter: Skapa överföringsposter

Det informationsmaterial som ska överföras identifieras i en **överföringsplan** mellan Riksarkivet och Överlåtaren. Planen lämnas till Riksarkivet av Överlåtaren. Det material som ska överföras enligt avtalet om överföring och överföringsplanen sammanställs i en eller flera **överföringsposter**. En enskild överföringspost är en logisk, integrerad och färdig helhet, som inte längre behöver redigeras i fråga om material eller data.

När det finns flera överföringsposter ska Överlåtaren se till att allt material som ska arkiveras i var och en av överföringsposterna faktiskt överförs.

En överföringspost indelas i en eller flera helheter som ska överföras. Dessa lagras i ett format och en katalogstruktur i enlighet med anvisningarna och komprimeras till filer för överföring (**överföringspaket**), se kapitel 3.

Riksarkivet tillhandahåller inga verktyg för att bilda överföringspaket, utan Överlåtaren skapar överföringspaketet med de verktyg som hen förfogar över eller tar fram den lösning som behövs med sin programleverantör/tjänsteleverantör eller någon annan teknisk partner.



Bild 1 Det informationsmaterial som ska överföras utgör en eller flera logiska överföringsposter som sammanställs till ett eller flera överföringspaket.

Obs!

- För att mottagningen av informationsmaterial ska lyckas ska ett överföringspaket skapas enligt anvisningen.
- Paketet får inte innehålla något annat än de kataloger med tillhörande filer som anges i anvisningen.
- Läs anvisningen omsorgsfullt så slipper du onödigt arbete.
- Du kan kontakta Tjänsten för digital arkivering med frågor i anslutning till tolkningen av anvisningen.

3. Skapa överföringspaket

Katalogstrukturer

De filer som produceras i digitaliseringsprocessen enligt digitaliseringsspecifikationen ska lagras i katalogstrukturen enligt figur 2. En förteckning över de detaljerade metadatakraven och övriga krav på material som digitaliserats i enlighet med Riksarkivets specifikationer finns i Riksarkivets digitaliseringskrav (se tabell 1).

Rotkatalogen namnges med den unika identifikationskod som Överlåtaren har gett överföringspaketet. Identifikationskoden används också i överföringsgränssnittet i samband med att kontextmetadata skapas för överföringspaketet. Överföringspaketet kopplas till tillhörande metadata med hjälp av identifikationskoden. Om den unika identifikationskod som överföringspaketet har fått och den identifikationskod som fastställts i överföringsgränssnittet inte är identiska går det inte att överföra informationsmaterialet i gränssnittet.

Det informationsmaterial som ska överföras (master), tekniska metadata för bildfiler (mix) och eventuella filer som skapats med textigenkänning (ocr) lagras i sina egna **underkataloger**. Överföringspaketet får innehålla endast kataloger enligt den katalogstruktur som beskrivs i den tillämpade digitaliseringsspecifikationen. Därför kan en ocr-katalog endast inkluderas i överföringspaketet skapade enligt digitaliseringsspecifikationerna från 2019 (KA/3357/07.01.02.04.02/2019).

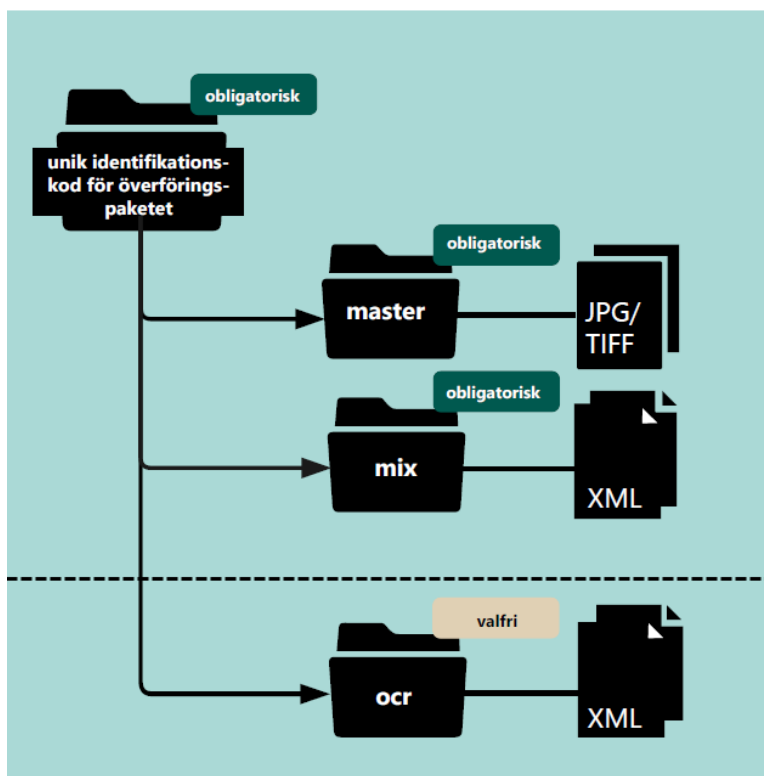


Bild 2 Katalogstruktur

En ocr-katalog kan inkluderas endast i överföringspaket skapade enligt digitaliseringsspecifikationen KA/3357/07.01.02.04.02/2019

Informationsmaterialet kopplas till tillhörande tekniska metadata och eventuella data som skapats med textigenkänning med hjälp av samma filnamn baserat på en talföljd (se figur 3). Endast filändelsen (t.ex. .jpg eller .xml) skiljer dem åt. Med andra ord ska filnamnen på filerna som finns i mix- och ocr-katalogerna motsvara namnen på bildfilerna som finns i master-katalogen. Den talföljd som motsvarar bildfilen kopplar bildfilen som skapats vid digitaliseringen och tillhörande mix-metadata samt innehåll som lästs in med textigenkänning till varandra.

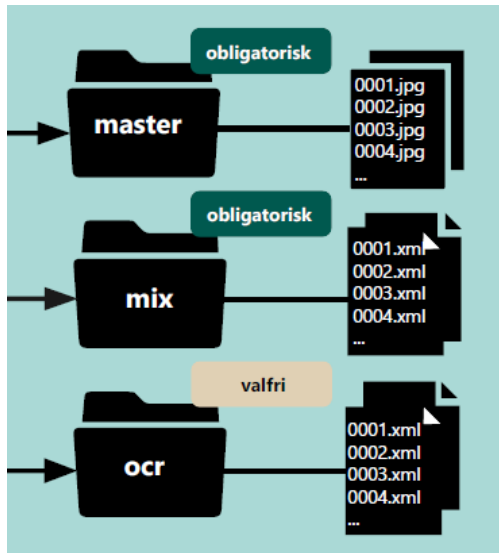


Bild 3 Exempel på namngivning

i master-katalogen finns bildfilen **0001.jpg**.

i mix-katalogen **bör** det finnas tekniska metadata för bildfilen i fråga (xml-filen **0001.xml**).

i ocr-katalogen **får** det finnas bildfilens innehåll inläst med textigenkänning (xml-filen **0001.xml**). Textigenkänning gör det möjligt att i ett senare skede göra innehållssökning i materialen i överföringspaketet.

Katalogernas namn och innehåll beskrivs i tabellerna 2 och 3. Katalogernas namn är skiftlägeskänliga. Namnen på underkatalogerna skrivs med liten begynnelsebokstav.

Tabell 2. Rotkatalogens namn och innehåll

Katalog	Beskrivning/innehåll	Gäller digitaliseringsspecifikationen
rotkatalog	Rotkatalogen ska namnges med överföringspaketets identifikationskod, t.ex. <i>paket1</i> eller <i>aren9195</i> . Överlåtaren fastställer paketets identifikationskod och använder den också i överföringsgränssnittet i samband med att kontextmetadata skapas för överföringspaketet. Det är viktigt att identifikationskoden skiljer överföringspaketet från andra paket i överföringsposten.	KA/3357/07.01.02.04.02/2019 KA/15780/07.01.02.04.02/2020



	<ul style="list-style-type: none">• Identifikationskoden får innehålla följande tecken: a–z, A–Z och 0–9.• Den metadatakod som Överlåtaren fått för att skapa kontextmetadata får inte användas som identifikationskod för överföringspaketet.	
--	---	--

Principen för namngivning av filer är alltid densamma för följande underkataloger (master, mix och ocr). Filerna i katalogerna namnges med en talföljd. I princip ska talföljden bestå av fyra siffror. Numreringen är löpande på så sätt att den första filen får namnet 0001, den andra 0002 osv. Underkatalogerna får namn enligt innehållet: master, mix eller ocr.

Tabell 3. Underkatalogernas namn och innehåll

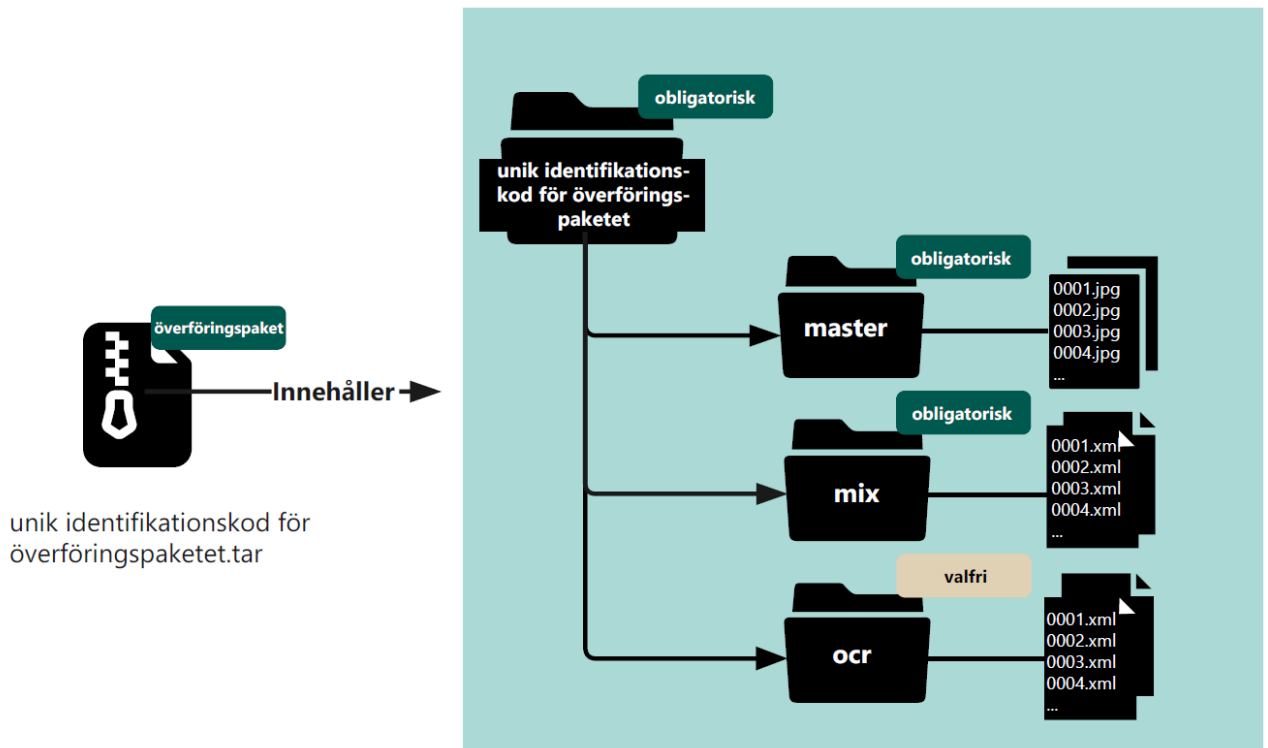
Underkatalog	Beskrivning/innehåll	Gäller digitaliseringsspecifikationen
master	<p>En underkatalog innehåller bildfiler i form av enskilda filer.</p> <ul style="list-style-type: none">• Filerna namnges med en talföljd (0001, 0002, 0003, 0004).• Filnamnen på filerna som finns i mix- och ocr-katalogerna ska motsvara namnen på bildfilerna som finns i master-katalogen.	KA/3357/07.01.02.04.02/2019 KA/15780/07.01.02.04.02/2020
mix	<p>Underkatalogen innehåller obligatoriska tekniska metadata för alla bildfiler i master-katalogen.</p> <ul style="list-style-type: none">• Filerna namnges med en talföljd (0001, 0002, 0003, 0004).• Filnamnen ska motsvara namnen på bildfilerna som finns i master-katalogen.	KA/3357/07.01.02.04.02/2019 KA/15780/07.01.02.04.02/2020

Underkatalog	Beskrivning/innehåll	Gäller digitaliserings-specifikationen
ocr	<p>Underkatalogen innehåller AltoXML- filer med eventuell text (ocr eller htr) som lästs in med textigenkänning från bildfilerna.</p> <ul style="list-style-type: none">• Filerna namnges med en talföljd (0001, 0002, 0003, 0004).• Filnamnen ska motsvara namnen på bildfilerna som finns i master-katalogen.	KA/3357/07.01.02.04.02/2019

Överföringspaket

Rotkatalogen namnges med överföringspaketets identifikationskod och **komprimeras** till ett överföringspaket, dvs. till en TAR-fil. TAR-filen får dessutom **komprimera till ett förlustfritt** GZIP-format (.tar.gz) eller BZIP2-format (.tar.bz2).

När överföringspaketet laddas upp i överföringsgränssnittet styr identifikationskoden som är paketets namn materialet till rätt beskrivningshelhet i Riksarkivets metadatalager.





Obs! Före överföring är det lämpligt att spara överföringspaketen, dvs. TAR-filerna, så att de överföringspaket som ingår i samma serie (eller en annan större helhet) finns i samma katalog. Katalogstrukturer med fler nivåer än detta kan göra sändningen av överföringspaketen betydligt långsammare, när användaren försöker välja flera paket som ska överföras samtidigt.

4. Beskriva informationsmaterial

Arkivering av informationsmaterial som ska överföras kräver alltid en särskild beskrivning. Med hjälp av beskrivningen säkerställs att informationsmaterialet bevaras i tillgänglig och begriplig form, även om användaren inte känner till materialet eller dess ursprungliga användningssyfte och historia.

Överföringspaketen eller arkivenheterna beskrivs i överföringsgränssnittet före mottagningen. Överlåtaren skickar också det informationsmaterialpaket som ska överföras och metadata som beskriver de tillhörande överföringsposternas arkivkontext, användningsbegränsningar och aktörer med anknytning till materialet till Tjänsten för digital arkivering på en separat metadatablankett.