

RIKSARKIVETS SPECIFIKATIONER FÖR DIGITALISERING AV SPECIALMATERIAL SOM SKA ARKIVERAS

- Innehåll** Riksarkivets specifikationer för digitalisering av specialmaterial som ska arkiveras.
- Syftet** Den här anvisningen definierar digitaliseringsprocessen och det slutresultat som säkerställer bevarandet och användbarheten av informationsinnehållet i det arkiverade specialmaterialet i digitalt format.
- Målgrupp** Anvisningen är avsedd för Riksarkivet samt de aktörer som utför digitalisering av analoga handlingar som ska arkiveras med stöd av arkivlagen (831/1994).
- Avgränsningar** Anvisningen gäller inte skanning som utförs i samband med att handlingen utarbetas eller ärendet behandlas eller när det gäller konvertering till digitalt format av handlingar som anländer till organisationen i analogt format. Specialmaterialet omfattar kartsamlingar, ritningssamlingar, fotografisamlingar, negativ, diabilder, mikrofilmer samt band, där informationsinnehållet i banden presenteras per uppslag (Tabell 1). Anvisningen gäller inte enskilda handlingar som klassas som specialmaterial och som har arkiverats bland handlingar i A4- eller foliestorlek.

Tabell 1, Tillämpning av anvisningen på olika slag av handlingar.

Material	Riksarkivets specifikationer för digitalisering av specialmaterial som ska arkiveras (2021)	Riksarkivets specifikationer för digitalisering som syftar till förstöring av det analoga uttrycket (2019)
Kartsamling	x	
Ritningssamling	x	
Fotografisamlingar	x	
Glasnegativ, andra negativ och dior	x	
Mikrofilmer	x	
Band där informationsinnehållet i banden presenteras per uppslag	x	
Band		x
Handlingar i A4- och foliestorlek		x
Enskilda fotografier, ritningar och kartor bland handlingar i A4- och foliestorlek		x



Slutresultatet av digitaliseringsprocessen enligt denna anvisning är inte ett arkivpaket för långtidsbevarande som lagras i systemet för långtidsbevarande. Det arkivpaket för långtidsbevarande som ska lagras kan dock bildas utgående från material som digitaliserats i enlighet med anvisningen. Den här anvisningen tar inte ställning till hur det analoga uttrycket ska förstöras, utan den beskriver de kriterier som ska uppfyllas för att det ska vara möjligt att förstöra den analoga handlingen.

Författning som ligger till grund för befogenheten att ge anvisningen

Lag om Riksarkivet 1145/2016 2 §

Giltighetstid

Tills vidare, från anvisningens datum.

Upphäver anvisningen och rekommendationen

Arkiverkets rekommendation för kvalitetskriterier vid digitalisering 24.8.2016 och 15.8.2008, AL/11130/07.01.02.04.02/2008.

Innehåll

1	Termer och begrepp	1
2	Inledning	3
3	Allmänna rekommendationer om digitaliseringsprocessen och god praxis vid digitalisering	4
4	De allmänna specifikationerna för digitaliseringsprocessen	5
5	Definitioner för respektive typ av material	5
5.1	Kart- och ritningsamlingar	6
5.2	Band vars informationsinnehåll presenteras per uppslag	6
5.3	Fotografisamlingar	8
5.4	Glasnegativ, andra negativ och dior	8
5.5	Mikrofilmer	9
6	Godkända format	10
6.1	Bildfil	10
6.2	Metadata och struktur som beskriver bildfilen och behandlingen av bildfilen	14
7	Leveranspaket som bildas genom digitaliseringsprocessen	16
8	Underskrifter	17

1 Termer och begrepp

Den terminologi som används i anvisningen baseras på definitionerna [RFC 2119] som sammanställts av Internet Engineering Task Force.¹ I tabellen nedan (Tabell 2) beskrivs översättningarna av termerna som används i denna anvisning.

Tabell 2, De översättningar som används i detta dokument

ENGELSKA	SVENSKA
must	ska
must not	får inte
required	obligatorisk
should	bör
should not	bör inte
may	får
optional	valfri

I tabellen nedan (Tabell 3) presenteras de begrepp som används i dokumentet inklusive förklaringar:

Tabell 3, Begrepp

BEGREPP	FÖRKLARING
Analog handling	En handling som uppgjorts, förvarats och används på papper eller i något annat konkret format.
Analogt uttryck	Den analoga formen av en helhet av analoga handlingar som man beslutat digitalisera.
Det analoga bevarandeformatets kulturhistoriska värde	Den arkiverade analoga handlingens (dokumentinformationens) egenskap, kulturhistoriska värde, på grund av vilken den analoga handlingen trots digitaliseringen även förvaras i analog form.
Arkivering	Arkivering sker efter förvaringstidens utgång (21 § i informationshanteringslagen 906/2019). Arkivering avser även varaktig förvaring (8 § 3 mom.) enligt arkivlagen.
Arkiveringsgrund	Arkiveringsgrunden bestämmer den inbördes ordningen mellan analoga handlingar inom en enskild serie.
Handling	En handling är beviskraftig information. Aktören har producerat eller tagit emot denna information som en del av sina uppgifter och förvarar informationen både som datalager och verksamhetsdokumentation. En handling är en helhet av innehåll, struktur/uttryck och kontext. Metadata ingår i eller anknyter till dokumentinformationen. (SFS-ISO 15489-1: Dokumentation – Hantering av verksamhetsinformation. Del 1. Grunder och principer). I lagstiftningen avses med handling utom en framställning i skrift eller bild även ett meddelande som avser ett visst objekt eller

¹ <https://www.ietf.org/rfc/rfc2119.txt> Hämtad 11.1.2021

ANVISNING

19.1.2021

RA/15780/07.01.01.03.01/2020

BEGREPP	FÖRKLARING
	ärende och uttrycks i form av tecken som på grund av användningen är avsedda att höra samman och vilket kan uppfattas endast med hjälp av automatisk databehandling eller en ljud- eller bildåtergivningsanordning eller något annat hjälpmedel. (5 § i offentlighetslagen 621/1999).
Digitalt uttryck	Den digitala formen av en helhet av analoga handlingar som man beslutat digitalisera.
Digitalisering	Handlingar som förvaras eller arkiveras i analogt format omvandlas till digitalt format. En större grupp funktioner, inom vilka skanning är ett steg.
Digitaliseringsapparat	Enhet med vilken en analog handling konverteras till digitalt format.
Digitaliseringsprocess	Ett antal funktioner med vilka ett analogt uttryck konverteras till ett digitalt uttryck.
Digitaliserad handling	En digital version av en analog handling som producerats genom digitaliseringsprocessen. Kan bestå av n stycken digitala filer.
Specialmaterial	Utgörs av kartor, ritningar, fotografier, negativ, mikrofilmer och mikrokort samt band, där bandens informationsinnehåll presenteras per uppslag, som har arkiverats som egna helheter/samlingar.
Förstöring av handlingar	Fysisk (inkl. teknisk) förstöring av en handling som ska förvaras en viss tid samt förstöring av den analoga formen, dvs. uttrycket av den handling som ska arkiveras.
Bildyta	Den information som ses när man visuellt betraktar en bildfil.
Bildfil	En bitmap-bild som producerats genom digitaliseringsprocessen. Specificerar formatet på en digital fil till en bitmap-bild.
Bruksexemplar²	En fil som tillhandahålls för användning via olika gränssnitt. Jämför arkivexemplar.
Varaktig förvaring	Med varaktig förvaring avses vid behandlingen av personuppgifter att uppgifterna förvaras för sitt ursprungliga användningsändamål så att de inte förstörs eller överhuvudtaget arkiveras i egentlig mening. För förvaringen ska då finnas en varaktig grund som anger intressen, rättigheter eller skyldigheter. (Informationshanteringslagen, RP 284/2018 rd, s. 104). I arkivlagen avses med termen varaktig förvaring arkivering enligt EU:s dataskyddsförordning (EU) 2016/679 och informationshanteringslagen.
Huvudläsriktning	Gör det möjligt att tolka handlingens informationsinnehåll utan att vända bildfilen. Om det finns informationsinnehåll med flera läsriktningar i handlingen, avses med huvudläsriktningen den riktning i vilken största delen av handlingens informationsinnehåll kan läsas.
Gallringsbeslut	Riksarkivets beslut om handlingar som med stöd av 8, 11 och 14a § i arkivlagen ska förvaras varaktigt (arkiveras) och om bevarandeformatet för dem.

² FADGI, användningsexemplar: <http://www.digitizationguidelines.gov/term.php?term=derivativefile> Hämtad 11.1.2021

BEGREPP	FÖRKLARING
Leveranspaket	En helhet som bildas i digitaliseringsprocessen och som kan överföras till Riksarkivets datasystem.
Skanning	I denna anvisning behandlas skanning som en del av digitaliseringsprocessen. Skanningen kan göras antingen med kamera eller med skanner. På grund av läsbarheten används termen skanning i denna anvisning.
Arkivexemplar³	En fil (bitmap-bild) som producerats genom digitaliseringsprocessen och vars tekniska karakteristika gör den till den mest högklassiga bildfil som producerats vid digitalisering. Arkivexemplaret överförs i regel till systemen för långtidsbevarande.
Fil	Bildar antingen ensam eller tillsammans med andra filer en digitaliserad handling. Med fil avses i det här dokumentet antingen en bitmap-bild eller en XML-fil.
Produktionsdygn	Dygn under vilket digitala filer produceras med utrustningen.

2 Inledning

Digitaliseringen av analoga handlingar till digitalt format bidrar till digitaliseringen och bevarandet och tillgängligheten av kulturarvsmaterial. De specifikationer som beskrivs i denna anvisning gäller digitaliseringsprocessen för specialmaterial. De angivna specifikationerna **ska** iakttas när en aktör inom den offentliga förvaltningen digitaliserar analoga specialmaterial som ska arkiveras. Anvisningen **ska** följas även vid digitalisering av specialmaterial, vars analoga bevarandeformat är kulturhistoriskt värdefulla.

Handlingar som är föremål för digitalisering **ska** ha Riksarkivets gallringsbeslut innan digitaliseringen inleds. I gallringsbeslutet fastställs de analoga handlingarnas kulturhistoriska värde och bevarandeformatet efter digitaliseringen. Gallringsbeslutet fastslår huruvida det analoga bevarandeformatet har ett kulturhistoriskt värde som är en grund för att det inte får förstöras efter digitaliseringen. Om inte ett sådant kulturhistoriskt värde föreligger och om det digitala uttryckets kvalitet, hållbarhet och användbarhet har säkerställts, kan originalhandlingarna förstöras. Analoga handlingar **får inte** förstöras innan ett gallringsbeslut har fattats.

För att informationsinnehållet i analoga handlingar ska kunna arkiveras enbart i digitalt format förutsätts att konverteringen till digitalt format har utförts med metoder som inte försämrar handlingens bevisvärde, integritet och autenticitet. Iakttagandet av denna anvisning bidrar till att i digitalt format säkerställa bevarandet av informationsinnehållet i analoga handlingar som digitaliseras.

Mottagandet av digitala uttryck i Riksarkivets datasystem förutsätter att de digitala uttrycken uppfyller de specifikationer som anges i denna anvisning. Vid uppgörandet av denna anvisning har de allmänna standarderna som används inom arkivsektorn och riksarkivens övriga kvalitetskrav på digitaliseringen beaktats. Dessutom har det i kapitlen 6 Godkända format och 7 Leveranspaket

³ FADGI, arkivexemplar: <http://www.digitizationguidelines.gov/term.php?term=productionmasterfile> Hämtad 11.1.2021

som bildas genom digitaliseringsprocessen beaktats de specifikationer som de nationella långtidsbevarandetjänsterna (PAS-tjänsterna) ställt upp för de material som ska förvaras.⁴

3 Allmänna rekommendationer om digitaliseringsprocessen och god praxis vid digitalisering

I detta kapitel beskrivs allmänna bästa praxis. I det här kapitlet används inte termerna i tabell 2.

Digitaliseringen uppfattas på det allmänna planet som en process som omfattar de skeden som presenteras på bilden nedan (bild 1).

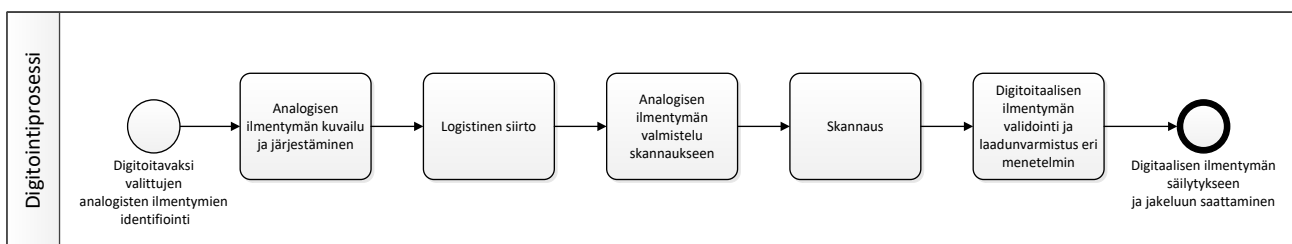


Bild 1, Process för konvertering till digitalt format - allmänt.

Kvalitetssäkringen i digitaliseringsprocessen kan inte endast omfatta ett visst skede av processen, utan kvalitetssäkring ska göras i de olika skedena av processen. I den här anvisningen ligger fokus särskilt på kvalitetssäkring i anslutning till skanningen.

De handlingar som ska digitaliseras bör gärna beskrivas i systemet för hanteringen av metadata innan konverteringen till digitalt format påbörjas. Då kan behandlingskedjan för det analoga uttrycket och till exempel uppgifter om beredningen av materialet dokumenteras på ett mer heltäckande sätt. Efter skanningen kan metadatan om materialet kompletteras antingen för hand eller med automatiska metoder.

Kvalitetssäkringen av skanningen kan grovt delas in i det som görs före skanningen och till kvalitetssäkringen efter skanningen, dvs. validering, där man med visuell granskning strävar efter att upptäcka kvalitetsavvikelser.

Skanningsinfrastrukturens prestanda optimeras före skanningsåtgärden på så sätt att det digitala uttrycket som är resultatet av åtgärden är av den bästa möjliga kvalitet som kan produceras med den tekniska sammansättningen i fråga. Efter optimeringen bör infrastrukturens prestationsförmåga kontrolleras planmässigt så att kvaliteten på de digitala uttryck som produceras i processen förblir jämn. För kontrollen behövs i allmänhet en måttavla, referensvärden för måttavlan och en analysprogramvara. Utöver bildkvaliteten bör man i utrustningsinfrastrukturen säkerställa att de analoga handlingarna konverteras till digitalt format med komplett informationsinnehåll. I skanningsprocessen ska man undvika att till exempel två överlappande analoga handlingar går genom apparaten (skanner med genommatning eller med öppen bana och andra skanningslösningar där handlingarna skannas in som "massa"). Den bildfil som infrastrukturen producerar får inte heller vara förvrängd eller annars ha ändrade proportioner, så att inte tolkningen av informationsinnehållet äventyras.

⁴ <http://digitalpreservation.fi/files/PAS-tiedostomuodot-1.8.0.pdf> Hämtad 11.1.2021

Valideringen efter skanningen kan utföras med urval. Urvalsstorleken beror på skanningsprocessens pålitlighet. Syftet med valideringen är att säkerställa att de krav som anges i den här anvisningen uppfylls.

Om bildfilerna behandlas efter skanningen, bör en detaljerad bildhanteringshistorik sparas åtminstone i bildfilernas metadata, om möjligt även i de XML-uppgifter (MIX) som beskriver bildfilens uppkomst.

Det bör även observeras att varje migration och filkonversion (förändrat filformat) medför risker för informationens integritet. Därför är det viktigt att digitalisera det analoga materialet till det lagringsformat som anges i den här anvisningen. Det är inte möjligt att av ett lagringsformat av sämre kvalitet producera en version av bättre kvalitet vid senare behandling. Det är också förbjudet enligt kraven i denna anvisning.

4 De allmänna specifikationerna för digitaliseringsprocessen

Faserna i konverteringen av analoga uttryck till digital form (digitaliseringsprocessen) **ska** dokumenteras på det sätt och med den noggrannhet som beskrivs i denna anvisning. Med dokumentering av processen avses lagring av metadata som dokumenterar skanningen samt eventuell behandling av bilderna. Man **får** lagra andra åtgärder som metadata i processen.

I digitaliseringsprocessen **ska** man vara säker på att den helhet som ska digitaliseras blir digitaliserad i sin helhet, enligt arkiveringsgrunden och innehållsmässigt komplett. Detta innebär att alla analoga handlingar i helheten som ska digitaliseras **ska** digitaliseras så att ingen information förblir okonverterad på grund av problem med tekniken eller funktionerna.

Alla enskilda bildfiler i en helhet som ska digitaliseras **ska** med en visuell granskning innehålla samma information som den analoga motsvarigheten. En bildfil **får inte** innehålla element som inte finns i den analoga motsvarigheten. Ett undantag till detta är eventuella skanningstekniska måttavlor som skannas eller fotograferas till samma bildfil och som verifierar bildfilens färger, gråtoner, proportioner och upplösning. Dessa måttavlor **ska** placeras så att de inte täcker det objekt som digitaliseras.

Sidor som innehåller anteckningar **får inte** avlägsnas i digitaliseringsprocessen. De bildfiler som produceras i digitaliseringsprocessen **ska** vara vända i huvudläsriktningen. Bildfiler som produceras i digitaliseringsprocessen **får** vändas endast i steg om 90 grader efter att de skannats in.

5 Definitioner för respektive typ av material

I det här kapitlet beskrivs kraven för digitaliseringsprocesserna för respektive typ av material och för de arkivexemplar som produceras i processen. Allmänna krav för digitala filer presenteras i kapitel 6. Materialtyperna framgår av tabell 1. I kapitel 5 har materialtyperna kombinerats, om de krav som hänför sig till materialtyperna är sinsemellan likadana.

5.1 Kart- och ritningssamlingar

Kart- och ritningssamlingar som utgör separata samlingar **ska** digitaliseras till bildfiler så att den bildfil som skapas till sina fysiska mått (n cm x n cm) och den analoga handlingen har storleksförhållandet 1:1 med upplösningen 300 ppi. För att säkerställa integriteten för informationsinnehållet i analoga handlingar som digitaliseras **ska** den analoga handlingen digitaliseras så att bildytan överskrider kanterna på sidorna. Handlingen **ska** vara slät vid avbildningsögonblicket. Om glas eller motsvarande hjälpredskap används för att göra handlingen slät **ska** detta beaktas vid färghanteringen. Kartor och ritningar digitaliseras till färgfiler (RGB, 24 bit).

Före skanningen **ska** prestandan hos den infrastruktur som används vid digitaliseringen optimeras med måttavlor och program som är lämpliga för detta syfte. Efter optimeringen **bör** kvaliteten på de bildfiler som infrastrukturen producerar verifieras en gång per produktionsdygn med hjälp av måttavlor för detta syfte.

Av tabellen nedan (tabell 4) framgår kraven för digitalisering av kart- och ritningssamlingar.

Tabell 4, Krav för kart- och ritningssamlingar

Element	Specificerat värde
Arkivexemplarets format	TIFF 6.0
Arkivexemplarets kompression	Okomprimerad eller LZW
Upplösning	300 ppi
Färgrymd	2 = RGB
Bitdjup	8 bitar per kanal, 3 kanaler (24 bit)
Arkivexemplarets ICC-profil	eciRGB v2, ProPhoto RGB, AdobeRGB (1998)

5.2 Band vars informationsinnehåll presenteras per uppslag

Om informationsinnehållet i band presenteras per uppslag **ska** man vid digitalisering av banden tillämpa de krav som anges i detta kapitel. Sådant innehåll kan vara till exempel tabeller som överskrider en sida eller något annat informationsinnehåll som ursprungligen har producerats i en analog handling på ett sådant sätt att de enskilda sidorna av den analoga handlingen måste granskas parallellt, dvs. per uppslag (Bild 2). För att säkerställa informationsinnehållets integritet **ska** analoga handlingar digitaliseras i sin helhet så att bildytan överskrider kanterna på sidorna.

19.1.2021

RA/15780/07.01.01.03.01/2020

Aukeama												
Sivu					Sivu							
<i>Turun ja Pöytä</i>					<i>laissa vuosi 1907</i>					8		
Periätökirja		Kirkkunta	Pitäjä	Kylä	Tilaa		Periätönsatukinta					
N:o	Astajin				N:o	Talokki	Mastall	Alta jätin se on tilin re- husten aari- tella	Koko peri- tönsatukin- ta	1/2 osa pe- riätönsatukin- ta	1/4 osa pe- riätönsatukin- ta	
1907					1907							
1.	<i>Huhtala</i>	<i>Machon</i>	<i>Siedon</i>	<i>Ingvisten</i>	2.	1.	13	<i>Ingvisten</i>	<i>Talokki</i>	<i>Kaarle Tuoma</i>	<i>Laakso</i>	93.58
2.	<i>Harjo</i>	<i>Thorsen</i>	<i>Lemlandin</i>	<i>Teterängän</i>	3.	1/4	67	<i>Storgård</i>	<i>Sten Karl Johan</i>	<i>Eskelinen</i>	<i>Laakso</i>	3.78

Bild 2, Informationsinnehållet visas per uppslag

Av tabellen nedan (tabell 5) framgår kraven för digitalisering av band som digitaliseras per uppslag.

Tabell 5, Krav för band

Element	Specificerat värde
Arkivexemplarets format	TIFF 6.0 eller JPEG part 1 version
Arkivexemplarets kompression	TIFF = LZW eller JPEG = JPEG-kvalitetsnivå 90 %
Upplösning	300 ppi
Färgrymd	2 = RGB

⁵ Bildfilen är inte en kvalitativ referens, utan ett exempel på ett informationsinnehåll som är uppdelat på uppslag.

Element	Specificerat värde
Bitdjup	8 bitar per kanal, 3 kanaler (24 bit)
Arkivexemplarets ICC-profil	eciRGB v2, ProPhoto RGB, AdobeRGB (1998), sRGB

5.3 Fotografisamlingar

För att säkerställa integriteten för informationsinnehållet i fotografier **ska** analoga handlingar digitaliseras i sin helhet så att bildytan överskrider kanterna på sidorna. Fotografier **ska** digitaliseras i färg till bildfiler (RGB, 24 bit). Av tabellen nedan (Tabell 6) framgår kraven för digitalisering av fotografier.

Tabell 6, Krav för fotografisamlingar

Element	Specificerat värde
Arkivexemplarets format	TIFF 6.0
Arkivexemplarets kompression	LZW
Upplösning	Enligt det analoga uttrycket, se tabell 7 Upplösningstabell
Färgrymd	2 = RGB
Bitdjup	RGB = 8 bit kanal, 3 kanaler (24 bit)
Arkivexemplarets ICC-profil	RGB: eciRGB v2, ProPhoto RGB, AdobeRGB (1998)

När handlingen som ska digitaliseras är liten **ska** upplösningen höjas enligt tabellen nedan (tabell 7). Om den analoga handlingen till sina fysiska mått är mindre än vad som anges i tabell 7, **ska** upplösningen vid digitaliseringen ökas genom tillämpning av samma beräkningsformel. Upplösningen för arkivexemplarsfilen **får inte** ändras när bildfilen behandlas, eftersom det ska vara möjligt att uppskatta den analoga handlingens fysiska mått utifrån det digitala arkivexemplaret. Målet är alltid ett arkivexemplar i ~A3-storlek (cm x cm) med upplösningen 300 ppi.

Tabell 7, Upplösningstabell

Storlek på den analoga handlingen	Upplösning
6 x 9 cm	1 600 ppi
9 x 12 cm	1200 ppi
13 x 18 cm	900 ppi
18 x 24 cm	600 ppi
A3 (eller större)	300 ppi

5.4 Glasnegativ, andra negativ och dior

För att säkerställa integriteten för informationsinnehållet i glasnegativ, andra negativ och diabilder **ska** analoga handlingar digitaliseras i sin helhet så att bildytan överskrider kanterna på det analoga negativ eller den analoga diabild som digitaliseras. Om ett negativ eller en diabild har satts i ramar, **ska** man vid avgränsningen iaktta extrem noggrannhet för att inte informationsinnehåll avgränsas från bildfilen. Diabilder och negativ **ska** digitaliseras i färg till bildfiler (RGB, 24 bit). Av tabellen nedan (Tabell 8) framgår kraven för digitalisering av glasnegativ, andra negativ och diabilder.

Tabell 8, Krav för glasnegativ, övriga negativ och diabilder

Element	Specificerat värde
Arkivexemplarets format	TIFF 6.0
Arkivexemplarets kompression	LZW
Upplösning	Enligt det analoga uttrycket, se tabell 7 Upplösningstabell
Färgrymd	2 = RGB
Bitdjup	RGB = 8 bit kanal, 3 kanaler (24 bit)
Arkivexemplarets ICC-profil	RGB: eciRGB v2, ProPhoto RGB, AdobeRGB (1998)

Om den analoga handling som ska digitaliseras är liten, **ska** upplösningen ökas enligt upplösningstabellen (Tabell 7). Om den analoga handlingen till sina fysiska mått är mindre än det som anges i tabell 7, **ska** upplösningen vid digitaliseringen ökas genom tillämpning av samma beräkningsformel. Upplösningen för arkivexemplarsfilen **får inte** ändras när bildfilen behandlas, eftersom det ska vara möjligt att uppskatta den analoga handlingens fysiska mått utifrån det digitala arkivexemplaret. Målet är alltid ett arkivexemplar i ~A3-storlek (cm x cm) med upplösningen 300 ppi.

5.5 Mikrofilmer

Mikrofilmens förminskningsfaktor **ska** kontrolleras i filmrullens början. Om ingen förminskningsfaktor anges **ska** en tillräckligt stor förstöringsfaktor ställas in för att informationsinnehållet i den digitaliserade handlingen ska vara läsbart. I fråga om mikrofilmer kan man med hjälp av förstöringsfaktorn på filmen få ett tillräckligt läsbart material med upplösningen 300 ppi. Vid skanning ska en upplösning användas som är förstöringsfaktorn x den önskade upplösningen (till exempel $12 \times 300 = 3\,600$). Mikrofilm är i sig ett lagringsmedium med hög kontrast, och vid digitaliseringen **ska** man därför vara försiktig så att inte kontrasten accentueras och se till att kraven i kapitel 4 uppfylls.

Vid verifieringen av originalmaterialets rätta storlek har även andra metoder (till exempel måttband) kunnat användas. De använda metoderna **ska** granskas från fall till fall. Samtidigt **ska** man sträva efter att bedöma de fysiska måtten på den analoga handling som avbildats på mikrofilm, även om ingen av de ovan nämnda metoderna har använts vid mikrofilmning.

Av tabellen nedan (Tabell 9) framgår kraven för digitalisering av mikrofilmer.

Tabell 9, Krav för mikrofilmer

Element	Specificerat värde
Arkivexemplarets format	TIFF 6.0 eller JPEG part 1 version 1.00 eller 1.01 eller 1.02
Arkivexemplarets kompression	TIFF = LZW eller JPEG = JPEG-kvalitetsnivå 90 %
Upplösning	300 ppi
Färgrymd	Gråton
Bitdjup	8 bitar per kanal, 1 kanal (8 bit)
Arkivexemplarets ICC-profil	Gray gamma 2,2

Om det rör sig om en färgmikrofilm, **ska** man på den tillämpa de krav som anges i det här kapitlet, men arkivexemplarets färgrymd **ska** vara RGB, bitdjupet **ska** vara 24 bit och arkivexemplarets ICC-profil **ska** vara eciRGB v2, ProPhoto RGB, AdobeRGB (1998) eller sRGB.

6 Godkända format

Avsnittet om format har indelats i två underkategorier:

1. Bildfil
2. Metadata och struktur som beskriver bildfilen och behandlingen av bildfilen

6.1 Bildfil

Den bildfil som produceras i digitaliseringsprocessen **får inte** i något skede av behandlingen ha en kvalitet som underskrider kraven i det här kapitlet. Arkivexemplaret **ska** sparas antingen i okomprimerat/förlustfritt komprimerat TIFF-format eller i JPEG-format som medför förlust (inte i båda).⁶ Lagringsformatet för arkivexemplaret är beroende av formen för det analoga materialet. Definitionerna för respektive typ av material framgår av kapitel 5 och dess avsnitt. Om den bildfil (arkivexemplaret) som skapas först i processen är icke-förlustfritt komprimerad, **får den inte** behandlas efter den första lagringsgången och därefter sparas som arkivexemplar på nytt.

I tabellerna nedan (Tabell 10 och Tabell 11) anges de obligatoriska uppgifter som **ska** finnas i bildfilen i maskinläsbart format för varje filformat. Om kolumnen "Element" i tabellerna inte specificeras **ska** uppgiften anges, men för denna uppgift har i detta sammanhang inte specificerats ett obligatoriskt fält. Utöver de uppgifter som anges i tabellen **får** bildfilen innehålla också andra metadatafält. Metadata som definieras som personuppgifter **får inte** sparas i bildfilerna.

I tabellerna 10 och 11 anges alla de möjliga värden som **ska** uppges för alla lagringsformat för bildfiler som definierats i den här anvisningen. Med andra ord **ska** värdena i tabellerna 10 och 11 tolkas tillsammans med de krav som gäller för materialtyperna i kapitel 5 och dess avsnitt. De ovan nämnda punkterna anges i kolumnen "Specificerat värde" med värdet "Beroende av materialtyp".

Om värdet i kolumnen "Specificerat värde" är "Bildfilsspecifikt", har varje bildfil ett individuellt värde. Om värdet är "Beroende av produktionsprocessen" i kolumnen är värdet beroende av den process genom vilken de digitala bildfilerna har producerats vid respektive tidpunkt.

Tabell 10, Obligatoriska metadata för bildfil (TIFF)

Element	Identifierare	Specificerat värde	Schema för metadata	Fält för metadata
Format	TIFF	image/tiff		MIME Type
Version	Den specificerade versionen av TIFF-fil	6.0		
Bildens namn	Bildens namn anges med fyra siffror 0001.tif, 0002.tif, 0003.tif osv. Se kapitel 7.			
Bildfilens storlek				
Färgrymd	Bildfilens färgrymd	Beroende av materialtyp. 0 = WhiteIsZero eller 1 = BlackIsZero eller 2 = RGB	Exif.Image	PhotometricInterpretation (262)

⁶ JPEG: <https://jpeg.org/jpeg/index.html> Hämtad 11.1.2021

TIFF: <https://www.itu.int/itudoc/itu-t/com16/tiff-fx/docs/tiff6.pdf> Hämtad 11.1.2021

ANVISNING

19.1.2021

RA/15780/07.01.01.03.01/2020

Element	Identifierare	Specificerat värde	Schema för metadata	Fält för metadata
ICC-profil	Färgprofil som sparats (embedded) i bildfilens metadata.	Beroende av materialtyp. sRGB, eciRGB v2, ProPhoto RGB, AdobeRGB (1998) eller Gray gamma 2,2.	ICC	profileDescription
Bitdjup	Antalet bitar i pixelns kanalvärde	8	Exif.Image	BitsPerSample (258)
	Antalet kanalvärden per pixel	Beroende av materialtyp. 3 eller 1	Exif.Image	SamplesPerPixel (277)
Komprimering av filen		Beroende av materialtyp: 1 = Uncompressed eller 5 LZW	Exif.Image	Compression (259)
Bildens bredd	Anger bildens bredd i antal pixlar per rad	Bildfilsspecifik	Exif.Image	ImageWidth (256)
Bildens höjd	Anger bildens höjd med antalet pixelrader på bilden	Bildfilsspecifik	Exif.Image	ImageLength (257)
Den digitala bildfilens producent	Den organisation som skapat bildfilen av det analoga uttrycket	Beroende av produktionsprocessen	Exif.Image	Artist (315)
Digitaliseringsapparat (skanning eller fotografering)	Anger med vilken tillverkarens apparat den analoga förekomsten har konverterats till digitalt format (tillverkarens namn)	Beroende av produktionsprocessen	Exif.Image	Make (271)
Digitaliseringsapparatens modell (skanning eller fotografering)	Specificerar digitaliseringsapparatens genom att ange namnet på tillverkarens modell	Beroende av produktionsprocessen	Exif.Image	Model (272)
Serienumret för den apparat som använts vid digitaliseringen	Identifierar den använda apparaten	Beroende av produktionsprocessen	Exif.Image	CameraSerialNumber (50735)
Program som användes för att skapa bildfilen	Applikation och version med vilken den digitala filen skapades	Beroende av produktionsprocessen	Exif.Image	Software (305)
Datum och klockslag när bildfilen skapades (skanningsdatum)	Anges i formatet: YYYY:MM:DDTHH:MM:SS	Bildfilsspecifik	Exif.Image	DateTimeOriginal (36867)
Läsriktning	Läsriktning för filen (liggande eller stående). Läsriktningen tar inte ställning till läsriktningen för informationsinnehållet i bilden, utan den anger i vilken riktning filen är.	Beroende av produktionsprocessen/ Bildfilsspecifik	Exif.Image	Orientation (274)
Mätenheten för upplösning	Mätenhet för tolkning av X- och Y-upplösningen	2 = inch	Exif.Image	ResolutionUnit (296)

ANVISNING

19.1.2021

RA/15780/07.01.01.03.01/2020

Element	Identifierare	Specificerat värde	Schema för metadata	Fält för metadata
XUpplösning	Antalet pixlar per upplösningens måtenhet i bildens breddriktning.	Beroende av materialtyp	Exif.Image	XResolution (282)
YUpplösning	Antalet pixlar per upplösningens måtenhet i bildens höjdriktning.	Beroende av materialtyp	Exif.Image	YResolution (283)
Byteordning		big endian eller little endian		ByteOrder
Program för behandling av bildfilen	Om man i digitaliseringsprocessen först skapar en okomprimerad fil som behandlas efter skanningen, sparas behandlingsprogrammets namn och version.	Beroende av produktionsprocessen	Exif.Image	Image.ProcessingSoftware (11)

Tabell 11, Obligatoriska metadata för bildfil (JPEG)

Element	Identifierare	Specificerat värde/förklaring	Schema för metadata	Fält för metadata
Format	JPEG	image/jpeg		MIME Type
Version	JPEG part 1 version	1.00 eller 1.01 eller 1.02		JFIF Version
Bildens namn	Bildens namn anges med fyra siffror 0001.jpg, 0002.jpg, 0003.jpg osv. Se kapitel 7.			
Bildfilens storlek				
Färgrymd	Bildfilens färgrymd	Beroende av materialtyp. 0 = WhiteIsZero eller 1 = BlackIsZero eller 2 = RGB	Exif.Image	PhotometricInterpretation (262)
ICC-profil	Färgprofil som sparats (embedded) i bildfilens metadata.	Beroende av materialtyp. sRGB, eciRGB v2, ProPhoto RGB, AdobeRGB (1998) eller Gray gamma 2,2.	ICC	profileDescription
Bitdjup	Antalet bitar i pixelns kanalvärde	8	Exif.Image	BitsPerSample (258)
	Antalet kanalvärden per pixel	Beroende av materialtyp. 3 eller 1	Exif.Image	SamplesPerPixel (277)
Komprimering av filen	JPEG	6 = JPEG (old-style) eller 7 = JPEG	Exif.Image.	Compression (259)
JPEG-kvalitet	JPEG-komprimeringens kvalitet på skalan 0–100 %	90 %		

ANVISNING

19.1.2021

RA/15780/07.01.01.03.01/2020

Element	Identifierare	Specificerat värde/förklaring	Schema för metadata	Fält för metadata
Bildfilens producent	Den organisation som skapat bildfilen av det analoga uttrycket	Beroende av produktionsprocessen	Exif.Image	Artist (315)
Bildens höjd	Anger bildens höjd med antalet pixelrader på bilden	Bildfilsspecifik	Exif.Image	ImageLength(257)
Bildens bredd	Anger bildens bredd med antalet pixelrader på bilden	Bildfilsspecifik	Exif.Image	ImageWidth(256)
Digitaliseringsapparat	Anger med vilken tillverkarens apparat den analoga handlingen har konverterats till digitalt format (tillverkarens namn)	Beroende av produktionsprocessen	Exif.Image	Make (271)
Modell på digitaliseringsapparaten	Specificerar digitaliseringsapparaten genom att ange namnet på tillverkarens modell	Beroende av produktionsprocessen	Exif.Image	Model (272)
Serienumret för den apparat som använts vid digitaliseringen	Specificerar modellen och individualiserar apparaten med vilken den analoga handlingen har konverterats till digitalt format	Beroende av produktionsprocessen	Exif.Image	CameraSerialNumber (50735)
Program som användes för att skapa bildfilen	Applikation och version med vilken den digitala filen skapades	Beroende av produktionsprocessen	Exif.Image	Software (305)
Datum och klockslag när bildfilen skapades	Anges i formatet: YYYY:MM:DD HH:MM:SS	Bildfilsspecifik	Exif.Image	DateTimeOriginal (36867)
Läsriktning	Filens läsriktning (horisontell eller vertikal)	Beroende av produktionsprocessen/ Bildfilsspecifik	Exif.Image	Orientation (274)
Mätenheten för upplösning	Mätenhet för tolkning av X- och Y-upplösningen	2 = inch	Exif.Image	Image.ResolutionUnit (296)
XUpplösning	Antalet pixlar per upplösningens måtenhet i bildens breddriktning	Beroende av materialtyp/ Bildfilsspecifik	Exif.Image	Image.XResolution (282)
YUpplösning	Antalet pixlar per upplösningens måtenhet i bildens höjdriktning	Beroende av materialtyp/ Bildfilsspecifik	Exif.Image	Image.YResolution (283)
Program för behandling av bildfilen	Om man i digitaliseringsprocessen först skapar en okomprimerad fil som behandlas efter skanningen, sparas behandlingsprogrammets namn och version.	Beroende av produktionsprocessen	Exif.Image	Image.ProcessingSoftware (11)

6.2 Metadata och struktur som beskriver bildfilen och behandlingen av bildfilen

De metadata som beskrivs i det här kapitlet beskriver bildfilens uppkomst. Detta bidrar till att verifiera autenticiteten hos det digitala uttrycket som uppkommit i processen. Bildfilernas obligatoriska tekniska metadata **ska** presenteras i enlighet med metadataformatet MIX version 2.0.⁷

I tabellen nedan (Tabell 12) anges MIX-fältets namn, syftet med fältet i fri översättning samt skyldigheten. Uppgifterna presenteras från vänster till höger. I fältet för skyldighet meddelas huruvida fältet och uppgiften i enlighet med schemat är obligatoriska på följande sätt:

- O = Obligatorisk – den här uppgiften **ska** beskrivas
- F = Frivillig – den här uppgiften **bör** beskrivas, men det är inte obligatoriskt

I metadataformatet MIX finns fält av två slag: behållare och dataelement. Dataelementen innehåller ett visst värde, medan behållarna innehåller ett eller flera dataelement och de kan innehålla andra behållare jämte dataelement. I tabell 12 meddelas endast fält, dvs. dataelement, som innehåller ett visst värde.

Tabell 12, Metadata som beskriver bildfilen och dess behandling (i tabellen anges endast fält som innehåller information som SKA presenteras i strukturen enligt metadataformatet MIX version 2.0)

MIX-fältets namn	Syftet med fältet	Skyldighet
objectIdentifierType	Dataelement som anger det system eller den domän där den digitala handlingens specificerande ID är unikt.	O
objectIdentifierValue	Serie av tecken som identifierar den digitala filen.	O
fileSize	Filens storlek i byte, t.ex. 72839.	O
formatName	Filens format. Det specificerade värdet är image/jpeg eller image/tiff	O
formatVersion	Filens version, t.ex. 1.01 ELLER 6.0	F
byteOrder	Dataelement som definierar ordningsföljden då byte sparas. Det specificerade värdet är antingen big endian eller little endian.	O
compressionScheme	Den använda komprimeringen. Det specificerade värdet är JPEG (när arkivexemplaret är i jpeg-format) eller LZW (när arkivexemplaret är i TIFF-format).	O
compressionRatio	Dataelement som meddelar den använda komprimeringens nivå. Uppges med numerator "90" och denominator "100".	O (endast JPEG)
messageDigestAlgorithm	Dataelement som identifierar den algoritm med vilken värdet i fältet messageDigest har bildats. Fältet ska ha något av de följande värdena: MD5, SHA-1, SHA256, SHA384, SHA512.	O
messageDigest	Den teckenserie som den algoritm som definierar fältet messageDigestAlgorithm producerat, t.ex. e8064dc0.	O
imageWidth	Bildens bredd i pixlar, t.ex. 1330.	O

⁷ The Library of Congress » Standards » MIX. Webbplats för kongressens bibliotek <http://www.loc.gov/standards/mix/> och ANSI/NISO Z39.87-2006 (R2017) Data Dictionary - Technical Metadata for Digital Still Images <https://www.niso.org/publications/ansiniso-z3987-2006-r2017-data-dictionary-technical-metadata-digital-still-images>
Hämtad 11.1.2021

ANVISNING

19.1.2021

RA/15780/07.01.01.03.01/2020

MIX-fältets namn	Syftet med fältet	Skyldighet
imageHeight	Bildens höjd i pixlar, t.ex. 1600.	O
colorSpace	Dataelement som definierar bildens färgrymd. Det specificerade värdet är RGB eller WhitesZero eller BlacksZero.	O
iccProfileName	Dataelement som definierar den allmänt använda ICC-profilens namn. Det specificerade värdet är sRGB, eciRGB v2, ProPhoto RGB, AdobeRGB (1998) eller Gray gamma 2,2.	O
iccProfileVersion	Dataelement som anger den använda ICC-profilens version, t.ex. v4 [t.ex. sRGB v4].	O
iccProfileURL	Dataelement i vilket ICC-profilens URL/URN sparas om ICC-profilen inte är väldokumenterad.	F
dateTimeCreated	Dataelement som anger tidpunkten då bildfilen skapades. Anges i formatet: YYYY-MM-DDTHH:MM:SS.	O
imageProducer	Dataelement som identifierar den organisation som skapade den digitala filen.	O
scannerManufacturer	Dataelement som anger namnet på tillverkaren av den apparat som användes vid skanningen.	O
scannerModelName	Dataelement som anger namnet på modellen av den apparat som användes vid skanningen.	O
scannerModelNumber	Dataelement som preciserar namnet på digitaliseringsapparatens modell med dess typnummer.	O
scannerModelSerialNo	Digitaliseringsapparatens serienummer med vilken apparaten kan individualiseras.	O
scanningSoftwareName	Namnet på den skanningsprogramvara som använts.	O
scanningSoftwareVersionNo	Versionsnumret för den skanningsprogramvara som använts.	O
orientation	Dataelement som anger bildens läsriktning.	O
samplingFrequencyUnit	Dataelement som anger den måttenhet som används för att tolka X- och Y-upplösningen. Det specificerade värdet är "in."	O
xSamplingFrequency	Antalet pixlar per upplösningens måttenhet i bildens breddriktning. Det specificerade värdet är beroende av materialtypen.	O
ySamplingFrequency	Antalet pixlar per upplösningens måttenhet i bildens höjdriktning. Det specificerade värdet är beroende av materialtypen.	O
bitsPerSampleValue	Dataelement som definierar antalet bitar i varje kanal. Det specificerade värdet är 8.	O
bitsPerSampleUnit	Dataelement som definierar hur bitarna tolkas. Värdet är antingen integer eller floating point.	O
samplesPerPixel	Dataelement som definierar antalet kanalvärden per pixel. Det specificerade värdet är antingen 3 eller 1.	O
targetType	Dataelement som anger huruvida måttavlan som verifierar skanningens kvalitet är en del av bilden eller om den skannats i en separat bild.	F
targetManufacturer	Dataelement där måttavlans tillverkare antecknas.	F
targetName	Dataelement som identifierar den använda måttavlans namn.	F
targetNo	Dataelement som innehåller den använda måttavlans serienummer.	F
externalTarget	Dataelement som anger var den digitala bilden på TargetID-behållarens specificerade måttavla finns.	F
performanceData	Dataelement som anger var mätdata om TargetID-behållarens specificerade måttavla finns.	F

7 Leveranspaket som bildas genom digitaliseringsprocessen

De i kapitel 6 och dess avsnitt nämnda digitala filer som produceras i digitaliseringsprocessen **ska** sparas i en sådan katalogstruktur som beskrivs på bilden nedan (bild 3), så att de kan tas emot i Riksarkivet. Ett digitalt uttryck **ska** produceras i katalogstrukturen oberoende av när det överförs till Riksarkivet. När materialet överlämnas **får det inte** innehålla något annat material än de kataloger som anges i bild 3.

Utöver den katalogstruktur som definieras här **får** organisationen spara till exempel bruksexemplar i sina egna datasystem i den datastruktur som det aktuella systemet förutsätter. Den struktur som definieras i den här handlingen utesluter således inte användningen av eventuella andra registreringsstrukturer.

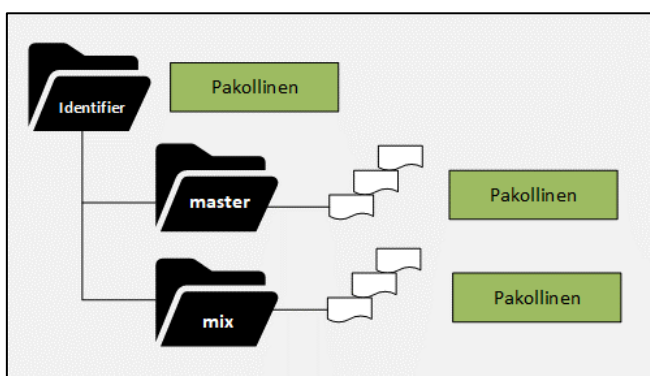


Bild 3, Den leveranspaketstruktur som krävs i digitaliseringsprocessen

Av tabellen nedan (Tabell 13) framgår hur de digitala filerna **ska** namnges inuti den presenterade katalogstrukturen. De digitala filer som producerats i processen **ska** mötas sinsemellan, med andra ord en 0001.xml-fil enligt metadataschemat MIX **ska** innehålla metadata som beskriver bildfilen 0001.jpg eller 0001.tif.

Tabell 13, Innehållet i katalogerna för leveranspaket

Mapp	Förklaring
identifier	Avser en beteckning som individualiserar ett digitalt uttryck som ska göra det möjligt att identifiera den aktuella handlingshelheten (till exempel en arkivenhet). ⁸ Katalogen innehåller "de digitala objektens kataloger".
master	De bildfiler som beskrivs i tabellerna (Tabell 10, Obligatoriska metadata för bildfil (TIFF) eller Tabell 11, Obligatoriska metadata för bildfil (JPEG) ska sparas i katalogen som separata filer. Filerna ska namnges med fyra siffror med början från 0001.tif eller 0001.jpg.
mix	I katalogen ska sparas de obligatoriska uppgifter som meddelas i tabell 12 om alla bildfiler som finns i master-katalogen. Man får spara också andra uppgifter i enlighet med metadataschemat MIX i filen i en struktur som är förenlig med schemat. Filerna ska ges ett namn med fyra siffror med början i 0001.xml.

Om materialet levereras till Riksarkivet **ska** varje leveranspaket paketeras till ett TAR- eller ZIP-paket. TAR- eller ZIP-paketets innehåll **får inte** komprimeras i det här skedet. För TAR- eller ZIP-paket

⁸ Ett analogt uttryck som bestämts att ska digitaliseras bör vara beskrivet (beskrivande metadata har producerats) innan det digitaliseras. Det ska vara möjligt att med hjälp av identifier koppla de digitala uttryck som skapats genom digitaliseringsprocessen till de metadata som beskrivs ovan.

ANVISNING

19.1.2021

RA/15780/07.01.01.03.01/2020

ska beräknas en kontrollsumma i MD5-format och uppgiften **ska** lämnas i anslutning till överföringen. När material levereras till Riksarkivet avses med identifier-katalogen materialbeteckningen Ai01 för material i AHAA-systemet (i rollen ahaa teknisk).

Exempelpaket

1. EsimerkkipakettiJPEG2020.tar
2. EsimerkkipakettiTIF2020.tar

Bildfilerna i master-katalogerna är inte referenser avseende bildkvaliteten. I bildfilerna ingår de metadata som definieras som obligatoriska i det här dokumentet. Filerna i mix-katalogerna är exempel på master-filerna i bilagepaketet frånsett eventuella element för vars del annat konstateras.

8 Underskrifter

Generaldirektör

Jussi Nuorteva

Forskningschef

Päivi Happonen