

# Vägledning om digitalt bevarande

## Innehåll

Inledning .....	2
Så här läser du vägledningen .....	4
Regelverk .....	5
Lagar och förordningar .....	5
Tryckfrihetsförordningen .....	5
Offentlighets- och sekretesslagen .....	6
Arkivlagen .....	6
EU:s dataskyddsförordning .....	7
Lagen om vidareutnyttjande av handlingar från den offentliga förvaltningen .....	7
Upphovsrättslagen .....	8
Föreskrifter .....	8
Riksarkivets generella föreskrifter och allmänna råd .....	8
Metadata .....	9
Metadatastandarder .....	10
Metadatatyper .....	11
OAIS-modellen .....	12
Vad är ett OAIS? .....	12
OAIS och målgrupper .....	12
OAIS informationsmodell .....	15
Bevarandemetadata .....	16
Implementation av arkivpaket .....	17
Bevarandefaser .....	18
Skapandefasen .....	18
Aktiva fasen .....	18
Förberedelsefasen .....	19
Bevarandefasen .....	20
Tillgängliggörandefasen .....	20
Bevarandenaspekter .....	21
Lagring och bitbevarande .....	21
Logiskt bevarande .....	22

Innehållsmässigt bevarande .....	23
Äkthetsbevarande .....	24
Bevarandeformat.....	25
Grundläggande kriterier för val av filformat.....	25
Bedömning av lämpligt bevarandeformat.....	25
Bevarandeformat i praktiken.....	26
Vidare läsning .....	27
Organisationer .....	27
Andra resurser .....	28
Bilaga: Begrepp.....	29
Bilaga: Standarder .....	32
Bilaga: Val av bevarandeformat.....	35
Bild.....	35
Dokument.....	39
Ljud.....	41
Rörlig bild.....	43
Strukturerade data .....	44
System .....	45
Bilaga: Verksamhetsmodell .....	45
Införande av verksamhetsmodell för digitalt bevarande .....	46
Verksamhetsmodellens delar .....	48
Bilaga: Översättning av Levels of Preservation.....	50

## Inledning

Den totala mängden digital kulturarvsinformation i samhället ökar hela tiden och innebär helt nya möjligheter. Genom att informationen görs tillgänglig digitalt kan den nå både nya och fler användare inom offentlig verksamhet, näringsliv och inte minst en intresserad allmänhet. Och genom att den används på nya sätt kan det i sin tur leda till både kunskaps spridning, bättre samhällsservice, innovativa lösningar och en stärkt demokrati.

När kulturarvsinformation digitaliseras kan den i större utsträckning möta både samtida och framtida användares behov. Redan idag kan digital information användas av både människor och maskiner. Informationens potential är heller inte begränsad till vad som idag är tekniskt möjligt. Utvecklingen går snabbt framåt och effektivisering och automatisering kan innebära att det vi idag ser som en omöjlighet snart är en möjlighet.

Det går inte att med säkerhet förutsäga vilken information som kommer att användas framöver eller på vilka nya sätt, och med det kommer ansvaret att se till att dagens information tas med in i framtiden. För att åstadkomma detta behöver ett digitalt och långsiktigt bevarande av informationen prioriteras och tas om hand. Denna vägledning ger en introduktion till ett sådant arbete.

### Vad är digitalt bevarande?

Med digitalt bevarande som begrepp avses det att digital information för framtiden hålls tillgänglig och förstälig med beaktande av dess autenticitet och dataintegritet.

Autenticitet innebär i sammanhanget i vilken utsträckning en person eller ett system kan påvisa att informationen är vad den utger sig för att vara. Autenticitet är beroende av olika sorters bevis som behöver dokumenteras och bevaras. Det kan till exempel vara avgörande att kunna fastställa identiteten hos den som utfärdat ett dokument för att kunna avgöra om det är autentiskt eller inte. Det är viktigt att notera att bara för att något är autentiskt är det inte säkert att det är sant. Autenticiteten handlar bara om att något är på samma sätt som när det ursprungligen skapades. Desinformation kan vara autentisk utan att innehålla minsta gram sanning.

Dataintegritet innebär att informationen som bevaras är fullständig och inte har förvanskats. För att lyckas med det digitala bevarandet krävs att vi så långt det är möjligt säkerställer att vi varken förlorar dataobjektet, det vill säga förlust av 1:or och 0:or, eller möjligheten att tolka dessa 1:or och 0:or till meningsfull och autentisk information.

#### Kontrollsummor som bevis

En metod för att säkerställa dataintegritet är att beräkna kontrollsummor för de digitala filerna och jämföra dessa. Kontrollsumman blir då ett bevis för att innehållet är oförändrat och stärker därmed den digitala informationens autenticitet.

#### Bevarande av digitalt kulturarv

När allt mer information blir digital ställs nya krav på de kulturarvsinstitutioner som ansvarar för att informationen bevaras och blir långsiktigt användbar. Utmaningarna i det långsiktiga bevarandet blir annorlunda, och ibland mer komplexa, när det handlar om digital information jämfört med det vi är vana vid när vi arbetar med fysiska handlingar eller objekt.

Utvecklingen på området går också så snabbt att det kan upplevas svårt att fatta rätt beslut i bevarandefrågor. Här behövs hela tiden ny kunskap och nya överväganden kring hur bevarande sker bäst.

Det digitala bevarandet behöver uppmärksammas både innan, under och efter ett digitaliseringsarbete och kräver medvetna beslut och handlingsplaner. Digitalt bevarande uppfattas ofta som en IT-fråga men förutom teknisk lagring handlar bevarande också om att det finns rutiner och processer för informationsförvaltning på plats så att informationen kan tolkas över tid. Därmed är frågan om långsiktigt bevarande en fråga som oftast rör flera olika roller inom en organisation.

#### Kulturarvsinstitutionernas särskilda förutsättningar

Kulturarvsinstitutioner har ett unikt material som visar hur dåtid, samtid och framtid hänger ihop. Institutionerna ägnar sig åt ett dubbelt bevarande på så sätt att de bevarar dels fysiska objekt såsom museiföremål, böcker, filmrullar med mera, dels digitala representationer av materialet såsom 3D-scannade objekt, e-böcker, digitaliserad film med mera. Dessutom tillkommer information som är digitalt född, det vill säga information som har skapats i digital form och inte är ett resultat av att ett fysiskt objekt har digitaliserats.

Rutiner och processer för bevarandet av fysiska objekt har utvecklats över lång tid, men många kulturarvsinstitutioner saknar rutiner och processer för hantering av digital kulturarvsinformation som ska bevaras långsiktigt. Kulturarvsinstitutioner kan också ha delvis olika syften med sitt digitala bevarande och genomför det därför på olika sätt. Ibland är det till exempel viktigast att den digitala informationen är så lik sin förlaga som möjligt, medan det ibland är viktigare att den digitala informationen uppdateras utifrån den senaste tekniken så att den blir mer tillgänglig för användaren. Ibland är det bråttom att bevara information som finns på exempelvis äldre magnetband eller tidiga hårddiskar innan tidens tand gör det omöjligt att nå den.

Det finns ingen formel för digitalt bevarande som passar alla organisationer, men det går att utvärdera och utveckla bevarandet utifrån ett antal perspektiv som mer eller mindre alla aktörer ställs inför. Denna vägledning tar avstamp i Digisams vägledande principer för arbetet med digitalt kulturarv<sup>1</sup>, som pekar ut det digitala bevarandet som en nyckelfråga. Oavsett vilken roll du har och hur förutsättningarna ser ut i din organisation är förhoppningen att vägledningen gör det enklare att kartlägga nuläget, sätta rätt ambitionsnivå och se till att den digitaliserade informationen ges rätt förutsättningar att kunna nås och förstås även i framtiden.

### Så här läser du vägledningen

Vägledningen är uppbyggd kring olika teman som är centrala i arbetet med digitalt bevarande. I varje avsnitt ges en överblick över ett specifikt tema och i slutet av varje avsnitt lyfts några sammanfattande punkter. I slutet av vägledningen hittar du ett avsnitt med tips för vidare läsning samt bilagor. Vägledningen är uppdelad i följande avsnitt:

#### **Regelverk**

Avsnittet ger en orientering av lagstiftning och andra regelverk som berör digitalt bevarande, till exempel arkivlagen och Riksarkivets generella föreskrifter och allmänna råd.

#### **Metadata**

Metadata tas upp i flera delar av vägledningen, och detta avsnitt syftar därför till att ge en översiktlig sammanfattning av vad metadata är och hur det används. Avsnittet tar upp metadatastandarder och olika typer av metadata. Bevarandemetadatan behandlas senare i avsnittet om OAIS-modellen.

#### **OAIS-modellen**

OAIS-modellen är en ISO-standard och ett ramverk för arkiv som praktiskt taget alla som arbetar med digitalt bevarande förhåller sig till. Avsnittet ger en övergripande förståelse för vissa nyckelkoncept i standarden som har bäring på övriga avsnitt i vägledningen.

#### **Bevarandefaser**

Ett bra sätt att ta reda på och förstå vad som är viktigt att tänka på vid digitalt bevarande är att följa det enskilda digitala objektets livscykel, från att objektet tas emot i ett system för digitalt bevarande till att tillgången till materialet säkras över tid genom olika åtgärder. För digitalt fött material börjar processen redan innan objektet tas emot, med krav på medföljande metadata och filformat. I detta avsnitt beskrivs olika faser för bevarandet delvis utifrån OAIS-modellen.

#### **Bevarandenaspekter**

Avsnittet tar upp olika aspekter av digitalt bevarande såsom lagring och bitbevarande, logiskt bevarande, innehållsmässigt bevarande och äkthetsbevarande. Kunskap om dessa kan bland annat

---

<sup>1</sup> <http://www.digisam.se/vagledande-principer-for-arbetet-med-digitalt-kulturarv/>

vara till nytta om du ska ta fram en strategi för digitalt bevarande eller ställa krav på och utvärdera olika alternativ för e-arkivering.

#### **Bevarandeformat**

Avsnittet ger vägledning kring vad du bör tänka på vid val av filformat för digitalt bevarande, vare sig det handlar om val av filformat vid skapande av digitalt material, vid konvertering från ett filformat till ett annat, eller vid digitalisering av analogt material.

#### **Vidare läsning**

Avsnittet ger vägledning till vidare och fördjupad läsning inom området digitalt bevarande.

## Regelverk

För att lyckas med digitalt bevarande är det viktigt att ha överblick över vilka lagar och andra externa regelverk som gäller för den egna verksamheten och i vilket syfte de finns. Nedan följer en orientering kring regelverk inom området digitalt bevarande.

Notera att kulturarvsinstitutioner även kan ges separata uppdrag, via sina instruktioner, som rör digitalt bevarande.

#### Lagar och förordningar

Den lag som uttalat ställer krav på att information ska bevaras är arkivlagen<sup>2</sup> (1990:782). Övriga lagar och förordningar som tas upp i vägledningen berör främst situationer där informationen ska göras tillgänglig.

#### Tryckfrihetsförordningen

Tryckfrihetsförordningen<sup>3</sup>, TF, är en av Sveriges fyra grundlagar som förutom att reglera friheten för var och en att i tryckt skrift uttrycka tankar, åsikter och känslor även innehåller grundläggande bestämmelser om handlingsoffentligheten och begränsningar av densamma.

Tryckfrihetsförordningen definierar begrepp som handling, allmän handling samt kriterierna för att en handling ska betraktas som allmän. I 2 kap 13-14 §§ TF finns bestämmelser om undantag från huvudreglerna om vilka handlingar som är att anse som allmänna. För kulturarvsinstitutioner är 14 § första stycket särskilt relevant då det undantar handlingar som ingår i ett bibliotek, eller som från en enskild har tillförts ett allmänt arkiv uteslutande för förvaring och vård eller forsknings- och studieändamål, eller privata brev, skrifter eller upptagningar som annars har överlämnats till en myndighet uteslutande för något av de angivna ändamålen.

<sup>2</sup> [https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/arkivlag-1990782\\_sfs-1990-782](https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/arkivlag-1990782_sfs-1990-782)

<sup>3</sup> [https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/tryckfrihetsforordning-1949105\\_sfs-1949-105](https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/tryckfrihetsforordning-1949105_sfs-1949-105)

#### Teknikneutrala regler

Reglerna i såväl tryckfrihetsförordningen, arkivlagen och dess följdförfattningar är teknikneutrala, det vill säga att de gäller oavsett om det handlar om analog eller digital information.

#### Offentlighets- och sekretesslagen

En begränsning av rätten att ta del av allmänna handlingar ska enligt tryckfrihetsförordningen anges noga i en bestämmelse i en särskild lag. Denna lag är offentlighets- och sekretesslagen<sup>4</sup> (2009:400), OSL, som förutom att reglera begränsningar i rätten att ta del av allmänna handlingar, tystnadsplikt och vissa sekretessbrytande bestämmelser även reglerar myndigheternas handläggning vid registrering, utlämnande och övrig hantering av allmänna handlingar. Den ställer även krav på att myndigheten beskriver sina allmänna handlingar med information om:

- myndighetens organisation och verksamhet i syfte att underlätta sökande efter allmänna handlingar
- register, förteckningar eller andra sökmedel till myndighetens allmänna handlingar
- tekniska hjälpmedel som enskilda själva kan få använda hos myndigheten för att ta del av allmänna handlingar
- vem hos myndigheten som kan lämna närmare upplysningar om myndighetens allmänna handlingar, deras användning och sökmöjligheter
- vilka bestämmelser om sekretess som myndigheten vanligen tillämpar på uppgifter i sina handlingar
- uppgifter som myndigheten regelbundet hämtar från eller lämnar till andra samt hur och när detta sker
- myndighetens rätt till försäljning av personuppgifter.

#### Arkivlagen

I arkivlagen 5(1990:782) definieras begreppen arkiv och arkivbildning och lagen anger de grundläggande bestämmelserna om myndigheter och vissa andra organs arkivvård. Med arkivvård avses alla de nödvändiga åtgärder såsom redovisning, organisering, skydd och gallring som utförs för att säkerställa att den information som myndigheten bevarar är användbar. Huvudregeln i arkiv- och offentlighetslagstiftningen är att alla allmänna handlingar ska bevaras och att de är en del av det nationella kulturarvet. Men information får ändå gallras under förutsättning att den inte längre behövs för att tillgodose arkivlagens tre bevarandemål:

1. rätten att ta del av allmänna handlingar
2. behovet av information för rättskipning och förvaltning
3. forskningens behov.

<sup>4</sup> [https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/offentlighets-och-sekretesslag-2009400\\_sfs-2009-400](https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/offentlighets-och-sekretesslag-2009400_sfs-2009-400)

<sup>5</sup> [https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/arkivlag-1990782\\_sfs-1990-782](https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/arkivlag-1990782_sfs-1990-782)

#### Arkiv och arkivbildning

Arkivlagens definition av arkiv är att arkivet bildas av dess allmänna handlingar. Med arkivbildning menas ett arkivs kontinuerliga tillväxt genom att allmänna handlingar förs samman i planerad och strukturerad form.

#### EU:s dataskyddsförordning

Inom EU-samarbetet finns olika typer av rättsakter. EU-direktiv ska införlivas i den nationella lagstiftningen, medan EU-förordningar gäller direkt i alla medlemsländer oberoende av nationell lagstiftning. Dataskyddsförordningen<sup>6</sup> (Europaparlamentets och rådets förordning (EU) 2016/679, ofta kallad GDPR efter det engelska namnet General Data Protection Regulation) är till för att skydda enskildas grundläggande rättigheter och friheter, särskilt deras rätt till skydd av personuppgifter.

Personuppgifter är all slags information som kan knytas till en levande person. Dataskyddsförordningens regler är alltså inte tillämpliga på avlidna personers uppgifter.

Även om en EU-förordning innebär att reglerna kommer att gälla som lag direkt i alla EU:s medlemsländer lämnar dataskyddsförordningen utrymme för medlemsländerna att stifta nationella kompletterande lagar inom vissa av dataskyddsförordningens områden. För att bland annat säkerställa att den svenska offentlighetsprincipen ska fungera har en rad anpassningar gjorts genom SFS 2018:218 Lag med kompletterande bestämmelser till EU:s dataskyddsförordning (Dataskyddslagen)<sup>7</sup>. I sammanhanget är 3 kap 6 § särskilt relevant då den stadgar att känsliga personuppgifter får behandlas för arkivändamål av allmänt intresse om behandlingen är nödvändig för att den personuppgiftsansvarige ska kunna följa föreskrifter om arkiv.

#### Begränsa tillgången

Även om information med personuppgifter arkiveras med stöd av arkivlagen bör tillgången till informationen begränsas. Medarbetare bör bara ha tillgång till de personuppgifter som krävs för att de ska kunna utföra sina arbetsuppgifter.

#### Lagen om vidareutnyttjande av handlingar från den offentliga förvaltningen

Syftet med lagen om vidareutnyttjande av handlingar från den offentliga förvaltningen<sup>8</sup> (PSI-lagen) är att främja utvecklingen av en informationsmarknad genom att underlätta enskildas användning av

<sup>6</sup> <https://eur-lex.europa.eu/legal-content/SV/TXT/PDF/?uri=CELEX:32016R0679&rid=1>

<sup>7</sup> <https://www.datainspektionen.se/lagar--regler/dataskyddslagen/>

<sup>8</sup> <http://rkrattsbaser.gov.se/sfst?bet=2010:566>

handlingar som tillhandahålls av myndigheter. Med vidareutnyttjande avses användning av handlingar för andra ändamål än vad de var avsedda för från början.

Lagen innehåller bland annat regler som ska förhindra att myndigheter beslutar om sådana villkor för vidareutnyttjande av handlingar som begränsar konkurrensen. Med andra ord ska den offentliga förvaltningen möjliggöra vidareutnyttjande av information i de fall det är möjligt, det vill säga när varken upphovsrätt hos tredje man, sekretess eller i viss mån personuppgifter hindrar detta.

### Upphovsrättslagen

Lagen om upphovsrätt till litterära och konstnärliga verk (1960:729)<sup>9</sup> kallas populärt för upphovsrättslagen. Upphovsrätten är ett skydd för konstnärliga och litterära verk. Skyddet uppstår hos den person som har skapat verket eller närstående rättighetshavare som skapar framträdanden eller prestationer av olika slag<sup>10</sup>. Det som skyddas i lagen är i praktiken en hel rad verkskategorier (texter, datorprogram, musikaliska eller sceniska verk, filmverk, fotografiska verk eller annan bildkonst, byggnadskonst eller brukskonst, samt andra typer av verk). I upphovsrättslagen finns det även skydd för närstående rättigheter som kan vara olika framföranden och prestationer. Till den kategorin hör fotografiska bilder.

Enligt direktiv (EU) 2019:790, populärt kallat DSM-direktivet, som ska implementeras i upphovsrättslagen senast den 7 juni 2021, får kulturarvsinstitutioner (specifikt museer, bibliotek och arkiv) skapa digitala kopior i bevarandesyfte utan att ha bitt upphovspersonen till originalen om tillåtelse. Med bevarandesyfte avses i detta fall dock enbart arkivering, inte tillgängliggörande.

För fortsatt läsning om upphovsrätt, se Digisams vägledning om upphovsrätt och angränsande lagstiftning<sup>11</sup>.

### Föreskrifter

#### Riksarkivets generella föreskrifter och allmänna råd

Riksarkivets generella föreskrifter och allmänna råd publiceras i Riksarkivets författningssamling (RA-FS)<sup>12</sup>. Föreskrifterna ska i regel tillämpas av statliga myndigheter och de enskilda organ som förvarar allmänna handlingar.

Den RA-FS som omfattar allmänna regler, oavsett medium, för alla områden inom myndigheternas arkivhantering är Riksarkivets föreskrifter och allmänna råd om arkiv hos statliga myndigheter, RA-FS 1991:1 (senast ändrad och omtryckt genom RA-FS 2019:2). Här finns regler om allt ifrån avgränsning och organisation av arkiv till arkivredovisning och gallring samt framställning, förvaring, skydd och överlämnande av handlingar.

<sup>9</sup> [https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-1960729-om-upphovsratt-till-litterara-och\\_sfs-1960-729](https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-1960729-om-upphovsratt-till-litterara-och_sfs-1960-729)

<sup>10</sup> [http://www.digisam.se/wp-content/uploads/2015/06/Vagledning\\_om\\_upphovsratt\\_och\\_angransande\\_lagstiftning\\_ver-2.0.pdf](http://www.digisam.se/wp-content/uploads/2015/06/Vagledning_om_upphovsratt_och_angransande_lagstiftning_ver-2.0.pdf)

<sup>11</sup> [http://www.digisam.se/wp-content/uploads/2020/07/Vagledning\\_om\\_upphovsratt\\_3.1.pdf](http://www.digisam.se/wp-content/uploads/2020/07/Vagledning_om_upphovsratt_3.1.pdf)

<sup>12</sup> <https://riksarkivet.se/generella-foreskrifter>



Övriga föreskrifter i serien RA-FS förtydligar och kompletterar de allmänna reglerna inom olika områden. Föreskrifterna kan grupperas på följande sätt:

- Mediespecifika föreskrifter är regler för handlingar på olika typer av medier. Det finns exempelvis föreskrifter om handlingar på papper och mikrofilm. I RA-FS 2009:1 om elektroniska handlingar finns bestämmelser gällande strategi, planering och dokumentation samt relevanta åtgärder som behöver tas för att de elektroniska handlingarna ska kunna framställas, överföras, dokumenteras, hanteras, förvaras och vårdas så att de kan presenteras upprepat under den tid som de ska bevaras.
- Föreskrifter med tekniska krav behandlar tekniska krav för olika medier, exempelvis krav på bevarandeformat för elektroniska handlingar (RA-FS 2009:2) eller krav på arkivbeständigt papper. De mediespecifika föreskrifterna och föreskrifterna med tekniska krav ska alltid läsas tillsammans.
- Föreskrifter om arkivlokaler finns i en särskild RA-FS om planering, utförande och drift av arkivlokaler (RA-FS 2013:4).
- Generella gallringsföreskrifter i serien RA-FS är i regel sådana som behandlar allmänt förekommande handlingar som finns hos alla eller många statliga myndigheter. Inom stödverksamheter finns föreskrifter om gallring av räkenskaps-, personal-, och upphandlingshandlingar. Inom kärnverksamheter finns föreskrifter för handlingar inom forskning, EU-samarbete samt för handlingar inom universitet och högskolor.
- Myndighetsspecifika föreskrifter och beslut utfärdas av Riksarkivet i fall där de generella föreskrifterna inte går att tillämpa eller vid särskilda behov. Handläggningen inleds med en framställan från myndigheten, varefter Riksarkivet utreder och fattar beslut.

### Sammanfattning

- Digitalt bevarande kräver god överblick över lagar och andra regelverk. Den lag som uttalat ställer krav på att information ska bevaras är arkivlagen.
- Riksarkivets generella föreskrifter och allmänna råd tar upp både allmänna och mer specifika regler för områden inom myndigheters arkivhantering.

Kommenterad [HS(S1): Bör stå: Statliga myndigheter

### Metadata

Metadata är en viktig del i arbetet med digitalt bevarande. Metadata tas upp i flera delar av vägledningen, och detta avsnitt syftar därför till att ge en sammanfattning av vad metadata är och hur det används. Avsnittet tar upp metadatastandarder och olika typer av metadata. Bevarandemetadata behandlas senare i avsnittet om OAIS-modellen.

Metadata är strukturerade data som beskriver och definierar annan data eller information för att underlätta förståelse, sökning och användning. Exempel på metadata är: namn, placering, ämnesord, information om när objektet skapades, eller specifika egenskaper som objektets dimensioner eller användningsområde. Metadata är nyckeln till att kunna återsöka digitala informationsobjekt och att lagra viktig information om dem. Metadata är användbara för att sammanställa och analysera information. Ibland kan det vara svårt att skilja information från dess metadata, då metadata i många fall kan anses vara en del av själva informationsobjektet, samt eftersom metadata också är en form av information.

Metadata kan lagras och delas på många sätt. I system kan de sparas som fält i en relationsdatabas eller ett register, där en samling metadata som beskriver samma objekt kan beskrivas som en post. Metadata kan också vara inbäddade i själva filen, lagras i Excelark eller fristående filer, eller uttryckas som XML. XML, Extensible Markup Language, är ett märkspråk vars kod ger datorprogram instruktioner för hur information ska hanteras och presenteras.

Inom kulturarvsvärlden finns olika traditioner för hur metadata används. Biblioteksvärlden har väl utarbetade system för att katalogisera publikationer och exempelvis beskriva ämnesord och använda auktoritetslistor för författare och utgivare. Museer använder metadata för att beskriva föremålssamlingar och artefakter, till exempel utifrån föremålets egenskaper (utformning, material, storlek, typ) eller vad som har hänt under objektets livstid (hur det skapades, hur det har använts, var det har förvarats, om det har konserverats, lånats ut eller ställts ut). Arkiven har traditionellt fokuserat på sökingångar, arkivbeskrivningar och att dokumentera inom vilken kontext arkivbestånden har uppstått.

### Metadatastandarder

Att använda sig av standarder för metadata underlättar arbetet vid till exempel registrering och exportering eller delning av information. Anledningen till att det blir lättare att dela information är att standarder är dokumenterade och definierade samt att betydelsen och användningen av olika metadataelement är överenskommen och ska användas på samma sätt av alla som använder standarden. De som använder samma standard kan utbyta information direkt. Om öppna standarder används är det också lättare att transformera informationen från ett format till ett annat.

Det finns en uppsjö standarder för metadata som används av kulturarvsinstitutioner. Standarderna kan till exempel vara anpassade för att beskriva olika materialtyper (arkiv, konstverk, fotografier, böcker, museiföremål m.m.) eller utifrån de förutsättningar eller den bransch där informationen hanteras (arkiv, bibliotek, museer, bildbyråer med flera). Metadatastandarder kan avse både hur innehållet ska beskrivas, till exempel genom listor med tillåtna termer (vokabulärer) och hur metadata ska struktureras och vilka dataelement den ska innehålla.

Valet av metadatastandard för att beskriva den information som ska förvaltas och bevaras påverkas av vilken typ av information det handlar om och vem som ska kunna ta del av den. Inom biblioteksvärlden är det till exempel fördelaktigt att använda sig av den inom biblioteksvärlden utbredda standarden MODS<sup>13</sup>, medan det inom samlingsförvaltning kan vara bättre att använda sig av CIDOC-CRM<sup>14</sup> som används av många museer. Vid hantering av ett mycket specifikt material, till exempel enbart fotografier, kan det vara fördelaktigt att använda en standard som har specifika fotografiska metadata såsom registrerad slutartid och bränntid. Vid förvaltning av ett spretigt material som innehåller allt ifrån exempelvis fotografier och kontorsdokument till ritningar och ljudupptagningar, kan det istället finnas behov av en bredare standard. Vid valet av ett mer generellt sätt att beskriva metadata på kan det ibland hända att informationen inte kan beskrivas på den detaljnivå som önskas.

Standarder anger oftast en miniminivå för vilka metadata som alltid måste finnas med på alla informationsobjekt, d.v.s. ett antal obligatoriska element. Sådana metadata är ofta titel och upphovsman till objektet, datum då objektet skapades, arkivbildare eller utgivare av objektet samt identifierare för objektet. Metadatastandarder skiljer sig åt när det kommer till hur många element som är obligatoriska. I Dublin Cores DCMI Metadata Terms<sup>15</sup> är till exempel inte något av de 15

<sup>13</sup> Metadata Object Description Schema (MODS), <http://www.loc.gov/standards/mods/>

<sup>14</sup> CIDOC Conceptual Reference Model (CRM), <http://www.cidoc-crm.org/>

<sup>15</sup> DCMI Metadata Terms, <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

grundelementen obligatoriska. Även standarder för katalogisering och metadataregistrering kan ange en miniminivå för metadata. Exempel på sådana standarder är Spectrum (se avsnitt Regelverk och standarder) och CCO (Cataloging Cultural Objects)<sup>16</sup>.

Vid avsaknad av ett system för digitalt bevarande som stödjer en specifik metadatastandard är det viktigt att använda en standard som är lämplig för den information som ska bevaras och som minst används av de obligatoriska elementen.

### Metadatatyper

För att kunna återsöka, förstå och bevara digital information behövs både beskrivande metadata, administrativa metadata, tekniska metadata och strukturella metadata.

#### Beskrivande metadata

Metadata som rör innehållet, och som fungerar som ett stöd i att återsöka och förstå innehållet, brukar grupperas som beskrivande metadata. Inom kulturarvssammanhang är beskrivande metadata centralt. Exempel på beskrivande metadata är genre, motivbeskrivningar, användningsområde, ämnesord, benämningar och titlar.

#### Administrativa och tekniska metadata

Administrativa metadata är metadata som krävs för att hantera ett digitalt objekt medan tekniska metadata behövs för att programvara ska kunna öppna och visa filer på ett korrekt sätt genom att rätt filtyp är angiven. Metadata som styr åtkomst och användning av digitala objekt, och till exempel anger sekretess eller personuppgifter är viktiga för hur objektet får hanteras. Här ingår också metadata som är viktiga i bevarandesyften såsom checksummor som hindrar förändring av det digitala objektet eller information om migrering och konvertering. En stor del av tekniska och administrativa metadata skapas av system eller program. Ofta ligger tekniska metadata inbäddade i en fil, men i ett system för digitalt långtidsbevarande behövs troligtvis också funktionalitet för att kunna logga och lagra information kring vad som har skett med ett objekt, till exempel information om tidpunkt för konvertering och tidigare format.

#### Strukturella metadata

Strukturella metadata beskriver relationer mellan olika digitala objekt, eller delar inom samma digitala objekt. Det kan till exempel handla om sidor i en publikation eller ett katalogsystem med filer.

### Sammanfattning

- Metadata är nyckeln till att kunna återsöka digitala informationsobjekt och att lagra viktig information om dem. Vid arbete med metadata, utgå ifrån frågan: Vilka metadata behövs för att informationen ska gå att förstå i framtiden?
- Att utgå från en metadatastandard är rekommenderat för att underlätta arbete vid registrering och för att skapa förutsättningar för att exportera eller dela vidare sin information. Välj en standard som passar den information och typ av verksamhet där informationen hanteras.

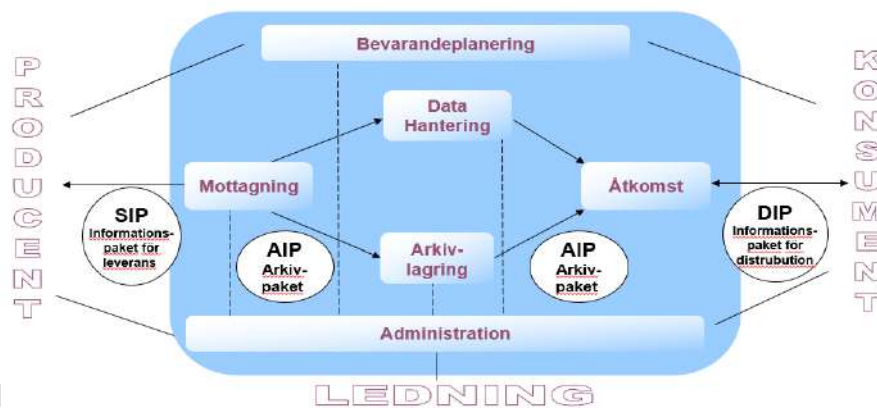
<sup>16</sup> Cataloging Cultural Objects (CCO), <https://vraweb.org/resources/cataloging-cultural-objects/>

## OAIS-modellen

OAIS (oberoende arkivinformationssystem) är en standard och referensmodell för digitalt bevarande. OAIS är både antagen som internationell och svensk ISO-standard men den finns också fritt tillgänglig via The Consultative Committee for Space Data Systems (CCSDS) som rekommenderad praxis (Magenta book CCSDS 650.0-M-2).

Vad är ett OAIS?

### OAIS Funktionsmodell



Arkiv kan ha flera olika betydelser. I arkivlagen är definitionen av arkiv som bekant en myndighets allmänna handlingar. Arkiv kan också vara en plats, en byggnad eller ett rum eller också en organisation som exempelvis Riksarkivet. OAIS definition av arkiv är det senare, nämligen: "En organisation som har för avsikt att bevara information för åtkomst och användning av en utpekad målgrupp".

Ett OAIS är i sin tur ett arkiv som består av en organisation av människor och system som har åtagit sig ansvaret att bevara information och göra den åtkomlig för en utpekad målgrupp. Ett OAIS är med andra ord inte bara en teknisk lösning utan omfattar alla resurser en organisation behöver för att kunna bevara information långsiktigt. Se även bilaga Verksamhetsmodell som utgår från grundläggande delar i standarden Ledningssystem för kvalitet - Krav (ISO 9001:2015) och är tänkt att kunna tillämpas vid digitalt bevarande hos kulturarvsinstitutioner av olika storlek och med olika förutsättningar.

#### OAIS och målgrupper

Ett av OAIS viktigaste begrepp är den utpekade målgruppen (på engelska Designated Community), en identifierad grupp konsumenter som ska förstå en specifik uppsättning information. Det är målgruppens behov som styr strategiska val kring exempelvis åtkomst och bevarandeåtgärder.

Att bestämma vilken grupp eller grupper som ska vara arkivets målgrupp är också ett av de obligatoriska åtaganden som ett OAIIS har enligt standarden. För myndigheter kan arkivlagens bevarandemål vara ett första steg i att identifiera grupper av konsumenter.

#### Vem är vem i ett OAIIS

**Producent:** Person/system som tillhandahåller den information som ska bevaras.

**Konsument:** Person/system som interagerar med OAIIS tjänster för att hitta bevarad information.

**Designated Community:** Identifierad grupp konsumenter som ska förstå en specifik uppsättning information.

**Administration:** Funktion som har den löpande kontrollen över övriga funktioner i OAIIS.

**Ledning:** Fastställer övergripande regler och ansvarar för bl.a. finansiering.

**Övriga intressenter:** Andra intressenter utöver de ovan angivna.

#### Rätten att ta del av allmänna handlingar

Enligt tryckfrihetsförordningens 2 kap 1§ gäller rätten att ta del av allmänna handlingar för var och en. Den utpekade målgruppen som vi bevarar för blir alltså allmänheten i stort. En sätt att tillgodose detta behov i det digitala bevarandet är att säkerställa att information kan tillhandahållas i ett för tiden vanligt förekommande format. Det vill säga, att det finns vanligt förekommande program och applikationer där formatet går att öppna och läsa.

#### Behovet av information för rättskipningen och förvaltningen

Målgruppen kan alltså vara den egna myndigheten, dess verksamhetsområden och yrkeskompetenser. Viss speciallagstiftning och vissa författningar ställer också krav på hur information ska bevaras. Till exempel kan det finnas krav på att information för en viss tid ska bevaras i ursprungligt skick.

#### Forsknings behov

Att sätta om forskningens framtida behov är överlag en svår utmaning. En övergripande princip är förstås att möjligheten att ställa nya frågor utifrån informationen inte ska förstöras. Att använda forskningen i allmänhet som målgrupp kan vara för vagt för att vara till någon större hjälp även om vissa generaliseringar kan göras.

### FAIR-principerna

FAIR är en akronym som står för Findable, Accessible, Interoperable och Reusable. FAIR-principerna innebär att forskningsdata ska gå att hitta, det ska finnas information om hur det går att få tillgång till den, de ska vara kompatibla med andra data och de ska vara möjliga att återanvända. FAIR-principerna spelar en viktig roll i arbetet för öppen vetenskap och beskriver några av de mest centrala riktlinjerna för god datahantering och öppen tillgång till forskningsdata.

De internationella FAIR-principerna kan användas för att skapa förutsättningar för att forskningsdata och metadata hanteras på ett sätt som gör forskningsdata återanvändningsbara för forskare. Principen för återanvändningsbarhet (R i FAIR, Reusable) säger bland annat att den som erbjuder data inte själv ska försöka förutse hur konsumenten tänkt använda dessa data, utan istället erbjuda så rika metadata som möjligt för att skapa större användbarhet. FAIR-principerna handlar också om sökbarhet, tillgänglighet och interoperabilitet.

FAIR-principerna har de senaste åren fått starkt genomslag och blivit ledord för god datahantering och öppen tillgång till forskningsdata.

Olika bevarandestrategier kan behöva tillämpas på samma information för att tillgodose olika målgruppers behov. Som ett exempel på detta kan vi tänka oss ett arkiv eller museum som förvarar arkeologisk dokumentation, där dokumentation från arkeologiska undersökningar kan vara av intresse för flera målgrupper: Arkeologer som behöver underlag i sin yrkesutövning för att genomföra förundersökningar och utredningar, handläggare på myndigheter som behöver underlag för att göra korrekta bedömningar och beslut, forskare som vill analysera och studera dokumentation och data för att göra nya tolkningar av resultaten, eller en hembygdsforskare eller privatperson som vill lära sig historien kring deras trakt eller vad som finns på deras tomt. Dessa målgrupper kan ha olika behov och förkunskaper som styr hur de vill ta del av informationen, i vilka format, hur de vill kunna använda eller behandla informationen och vad de behöver för att förstå informationen.

Den historieintresserade allmänhetens intressen kan kanske tillgodoses genom att materialet tillhandahålls i läsbar form, till exempel rapporter i PDF-format och bilder i JPG-format. Yrkesutövarna kan ha användning för både läsbara och bearbetningsbara format, exempelvis PDF-format för läsning, tabellformat för listor över fynd och anläggningar eller tillgång till GIS-data. För framtida forskning kan ytterligare krav behövas för att kunna sammanställa och analysera data och kanske till och med för att göra sammanslagningar av data från flera källor. Den intresserade allmänheten har troligtvis inte samma tillgång till eller vana av att använda GIS-data som en arkeolog eller forskare.

För att tillgodose vissa målgruppers behov kan alltså data behöva anpassas (till exempel genom konvertering av filformat). Det är viktigt att inte råka stänga ute vissa målgrupper för att data anpassas utifrån behovet hos en specifik målgrupp. Ett sätt att undvika detta är att försöka hitta filformat som fyller flera användargrupperns behov, eller erbjuda flera filformat av samma information för att täcka olika användningsområden.

### OAIS funktioner

OAIS-standarden beskriver två olika modeller, dels den funktionsmodell som illustreras ovan som definierar övergripande informationsflöden och uppgifter som de flesta arkiv behöver och dels en informationsmodell som beskriver beståndsdelarna av de objekt ett arkiv tar emot, vårdar och tillgängliggör vilket innefattar de metadata som behövs för att informationen ska kunna förstås och användas nu och i framtiden. Modellen specificerar också hur olika typer av informationsobjekt relaterar till varandra och hur de är strukturerade genom självbeskrivande informationspaket som kan hanteras och förstås utan en specifik teknisk lösning. Dessa informationspaket kommer i olika former beroende på i vilken fas i bevarandecykeln de befinner sig:

- SIP Informationspaket för leverans (på engelska Submission Agreement Package)  
Mottagning (på engelska Ingest)
- AIP Arkivpaket (på engelska Archival Information Package)  
Åtkomst (på engelska Access)
- DIP Informationspaket för distribution (på engelska Dissemination Information Package)

OAIS funktionsmodell definierar sex funktionella entiteter som i sin tur beskriver ytterligare funktioner, arbetsuppgifter och andra aktiviteter i detalj. Mottagning, Arkivlagring, Datahantering och Åtkomst motsvarar ungefärligt ett informationspakets rörelser inom ett arkiv. 17 och Bevarandeplanering och Administration omfattar det dagliga arbetet med styrning, planering, och aktiviteter som utförs under informationens bevarandecykel. Ytterligare en funktion som ligger utanför OAIS kärna på samma sätt som Producent och Konsument är Ledning vars roll är att fastställa övergripande regler och ansvara för finansiering.

#### OAIS informationsmodell

OAIS-modellens informationsmodell är till stora delar en beskrivning av olika typer av relevant metadata. Enkelt uttryckt finns metadata gällande det logiska bevarandet i representationsinformation, metadata gällande det innehållsmässiga bevarandet i information om bevarande samt till viss del i representationsinformation och metadata gällande äkthetsbevarande i information om beständighet. För mer information om bevarandeaspekter, se avsnitt Bevarandeaspekter.

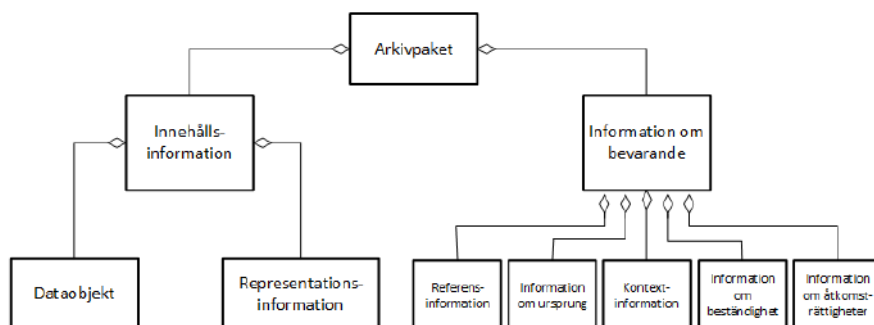
De metadata som är särskilt viktiga för bevarandet kallas i OAIS-modellen för information om bevarande (PDI)<sup>18</sup>. Dessa metadata behövs för att kunna garantera att informationen i bevarandeobjektet är tillförlitlig över lång tid. Ett specifikt bevarandeobjekt ska kunna identifieras, och förändring över tid i informationen ska dokumenteras, för att försäkra att innehållet inte har bytts ut, förvanskats eller gått förlorat. Information om bevarande har tydliga kopplingar till de aspekter som tas upp i avsnitt Bevarandeaspekter.

---

<sup>17</sup>

[https://www.cessda.eu/content/download/496/4465/file/CESSDA%20User%20Guide%20for%20digital%20preservation\\_2\\_OAIS.pdf](https://www.cessda.eu/content/download/496/4465/file/CESSDA%20User%20Guide%20for%20digital%20preservation_2_OAIS.pdf)

<sup>18</sup> Preservation Description Information, Reference Model for an Open Archival Information System (OAIS) avsnitt 4.2.1.4.2



Modellen ovanför visar hur ett arkivpaket i OAIS-modellen är uppbyggt. Arkivpaketet innehåller både innehållsinformationen - det vill säga det innehåll som ska bevaras – och Information om bevarande. Innehållsinformationen består av dataobjektet och dess representationsinformation. Dataobjektet kan vara ett digitalt objekt eller ett fysiskt objekt i OAIS informationsmodell. Arkivpaketet kan i sin tur beskrivas av Paketbeskrivning och identifieras genom Paketeringsinformation. Se vidare avsnitt 4.2 i Reference Model for an Open Archival Information System (OAIS), samt figur 4-18 i samma avsnitt.

Det är viktigt att definiera det innehåll som ska bevaras, det vill säga vad som utgör själva innehållsinformationen (se modell under avsnittet OAIS-modellen och bevarandemetadata), innan beslut tas om vilka metadatas som behövs.<sup>19</sup> En nödvändig fråga att ställa sig vid bevarande av information är: Vilka metadatas behövs för att informationen ska gå att förstå? Enligt OAIS-modellen bör den tilltänkta målgruppen, det vill säga den målgrupp som ska kunna tolka och förstå innehållsinformationen, vara utgångspunkt vid besvarandet av frågan.<sup>20</sup>

Ett vanligt exempel på vad som kan behöva förklaras närmare är förkortningar och koder. Detta på grund av att kunskapen om betydelsen av förkortningar och koder kan gå förlorad över tid, eller på grund av att koderna bara är allmän kännedom inom en viss yrkesgrupp eller bransch men okända för en bredare allmänhet. I innehållsinformationens metadatas kan det finnas behov av en beskrivning av vilket kodsysteem som har använts. Ett metadatafält med uppgiften "59.018" säger till exempel inte så mycket förrän vi vet att uppgiften rör en koordinat. Dessutom behövs metadatas kring vilket koordinatsysteem som har använts, och om det är x- eller y-axelns punkt som avses, innan betydelsen kan tolkas.

Innehållsinformationen behöver kunna förstås fristående. Med andra ord behöver innehållsinformationen innehålla alla nödvändiga metadatas för att de ska gå att tolka och förstå oavsett om kopplingar tappas mellan system eller om objektet flyttas till en ny plats.

### Bevarandemetadata

I OAIS-modellen delas information om bevarande in i fem kategorier:

- Information om ursprung (Provenance). Här beskrivs ursprungskällan för informationen samt historiken kring vem som har förvaltat informationen och hur den har förändrats. För att garantera informationens autenticitet behöver förändringar kunna spåras bakåt i tiden hela vägen till ursprunget.

<sup>19</sup> "Content Information", Reference Model for an Open Archival Information System (OAIS), s. 2-6 och avsnitt 4.2.1.4.1.

<sup>20</sup> Reference Model for an Open Archival Information System (OAIS), s. 2-4



- Kontextinformation (Context). Här beskrivs hur informationen relaterar till omvärlden. I vilket syfte skapades informationen och hur relaterar den till annan information?
- Referensinformation (Reference). Här anges en eller flera identifierare som används för att identifiera den specifika informationen. Referensinformation inkluderar också identifierare som möjliggör för externa system att hänvisa till denna specifika information. Ett exempel på detta är ISBN-numret på en bok.
- Information om beständighet (Fixity). Här dokumenteras mekanismer som ska skydda bevarandeobjektet från odokumenterade förändringar. Oftast rör det sig om tekniska metadata såsom checksummor.
- Information om åtkomsträttigheter (Access rights). Här anges hur bevarandeobjektet får användas, hur åtkomst ska ske samt vilka som har åtkomst till det, exempelvis utifrån legala ramverk, licensieringsvillkor och åtkomstkontroll. Angivelserna rör både hur bevarandeobjektet ska hanteras i arkivet och hur det får användas av slutanvändaren.

### Implementation av arkivpaket

OAIS-modellen talar inte om hur ett informationspaket ska implementeras rent tekniskt utan talar endast om vad som ska implementeras. Det finns fler än ett sätt att tekniskt implementera ett informationspaket och fler än en standard att följa. Den standard som kommit att dominera beskrivning av strukturella, administrativa och till viss del teknisk metadata är METS vilken används av Riksarkivet för FGS-Paket såväl som av Kungliga Biblioteket och EU:s byggnadsblock eArchivings specifikationer. I dessa implementationer kompletteras METS med PREMIS för bevarandemetadatan och ofta överlappar standarderna varandra. Till exempel kan PREMIS användas för att beskriva strukturella relationer mellan informationsobjekt och information om beständighet kan dokumenteras i båda. Viss redundans av metadata i olika standarder behöver inte innebära några problem och kan vara svår att undvika. Det är dock av vikt att bestämma vad och vilken standard som huvudsakligen används för vilken typ av metadata (se även avsnitt Metadata).

Standard	Beskrivning	Huvudsakligt Användningsområde
METS		Strukturell, administrativ och teknisk metadata
PREMIS		Bevarandemetadatan
W3C-PROV		Bevarandemetadatan (Proveniens)
MPEG-21 DID/DIDL (ISO21000-2:2005)		Strukturell-, Teknisk metadata
BAGIT (IETF-Specification RFC 8493)		Administrativ-, Teknisk metadata (beständighet)
XFDU (ISO 13527:2010)	XML formatted data unit	Strukturell-, administrativ, teknisk metadata
SIRF (ISO 23681:2019)	Self Contained Information Retention Format	Strukturell, administrativ, teknisk, bevarandemetadatan

Figur: Översikt av metadatastandarder för paketering av arkivinformaton

### Sammanfattning

- Bevarandemetadatan behövs för att kunna garantera att informationen i bevarandeobjektet är tillförlitlig, åtkomlig och användbar över lång tid. Beskrivande, administrativa och strukturella metadata behövs alla för att bevarandeobjektet ska gå att sökas upp, tolkas, och förvaltas över tid.

→ Ett OAIS är inte bara en teknisk lösning utan omfattar alla resurser en organisation behöver för att kunna bevara information långsiktigt.

## Bevarandefaser

Digitalt bevarande handlar om att säkra tillgänglighet till digitala objekt över lång tid. För detta behövs kunskaper om det digitala objektets olika bevarandefaser, faser som även har cykliska inslag. De bevarandefaser som tas upp här följer OAIS-modellen (se avsnitt OAIS-modellen) men är också utbyggda för att omfatta hela livscykeln.

### Här ska det vara en bild som illustrerar livscykeln

Digitala objekt som ska bevaras kan vara objekt som är digitalt födda, digitaliserade för tillgängliggörande eller digitaliserade av bevarandeskäl och alltså inte främst skapade för att visas upp i ett gränssnitt. Innan digitala objekt tas om hand i processen för digitalt bevarande kan andra objekt ha gallrats, det vill säga avlägsnats och förstörts enligt fastställda kriterier.

Nedan följer ett antal faser som det digitala objektets livscykel består av. Faserna har nära kopplingar till de bevarandenaspekter, som tas upp i den här vägledningens avsnitt Bevarandenaspekter, och kan vara till bra hjälp när en verksamhet ska identifiera vad som behöver tas om hand i en verksamhetsmodell för digitalt bevarande (se bilaga Verksamhetsmodell).

### Skapandefasen

Ett digitalt objekt kan antingen uppstå i digital form från första början eller bli till genom att ett fysiskt objekt digitaliseras. Det är viktigt att verksamheten tar höjd för det digitala bevarandet, bland annat vad gäller krav på metadata och filformat, redan vid planering av digitaliseringsprojekt (se avsnitt Bevarandenaspekter gällande innehållsmässigt bevarande och logiskt bevarande samt avsnitt Metadata). Även vid hantering av digitalt fött material kan krav behöva ställas på metadata och filformat.

Ett digitalt objekt, som har uppstått genom att ett fysiskt objekt har digitaliserats, har en stor andel metadata gemensamt med det fysiska objektet. Viktigt blir då att besluta om all metadata behöver kopieras eller om framtidssäker hänvisning kan ske till det fysiska objektets metadata. Om metadata för det digitala objektet ska kopieras från det fysiska objektets metadata är det också viktigt att, i de fall det är relevant, komma ihåg proveniensmetadata (det vill säga uppgifter om objektets ursprung, ägarhistoria samt annan historia).

### Aktiva fasen

Ett digitalt objekt kan innan överföring till ett e-arkiv befinna sig i en aktiv förvaltningsfas. Det betyder att det hanteras hos en kulturarvsinstitution utan att det ännu ses som en arkivhandling. Under den aktiva fasen är det viktigt att metadata samlas i den omfattning som krävs. Särskilt centrala är här de metadata som rör objektets kontext och historik.

Ibland förs en kopia av ett digitalt objekt direkt till ett e-arkiv i samma stund som det digitala objektet skapas. I dessa fall är det viktigt att lägga märke till om kopiorna över tid får "eget liv" genom att fyllas på med information som det digitala originalobjektet saknar. I så fall bör de bevaras som egna original och inte enbart kompletteras med de metadata som över tid tillkommer det digitala originalobjektet.

Lagring och bitbevarande (se avsnitt Bevarandenaspekter gällande lagring och bitbevarande) är ungefär lika relevant under hela livscykeln och uppkommer som en viktig bevarandenaspekt redan under den aktiva fasen.

En särskilt betydande fråga under den aktiva fasen är den om urval. Att säkerställa tillgänglighet av digitala objekt över lång tid kan bli resurskrävande och ekonomiskt kostsamt. Beslut kan behöva tas om att vissa digitala material inte kommer att sparas i ett system för digitalt bevarande eller vara föremål för bevarandeåtgärder. Notera att urvalsbeslut kan förekomma även under andra delar av livscykeln.

För allmänna handlingar gäller speciella gallringsregler, och urvalet behöver beslutas utifrån fastställda rutiner. Detta kan även bli aktuellt vid byte av format vid konvertering (se bevarandefasen nedan). Utarbetade policyer (se bilaga Verksamhetsmodell) kan vara ett bra stöd i urvalsarbetet genom att principer för verksamhetens digitala bevarande lyfts i dokumenten.

### Förberedelsefasen

I den här fasen förbereds de digitala objekten för att tas om hand i ett system för digitalt bevarande. Metadata har här central betydelse för att de digitala objekten ska kunna förstås, nu och i framtiden (se avsnitt Bevarandenaspekter gällande innehållsmässigt bevarande samt avsnitt Metadata). Enligt OAIS-modellen är det viktigt att utgå från den tilltänkta målgrupp som ska kunna förstå informationen när beslut tas om vilka metadata som behövs.

Vad gäller äkthetsbevarandet specifikt (se avsnitt Bevarandenaspekter gällande äkthetsbevarande) är det redan under förberedelsefasen viktigt att metoder av teknisk art som behöver tillämpas för detta ändamål ses över. Detta på grund av att äkthetsbevarandet i den aktiva förvaltningsfasen hanteras av ursprungssystemet i sig, men från och med nu behöver tas om hand på andra sätt.

Innan objekten omhändertas är det också lämpligt att de om möjligt katalogiseras på objektsnivå. En annan slags åtgärd i denna fas är filformatsvalidering, som förvisso också kan ske under mottagandefasen (se nedan) om det finns stöd för detta i bevarandesystemet. Om det finns bestämda filformat kan en så kallad normalisering av filerna behöva göras. Det betyder att filerna konverteras till ett förutbestämt filformat.

Efter det skickas filer antingen direkt in i ett system för digitalt bevarande, eller så packas de till så kallade SIP:ar, Submission Information Packages eller på svenska informationspaket för leverans till ett OAIS-arkiv. SIP:ar är paket med filer och metadata och ibland har det även påförts checksummor för att säkerställa att paketen inte har förändrats under överföringen till bevarandesystemet.

**Kommenterad [HS(S2)]:** Bör stå: e-arkiv som följer OAIS-standarden

### Mottagandefasen

Under mottagandet av filerna till bevarandesystemet kontrolleras de överförda filerna. Sedan påförs tekniska bevarandemetadata och så kallade AIP:er, Archival Information Packages eller arkivpaket, skapas. AIP:erna är helt enkelt de paket som ska bevaras för framtiden. De kan över tid konverteras till nya AIP:er efter olika bevarandeåtgärder, men original-AIP:erna sparas ofta oavsett. I denna fas behöver även informationens äkthet och filformaten valideras.

Både förberedelsefasen och mottagandefasen kan ibland upprepas flera gånger. Ett exempel är när en verksamhet först väljer att bevara ett digitalt objekt i egen regi, för att senare arrangera så att det digitala bevarandet av objektet sker hos en annan institution. Återigen fyller framtagandet av tillräckliga metadata en avgörande funktion för det digitala bevarandet (se avsnitt Bevarandenaspekter gällande innehållsmässigt bevarande och logiskt bevarande samt avsnitt

Metadata), vare sig situationen handlar om att information behöver skiljas från ursprungssystemet eller att den ska levereras till ett e-arkiv med en annan huvudman. När en verksamhet vill föra över handlingar från ett verksamhetssystem till ett arkiv, samtidigt som verksamheten fortfarande vill ha tillgång till handlingarna, hanteras handlingarna i ett så kallat mellanarkiv. När handlingar förs över till ett slutarkiv är det istället en arkivmyndighet som har tagit över ansvaret för de arkiverade handlingarna. Det står dock fortfarande verksamheten som levererade handlingarna fritt att vid behov begära ut dem från slutarkivet.

PAIMAS, Producer Archive Interface Methodology Abstract Standard (se bilaga Standarder), är en ISO-standard som berör de första stegen i mottagningen till ett OAIS-arkiv (se avsnitt OAIS-modellen).

### Bevarandefasen

För att digitala objekt ska säkras över tid krävs att lagringen och bitbevarandet sker tillfredsställande (se avsnitt Bevarandenaspekter gällande lagring och bitbevarande). Vad gäller externa databärare och utsträckt tid bör till exempel beständighetsfrågor och klimat beaktas. I mångt och mycket är det en datasäkerhetsfråga.

Utöver lagring och bitbevarande behöver också andra bevarandeåtgärder utföras för att den allmänna tilltron till e-arkivet ska hållas intakt över tid. Vid bevarandeåtgärder genereras till exempel nya AIP:er och kontroller av filer görs med jämna mellanrum för att ersätta obsoleta filformat. Filkontrollerna kan ske automatiskt eller manuellt genom att de helt enkelt öppnas. Kontroll av checksummor, som antingen sker automatiskt i bevarandesystemet eller via andra verktyg, är en metod för att kontrollera att AIP:erna inte har förändrats och därmed kunnat få förändrat informationsvärde eller till och med blivit oläsbara. Över tid kan också migrering behöva ske från ett bevarandesystem till ett annat. Detta behöver planeras noggrant och i god tid.

I allmänhet behövs inga nya metadata, förutom viss dokumentation av utförda bevarandeåtgärder. Samtidigt krävs bevakning i frågan utifall långa tidsrymder gör att ytterligare metadata behövs. Ett exempel på detta är behov av förklaringar av allmänna begrepp som har slutat användas. Standarden PREMIS, Data Dictionary for Preservation Metadata (se bilaga Standarder), är en specifikation för hur OAIS-modellens (se avsnitt OAIS-modellen) konceptuella struktur för informationsobjekt och arkivpaket kan implementeras i form av metadata.

Även om objekt som finns i ett system för digitalt bevarande behöver vara sig själva nog, det vill säga bära tillräckliga metadata för att kunna förstås över tid, är det samtidigt viktigt att vårda de register eller kataloger där objekten finns eller är beskrivna. Ibland kan det vara värt att överväga om även dessa register och kataloger bör bevaras i ett system för digitalt bevarande.

### Tillgängliggörandefasen

Tillgängliggörandefasen innebär att digitala objekt tillgängliggörs för olika målgrupper. Utifrån OAIS-modellen (se avsnitt OAIS-modellen) är det möjligt att efter utförd beställning till bevarandesystemet leverera så kallade DIP:ar, Dissemination Information Packages eller informationspaket, för distribution.

Bevarandeaktiviteter som sådana kan dock vara oavhängiga visningsgränssnitt eller andra sätt att presentera materialet. Med andra ord kan verksamheter i vissa lägen välja att lösa frågan om hur det digitala materialet ska visas upp eller användas vid ett senare tillfälle.

### Sammanfattning

- Digitala objekt som ska bevaras kan vara objekt som är digitalt födda, digitaliserade för tillgängliggörande, eller digitaliserade av bevarandeskäl.
- Digitalt bevarande handlar om att säkra tillgänglighet av digitala objekt över lång tid. För detta behövs kunskaper om det digitala objektets olika bevarande-faser, faser som också har cykliska inslag.
- De bevarande-faser som har beskrivits i avsnittet följer bland annat OAIS-modellen (se avsnitt OAIS-modellen) och innehåller följande faser: skapandefasen, aktiva faser, förberedelsefasen, leveransfasen, mottagandefasen, bevarandefasen och tillgängliggörandefasen.

## Bevarandenaspekter

Ett sätt att beskriva utmaningarna vid digitalt bevarande är att tala om olika aspekter av bevarande. En sådan uppdelning förkommer i flera framställningar av digitalt bevarande men kan se lite olika ut från fall till fall<sup>21</sup>. I detta avsnitt presenteras fyra aspekter på bevarande: Lagring och bitbevarande, logiskt bevarande, innehållsmässigt bevarande och äkthetsbevarande. Aspekterna ringar in perspektiv som har avgörande betydelse för det digitala bevarandet och kunskap om dem ger en grundläggande förståelse för vad digitalt bevarande handlar om. En femte aspekt som ibland nämns är organisation. Organisationsaspekten behandlas i bilagan Verksamhetsmodell.

Kännedom om bevarandenaspekterna kan bland annat vara till nytta vid utarbetandet av strategier och processer för digitalt bevarande eller vid kravställning och utvärdering av olika alternativ för e-arkivering. Bevarandenaspekterna kan också vara bra att känna till vid införande och drift av e-arkivsystem. Vid bedömningen av ett konkret system kan det till exempel bli tydligt att stöd för tekniskt bevarande i form av konvertering av filformat saknas. Eller så kanske systemet saknar förmågan att skapa tillräckligt beskrivande metadata för att det innehållsliga bevarandet ska kunna utföras tillfredsställande. Vid upptäckter av detta slag blir det tydligt vad som behöver åtgärdas för ett lyckat digitalt bevarande. Likartade resonemang kan föras vid egenutveckling av system för digitalt bevarande.

### Lagring och bitbevarande

Lagring och bitbevarande har en tydlig koppling till databäraren, det vill säga det fysiska mediet som håller informationen. Bevarandet handlar här främst om att kopiera informationen i tid till en annan lämplig databärare, oavsett om det beror på risk för läsproblem på grund av fysiskt åldrande hos den tidigare databäraren eller risken för att den tekniska utvecklingen gör att det kommer att saknas utrustning som kan läsa den tidigare databäraren. Informationen ska utifrån denna bevarandenaspekt bevaras och kopieras i ursprungligt skick bitmässigt (en bit är den minsta räkneenheten i en dator, och har antingen värdet ett eller noll).

Om lagringen hanteras i något slags integrerat system, till exempel en upphandlad lagrings- eller e-arkivprodukt, gäller det att se till att systemet är uppbyggt med metoder som kan hantera de risker som nämnts ovan och att systemet i sig förnyas i den omfattning som behövs. Vissa läsfel på bitnivå brukar vara regel men dessa hanteras normalt automatiskt av läsutrustningen, till exempel genom att informationsinnehållet säkras genom checksummor. Lagringssystem brukar då notera och larma om mängden bitfel blir för hög, vilket exempelvis kan innebära att en hårddisk behöver bytas ut.

<sup>21</sup> <https://www.digime.fi/sv/langsiktigt-digitalt-bevarande/>

Nedan följer ett par delaspekter av lagring och bitbevarande.

#### Fysiskt bevarande

Det fysiska bevarandet handlar om hur länge databäraren är läslig med tanke på det fysiska åldrandet. Här gäller det att ringa in under hur lång tidsperiod en viss typ av databärare i en given förvaringsmiljö kan bedömas vara fysiskt beständig, samt när behov uppstår göra en kopiering av aktuella filer till nya databärare.

En metod som ofta används för att öka säkerheten kring det fysiska bevarandet är att bevara informationen i två eller flera kopior på olika databärare. Det kan då finnas en poäng i att välja databärare av olika fabrikat. Det har nämligen förekommit att en del tillverkare inte ordentligt kvalitetssäkrat ny materialteknik avseende långtidsbeständighet.

#### Tekniskt bevarande

Det tekniska bevarandet kan på ett sätt sägas vara en delaspekt av det logiska bevarandet (se nästa bevarandenaspekt i ordningen). Men eftersom det tekniska bevarandet har att göra med utvecklingen av hård- och mjukvara blir det också aktuellt för lagringen med dess tekniker för databärare och utrustning. Det handlar om att hela tiden kunna spela upp databärarna i relevant utrustning. Här, såväl som vid det fysiska bevarandet, handlar det om att kopiering ska ske i tid, men i detta fall specifikt om behovet av att gå över till mer aktuell lagringsteknik.

Värt att nämna är de kompatibilitetsproblem som kan uppstå mellan olika skrivare och läsare av en viss typ av datamedia. Genom att arbeta med kvalitetssäkring kan situationer undvikas där databärare inte fullt ut kan läsas av **annan** utrustning av till exempel annat fabrikat.

#### Logiskt bevarande

Det logiska bevarandet är en process för att säkerställa att materialet bevaras tekniskt förståeligt och läsbart oberoende av teknologiska förändringar. Nedan följer några delaspekter av logiskt bevarande.

#### Tekniskt bevarande

Denna delaspekt handlar om det tekniska bevarandet av filer, val av filformat och förmågan att hålla filerna läsbara i takt med den tekniska utvecklingen.

#### Migreringsstrategi

Migreringsstrategin, som har kommit att bli den dominerande tekniska strategin, bygger på att datafiler exporteras från ursprungssystemet och vid behov konverteras löpande för att de ska finnas i format som är bearbetningsbara med aktuell mjukvara. Normalt gäller att filerna redan vid export konverteras till standardiserade och spridda format som underlättar framtida konvertering och förlänger tidsfönstret innan konvertering blir nödvändig.

Systemens ursprungliga funktionalitet kan till viss del beskrivas genom metadata och utifrån dessa återskapas i tillgänglig programvara. Omfattningen av återskapandet varierar beroende på informations- och systemtyp. En viktig faktor är att de filer som skapas fullt ut följer aktuell standard och av denna anledning har olika validatorer<sup>22</sup> för filformat tagits fram. Det har också upprättats speciella register som stöd för val och hantering av filformat för bevarande, se avsnittet Vidare läsning.

<sup>22</sup> Se t.ex. validatorer på <https://libguides.bodleian.ox.ac.uk/digitalpreservation/validation>

### Emuleringsstrategi

Ett alternativ till migreringsstrategin är emuleringsstrategin, som bygger på att filerna bevaras i ursprungsformat, alltså som bitbevarande. Beroende på vilka alternativ som finns utförs sedan emulering i framtida tekniska miljöer. Ett första alternativ är att ny läsmjukvara framställs, ett andra alternativ att läsmjukvaran bevaras medan det gamla operativsystemet emuleras i den nya tekniska miljön och ett tredje alternativ att även operativsystemets mjukvara bevaras medan den gamla tekniska plattformen emuleras till den nya.

Emuleringsstrategin förordas ibland framför migreringsstrategin för att det genom den förra blir möjligt att undvika återkommande konverteringar där informationsförluster riskerar att ske. Strategin har också fördelen att mjukvara och funktionalitet kan bevaras, något som normalt inte sker med migreringsstrategin. Dock, trots ett antal föreslagna kompletterande metoder, har det ifrågasatts om emuleringsstrategin kommer att fungera långsiktigt med ett antal nödvändiga återkommande emuleringar allt eftersom tiden går.

Vilken typ av digitalt objekt som ska bevaras får också stor betydelse. Bevarande av dataspel lär främst ske genom någon form av emulering. Strategierna kan också mycket väl användas parallellt, det går att välja olika beroende på typ av digitalt objekt eller till och med använda båda för samma ursprungsobjekt för att få större möjligheter och säkerhet i bevarandet.

### Tekniska metadata

I och med genomslaget för OAIS-modellen, se avsnitt OAIS-modellen, har fokus allmänt flyttats något från det tekniska bevarandet till den metod som handlar om att bevara information och tillhörande metadata i så kallade paket. När paket för digitalt bevarande skapas på bästa möjliga sätt, samt inkluderar en stor mängd metadata som även beskriver tekniska förhållanden, kan ibland frågan om det tekniska bevarandet bli något skjuten på framtiden.

För att lära mer om metadata för digitalt bevarande, se avsnitt Metadata.

### Innehållsmässigt bevarande

Innehållsmässigt bevarande handlar om att informationen ska kunna förstås och tolkas nu och i framtiden. För detta räcker inte alltid att filerna kan läsas fullt ut, utan andra insatser kan krävas. Sättet att hantera frågan på sker primärt med hjälp av metadata.

### Koder och andra förklaringar

En delaspekt av det innehållsmässiga bevarandet är koder och andra förklaringar. Information som består av koder är oftast av äldre slag eftersom datalagringen var mycket dyrare förr och det var viktigt att spara in på varje bit. För att kunna tolka kod krävs kodöversättningar. Här sker alltid en bedömning gällande om en viss tolkning kan anses vara allmänt känd och hållbar på lång sikt. Exempel på detta är på vilket sätt postnummer översätts eller hur personnummer är uppbyggda och kontrollsiffrorna uträknade.

En annan fråga är om de begrepp som används i informationen är allmänt förstådda av den tilltänkta målgruppen, något som betonas av OAIS-modellen. Om du till exempel har en medicinsk databas som ska bevaras digitalt kan generella medicinska begrepp anses vara allmänt kända. Samtidigt är det viktigt att identifiera om ett begrepp till exempel enbart har använts i ett specifikt projekt och därför även behöver definieras i de metadata som tillhör informationen. Om målgruppen även består av personer i en fjärran framtid kanske exempelvis en beskrivning av det svenska språket behöva ingå i metadata. Troligen kommer då metadata om ett helt språk inte finnas i varje AIP utan samlat för delar av beståndet i ett e-arkiv.

Gränsdragningen mellan information och dess metadata kan ibland vara lite flytande. I exempelvis en relationsdatabas, det vill säga en databas där informationen är organiserad i tabeller, har ofta översättningar av koder egna tabeller i databasen och är därför en naturlig del av den samlade informationen. I andra fall kan en sådan översättning istället hanteras som och ses som en del av informationens metadata.

### Kontexten

Förutom den direkta tolkningen av informationen är också det sammanhang den uppstått i, kontexten, viktig att känna till för att fullt ut kunna förstå och bedöma informationen. Metadata som beskriver kontexten kan i högre eller lägre grad vara integrerade i informationsobjektet som sådant. Ofta behöver sådana metadata även bevaras parallellt med datafilerna.

För att lära mer om metadata för digitalt bevarande, se avsnitt Metadata.

### Äkthetsbevarande

Denna aspekt på digitalt bevarande handlar om att på olika sätt se till att informationen går att lita på - helt enkelt att den är äkta.

### Autenticitet och integritet

Två begrepp som ofta återkommer i sammanhanget är autenticitet och integritet. Integriteten handlar om att informationsinnehållet i en fil inte har ändrats eller förvanskats på något betydelsebärande sätt, till exempel genom att det inte har drabbats av några förändringar vid konvertering. Autenticitet handlar i sin tur om att den information som går att koppla till kontexten, såsom upphovsman, framställningstid och tillkomstförhållanden för filen, är känd och korrekt.

### Tekniska metoder

Äkthetsbevarande har ofta hanterats med olika tekniska metoder såsom checksummor och elektroniska signaturer. Dock är det svårt att se att dessa metoder kan ge några fullständiga garantier för äkthetsbevarandet. När konvertering av datafiler sker, som en del av det tekniska bevarandet, ändras ofta bitmönstret och ursprungliga checksummor slutar att fungera. Långsiktiga problem med certifikat och förändrade mjukvaruplattformar kan också försvåra användning av elektroniska signaturer.

### Allmän trovärdighet

Med anledning av brister i de tekniska metoderna har andra metoder för att bygga upp en allmän trovärdighet i det digitala bevarandet skapats. För att erhålla trovärdighet kring filernas äkthet är det till exempel viktigt att ha så bra och tydliga arbetsrutiner som möjligt. En ISO-standard som tar upp ett sådant arbete är ISO 16363/TDR (Trusted Digital Repositories), se bilagan Standarder.

### Sammanfattning

- Det finns flera olika aspekter på digitalt bevarande. I detta avsnitt har fyra viktiga aspekter tagits upp: lagring och bitbevarande, logiskt bevarande, innehållsmässigt bevarande samt äkthetsbevarande.
- Kännedom om olika bevarandeaspekter kan bland annat vara till nytta vid utarbetande av strategier och processer för digitalt bevarande eller vid kravställning och utvärdering av olika alternativ och system för e-arkivering.



## Bevarandeformat

Detta kapitel ger vägledning kring vad som är viktigt att tänka på när det gäller val av filformat för digitalt bevarande, vare sig det handlar om val av till exempel filformat vid skapande av digitalt födda dokument, vid konvertering från ett filformat till ett annat, eller val av format vid digitalisering av fysiska museiföremål.

### Grundläggande kriterier för val av filformat

Hur vet man att ett format är lämpligt för digitalt långtidsbevarande? Några grundläggande kriterier är att filformatet:

- följer en öppen standard och har publikt tillgängliga specifikationer, det vill säga att specifikationer och dokumentation för formatet inte är inlåsta av licenskostnader. Detta skapar förutsättningar för att filformatet ska kunna återskapas även om formatets ursprungliga skapare slutar stödja formatet i sina programvaror.
- är leverantörsberoende. Leverantörsberoende format kan bara öppnas och användas i en specifik leverantörs programvara. Detta kan medföra att licens för programvara behöver köpas för att formaten ska kunna öppnas. Om leverantören upphör med sin verksamhet kanske det heller inte finns någon annan leverantör på marknaden som står redo att ta över arbetet med att förvalta formatet och att erbjuda programvara som kan öppna formatet.
- är fritt från kryptering och DRM-kopieringsskydd (Digital Rights Management eller digital rättighetshandling).
- är vanligt förekommande. Detta ökar chanserna att formatet kommer att fortsätta användas och stödas av programvara.
- om möjligt är okomprimerade eller icke-destruktivt komprimerade (gäller bild, ljud och video). Destruktivt komprimerade format förstör gradvis informationsinnehållet genom att kvaliteten gradvis blir sämre varje gång filen konverteras om.

### Bedömning av lämpligt bevarandeformat

Val av filformat bör också göras utifrån de bevarandaspekter som tas upp i avsnittet Bevarandaspekter. Filformatet behöver till exempel vara långsiktigt hållbart och stödja funktionalitet som garanterar autenticitet och integritet. För att säkra att informationen i filen inte har ändrats kan ett filformat som stödjer skrivskydd behövas. En annan åtgärd är att se till att filformatet inte innehåller exekverbar, det vill säga av en processor körbar, kod. Detta förhindrar nämligen att skadlig kod läggs i filen och potentiellt förstör filens eller andra filers information.

Samtidigt som det finns krav på långsiktig hållbarhet hos informationen finns krav på att informationen ska vara tillgänglig för den tilltänkta målgruppen. I OAIS-modellen (se avsnitt OAIS-modellen) framgår att den tilltänkta målgruppen ska kunna förstå och använda informationen. Kraven på hållbarhet respektive tillgänglighet behöver inte alltid stå i motsättning till varandra, men det är bra att vara medveten om att alla krav på den långsiktiga hållbarheten inte säkrar tillgängligheten eller vice versa. Ibland krävs olika avvägningar när kraven ser olika ut och ibland behöver exempelvis två olika versioner av samma information bevaras men i olika filformat.

För varje specifik informationstyp bör en bedömning göras kring vilket bevarandeformat som är lämpligt. Ibland kan olika format vara lämpliga för samma typ av informationsmängd, beroende på hur informationen är tänkt att användas eller förstås.

Som stöd i bedömningen av formatval kan följande frågor ställas:

### **Vad är syftet med informationen?**

Vad ska informationen användas till? Ska den gå att bearbeta om till nya versioner? Ska den kunna analyseras och sammanställas maskinellt? Är den tilltänkta användaren en privatperson som ska läsa eller titta på informationen på sin dator och bara har tillgång till de vanligaste programmen för att kunna öppna filer? Eller är den tilltänkta användaren en programvara som ska analysera och tolka informationen maskinellt? Är informationen tänkt att vara tillgänglig via webben? Ska den kunna ersätta ett fysiskt original som kopia?

### **Hur mycket av informationen ska bevaras och i vilken kvalitet?**

Hur mycket av den innehållsmässiga informationen som kan bevaras styrs av hur mycket fysiskt lagringsutrymme som finns och hur mycket data det finns resurser till att lagra. Vid stora datamängder påverkas givetvis behovet av lagringsutrymme. Om filformaten dessutom ger hög återgivning av och hög kvalitet på informationen kommer datamängden att öka. I sådana fall kan det bli tvunget för en verksamhet att kompromissa med kvaliteten på åtminstone vissa av filerna. Huvudsaken är att varje verksamhet tar medvetna beslut utifrån egna förutsättningar för att hitta en för sin verksamhet acceptabel nivå.

Återigen är det viktigt att utgå ifrån den tilltänka användarens behov. Om informationen är av sådan karaktär att den till exempel ska användas av specialister i forskningssyfte, eller att den behöver ha så pass stor upplösning att användare kan göra uppförstoringar och trycka upp stora fysiska exemplar av informationen, kanske det inte är möjligt att göra avkall på kvaliteten. Då måste helt enkelt en större filstorlek och ett större lagringsutrymme (eller en mer komplicerad datamodell) accepteras. I andra fall kanske det räcker att informationens bildupplösning och kontrast är så pass bra att användaren kan se och läsa texten eller bilden på sin datorskärm. Då kan också filstorleken göras mindre.

### **Är informationen maskinberoende eller programberoende?**

Om proprietär programvara används och det krävs licenser och särskilda program för att öppna eller visa informationen påverkar detta den tilltänkta användarens möjligheter att kunna ta del av den. Vid uppdateringar av programvaran kan dessutom formatet ha hunnit bli föråldrat, vilket gör att det i efterhand blir kostnadskrävande att försöka återskapa informationen. I vissa fall behöver helt enkelt ett byte av bevarandeformat för informationen göras.

Ibland är informationen tänkt att användas i applikationer som kräver att den är lagrad i ett specifikt format, för att exempelvis kunna bearbetas maskinellt genom strukturerade data. Vid konvertering till format lämpliga för långtidsbevarande kan då denna möjlighet gå förlorad. Är så fallet är det bra att ta reda på om det finns andra format som lämpar sig bättre för visning av informationen för den tilltänkte användaren.

### **Är informationsförlust tillåtet och i vilken omfattning?**

Är det möjligt att avvara funktionalitet i ett system, sammanställningsmöjligheter i en databas, eller lagerinformation i bilder? Eftersom konvertering kan innebära informationsförlust är det viktigt att vara medveten om att detta innebär gallring där formella regler kan gälla.

### **Bevarandeformat i praktiken**

Området filformat är ett område under ständig förändring. Filformat föråldras, formatspecifikationer förändras och stödet för filformat i programvara kan ändras eller upphöra. Du som läser vägledningen rekommenderas att undersöka vilka standarder och rekommendationer som finns inom det område som rör de typer av information som förvaltas hos just din organisation.

Exempelvis behöver statliga myndigheter följa Riksarkivets föreskrifter rörande bevarandeformat (se avsnitt Regelverk). I bilaga Vägledning för val av bevarandeformat hittar du förslag på bevarandeformat för olika typer av information som ska bevaras. Bilagan är inte en heltäckande guide för området filformat, utan ska ses som en guide till hur de generella principerna för val av filformat kan tillämpas.

### Sammanfattning

- Använd filformat som har öppna specifikationer, är leverantörsoberoende och vanligt förekommande och om möjligt okomprimerande eller icke-destruktivt komprimerande.
- Ha alltid den tilltänkta målgruppen i åtanke vid val av format. Hur ska informationen kunna användas och förstås av användarna i framtiden?
- Det finns organisationer, projekt och initiativ som arbetar specifikt med filformatfrågor. Undersök om de större institutionerna inom ditt område har rekommendationer och vägledningar att följa.

### Vidare läsning

Den här vägledningen bygger i stora delar på råd och rekommendationer från internationella organisationer så väl som på etablerad litteratur inom digitalt bevarande. Tre skrifter som har använts särskilt i framtagandet av vägledningen är Adrian Browns *Practical Digital Preservation: A how-to Guide for Organizations of Any Size*, Facet Publishing, London, 2013, Digital preservation Coalitions *Digital Preservation Handbook*<sup>23</sup> samt National Digital Stewardship Alliances *Levels of Digital Preservation*<sup>24</sup>, vilka är en utmärkt ingång till vidare och fördjupad läsning inom området digitalt bevarande. *Levels of Digital Preservation* är en resurs för alla som arbetar med digitalt bevarande och som vill utvärdera och vidareutveckla detta arbete. Som stöd för användning av *Levels of Digital Preservation* finns en översättning till svenska av dess grundläggande tabell som bilaga.

I resten av detta avsnitt listas organisationer som har en tongivande roll på området samt andra resurser som kan vara av värde i ert arbete med digitalt bevarande.

#### Organisationer

Preserving digital Objects With Restricted Resources (Digital POWRR) 25 är ett projekt finansierat av bland annat Institute of Museums and Library services 26 där ett stödmaterial tagits fram som riktar sig till framför allt små och medelstora minnesinstitutioner med små resurser. POWRR bidrar även med material till COPTR.

Open preservation foundation (OPF) är en fristående ideell organisation som skapades 2010 för att förvalta resultaten av det EU-finansierade projektet Planets. OPF har sedan Planets-projektets avslut varit en del av flera andra EU-finansierade projekt så som EARK, SCAPE och PREFORMA. Förutom att

<sup>23</sup> <https://www.dpconline.org/handbook>

<sup>24</sup> <https://nds.org/publications/levels-of-digital-preservation/>

25 <https://digitalpowrr.niu.edu/>

26 <https://www.ims.gov>

utveckla och förvalta verktyg inom digitalt bevarande som släpps under fria och öppna licenser erbjuder OPF ett rikt utbud av webinarier och annat utbildningsmaterial.

Digital preservation Coalition (DPC) <sup>27</sup> är en fristående ideell organisation som skapades 2002 som ett samarbete mellan myndigheter i Storbritannien och Irland men som idag har växt till ett internationellt samarbete med medlemsorganisationer som FN, Europeiska Centralbanken, CERN, Library of Congress och flera framstående universitet. DCP erbjuder fritt ett rikt utbud av stöd- och utbildningsmaterial, både för dig som vi lära dig mer om digitalt bevarande och för dig som vi lära andra. Bland annat har DPC tagit fram Digital Preservation Handbook som använts för delar av den här vägledningen.

Connecting Europe Facility (CEF Digital) <sup>28</sup> CEF är finansieringsmekanismen för genomförandet av de transeuropeiska näten inom områdena transport, energi och digitalisering och består av ett antal byggnadsblock där eArchiving är ett av dessa som baserats på resultaten från EARK-projektet (2014-2017) Byggnadsblocket erbjuder bland annat referensapplikationer <sup>29</sup> som stödjer de EU-gemensamma specifikationerna för digitalt bevarande som i sin tur förvaltas av DILCIS Board (The Digital Information LifeCycle Interoperability Standards Board) en organisation vars styrelse utses av DLM-forums årsmöte.

Document Lifecycle Management Forum (DLM-Forum) är en europeisk medlemsorganisation för offentliga arkiv och andra intressenter så som leverantörer av system för digitalt bevarande. Flera europeiska nationalarkiv är medlemmar men även mindre institutioner så som svenska Sydarkivera och Norska KDRS (Kommunearkivinstitusjonenes Digitale Ressurscenter)

National Digital Stewardship Alliance (NDSA) är ett konsortium med i huvudsak amerikanska medlemsorganisationer. NDSA startade sin verksamhet år 2010 på initiativ av Library of Congress. NDSAS huvudsakliga mål är att stötta sina medlemsorganisationer i professionsutveckling och adressera gemensamma utmaningar med digitalt bevarande vilket bland annat görs genom att ta fram rapporter och utbildningsmaterial. Bland annat har NDSA tagit fram Levels of digital preservation (LOP) som använts för delar av denna vägledning.

Digital Curation Centre (DCC) är ett resurscenter som tillhandahåller råd och praktisk stöd i digitalt bevarande av forskningsdata i form av fritt tillgängliga verktyg, vägledningar och utbildningsmaterial med teman som informationsvärdering, licensiering och digitalt långtidsbevarande.

Svensk Nationell Datatjänst (SND)

Svensk nationell datatjänst är en forskningsinfrastruktur som hjälper forskare att bevara, underhålla och dela forskningsdata på ett säkert och hållbart sätt. SND inrättades 2008 på initiativ av Vetenskapsrådet. Sedan 1 januari 2018 bedrivs verksamheten som ett konsortium under ledning av Göteborgs universitet. SND erbjuder en mängd stödmaterial för god datahantering och hur du rent praktiskt går till väga för att strukturera, bevara och tillgängliggöra data.

## Andra resurser

### COPTR

---

<sup>27</sup> <https://www.dpconline.org>

<sup>28</sup> <https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eArchiving>

<sup>29</sup> <https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/Sample+Software+Portfolio>

Community Owned digital Preservation Tool Registry (COPTR)<sup>30</sup> beskriver användbara verktyg för långsiktigt digitalt bevarande.

#### **Digitalbevaring.dk**

Webbsajten Digitalbevaring.dk<sup>31</sup> är en dansk kunskapsbank kring digitalt bevarande.

#### **PRONOM**

PRONOM<sup>32</sup> är ett filformatsregister som har tagits fram av brittiska National Archives.

#### **Recommended Formats Statement - Resources (Preservation, Library of Congress)**

Library of Congress rekommenderade format<sup>33</sup> för olika materialtyper, uppdelat i föredragna format (uppnår alla sju "sustainability factors") och godtagbara format (en minimumnivå som eventuellt inte uppnår alla krav).

#### **Sustainability of Digital Formats: Planning for Library of Congress Collections**

Library of Congress katalog<sup>34</sup> beskriver digitala filformat och innehåller bedömningar av formatens beständighet utifrån sju kriterier, så kallade "sustainability factors".

## Bilaga: Begrepp

Nedan finns en lista med begrepp och termer som är vanliga när det gäller digitalt bevarande och som tas upp i den här vägledningen. Se även Digital Preservation Coalitions ordlista<sup>35</sup>.

#### **AIP** (engelska Archival Information Package)

Arkivpaket som bevaras i ett OAIS-arkiv.

#### **Administration**

Har den löpande kontrollen över övriga funktioner i OAIS enligt OAIS funktionsmodell.

#### **Autenticitet**

Äkthet. Egenskap som styrker vad informationen är, dess skapare och tidpunkt. Som nämnts i inledningen av den här vägledningen så innebär inte autenticitet att informationen när den skapades är sann.

---

<sup>30</sup> <https://coptr.digipres.org/>

<sup>31</sup> <https://digitalbevaring.dk/>

<sup>32</sup> <https://www.nationalarchives.gov.uk/PRONOM>

<sup>33</sup> <https://www.loc.gov/preservation/resources/rfs/>

<sup>34</sup> <https://www.loc.gov/preservation/digital/formats/fdd/descriptions.shtml>

<sup>35</sup> <https://dpconline.org/handbook/glossary>

**Bitbevarande**

Bevarande av delarna (bits) i en digital resurs, alltså själva data-lagringen. Bitbevarande är en del av det digitala bevarande, men inte allt.

**Container-format**

Container-filformat kan liknas vid en låda som innehåller bild, ljud och/eller metadata. Det går också att säga att det är en filändelse, som till exempel AVI, MOV och MKV.

**De facto-standard**

Standard som fått bred spridning och tillämpas utan att ha föregåtts av en formell överenskommelse eller annan rättsgrund, och alltså heller inte har genomgått en process hos ett standardiseringsorgan. En de facto-standard kan ändå förvaltas och utvecklas av olika typer av organisationer men det är inget krav.

**Digitalt fött material**

Digitalt material som inte har en analog förlaga.

**DRM** (engelska Digital Rights Management)

Samlingsbenämning av olika tekniker som förhindrar att digitala objekt och program olovligt används och kopieras.

**DIP** (engelska Dissemination Information Package)

Informationspaket för leverans från ett OAIS-arkiv.

**Emulering**

En strategi för digitalt bevarande där det digitala objektet genom att man skapar ett gränssnitt eller ett program kan visa eller spela upp det i originalformatet. Emulering som strategi kantas av olika utmaningar i dagsläget, bland annat av upphovsrätt.

**FAIR**

FAIR-principerna<sup>36</sup> (Findable, Accessible, Interoperable och Reusable) innebär att forskningsdata ska gå att hitta, det ska finnas information om hur man får tillgång till dem, de ska vara kompatibla med andra data, och de ska vara möjliga att återanvända. FAIR-principerna spelar en viktig roll i arbetet för öppen vetenskap och beskriver några av de mest centrala riktlinjerna för god datahantering och öppen tillgång till forskningsdata.

**Förstörande kompression** (på engelska Lossy Compression)

Fil-storlek reduceras, originaldata tas bort.

**Gallring**

Arkiv-material tas bort från arkivet och förstörs.

Förstöra information eller vidta andra åtgärder med informationen som medför förlust av betydelsebärande data, möjliga sammanställningar, sökmöjligheter, eller möjligheter att bedöma handlingarnas autenticitet

**Ingest** (på svenska Mottagning)

Den process i ett OAIS-arkiv som gör om en SIP till en AIP.

<sup>36</sup> <https://www.force11.org/group/fairgroup/fairprinciples>

**Information om bevarande PDI** (engelska Preservation Description Information)

Information som behövs för ett adekvat bevarande kan sägas vara enligt OAIS-modellen - ursprung, referens, beständighet, kontext och åtkomsträttigheter.

**Integritet**

Riktighet. Egenskap att informationsinnehållet är fullständigt och inte har ändrats, till exempel vid filkonvertering i bevarande-syfte.

**Konsument** (engelska Consumer)

Person/system som interagerar med OAIS tjänster för att hitta bevarad information.

**Checksumma/kontrollsumma** (engelska Checksum)

Checksummor/kontrollsummor för filer räknas fram genom algoritmer och används för att kontrollera dem och jämföra kopior så att de inte har förändrats och därmed blivit förstörda.

**Konvertering**

Överföring av digitala filer från gamla till nya filformat.

**Ledning** (på engelska Management)

Inom OAIS funktionsmodell fastställer Ledning övergripande regler och ansvarar för bland annat finansiering.

**Identifierad målgrupp** (på engelska Designated Community)

Identifierad målgrupp i OAIS-sammanhang är de så kallade konsumenter som ska förstå en specifik uppsättning information.

**Icke-förstörande kompression** (på engelska Lossless Compression)

Filstorlek reduceras men all originaldata är kvar.

**Migrering/Migration**

Löpande konvertering av digitala filer till nya databärare, filformat och system i takt med att de befintliga blir omoderna, föråldrade eller oläsbara.

**Normalisering**

I digitalt bevarande-sammanhang menas med normalisering konvertering av fil-format eller metadata till förutbestämda format innan eller samband med att mottagande i ett bevarande-system.

**Open source**

Öppen källkod eller programvara.

**Outsourcing**

Företag och organisationer låter andra sköta någon del av verksamheten.

**Producent** (på engelska Producer)

I OAIS funktionsmodell är Producent person/system som tillhandahåller den information som ska bevaras.

**Proprietär**

Motsatsen till öppen standard.

**SIP** (på engelska Submission Information Package).

Informationspaket för leverans till ett OAIS-arkiv.

**Trovärdighet**

Ett trovärdigt digitalt arkiv har till uppdrag att över lång tid ge tillgång till hanterade digitala resurser

till dess användare. Termen "Trustworthy Digital Repository" används som ett samlande begrepp för magasinssystem som certifieras utifrån bland annat ett formellt certifieringsramverk, baserat på ISO 16363 och DIN 31644.

**Validering** (på engelska Validation)

Utvärdering/kontroll av att något verkligen har de egenskaper som det uppges eller förväntas ha. Validering kan syfta både på själva kontrollen och på intygandet av att kraven är upp-fyllda.

## Bilaga: Standarder

Syftet med standarder är att skapa enhetliga och transparenta rutiner som vi kan enas kring och kan fungera som ett verktyg för att leva upp till de krav som ställs i lagar och andra regelverk<sup>37</sup>. Standarder kan hjälpa dig och din organisation att göra rätt, samtidigt som de ibland kan vara resurskrävande att följa fullt ut. Varje verksamhet behöver därför överväga om, och i vilken utsträckning, olika standarder ska följas.

Standardiseringsorgan såsom Internationella standardiseringsorganisationen (ISO) tar fram standarder på en rad områden. Därutöver finns så kallade de facto-standarder som är standarder som fått bred spridning och tillämpas utan att ha föregåtts av en formell överenskommelse eller annan rättsgrund, och alltså heller inte har genomgått en process hos ett standardiseringsorgan.

Öppna standarder är standarder som, i motsats till en proprietär teknisk specifikation, tillåter vem som helst att implementera den utan att ägaren av standarden sätter upp orimliga eller diskriminerande hinder.

Nedan följer, i bokstavsordning, en översikt av de standarder som vanligen används inom digitalt bevarande.

**BAGIT**

BAGIT<sup>38</sup> kan användas för att packa ihop, överföra och lagra digitalt material.

**CCO** (Cataloging Cultural Objects)

CCO<sup>39</sup> är en katalogiseringsstandard för kulturarvsmaterial.

**CIDOC-CRM** (Conceptual Reference Model)

CIDOC Conceptual Reference Model<sup>40</sup> är en konceptuell modell, ontologi och ISO-standard (ISO 21127) för semantisk integration av digital kulturarvsinformation.

<sup>37</sup> <https://www.sis.se/standarder/vad-ar-en-standard/>

<sup>38</sup> <https://www.ietf.org/rfc/rfc8493.txt>

<https://blogs.loc.gov/thesignal/2019/04/bagit-at-the-library-of-congress/>

<sup>39</sup> <https://vraweb.org/resources/cataloging-cultural-objects/>

<sup>40</sup> <http://www.cidoc-crm.org/>



### **Dublin Core**

Dublin Core<sup>41</sup> är en metadatastandard för beskrivning av informationsresurser. Dublin Core består av femton beskrivande element och den används ofta som en slags minsta gemensamma nämnare.

### **EAC (Encoded Archival Context)**

EAC<sup>42</sup> är en XML-standard för auktoritetsposter som bygger på den internationella standarden ISAAR-CPF.

### **EAD (Encoded Archival Description)**

EAD<sup>43</sup> är en XML-standard för arkivbeskrivningar som bygger på den internationella standarden ISAD-G.

### **FGS (Förvaltningsgemensamma specifikationer)**

FGS:er<sup>44</sup> har i Sverige blivit ett sätt att ta fram specifikationer av innehåll och struktur för överföring och bevarande av digital informationen. FGS:erna bygger oftast på existerande standarder och handlar mycket om metadata gällande olika informationstyper.

### **ISAAR-CPF (International Standard Archival Authority Record for Corporate Bodies, Persons and Families)**

ISAAR-CPF<sup>45</sup> är en beskrivningsstandard.

### **ISAD-G (International Standard Archival Description)**

ISAD-G<sup>46</sup> definierar element som ska ingå i ett arkiveringshjälpmedel.

### **ISDF (International Standard for Describing Functions)**

ISDF<sup>47</sup> är en standard för att beskriva funktioner.

### **ISO30301 Ledningssystem för verksamhetsinformation**

Standarden ISO30301<sup>48</sup> anger krav som måste uppfyllas av ett ledningssystem för verksamhetsinformation.

### **ISO15498 Dokumentation - Hantering av verksamhetsinformation**

Standarden ISO15498<sup>49</sup> tar upp grunder och principer gällande hantering av verksamhetsinformation.

---

<sup>41</sup> <https://dublincore.org/>

<sup>42</sup> <https://eac.staatsbibliothek-berlin.de/>

<sup>43</sup> <https://www.loc.gov/ead/>

<sup>44</sup> <https://riksarkivet.se/intro-fgs>

<sup>45</sup> <https://www.ica.org/en/isaar-cpf-international-standard-archival-authority-record-corporate-bodies-persons-and-families-2nd>

<sup>46</sup> <https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>

<sup>47</sup> <https://www.ica.org/en/isdf-international-standard-describing-functions>

<sup>48</sup> <https://www.sis.se/produkter/terminologi-och-dokumentation/informationsvetenskap-publicering/informationsvetenskap/ss-iso-303012019/>

<sup>49</sup> <https://www.sis.se/produkter/terminologi-och-dokumentation/informationsvetenskap-publicering/informationsvetenskap/ssiso1548912016/>

### **ISO 16363/TDR** (Trusted Digital Repository Checklist)

ISO 16363/TDR<sup>50</sup> och föregångaren TRAC (Trustworthy Repositories Audit & Certification)<sup>51</sup> handlar om certifiering eller självgranskning av ett OAIS-baserat e-arkiv. Utveckling av metoden har främst skett inom amerikanska biblioteksvärlden med bistånd från amerikanska nationalarkivet men också resulterat i implementeringar på europainivå.

### **Ledningssystem för kvalitet**

Se bilaga Verksamhetsmodell som utgår från grundläggande delar i standarden Ledningssystem för kvalitet - Krav (ISO 9001:2015)<sup>52</sup>

### **MARC** (Machine-readable cataloging)

MARC<sup>53</sup> är en standard för katalogisering inom biblioteksväsendet.

### **METS** (Metadata Encoding and Transmission Standard)

METS<sup>54</sup> är en de facto-standard som används för att paketera beskrivande, administrativa och strukturella metadata för digitala objekt (för information om metadata typer, se avsnitt Metadata). METS är framtagen för användande inom biblioteksvärlden, men används även inom arkivvärlden - oftast vid skapande av SIP- och AIP-paket som ska överföras eller lagras i ett OAIS-system. Riksarkivets och Kungliga bibliotekets FGS:er innehåller specifikationer för SIP-paket och är bland annat uttryckta enligt METS-standarderna.

### **MODS** (Metadata Object Description Schema)

MODS<sup>55</sup> är en standard och ett schema för bibliografiska metadata.

### **OAI-PMH** (Open Archives Initiative Protocol for Metadata Harvesting)

OAI-PMH<sup>56</sup> är ett protokoll som används för överföring mellan olika "repositories".

### **OAIS** (Oberoende arkivinformatonssystem)

OAIS<sup>57</sup> är en standard och referensmodell för digitalt bevarande. OAIS är både antagen som internationell och svensk ISO-standard men finns också fritt tillgänglig via The Consultative Committee for Space Data Systems (CCSDS) som rekommenderad praxis (Magenta book CCSDS 650.0-M-2).

### **PAIMAS** (Producer Archive Interface Methodology Abstract Standard)

PAIMAS<sup>58</sup> är en ISO-standard som handlar om de första stegen i mottagningen (på engelska Ingest) i ett OAIS-arkiv.

### **PAIS**

PAIS<sup>59</sup> är en ISO-standard som stödjer effektiv överföring och validering av SIP-data.

<sup>50</sup> <https://www.iso.org/standard/56510.html>

<sup>51</sup> <https://www.crl.edu/archiving-preservation/digital-archives/metrics>

<sup>52</sup> <https://www.sis.se/produkter/foretagsorganisation/foretagsorganisation-och-foretagsledning-ledningssystem/ledningssystem/sseniso90012015/>

<sup>53</sup> <https://www.loc.gov/marc/>

<sup>54</sup> <https://www.loc.gov/standards/mets/>

<sup>55</sup> <http://www.loc.gov/standards/mods/>

<sup>56</sup> <https://www.openarchives.org/pmh/>

<sup>57</sup> <https://www.iso.org/standard/57284.html>

<sup>58</sup> [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=39577](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39577)

<https://public.ccsds.org/Pubs/651x0m1.pdf>

<sup>59</sup> <https://www.iso.org/standard/67056.html>

#### **PREMIS** (Preservation Metadata: Implementation Strategies)

PREMIS60 är en standard och specifikation för hur OAIS-modellens konceptuella struktur för informationsobjekt och arkivpaket kan implementeras i form av metadata gällande bevarandet. Den har fått ett utbrett användande och har byggts in i både kommersiella och open source-system, och kan ses som en de facto-standard.

#### **Spectrum**

Spectrum61 är en standard för samlingsförvaltning som är fritt tillgänglig. Standarden kan hjälpa museer eller andra organisationer med museiliknande samlingar med samlingsförvaltning. Spectrum är indelat i 21 processer - allt från dagliga arbetsuppgifter till saker som görs då och då. I Storbritannien, där standarden först togs fram, används den av alla ackrediterade museer.

#### **XML** (Extensible Markup Language)

XML<sup>62</sup> är en specifikation för att märka upp data.

## Bilaga: Val av bevarandeformat

Detta avsnitt innehåller generella råd och resonemang kring hur de generella principerna som beskrivs i avsnittet Bevarandeformat kan tillämpas. Området filformat är ett område under ständig förändring. Filformat föråldras, formatspecifikationer förändras, och stödet för filformat i programvara kan också ändras eller upphöra. Därför anges oftast inte specifika versioner av filformat i denna vägledning. Rekommendationerna i denna vägledning är inte heltäckande, utan kan ses som en utgångspunkt för att sätta egna krav för bevarandeformat.

Läsaren rekommenderas att undersöka vilka standarder och rekommendationer som finns inom det område som rör de typer av information som förvaltas inom läsarens organisation. Exempelvis så behöver statliga myndigheter följa Riksarkivets föreskrifter rörande bevarandeformat.

Formaten har delats in i övergripande grupper utifrån originalinformationens form och användningsområden.

För mer läsning kring bevarandeformat, se avsnitt Vidare läsning.

#### **Bild**

Valet av format för bilder påverkas av hur bilderna är tänkta att kunna användas. Ett långsiktigt hållbart format med möjlighet till återgivning i hög kvalitet (till exempel detaljåtergivning eller färgomfång) är önskvärt i bevarandesyften. Men bilder används ofta i tillgängliggörandesyfte till exempel genom publicering på nätet, och då kommer andra aspekter in i bedömningen, till exempel filstorlek, användbarhet, kompatibilitet. Olika filformat kan vara mer eller mindre lämpliga för till exempel publicering på nätet, publicering i tryck, bearbetningar, eller om filen är tänkt att kunna ersätta en analog bild vid förlust av originalet. Det är därför vanligt att välja att spara bilden i flera versioner i flera filformat för att representera samma bild för att uppfylla olika syften och användningar. Originalfilens filformat och kvalitet påverkar också valet av målformat för bevarande.

<sup>60</sup> <https://www.loc.gov/standards/premis/>

<sup>61</sup> <https://www.raa.se/museer/samlingsforvaltning/sa-fungerar-samlingsforvaltning/spectrum-en-standard-for-samlingsforvaltning/>

<sup>62</sup> <https://www.w3.org/XML/>

Till exempel, en vektoriserad bild som omvandlas till en rasterbild kommer att innebära en informationsförlust, och en bildfil som redan har sparats i ett komprimerande format kommer inte att öka i kvalitet av att konverteras till ett okomprimerande format.

Ursprungsformat	Ursprungsanvändningsområde	Resonemang kring format för bevarande
<p>Råformat för fotografisk bild (digitalt född)</p> <p>Exempel: 3FR (Hasselblad), DCR, K25, KDC (Kodak), CRW CR2 CR3 (Canon), ERF (Epson), NEF (Nikon), ORF (Olympus), PEF (Pentax), RW2 (Panasonic) and ARW, SRF, SR2 (Sony). DNG (Adobe)</p>	<p>Råformatfiler fyller samma funktion som negativ inom analogt fotografi: det vill säga att negativet är inte direkt användbart som bild utan innehåller informationen för att skapa bilden. Termen "framkallas" används också för att beskriva konverteringen av råformatfiler till ett format som går att visa. Syftet med råformatfilen är att med minimal informationsförlust spara all data som kamerans sensor har fångat, tillsammans med metadata från fototillfället.</p>	<p>Råformat kräver oftast proprietär programvara för att kunna öppnas. Därför bör ett annat format väljas för bevarandesyften, till exempel TIFF. Vid "framkallning" av råformat läses bilden i ett visst läge vilket innebär viss informationsförlust, till exempel i färgomfång. För att vara säker på att behålla all data, samt behålla möjligheter att skapa nya bilder utifrån rådatat, kan även råformatfilen sparas men det är viktigt att vara medveten om att hållbarheten ur ett långtidsbevarandeperspektiv inte är garanterad.</p>
TIFF	TIFF-filer används ofta som masterfiler vid bildbehandling, och som filformat för bevarande då det är ett okomprimerande filformat.	TIFF är ett lämpligt bevarandeformat då det är okomprimerande. Filstorleken för TIFF är dock hög, så om lagringsytan är begränsad kan detta vara något att ta i beaktande vid val av format. TIFF kan också innehålla flera lager, för att spara utrymme går det att välja att inte spara information om lager. TIFF-formatet tillåter inställningar för komprimering, och LZW komprimering kan användas utan större informationsförlust för att minska filstorleken. För allra högsta kvalitet gällande bitdjup rekommenderas 16 bitar, men även här går det att göra andra avväganden kring lagringsyta och användbarhet som kan göra att även 8 bitar kan ge tillräcklig kvalitet.
JPG	JPG används ofta som visningsbilder, både med hög upplösning såväl som låg upplösning (till exempel tumnaglar). JPG kan också vara	JPG-filer används till exempel för att skapa påsiktskopior av filer som har TIFF som originalformat.

	originalformat för fotografiska bilder eller analoga bilder som har digitaliserats. JPG-filer är mindre än TIFF-filer och är därför enklare att dela (till exempel som e-postbilaga eller posta på nätet).	JPG går att använda som bevarandeformat (vilket till exempel är lämpligt när originalformatet är JPG), men formatet är bildförstörande vilket betyder att dess komprimeringsmetod ger reducerad bildkvalitet.
PNG	PNG-filer har stöd för att bibehålla transparens i bilder. Det finns två varianter av PNG. PNG-8 lämpar sig bäst för enkel grafik med få färger, medan PNG-24 är jämförbart med JPG.	Om originalformatet är PNG kan detta användas även som bevarandeformat.
Rastergrafikformat (övriga format än TIFF, JPG, PNG) Exempel GIF, BMP med flera	I rastergrafik är bilder uppbyggda av rader med bildpunkter. Bildkvaliteten påverkas av vilken upplösning bilden har. En rasterbild kan inte skalas upp till en högre upplösning och behålla samma bildkvalitet, till skillnad från till exempel vektorgrafik. Rastergrafik brukar användas för foton och fotorealistiska bilder, men även till exempel grafiska element.	TIFF, JPG och PNG som tidigare nämnts, är rastergrafikformat som anses vara godkända bevarandeformat enligt Riksarkivet. Men det finns fler format som också är rastergrafikformat. Vid bevarande av dessa bör det alltså göras en bedömning av vilket bevarandeformat man ska konvertera till, om konvertering är möjlig.  Vid val av format för konvertering bör det bedömas bilden innehåller transparens eller lager som behöver bibehållas (transparens stöds ej av till exempel JPG), samt bildens storlek och kvalitet när format för bevarande väljs. Till exempel ökar behovet av lagringsutrymme om valet görs att spara i TIFF. Om en bild av redan låg kvalitet sparas i TIFF-format har heller inga vinster gjorts i kvalitet, men däremot ökar filstorleken.  När en bild ingår i en struktur tillsammans med annan information, till exempel en webbsida med länkar till bilder, är det viktigt att ha i åtanke att länken till den konverterade filen behöver

		ändras för att länken ska fortsätta fungera.
<p>Format för bildbehandlingsprogram</p> <p>Exempel: Photoshop (PSD), Adobe Illustrator (AI), Corel Draw (CDR), GIMP (XCF), AutoCAD (DWG) se även råformat från kamera</p>	<p>Bildbehandlingsprogram har ofta egna filformat som innehåller mer information eller möjlig funktionalitet än de format som rekommenderas för bevarande. Det är till exempel vanligt att det går att redigera lager eller lägga på effekter på bildlager som sedan kan justeras eller döljas. Ibland är det möjligt att blanda vektorgrafik och rastergrafik, eller lägga in textlager. I vissa program kan du också logga och spara ändringar som har gjorts under redigeringen. Filstorlekarna för dessa typer av filer kan bli väldigt stora på grund av all information de innehåller.</p>	<p>Bildbehandlingsprogrammets egna format kräver oftast proprietär programvara för att filerna ska kunna öppnas. Därför bör ett annat format väljas för bevarandesyften. Konvertering av bildbehandlingsformaten innebär viss informationsförlust, till exempel information om lager, möjlighet att blanda rastergrafik och vektorgrafik, eller möjlighet att backa fram och tillbaka i versioner.</p> <p>Vilket målformat som väljs beror till exempel på bildbehandlingsformatets möjligheter till konvertering, om det är en rasterbild eller vektorbild, om bilden innehåller transparens eller lager som behöver bevaras.</p>
<p>Vektorgrafik</p> <p>Exempel SVG, EPS, PS (PostScript), PDF, AI, SWF, DWG, DXF, DNG, WMF, CGM, CDR</p>	<p>Vektorgrafik består av geometriska primitiver, såsom punkter, linjer, cirklar, till skillnad från rastergrafik (bitmaps) som enbart består av bildpunkter/pixlar. Vektorgrafik är inte beroende av skärmens upplösning utan kan skalas upp och skalas ner utan att kvaliteten på grafiken ändras.</p> <p>Bland de uppräknade filformaten tillåter flera både vektor- och rastergrafik.</p>	<p>När det gäller att konvertera vektorbaserade format till bevarandeformat för vektorbaserad grafik gäller samma förhållningsregler som för rastergrafik. Det vill säga, att hänsyn behöver tas till vad bildens syfte är och vilka egenskaper man är beroende av att kunna bevara för att bilden ska kunna användas eller uppfattas så som den är tänkt. Till exempel, information om lager eller färger. Samt om formatet är beroende av kommersiell programvara för att gå att öppna eller om det har öppna specifikationer. Behöver vektorgrafiken bevaras i ett vektoriserat format eller kommer innehållet vara användbart även som rasterformat? Exempel på information som bör bevaras i vektoriserad form är typsnitt.</p>

<p>SVG (Scalable Vector Graphics)</p>	<p>SVG, är ett XML-baserat vektorgrafik-format för tvådimensionella bilder som stöder animationer och interaktivitet. SVG-specifikationen är en öppen standard och har utvecklats av W3C-konsortiet sedan 1999.</p> <p>SVG-bilder och deras egenskaper definieras i XML-textfiler.</p> <p>SVG tillåter tre typer av grafikobjekt: vektorgrafik, rastergrafik och text.</p>	<p>SVG brukar rekommenderas som bevarandeformat för vektorgrafik. Det är ett öppet icke-proprietärt format som har fått en utbredd användning, till exempel på webben. Att det är XML-baserat gör att det går att använda enkla textredigerare för att skapa och redigera filerna.</p> <p>Formatet nämns inte i Riksarkivets föreskrift RA-FS 2009:2.</p>
<p>DWG (AutoCAD Drawing) DXF (AutoCAD Drawing Exchange Format)</p>	<p>AutoCAD:s eget bildbehandlingsformat DWG är så dominerande bland användare, att det har blivit en de facto-standard för vektorgrafik.</p> <p>DXF-formatet utvecklades av Autodesk som ett dataformat som möjliggör utbyte av data mellan AutoCAD och andra program.</p>	<p>Om målet är att bibehålla all information i filerna kring lager, färger, och x-referenser behöver filerna sparas som DWG, dock är specifikationerna för DWG proprietära och har officiellt inte släppts som öppna. Fler program kan öppna och bearbeta en DXF-fil än en DWG-fil.</p> <p>DXF-formatet är ASCII-textbaserat och går därför att öppna i annan programvara än AutoCAD. Specifikationen för DXF ägs av Autodesk, men är fri att använda. Formatet används ofta för utbyte av vektordata mellan olika program och plattformar, men är bäst lämpad för relativt enkel linjefgrafik. I DXF-filen behålls bara ritningens grundläggande element, som linjer och text.</p> <p>Format som Riksarkivet i RA-FS 2009:2 rekommenderar för digitalt framställda ritningar är PDF/E eller CALS (som är ett rasterformat)</p>

"Dokument" används här som ett brett begrepp som innefattar många olika typer av dokumentation och handlingar som uppstår i en verksamhet. Rapporter, utredningar, redovisningar, anteckningar, protokoll, instruktioner, med mera. Vanligt är att dessa innehåller text. Dokument kan innehålla en blandning av text, bild, grafik, sammanställningar av data.

Ursprung	Ursprungsanvändningsområde	Resonemang kring format för bevarande
<p>Dokument som innehåller information i text</p> <p>Exempel DOC, DOCX, PDF, ODT, RTF, TXT, Pages, E-postmeddelanden, Wikisidor</p>	<p>Dokumentera, sammanställa, redovisa information i olika syften, till exempel kommunikation, delning, instruktioner, dokumentation. Det förekommer också format för dokumentmallar.</p> <p>Dokumenterna kan vara skapade digitalt eller ha digitaliserats från analoga format.</p>	<p>Många format är proprietära och bundna till särskilda operativsystem. Nya versioner av programvara stödjer inte alltid gamla versioner av filformat.</p> <p>Som bevarandeformat bör ett stabilt format väljas. Vid konvertering är det inte alltid möjligt att behålla all formatering, till exempel teckensnitt, sidlayout, disponering av text, infogade bilder och annan grafik, länkar, tabeller.</p> <p>PDF/A-1 anges av Riksarkivet (RA-FS 2009:2) som bevarandeformat för kontorsdokument. För dokument där det räcker att bevara enbart textinformationen kan filen sparas som en ren textfil, till exempel TXT.</p>
<p>Tryckoriginal</p> <p>Exempel INDD, PDF</p>	<p>Digitala tryckoriginal. Innehåller layout, formatering, text, bilder, information om beskärning, tryckfärger, färgrymd, med mera.</p>	<p>Se ovanstående resonemang om kontorsdokument. PDF/A-1 anges av Riksarkivet (RA-FS 2009:2) som bevarandeformat för kontorsdokument.</p> <p>Om behov finns av att fortsätta beställa tryck utifrån tryckoriginalet och redigering eventuellt behöver göras, kan även originalformatet sparas. Då bör det göras tydligt vilken version som är bevarandeverionen.</p>
<p>Kalkylark/ Spreadsheets</p>	<p>Kalkylark används ofta när beräkningar eller analyser av</p>	<p>Se ovanstående resonemang om kontorsdokument. PDF/A-</p>



Exempel XLS, XLSX, ODS	information behöver göras, till exempel i form av statistik eller grafer. Det kan också vara ett sätt att skapa enkla register eller användas för att presentera information på ett tydligt sätt.	1 anges av Riksarkivet (RA-FS 2009:2) som bevarandeformat för kontorsdokument.  Om möjligheter att fortsätta göra analyser av datat, eller att kunna sammanställa informationen på olika sätt behöver behållas, bör man betrakta informationen som strukturerad data, och välja bevarandeformat utifrån det.
Digitaliserade analoga dokument PDF, TIFF, JPG, TXT	Analoga dokument kan digitaliseras i flera syften och val av format görs utifrån dessa syften	Om det digitaliserade dokumentet ska bevaras, ska bevarandeformat ha valts redan innan digitalisering utifrån vilken information i dokumentet som ska bevaras: bildinformation, textinformation, formatering, layout.

## Ljud

Till skillnad mot dokument, bilder, kartor, märkspråk och kodning tar Riksarkivets föreskrift RA-FS 2009:2 inte upp digitala tekniska format för audiovisuella filer. Dock finns det flera andra rekommendationer för den som vill säkra bevarande och få tips på hur man ska tänka kring digitalt ljud. Generellt är det för ljud precis som det är för andra digitala format, att ju mer använt och accepterat ett format är desto säkrare och mer tillförlitligt är det över tid och därmed säkrare att använda för digitalt bevarande. För den som vill fördjupa sig mer kring detta så har International Association of Sound and Audiovisual Archives (IASA) publicerat flera generella och tekniska riktlinjer<sup>63</sup> för hantering av analoga och digitala ljudsamlingar. Även Svensk Nationell Datatjänst (SND) har publicerat riktlinjer kring digitalt ljud<sup>64</sup>.

Generellt är det bäst att långtidsbevара i okomprimerade, vanligt förekommande format. Det är också viktigt att ha med vissa tekniska metadata för att säkerställa långsiktigt bevarande av ljudfiler. Kortfattat handlar det om varaktighet (längd), bit-djup, samplingsfrekvens, antal kanaler och kanaltilldelning. Detta gäller oavsett val av format. Dock tar okomprimerade filer mer plats och ibland kan det även finnas andra skäl att välja bevarande av komprimerade filer.

Som exempel kan sägas att en vanlig ljud-CD (CD-DA) ofta är två kanaler (stereo) (L)PCM bitström, 16 bitar samplad i 44,1 kHz medan en DVD-A kan lagra två kanaler (L)PCM till 24 bitar samplad i 192 kHz men den kan också lagra fler kanaler till 24 bitar i 96 kHz (alltså högre kvalitet men då med kortare speltid). Samma data skulle kunna migreras till wave-filer som i dagsläget är det vanligaste bevarande (container)formatet för digitalt ljud för att sedan lagras på en disk, i en servermiljö eller

<sup>63</sup> <https://www.iasa-web.org/tc04/audio-preservation>

<sup>64</sup> [https://snd.gu.se/sites/default/files/legacy/SND%3B%20Digitalt%20ljud%3B%20En%20guide%20till%20god%20hantering\\_2017-06-01.pdf](https://snd.gu.se/sites/default/files/legacy/SND%3B%20Digitalt%20ljud%3B%20En%20guide%20till%20god%20hantering_2017-06-01.pdf)

på band. Wave-formatet har i sin tur utvecklats för att kunna hantera större datamängder per fil (senaste utvecklingen är RF64).

Förenklat uttryckt finns tre typer av format för digitalt ljud (precis som för bild, rörlig bild och multimedia) som är vanligt förekommande. Okomprimerade, "lossless"-komprimerade format (ickeförstörande komprimering) och "lossy"-komprimerade format (informationsförstörande komprimering). De flesta filformat för ljud är egentligen så kallade container-format. Dessa använder i sin tur en specifik kodek som kodar och avkodar dataflödet/signalen från digital till analog eller tvärtom (eller digital till digital) och då i vissa fall även komprimerar filen. Du kan alltså välja att bevara komprimerat digitalt ljud i exempelvis en wave-containerfil.

Precis som diskuteras i stycket nedan för rörlig bild skiljer det sig vanligtvis mellan master-/bevarandematerial och lyssnings-/spridningsmaterial när det kommer till digitalt ljud. Det förra är oftast okomprimerat ljud och när det gäller det senare är det relativt vanligt med olika komprimerade format. Dessutom finns det idag en rad olika applikationer och program för enklare ljudupptagningar som kan användas till exempel för att spela in en intervju och där olika format förekommer. Ljud förekommer även som deldatamängd i exempelvis film, tv/dator-spel, tv-program med mera.

Nedan listas några vanligt förekommande digitala ljudformat och hur de ofta används:

Ursprung	Användningsområde	Resonemang kring format för bevarande
Okomprimerade filformat. Exempelvis: wave, BWF (EBU TECH 3285, ITU-R BS.1352-3) och RF64 (EBU TECH 3306, ITU-R BS.2088), AIFF, Matroska.	Mastermaterial vid skapande av till exempel musikproduktioner, program för radio och TV, målformat vid digitisering av analoga band med mera.	Proprietära format (Microsoft respektive Apple, tilläggen utvecklade av EBU) men fri och öppen dokumentation samt finns stöd för båda formaten i flera operativsystem idag. Matroska är en öppen standard och container för bland annat ljud. Observera att ovan nämnda containerformat också kan användas för att bevara komprimerat ljud. Det är alltså fler aspekter än val av container som styr kvaliteten på bevarandet.
Lossless-komprimerade filformat Exempelvis: FLAC, ALAC, WMA Lossless, MPEG-4 ALS (ISO/IEC 14496-3)	Förekommer till exempel i olika applikationer för telefoner, TV och datorer. Vanligt förekommande vid spridning/delning av digitalt ljud med mera.	I de fall det inte finns eller inte finns möjlighet att ta fram ett okomprimerat originalmaterial är man hänvisad till ett komprimerat material. Dessa format går att bevara i ovan nämnda containerformat.
Lossy-komprimerade filformat. Exempelvis: MP3/MPEG-2 audio (ISO/IEC 13818-3), AAC/MPEG-4 audio (ISO/IEC 14496-3), Vorbis, Opus .	Används ofta vid spridning/delning av digitalt ljud.	Ofta öppna och standardiserade format. Går att använda som bevarandeformat om tillgång saknas till okomprimerat originalmaterial. Dessa format kan även bevaras för att de

		används vid tillgängliggörande men att originalmaterial samtidigt bevaras okomprimerat.
--	--	---

## Rörlig bild

Vad gäller rörlig bild så är det viktigt att veta skillnaden på visningsmaterial och bevarande- och mastermaterial. Visningsmaterial är det som vid varje givet tillfälle distribueras och visas i något fönster. Visningsmaterial skiljer sig beroende på kvalitetskrav och distributionsformer. Mastermaterial är ett material av högre kvalitet från vilket det går att framställa olika typer av visningsmaterial. Två olika typer av visningsmaterial är alltså ofta framställda från samma mastermaterial. Ett mastermaterial saknar dessutom ofta versioner (till exempel olika språk eller textningar).

För att på bästa sätt ha bevarat ett verk av rörlig bild är det alltså att föredra att bevara både visningsmaterial för de olika visnings- och distributionsformerna samt ett mastermaterial från vilket det går att skapa ett nytt visningsmaterial i framtiden då visnings- och distributionsformerna skiljer sig från dagens.

Den allmänna regeln är att visningsformat nästan uteslutande är komprimerade format och vad det gäller masterformat ska strävan vara att bevara ett okomprimerat format för att ha bästa möjliga förutsättningar att skapa ett nytt visningsmaterial av så bra kvalitet som möjligt i framtiden och skapa bästa förutsättningar för framtida migreringar.

För rörlig bild förekommer krypterat visningsmaterial till exempel för biografändamål. Det anses inte vara ett bra format att bevara för framtiden eftersom systemet bygger på att det går att ta fram så kallade visningsnycklar för att kunna spela upp ett krypterat material. Möjligheten att ta fram dessa nycklar i framtiden är högst osäkra eftersom de genereras av distributör eller rättighetsinnehavare för varje visning.

I dagsläget pågår ett standardiseringsarbete på EU-nivå för långtidsbevarande av rörlig bild ämnad för biografvisning. Standarden<sup>65</sup> planeras vara klar 2022.

Generellt bör strävan vara att använda filformat som är så standardiserade och ickeproprietära som möjligt. Inte helt sällan krockar denna strävan med så kallade branschstandarder, alltså vad som dagligen används och efterfrågas.

Exempel på format:

Ursprung	Användningsområde	Resonemang kring format för bevarande
Okomprimerat (eller oförstörande komprimering) mastermaterial (bild) Exempelvis: TIFF, JPEG2000 Lossless, DPX och FFV1 (För ljud se stycket ovan)	Mastermaterial för långtidsarkivering och för framtagande av nytt visningsmaterial.	Icke-proprietära, okomprimerade eller icke-förstörande komprimering av bildsekvenser eller av högsta möjliga kvalitet.
Komprimerat mastermaterial och mezzanineformat	För filmproduktioner med mindre budget och för TV-produktioner förekommer	I de fall det inte finns och inte finns möjlighet att ta fram ett okomprimerat mastermaterial

<sup>65</sup> <https://cen-pt457.atlassian.net/wiki/spaces/PUB/overview>

Exempelvis: ProRes, IMF, DnXHD, DnXHR	sällan ett okomprimerat mastermaterial.	är man hänvisad till ett komprimerat material. I de flesta fall även hänvisade till ett proprietärt format vilket är ett problem.
Visningsmaterial bio Okrypterad DCP	Visningsmaterial för biograf enligt standard.	Standardiserat.
Visningsmaterial för webb och streaming Exempelvis: H.264 och H.265	Visningsmaterial för olika typer av webb och streaming.	Här återfinns oftast relativt standardiserade och öppna format. Använd om möjligt dessa.

### Strukturerade data

Strukturerade data är lagrade i ett maskinläsbart format och är maskinellt tolkningsbara genom att datamängderna är beskrivna av metadata. Exempelvis: information som är strukturerad i en databas, data uppmärkta med märkspråk, register, tabeller. Strukturerade data kan bearbetas på olika sätt och användas till att skapa sammanställningar och beräkningar.

Vid bevarande bör en bedömning göras av vilka sammanställningsmöjligheter och analysmöjligheter som finns. Exempelvis, vid uttag ur en relationsdatabas som bygger på SQL kommer vissa sök- och sammanställningsmöjligheter som finns i databashanteraren att tappas om statiska bevarandeformat som XML eller CSV väljs. XML och CSV går att använda som överföringsformat vid flytt till nya databaser, men för att till exempel kunna återskapa sambanden mellan olika tabeller behöver databasen dokumenteras och förklaras.

Ursprung	Ursprungsanvändningsområde	Resonemang kring format för bevarande
Enkla tabeller och register  Exempel: XLS, XLSX, ODS, CSV, TXT, XML Tabeller i olika databasformat	Enkla register, till exempel kundregister, förteckningar över föremål, där bara en tabell behövs.	Enkla tabeller kan sparas som filer med teckenseparerade fält, till exempel kommaseparerade CSV-filer.
Databaser  Exempel: SQL-databaser Access-databaser		Databaser som består av tabeller kan exporteras så att varje tabell för sig sparas ner som till exempel en CSV fil. I detta fall behövs tydlig dokumentation kring tabellernas samband och hur informationen kan sökas fram. XML kan också användas som lagringsformat.
Data uppmärkt med märkspråk  Exempel: XML	XML är utformat för att kunna läsas både av maskiner och människor och används som ett överföringsformat och lagrar data som ren text.	XML är ett godkänt bevarandeformat enligt Riksarkivet (RA-FS 2009:2).  XML innehåller inte någon information om hur data ska presenteras. Som komplement till data i XML-filen kan du

		behöva stylesheets och scheman, samt beskrivning av vad elementen i XML-dokumentet betyder.
--	--	---

## System

System består av flera delar som kan behöva bevaras på olika sätt. Det finns ett eller flera gränssnitt som kan se olika ut för inloggade eller icke-inloggade användare. Det finns funktionalitet och interaktivitet. Om systemet används för att visa eller bearbeta data kan det också finnas en eller flera databaser och fillagring.

Vid bevarande av system bör det bedömas vilka delar av systemet som ska bevaras och i vilken form. Då är det viktigt att titta på huvudsyftet med systemet. Är syftet med bevarandet att kunna berätta om hur systemet såg ut och fungerade? Är det bara datamängderna i systemet som är intressanta? Är det själva funktionaliteten i sig och även källkoden? Bevarande av system innebär nästan alltid någon form av informationsförlust. Tillsammans med bevarandefiler för systemet bör därför alltid systembeskrivning och systemdokumentation bevaras.

Dokumentationen bör beskriva hur systemet användes, förklaring till tabeller och deras relationer och innehåll och hur information kan sökas fram, beskrivning av gränssnittet och hur systemet kunde användas.

Exempel	Resonemang vid bevarande
Administrativa system  Exempel Ärendehanteringssystem Ekonomisystem	Ska varje post bevaras för sig eller ska all information sparas i ett uttag? Hur komplex är datamodellen i databasen – hur behöver relationerna mellan olika tabeller i databasen beskrivas? Är det en statisk sammanställning av flera möjliga sammanställningar som ska arkiveras, eller måste möjligheter till nya sammanställningar bevaras?
Samlingsförvaltning Exempel Bilddatabas Samlingsdatabas	Ska varje post bevaras för sig eller ska all information sparas i ett uttag?
Söktjänster  Exempel: Sökgrenssnitt för samlingsdatabas	Behöver dokumentation och funktionalitet från söktjänsten sparas? Hur ser relationen mellan söktjänsten och systemet/databasen ut? Är gränssnittet fristående?

## Bilaga: Verksamhetsmodell

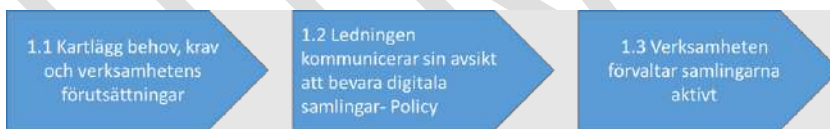
Digisam har tagit fram ett antal vägledande principer<sup>66</sup> som är ett stöd i arbetet med att digitalisera, digitalt tillgängliggöra och digitalt bevara kulturarvet. Principerna poängterar att digitaliseringsverksamhet ska styras av respektive institutions ledning samt att varje institution ska ansvara för de digitala objekt och metadata som den skapar. För att lyckas behöver verksamheten avsätta resurser så att den aktivt kan förvalta de digitala samlingar den skapar och förvärvar.

Arbetet med ett långsiktigt digitalt bevarande är en ständigt pågående process som löper genom hela den digitala livscykeln, från det digitala objektets födelse till dess slutarkivering. Det digitala bevarandet kräver inte bara att tekniken finns på plats utan även att verksamheten tar ett helhetsgrepp om frågan. I detta avsnitt ges förslag på hur en verksamhetsmodell för digitalt bevarande kan se ut. En verksamhetsmodell beskriver hur en verksamhet organiserar sina resurser för att se till att den löpande verksamheten stödjer strategiska mål. I avsnittet kan du också läsa om OAIS funktionsmodell och hur den kan vara ett stöd i arbetet med en verksamhetsmodell för digitalt bevarande. Som nämnts tidigare i den här vägledningen så är OAIS en ISO-standard och ett ramverk för digitalt bevarande.

Verksamhetsmodellen utgår från grundläggande delar i standarden Ledningssystem för kvalitet - Krav (ISO 9001:2015)<sup>67</sup>. Framställningen är mycket mer allmänt hållen än ISO-standarderna för att bli lättare att ta till sig och kunna tillämpas vid digitalt bevarande hos kulturarvsinstitutioner av olika storlek och med olika förutsättningar. Vidare baseras avsnittet delvis på Adrian Browns bok *Practical Digital Preservation, a how-to guide for organizations of any size*.

I detta avsnitt beskrivs först övergripande hur verksamhetsmodellen kan införas. Därefter beskrivs dess olika delar.

### Införande av verksamhetsmodell för digitalt bevarande



Figur 1. Införande av verksamhetsmodell.

#### Kartlägg behov, krav och verksamhetens förutsättningar

Ett första steg i införandet av en verksamhetsmodell för digitalt bevarande är att ringa in verksamhetens förutsättningar och kartlägga de behov och krav som verksamheten behöver hantera. Verksamheten bör i sin kartläggning ta höjd för både interna och externa behov och krav. Faktorer som kan påverka är till exempel regelverk och standarder (se avsnitt Regelverk och standarder), teknisk utveckling eller förväntningar från olika intressenter som berörs av det digitala bevarandet. Det kan också finnas stora vinster i att undersöka vilka krav och behov som andra organisationer har identifierat och hur dessa har hanterats. Kanske öppnas en dörr till samverkan och gemensamma lösningar?

<sup>66</sup> <http://www.digisam.se/vagledande-principer-for-arbetet-med-digitalt-kulturarv/>

<sup>67</sup> <https://www.sis.se/produkter/foretagsorganisation/foretagsorganisation-och-foretagsledning-ledningssystem/ledningssystem/ssenis090012015/>

Verksamheten bör så långt det är möjligt översätta behov till krav. Ett behov kan till exempel vara att endast behörig personal ska ha åtkomst till viss information. Kravet kan då formuleras som att olika användare ska tilldelas olika behörighetsnivåer. Det finns också olika typer av krav. Funktionella krav rör specifika funktioner i ett system och hur det interagerar med användarna. Ett exempel på detta är att en användare vid en ansökan om en inloggning till ett system får ett e-postmeddelande med information om nästa steg. Icke-funktionella krav rör istället andra egenskaper hos systemet såsom prestanda, säkerhet eller användbarhet. Ett exempel på detta är att systemet ska klara av minst 200 användare samtidigt. Slutligen finns servicekrav som handlar om stödfunktioner. Ett exempel på detta är en systemsupport.

Det digitala bevarandet innehåller många olika delar och möjliga krav. Därför kan det underlätta förståelsen och skapa överblick att gruppera och beskriva kraven utifrån teman. En utgångspunkt för gruppering kan till exempel vara OAIS-modellens funktionsmodell (se bilaga OAIS-modellen). Ett annat sätt att ta sig an frågan är att skapa sig en förståelse för olika aspekter på bevarande (se avsnitt Bevarandenaspekter) och det digitala objektets livscykel (se avsnitt Bevarandecykeln), från att förberedelser görs och det tas emot i ett system för digitalt bevarande till att tillgång till materialet säkras över tid genom olika åtgärder.

Ett användbart verktyg är också den vägkarta för digitalt bevarande som tagits fram av en rad organisationer på europeisk nivå och går under namnet DCH-RP, Digital Cultural Heritage Roadmap for Preservation<sup>68</sup>. Verktyget är ett stöd för kulturarvsinstitutioner för att kunna ta fram en egen aktivitetsplan för digitalt bevarande som bland annat inkluderar infrastruktur, tjänster, standarder och interoperabilitet.

Beskrivningarna av kraven för det digitala bevarandet kan anpassas till verksamhetens förutsättningar och ofta vara enkla i sin utformning. Det viktigaste är att de finns på plats och följs i arbetet.

#### **Ledningen kommunicerar sin avsikt att bevara digitala samlingar**

Nästa steg i införandet av verksamhetsmodellen är att utforma en övergripande avsikt med bevarandet, baserat på de förutsättningar, behov och krav som konstaterats. Verksamhetens ledning kommunicerar avsikten att aktivt bevara och förvalta de digitala samlingarna och en policy tas fram. En policy är ett dokument som klargör en organisations grundläggande principer i en specifik fråga - i det här fallet digitalt bevarande.



Figur 2. Steg i arbetet med att införa en policy.

Införandet av policyn kan därefter ske i fyra steg (se figur 2 ovan). Policyn bör under inventeringen knytas an till verksamhetens konstaterade förutsättningar och övriga styrande dokument. Den behöver också avgränsas. Den blir då mer slagkraftig och mindre möda behöver läggas på att justera

<sup>68</sup> <https://www.dchrp.eu/index.php?en/115/roadmap-for-preservation>

innehållet. Mer detaljerade beskrivningar, till exempel av hur olika arbetsmoment ska genomföras, anges istället i andra dokument såsom processbeskrivningar.

En policy för digitalt bevarande bör som minst innehålla information om syftet med policyn, dess omfattning, vem den gäller för och generella principer för det digitala bevarandet. Andra saker den kan innehålla är dess kontext, hur den relaterar till andra styrdokument, hur den ska granskas och revideras, samt standarder som ska följas. Policyn utvecklas och anpassas utifrån den specifika verksamhet den berör.

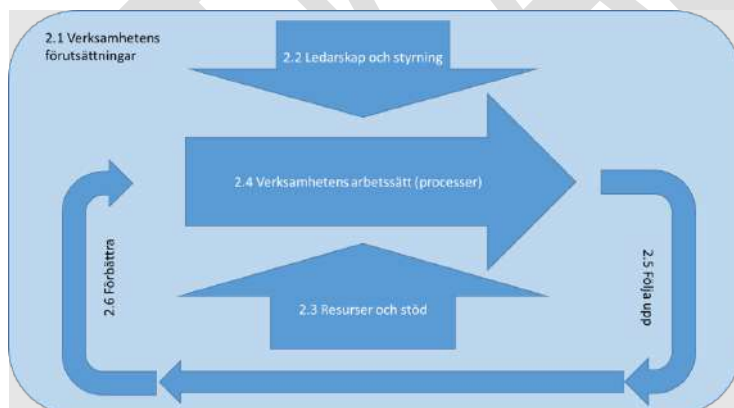
Slutligen behövs en plan för kommunikation kring och implementering av policyn för att den på allvar ska kunna efterlevas i verksamheten.

### Aktiv förvaltning och ständig förbättring

Då verksamhet och omvärld hela tiden är i förändring behöver arbetsätten ständigt anpassas till nya förutsättningar. Att omsätta en policy för digitalt bevarande i praktiken kräver med andra ord aktiv förvaltning och ständig förbättring.

NDSA, National Digital Stewardship Alliance, har tagit fram ett användbart verktyg<sup>69</sup> för utveckling och utvärdering av en verksamhets digitala bevarande. Med hjälp av verktyget kan din verksamhet notera på vilken nivå, av fyra nivåer där den fjärde nivån är den mest avancerade, den befinner sig vad gäller olika delar av digitalt bevarande (lagring, integritet, kontroll, metadata och innehåll). Utifrån resultatet kan verksamheten besluta om ett eventuellt nästa steg i arbetet.

### Verksamhetsmodellens delar



Figur 3. Verksamhetsmodellens övergripande delar och hur de förhåller sig till varandra.

### Verksamhetens förutsättningar

En kartläggning kring verksamhetens förutsättningar behöver regelbundet ses över för att vara i linje med tiden. Det är därför viktigt att dokumentera behov, krav och förutsättningar så att de går att följa över tid.

<sup>69</sup> <https://ndsa.org/publications/levels-of-digital-preservation/>



### **Ledarskap och styrning**

Ledningen behöver regelbundet följa upp att verksamheten ges rätt förutsättningar och stöd för att hållbart kunna bevara de digitala samlingarna och vid behov besluta om åtgärder och styra om resurser.

### **Resurser och stöd**

Med resurser och stöd menas här de tillgångar som finns inom organisationen för att stödja och utveckla verksamheten, såsom kompetens, ekonomiska resurser, lokaler och IT.

### **Kompetens**

Arbete med digitalisering och digitalt bevarande inom kulturarvsinstitutioner förutsätter tillgång till en bred kompetensbas. Beroende på organisationens storlek och behov kan relevant kompetens säkras antingen genom utbildning, nyrekrytering, supporttjänster eller projektanställningar. Mindre organisationer kan även dela kompetens i en gemensam lösning eller outsourca. Oavsett vilken lösning som är aktuell är ett viktigt första steg att ta reda på vilka kompetenser som redan finns inom organisationen och vilka luckor som behöver fyllas - tillfälligt eller mer permanent.

### **Ekonomiska resurser**

Kostnader för digitalt bevarande varken kan eller bör särskiljas från organisationens övriga ekonomi. Då i stort sett alla processer inom samlingsförvaltning är sammankopplade och påverkar varandra är det heller inte helt lätt att identifiera enskilda kostnader. Trots detta är realistiska kostnadsberäkningar, utifrån verksamhetens förutsättningar, av stor vikt för att skapa en resurseffektiv och hållbar verksamhetsmodell. För de flesta organisationer gäller att personal och kompetens innebär både de största tillgångarna och de största kostnaderna.

Vidare finns ett antal övergripande förutsättningar som verksamheten behöver ta hänsyn till vid budgetering, till exempel samlingarnas typ och storlek, bevarandegrad och tidsperspektiv samt krav på tillgänglighet. Viktigt att tänka på är också att samlingar vid digitalisering växer i högre takt, vilket medför ökade kostnader för samlingsförvaltning.

### **Infrastruktur**

Övriga resurser vid digitalt bevarande rör tekniskt och infrastrukturellt stöd som IT-system och lagringslösningar. Vid outsourcing kan själva bevarandetjänsten inklusive lagring, förvaltning och tillgängliggörande hanteras externt. I enlighet med OAIS-modellen (se bilaga OAIS-modellen) är dock en grundläggande förutsättning för digitalt bevarande att system och verktyg möjliggör förbindelser mellan leverantörer, digitalt arkiv och användare.

### **Verksamhetens arbetssätt (processer)**

Ett sätt att identifiera vilka arbetssätt som ska beskrivas i verksamhetsmodellen är att utgå ifrån identifierade krav och ställa sig frågan vad verksamheten behöver göra för att kraven ska tas om hand. Oftast är det enklast att först ta fram övergripande beskrivningar som anger vilka roller i verksamheten som har ansvar för olika delar av det digitala bevarandet och hur resultatet ska bli. Dessa beskrivningar kan sedan vid behov kompletteras med mer detaljerade beskrivningar av hur arbetet ska utföras.

Utöver detta finns det ett antal allmänna förutsättningar som särskilt rör kulturarvsinstitutioner. Viktiga och gemensamma frågor omfattar till exempel urval och gallring. Vad gäller digitala samlingar blir dessa frågor dessutom mer angelägna då digitala objekt har kortare hållbarhet än de flesta

fysiska objekt. Eftersom de flesta moment inom samlingsförvaltning påverkar och överlappar varandra kan allt arbete med samlingsförvaltning sägas få konsekvenser för det digitala bevarandet.

#### Följa upp

Uppföljning genomförs regelbundet och resultatet analyseras och presenteras för verksamhetens ledning, som beslutar om eventuella förbättringar av verksamhetsmodellen och dess arbetssätt.

#### Förbättra

Verksamhetsmodellens hållbarhet förutsätter att de förbättringar som beslutas genomförs. Genomförande, resultat och verkan av de förbättringar som genomförs följs upp av ledningen.

### Bilaga: Översättning av Levels of Preservation

Table 1: Version 1 of the Levels of Digital Preservation	Skydda din data Level 1 (Protect your data)	Känn din data Level 2 (Know your data)	Kontrollera din data Level 3 (Monitor your data)	Reparera din data Level 4 (Repair your data)
<b>Lagring och geografisk plats</b>  Storage and Geographic Location	<ul style="list-style-type: none"> <li>- Två kompletta kopior som inte förvaras ihop</li> <li>- För data på heterogena medier (optiska skivor, hårddiskar, etc.) importera innehållet från mediet till ditt lagringssystem</li> <li>- Two complete copies that are not collocated</li> <li>- For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system</li> </ul>	<ul style="list-style-type: none"> <li>- Minst tre kompletta exemplar</li> <li>- Minst en kopia på en annan geografisk plats</li> <li>- Dokumentera dina lagringssystem och lagringsmedier och vad du behöver för att använda dem</li> <li>- At least three complete copies</li> <li>- At least one copy in a different geographic location</li> <li>- Document your storage system(s) and storage media and what you need to use them</li> </ul>	<ul style="list-style-type: none"> <li>- Minst en kopia på en geografisk plats med ett annat katastrofhot</li> <li>- Övervakningsprocess för utgående teknik för dina lagringssystem och media</li> <li>- At least one copy in a geographic location with a different disaster threat</li> <li>- Obsolescence monitoring process for your storage system(s) and media</li> </ul>	<ul style="list-style-type: none"> <li>- Minst tre kopior på geografiska platser med olika katastrofhot</li> <li>- Ha en omfattande plan på plats som håller filer och metadata på aktuella tillgängliga media eller system</li> <li>- At least three copies in geographic locations with different disaster threats</li> <li>- Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems</li> </ul>
<b>Filbeständighet och dataintegritet</b>  File Fixity and Data Integrity	<ul style="list-style-type: none"> <li>- Kontrollera datas beständighet, ursprungliga skick vid mottagning om sådana metadata finns</li> <li>- Ta fram metadata om</li> </ul>	<ul style="list-style-type: none"> <li>- Kontrollera datas beständighet, ursprungliga skick vid alla mottagning</li> <li>- Använd skrivblockerare när du arbetar</li> </ul>	<ul style="list-style-type: none"> <li>- Kontrollera datas beständighet, ursprungligt skick med bestämda intervaller</li> <li>- Hantera loggar med beständig-hets-metadata; utföra</li> </ul>	<ul style="list-style-type: none"> <li>- Kontrollera att all data är beständig och i ursprungligt skick som svar på specifika händelser eller aktiviteter</li> <li>- Se till att ha möjlighet att</li> </ul>

	<p><b>beständighet om det inte ingick vid levereras</b></p> <ul style="list-style-type: none"> <li>- Check file fixity on ingest if it has been provided with the content</li> <li>- Create fixity info if it wasn't provided with the content</li> </ul>	<p><b>med originalmedia</b> - Viruskontrollera känsliga data</p> <ul style="list-style-type: none"> <li>- Check fixity on all ingests</li> <li>- Use write-blockers when working with original media</li> <li>- Virus-check high risk content</li> </ul>	<p><b>strukturerad granskning av dessa vid behov</b> - Se till att ha förmåga att upptäcka korrupta data - Viruskontrollera all data</p> <ul style="list-style-type: none"> <li>- Check fixity of content at fixed intervals</li> <li>- Maintain logs of fixity info; supply audit on demand</li> <li>- Ability to detect corrupt data</li> <li>- Virus-check all content</li> </ul>	<p><b>ersätta/reparera skadade data</b> - Se till att ingen har skrivbehörighet till alla kopior</p> <ul style="list-style-type: none"> <li>- Check fixity of all content in response to specific events or activities</li> <li>- Ability to replace/repair corrupted data</li> <li>- Ensure no one person has write access to all copies</li> </ul>
<p><b>Informationssäkerhet</b>  Information Security</p>	<ul style="list-style-type: none"> <li>- Identifiera vem som har behörighet att läsa, skriva, flytta och radera enskilda filer</li> <li>- Begränsa vilka som har dessa behörigheter till enskilda filer</li> <li>- Identify who has read, write, move and delete authorization to individual files</li> <li>- Restrict who has those authorizations to individual files</li> </ul>	<ul style="list-style-type: none"> <li>- Dokumentera åtkomst-begränsningar för data</li> <li>- Document access restrictions for content</li> </ul>	<ul style="list-style-type: none"> <li>- Underhåll och bevara loggar över vem som utförde vilka åtgärder på filer, inklusive borttagningar och bevarandeåtgärder</li> <li>- Maintain logs of who performed what actions on files, including deletions and preservation actions</li> </ul>	<ul style="list-style-type: none"> <li>- Utför strukturerad granskning av loggar</li> <li>- Perform audit of logs</li> </ul>
<p><b>Metadata</b>  Metadata</p>	<ul style="list-style-type: none"> <li>- Hantera ett beståndsregister inkluderande lagringsplats-information</li> <li>- Säkerställ säkerhetskopiering och skilda lagringsplatser</li> <li>- Inventory of content and its storage location</li> </ul>	<ul style="list-style-type: none"> <li>- Bevara administrativ metadata</li> <li>- Bevara metadata om migreringar och andra loggade åtgärder</li> <li>- Store administrative metadata</li> <li>- Store transformative</li> </ul>	<ul style="list-style-type: none"> <li>- Bevara tekniska och beskrivande metadata</li> <li>- Store standard technical and descriptive metadata</li> </ul>	<ul style="list-style-type: none"> <li>- Lagra bevarandemetadator</li> <li>- Store standard preservation metadata</li> </ul>

	- Ensure backup and non-collocation of inventory	metadata and log events		
<b>Filformat</b> File Formats	<p>- När du kan påverka skapandet av digitala filer uppmuntrar till att det användas en begränsad uppsättning kända öppna format och codecs</p> <p>- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs</p>	<p>- Ha register över filformat som används</p> <p>- Inventory of file formats in use</p>	<p>- Övervaka eventuella problem med föråldrade filformat</p> <p>- Monitor file format obsolescence issues</p>	<p>- Utför formatkonverteringar, emulering och liknande aktiviteter efter behov</p> <p>- Perform format migrations, emulation and similar activities as needed</p>