



Titel: Rekommendation - Arkivfunktionalitet för IT-system

Handläggare: Magnus Wählberg

Datum: 2010-01-27

Version 1.0

TAM 4:2010 REKOMMENDATION

Arkivfunktionalitet för IT-system

Version 1.0

Innehållsförteckning

1 INLEDNING	3
2 KRAV PÅ IT-SYSTEM	4
2.1 ALLMÄNT	4
2.2 ARKIVFUNKTIONALITET	5
2.2.1 VÅRD AV INFORMATION	5
2.2.2 GALLRING	6
2.2.3 ARKIVEXPORT	8
2.2.4 ÅTERSÖKNING OCH ARKIVIMPORT	9
3 SAMMANFATTNING AV KRAV FÖR ARKIVFUNKTIONALITET	10
4 BILAGOR	11
4.1 LITTERATURFÖRSLAG	11
4.2 TAM-ARKIVS REKOMMENDATIONER OCH RUTINER	12

1 Inledning

Denna rekommendation beskriver grundläggande krav på arkivfunktionalitet för ett godtyckligt IT-system. Kraven behandlas på en övergripande nivå med en del exempel. Målgrupperna för denna rekommendation är företrädesvis beställare av IT-system, projektledare, kravställare, arkivarier, systemarkitekter och systemutvecklare. Dokumentet är anpassat för de medlemsorganisationer som är knutna till TAM-Arkiv. Rekommendationen är framtagen inom projektet ”TAMAD3 – Förstudie för digitalt arkiv”. Ett stort tack går till Bo Johansson vid Sollentuna kommun och Viveka Carlsson vid Arbetsförmedlingen som varit vänliga att kommentera och ge förslag på förbättringar. Eventuella sakfel står undertecknad för. Rekommendationen är även granskad av Ylva Lindberg och Jim Löfgren, båda vid TAM-Arkiv.

För fördjupad information om kravställning för arkivfunktionalitet, se bilaga ”4.1. Litteraturförslag”. TAM 4:2010 är en av de rekommendationer och rutiner som tagits fram vid TAM-Arkiv och som har till syfte att hjälpa medlemsorganisationerna med deras informationshantering, se även bilaga ”4.2. TAM-Arkivs rekommendationer och rutiner”.

Magnus Wählberg

TAM-Arkiv

Alvik i januari år 2010

2 Krav på IT-system

2.1 Allmänt

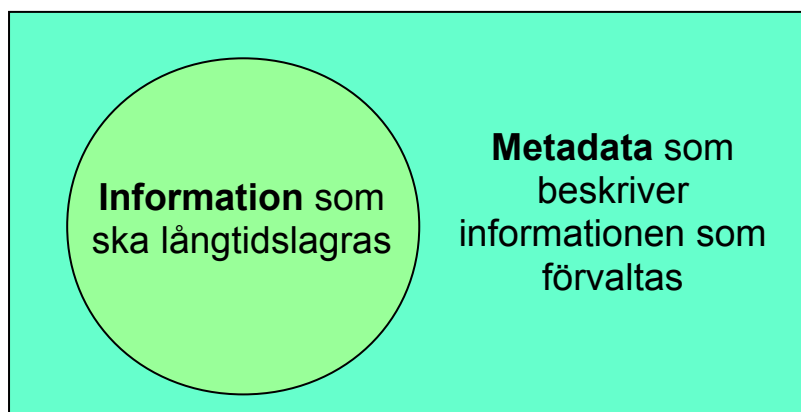
Vid utveckling, inköp och införande av IT-system är bland det viktigaste arbetet att systematiskt ta fram en kravbild för hur systemet ska vara och fungera. Kraven brukar indelas i funktionella och icke-funktionella krav (även kallat egenskapskrav). De funktionella kraven beskriver krav på systemfunktionalitet, vad som ska kunna göras. De icke-funktionella krav beskriver å andra sidan saker som hur många användare som ska kunna använda systemet, teknisk plattform, leveranstidpunkt och liknande. För mer information om bland annat generell kravställning se TAM-Arkivs ”TAM 1:2008 REKOMMENDATION - Vid förändring av ärende- och dokumenthantering” som går att ladda ner från TAM-Arkivs hemsida: <http://www.tam-arkiv.se>.

I avsnitt 2.2.1-2.2.3 nedan beskrivs vilken funktionalitet ett IT-system ska stödja för att ha en fungerande arkivhantering. 2.2.4 är aktuellt om en användare ska kunna komma åt arkiverad information via systemet. En sammanfattning av kraven i avsnitten 2.2.1-2.2.3 finns i kapitel 3.

2.2 Arkivfunktionalitet

2.2.1 Vård av information

Information behöver vårdas för att den inte ska förlora i värde. Information som inte vårdas blir svårare, och ibland omöjlig, att återsöka, använda och förvalta. För att kunna förvalta information på lång sikt är det viktigt att den har tillräckligt med beskrivande information (metadata) knuten till sig, inkluderande metadata som beskriver i vilket sammanhang som den skapats. Exempel på vanligt metadata kan vara kategorikod enligt dokumenthanteringsplan, beslutsdatum, senaste uppdateringsdatum, handläggare/författare, berörd part, titel, organisation, verksamhetsområde, projekt, sökord, format och version. Se även figur 1. Exakt vilket metadata som behöver tillföras beror på organisationens egna behov, informationsområde och krav på framtida återsökning och återanvändande. TAM-Arkiv planerar att arbeta vidare med denna fråga och förhoppningsvis återkomma med mer specifika råd vad gäller metadata.



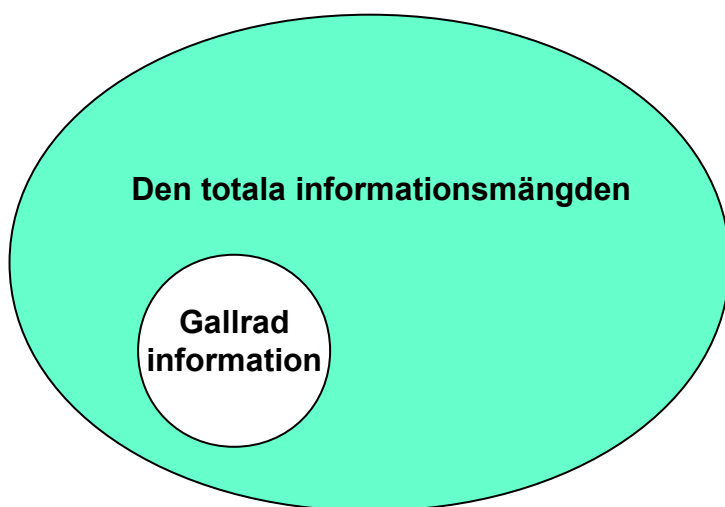
Figur 1. Information som ska långtidslagras och som kapslats in med beskrivande metadata.

Metadata kan behöva korrigeras om det är felaktigt eller kompletteras. Exempelvis upptäcker handläggaren Lisa att beslutsdatumet för ett godkännande är felaktigt angivet. Därför kompletterar hon med det korrekta beslutsdatumet och en kort kommentar om att hon ändrat ett tidigare felaktigt angivet beslutsdatum.

Detta är speciellt viktigt på längre sikt, då metadata ofta behöver tillföras för att kunna styra konvertering, gallring och arkivexport. Om möjligt bör man automatisera kompletteringar av metadata i så hög grad som möjligt. Detta för att undvika manuellt extraarbete. Ännu bättre är om tillräcklig metadata för att kunna styra framtida gallring, konvertering och arkivexport tillförs redan när informationen skapas i systemet. Utförd arkivvård, såsom ändring av metadata eller formatkonvertering, bör dokumenteras i logg. För mer information om format se separat rekommendation "TAM 5:2010 REKOMMENDATION - Format för Långtidslagring".

2.2.2 Gallring

Gallring innebär att man systematiskt förstör information som inte behöver eller ska finnas kvar. Se figur 2. Om ett IT-system innehåller information som behöver gallras måste det följaktligen finnas ett systemstöd för detta. Gallring görs primärt för att 1) förbättra prestanda i IT-system, 2) för att det ska vara lättare för användare att hitta aktuell information eller 3) på grund av lagkrav. Det är förvånansvärt många IT-system som inte har stöd för gallring, trots att de hanterar information som ska eller behöver gallras.



Figur 2. En informationsmängd där en delmängd gallrats bort.

Gallring består arbetsmässigt av tre steg¹:

Steg 1. Framtagande och formalisering av gallringsregler i dokumenthanteringsplan eller liknande.

Steg 2. Utvärderingen av vilken gallringsregel, som ska tillämpas för en viss informationsmängd.

Steg 3. När villkoren för en gallringsregel är infriade förstörs berörd information manuellt eller automatiskt.

Framtagningen av gallringsregler görs utifrån lagkrav och interna verksamhetsregler. Dokumentation av gallringsregler och informationstyper brukar göras i en dokumenthanteringsplan.

Det finns tre typer av gallringsregler:

1. *Statisk gallring.* Gallring utifrån en bestämd tidpunkt och informationstyp.
Ex. Förstör alla email i mailsystemet när ett år efter inkommandedatum har passerats.
2. *Händelsetyrd gallring.* Gallring är beroende av en yttre händelse.
Ex. Förstör dokumentet tio år efter att projektet har avslutats.

¹ De engelska arkivtermerna för gallring är appraisal följt av destruction.

3. *Hinderstyrd gallring.* Ingen gallring tills en yttre händelse skett. Ex. Ärendet är föremål för rättslig prövning och får inte förstöras innan prövning är klar.

Den vanligaste formen av gallring är ”Statisk gallring”. Denna form är också enklast att införa. Händelsetyrd och hinderstyrd gallring kan snarast liknas vid en form av undantagshantering. Båda dessa senare former av gallring kan vara svåra och i praktiken ofta omöjliga att utföra i redan driftsatta IT-system. Märk att system lätt kan bli svårförvaltade om man inte använder enkla och genomtänkta gallringsregler redan från början.

Gallring kan göras manuellt eller automatiskt. Små mängder kan oftast gallras manuellt av en användare eller administratör. Stora mängder brukar man däremot behöva gallra automatiskt (eller semi-automatiskt såsom att skriva ihop ett tillfälligt SQL-script). Om man använder automatisk gallring i ett IT-system kan det vara bra att ha någon form av kontrollmoment, såsom att en administratör aktivt måste initiera ett gallringsjobb. Att den gallring som utförs också är den önskade bör kontrolleras med stickprovskontroller eller genom annan form av revision såsom testkörningar av gallringsregler i ett system med testdata. Det är lätt att komplexa gallringsregler ger oväntade eller i värsta fall oönskade beteenden. Ett standardtips är här att använda enkla gallringsregler så långt som möjligt. Gallringsregler kan även prövas genom manuella tester. Exempelvis kan man kontrollera om gallringsregler kommer i konflikt med varandra, vilket inte är ovanligt vid mer komplex gallring. Utförd gallring bör dokumenteras i logg.

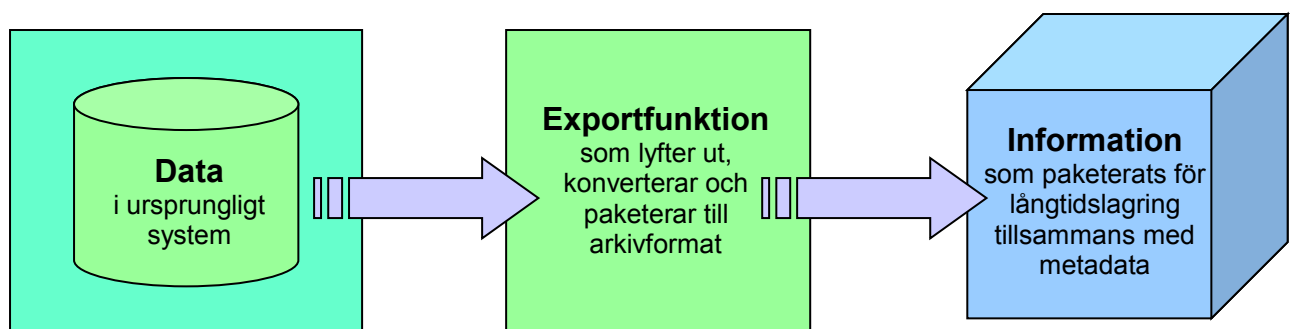
De flesta IT-system brukar ha någon form av backup-hantering där filer lagras på magnetband eller extern disk. När man gallrar information i ett system finns därför informationen kvar på tillhörande backup-media. Oftast brukar inte detta vara något problem, eftersom backup-media brukar skrivas över eller slängas inom ett år. Om man dock behåller backup-media för längre tidsperioder än ett år, så kan detta bli ett problem om det finns lagkrav på att information ska förstöras. Om man är rädd för att förlora information, och därför har en backup-eftersläpningstid på flera år, bör man istället överväga att införa regelbunden arkivexport eller avställning.

2.2.3 Arkivexport

TAM-Arkiv planerar att ta emot digitala leveranser i framtiden, men för att TAM-Arkiv ska kunna ta emot dessa, måste de vara enhetliga med avseende på struktur, filformat, beskrivningar och liknande. TAM-Arkiv kommer aldrig att ha möjlighet att ta emot godtyckliga digitala leveranser som inte följer standarder och rekommendationer. Exakt vilka krav TAM-Arkiv kommer att ställa kommer att utarbetas under de närmaste åren allt eftersom TAM-Arkiv bygger upp en verksamhet för digitala leveranser och återsökningar. För att få enhetlighet och underlätta framtida leveranser till TAM-Arkiv bör systemet ha arkivexportfunktionalitet. Detta avsnitt behandlar kraven på arkivexportfunktionalitet på ett övergripande plan.

Det är ett vanligt misstag att blanda ihop termen arkivexport med avställning. Arkivexport är inte samma sak som avställning. Avställning innebär bara att man ”lyfter ut” data tillfälligt ur ett system för att förbättra prestanda, för att senare läsa in informationen igen när den behövs. Avställning kan ske för både gallringsbart material och material som ska bevaras. Vid avställning brukar man inte konvertera informationen till en form som klarar långtidslagring, utan man brukar bara ”lägga upp filerna på disk”, eller göra en ”dump” av en databas. De flesta IT-systemen har bara stöd för avställning och inte för arkivexport.

Arkivexport å andra sidan handlar om att man lyfter ut information ur ett IT-system och omvandlar den till en form som klarar långtidslagring och där utlyft information dokumenterats med metadata. Se figur 3. Mycket få IT-system har stöd för detta i standardutförande.



Figur 3. Export ur ett IT-system med en exportfunktion till arkivbeständigt format.

Ett system som har stöd för arkivexport ska kunna göra följande:

1. En användare ska kunna välja ut vilken information som ska exporteras ut ur systemet till arkivform.
2. Teckentabeller och filformat ska konverteras till format som klarar långtidslagring.
3. Dataformat i fält och variabler ska följa standarder.

4. Filnamn ska vara genomtänkta och inte innehålla exempelvis mellanslag och specialtecken.
5. Metadata i XML-format ska finnas vilket beskriver och identifierar den information som exporteras. Vanliga XML-vokabulär som används för att beskriva filer är RDF med Dublin Core, PRISM, METS, PREMIS, MIX, MODS m.fl.
6. Den exporterade informationen inkl. metadata ska beskrivas med leveransmetadata i XML-format och paketeras. Ett vanligt sätt att beskriva och paketera en arkivleverans är att använda XML-vokabulären METS eller EAD.

Mer detaljer för punkterna 2, 3 och 4 kan fås i TAM-Arkivs rekommendation för format. Vad gäller punkt 5 & 6 har TAM-Arkiv än så länge inte tagit fram några mer detaljerade instruktioner, men planerar att göra det när organisationen successivt bygger ut sin verksamhet för digital arkivering.

2.2.4 Återsökning och arkivimport

Användare av system kan behöva komma återsöka och återanvända information som har arkivexporterats ut ur systemet. I så fall måste det finnas funktionalitet för att kunna:

1. Ställa frågor och hantera svar från ett digitalt arkiv.
2. Beställa och ta emot en informationsmängd.

Systemet bör även ha funktionalitet för att kunna fråga ett digitalt arkiv om vilken funktionalitet det erbjuder och vilka format och versioner som det digitala arkivet använder. TAM-Arkiv har än så länge inte tagit fram några mer detaljerade instruktioner för återsökning och arkivimport, eftersom TAM-Arkiv än så länge inte har ett digitalt arkiv i drift.

3 Sammanfattning av krav för arkivfunktionalitet

Ett IT-system med stöd för arkivhantering ska ha funktionalitet enligt nedanstående.

1. Funktionalitet för vård av information. Föråldrade filformat ska kunna konverteras till nya filformat. Metadata som behövs för att kunna gallra och långtidslagra information ska kunna tillföras eller förändras.
2. Gallringsfunktionalitet. Det ska finnas funktionalitet för att information lätt ska kunna gallras utifrån de gallringsregler som framgår av organisationens dokumenthanteringsplan. Detta ska kunna göras utan kompletterande systemutvecklingsinsatser. Utförd gallring bör dokumenteras i en logg.
3. Arkivexportfunktionalitet. Information ska kunna exporteras ut till arkivbeständigt format med beskrivande metadata. Teckentabeller och filformat ska fungera för långtidslagring, dataformat ska följa standarder, filnamn ska vara genomtänkta och metadata i XML-format ska finnas vilket beskriver och identifierar den information som exporteras. Se även TAM-Arkivs rekommendation för format. Den information som exporteras ska även paketeras med leveransmetadata som beskriver leveransen på ett övergripande plan. Även leveransmetadata lagras i XML-format.

4 Bilagor

4.1 Litteraturförslag

Bodleian Library, *Paradigm: Workbook on Personal Digital Archives*, (Oxford 2007).

Consultative Committee for Space Data Systems, *CCSDS 650.0-B-1BLUE BOOK Reference Model for an Open Archival Information System (OAIS)*, (Washington 2002).

Lazinger, Susan, *Digital Preservation and metadata*, (Englewood 2001).

4.2 TAM-Arkivs rekommendationer och rutiner

Följande rekommendationer och rutiner kan laddas ner från TAM-Arkivs webbplats,
<http://www.tam-arkiv.se>:

TAM 1:2008 REKOMMENDATION - Vid förändring av ärende- och dokumenthantering

TAM 2:2008 REKOMMENDATION - Informationsförvaltning

TAM 3:2008 REKOMMENDATION – Versionshantering

TAM 4:2010 REKOMMENDATION - Arkivfunktionalitetskrav för IT-system

TAM 5:2010 REKOMMENDATION - Format för Långtidslagring