Specifikation
Tillägg

FGS Databas (Alternativ 2, relationsdatabaser) Tillägg

Förvaltningsgemensam specifikation (FGS) för databaser (Alternativ 2, relationsdatabaser med hjälp av standarden SIARD)

RAFGSXVX.XA2020MMDD

UTKAST

Kontakta oss

Information om arbetet med FGS:er
hittar du på vår webbplats:
www.riksarkivet.se/fgs-earkiv

Du kan även kontakta oss via e-post:
ra-fgs@riksarkivet.se

Juni 2020

Riksarkivet

Riksarkivet

UTKAST

Riksarkivet

INNEHÅLLSFÖRTECKNING

1 Inledning 4

2 Att använda FGS Databas (Alternativ 2 Relationsdatabaser) 4

2.1 Att titta på informationen i uttaget 5

3 Att placera informationen i ett informationspaket 5

3.1 Fält att ange i FGS Paketstruktur 6

3.2 Placering av filer i ett paket enligt FGS Paketstruktur 7

UTKAST

Riksarkivet

ÄNDRINGSLOGG

Version	Datum	Namn	Beskrivning
RAFGSXVX.XA2020MMDD	2020-06-08	Kommunalförbundet Sydarkivera	Remissversion av tillägget

UTKAST

1 Inledning

Detta dokument kommer att förändras och därför även att versions hanteras för att underlätta att se vilken som är den senaste versionen.

Detta dokument innehåller teknisk information som är föränderlig, men påverkar inte och förändrar inte användandet av FGS Databas (Alternativ 2, relationsdatabaser). Dokumentet innehåller även beskrivning av metodiken och de programvaror som finns tillgängliga för arkivering av relationsdataser med hjälp av SIARD.

2 Att använda FGS Databas (Alternativ 2 Relationsdatabaser)

Uttaget av en relationsdatabas görs med hjälp av ett verktyg i form av någon programvara, antingen något av de tillgängliga fria verktygen, kommersiella alternativ eller egen utveckling. Att manuellt genomföra en mappning och export av en relationsdatabas är inte något som rekommenderas på grund av komplexiteten.

Verktyg kan bland annat hittas via det europeiska byggnadsblocket för eArkivering (Connection Europé Facility, CEF, eArchiving Building Block) https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eArchiving. Både specifikationen för SIARD och en exempelprogramvara för att skapa och kontrollera SIARD-filer är en del av byggnadsblocket. Allt som produceras i byggnadsblocket för eArkivering är fritt att använda för alla, både leverantörer och upphandlande myndigheter.

Programvaran behöver vara installerad på en PC (alternativt arbetsstation eller server) och det behövs åtkomst till den eller de databaser som programvaran ska ansluta till. För att köra programvaran direkt i producentens tekniska miljö så behöver man använda en programvara med stöd för den aktuella plattformen. Det går även att fjärransluta till en databas och köra uttaget på distans. Det är möjligt att arbeta mot en databaskopia (databasdump), men då behöver man läsa in kopian i en databashanterare först som programvaran kan ansluta till.

Det beror på användarens tekniska kunskaper och förutsättningar om det är mest lämpligt att använda programvara med grafiska användargränssnitt eller arbeta kommandobaserat. Grafiska användargränssnitt sänker tröskeln något när det gäller att hantera tekniken för den som ska genomföra uttaget. Det är i vilket fall en förutsättning att användaren har grundläggande kunskaper om SQL, både för att kunna analysera databasen och för att skapa vyer med hjälp av SQL-frågor.

Tekniskt så behöver programvaran känna till port för att nå den aktuella servern, databasinstans och konto med användarnamn och lösenord. Kontot behöver ha rättighet till att läsa alla tabeller i databasen. Det behövs inte skrivrättigheter.

Det är i grunden tre olika funktioner som det behövs programvaror för att hantera:

- Skapa SIARD-filen
- Validera (kontrollera) mot SIARD XML schema att SIARD-filen stämmer med specifikation
- Visning av SIARD-fil

För att skapa SIARD-filen är det ett fåtal uppgifter som användaren behöver ange manuellt: Det handlar om datum för uttaget, vem som skapat SIARD-filen, klient för att skapa SIARD-filen, databashanterarens produktnamn, och var data kommer ifrån när det gäller tidsperiod och organisation. Dessa fält preciseras i dokumentet FGS Databas, Alternativ 2, relationsdatabaser).

SIARD-filen packar innehåll och metadata från databasen i XML-format. Det finns en metadatafil som dokumenterar tabeller, kolumner och relationer. Metadatafilen anger till exempel både datatyp enligt SQL standard och ursprunglig datatyp. Varje tabell får en katalog som innehåller ett XML-dokument och eventuella bifogade filer (stora objekt). Innehållet i tabeller sparas i ett XML-dokument per tabell. SIARD-formatet har stöd för att hantera stora objekt i databasen (till exempel kontorsdokument eller Pdf) antingen som inbäddad kod i XML-dokumenten eller genom att läsa ut

UTKAST

och spara filerna i sitt dokument- eller bildformat eller som textfil. För bevarandeändamål rekommenderas att läsa ut dokumenten som då blir tillgängliga för att hantera filidentifiering, konvertering och validering av de förekommande formaten.

Validering (kontroll) av XML-filerna sker med automatik genom att använda medföljande XML-schema. Verktygssviterna kan användas både kontroll mot XML schema och andra kontrollpunkter som verktygen stödjer. XML filerna är fullt möjliga att validera direkt mot XML schema med oberoende verktyg.

Verktygen kan även på ett generellt sätt visualisera innehållet i databasen. Till exempel ge ett användargränssnitt för att navigera och söka i innehållet i databasen. SIARD-formatet och verktygen stödjer att skapa vyer för att visa valda handlingstyper. Vilka handlingstyper som ska sparas som vyer beror på överenskommelsen och vilka handlingstyper som vanligtvis behöver lämnas ut. Ett arbete som arkiv och systemförvaltningsorganisationerna behöver göra tillsammans. På det sättet går det att återskapa vyer från systemet för enklare sökning i den arkiverade databasen. SIARD-filerna kan även återanvändas med hjälp av valfri databashanterare. Det gör det möjligt att återanvända data och också visualisera och göra data sökbart med alla de möjligheter som finns tillgängliga.

Eftersom programvaran sköter packning och registrering av metadata med automatik så kräver själva uttaget minimal arbetsinsats. Det som kräver arbetsinsats är att identifiera handlingstyper som behöver sparas som vyer och analys av tabeller som kan gallras. Eftersom processen att skapa SIARD-filer kan automatiseras i hög grad så är det möjligt att planera och genomföra arbetet stegvis.

Till exempel:

- Steg 1: Spara en fullständigt SIARD-fil.
- Steg 2: Genomföra analyser av databasen.
- Steg 3: Skapa en ny SIARD-fil, och denna gång gallra de tabeller som inte behöver bevaras i processen.
- Steg 4: Skapa och spara vyer så att det går att spara en SIARD-fil med färdiga vyer för visning och sökning.

För spårbarhet behöver de olika stegen dokumenteras i bevarandesystemet, förslagsvis i form av bevarandemetadata i PREMIS-format.

2.1 Att titta på informationen i uttaget

De verktyg som skapar uttaget innehåller ofta även möjligheter att visualisera uttaget och att söka information i uttaget. Det går även att använda andra verktyg för att titta på uttaget.

Det går även att återskapa databasen med hjälp av olika verktyg.

3 Att placera informationen i ett informationspaket

När uttaget har genomförts placeras detta i ett informationspaket följande FGS Paketstruktur. De dataelement som beskrivs i FGS Databas (Alternativ 2, relationsdatabaser) och de dataelement som i FGS Paketstruktur som definieras nedan ska användas för överföring, bevarande och återsökning av informationspaket innehållande relationsdatabaser. Dataelementen innehåller dock inte sådan information att det går att göra en generell informationsåtersökning i överförd information då en relationsdatabas i sig ofta är av sådan storlek att för att effektivt söka och presentera informationen krävs att man återskapar eller via ett gränssnitt som kan hantera uttagsformatet för relationsdatabasen kan visa informationsinnehållet.

UTKAST

3.1 Fält att ange i FGS Paketstruktur

I FGS Paketstruktur definieras ett antal fält som kan användas för att ange att i informationspaketet finns en relationsdatabas skapad med hjälp av "FGS Databaser (Alternativ 2, relationsdatabaser". Att använda dessa fält som beskrives i tabellen nedan betyder att man inte tar fram en egen anpassning av paketstrukturen utan endast använder de fält som redan är definierade.

Tabell 1: Specifika fält att använda i FGS Paketstruktur version 1.2

Element namn	Elementets utseende beskriven med hjälp av XPath i METS	Värden
Beskrivning	<mets:LABEL=”[Beskrivning]”	Skriv i klartext att detta är ett uttag av en relationsdatabas med hjälp av formatet SIARD
Informationstyp	<mets:TYPE=”[Informationstyp]”	Ange värdet “Databases”
Informationstyps specification	<mets: ext:CONTENTTYPESPECIFICATION=”[Informationstypsspecifikation]”	Ange namnet på denna specifikation, ”FGS Databas (Alternativ 2, relationsdatabaser) version 1”
Systemtyp	<mets: ext:SYSTEMTYPE=”[Systemtyp]”	Ange värdet ”SIARD”
Startdatum	<mets: ext:STARTDATE=”[Datum]”	Om möjligt ange datum för första posten som återfinns i uttaget.
Slutdatum	<mets: ext:ENDDATE=”[Datum]”	Om möjligt ange datum för sista posten som återfinns i uttaget.
System namn	<agent: ROLE=”ARCHIVIST” TYPE=”OTHER” OTHERTYPE=“SOFTWARE”><name>[System Namn]	Ange namnet på systemet som uttaget skett ifrån
System version	<agent: ROLE=”ARCHIVIST” TYPE=”OTHER” OTHERTYPE=“SOFTWARE”><note>[System Version]	Om möjligt ange version på systemet som uttaget skett ifrån
Levererande System Namn	<agent: ROLE=”CREATOR” TYPE=”OTHER” OTHERTYPE=“SOFTWARE”><name>[Levererande System Namn]	Ange namn på programvaran som använts för

UTKAST

Element namn	Elementets utseende beskriven med hjälp av XPath i METS	Värden
		att skapa SIARD-filen
Levererande System Version	<agent: ROLE=”CREATOR” TYPE=”OTHER” OTHERTYPE=“SOFTWARE”><note>[Levererande System Version]	Ange version på programvaran som använts för att skapa SIARD-filen
Funktion	<file: USE=”[Funktion]”	Ange att funktionen är SIARD för att peka ut SIARD-filen i paketet

3.2 Placering av filer i ett paket enligt FGS Paketstruktur

De skapade SIARD-filerna placeras i paketet enligt instruktionerna givna i leveransöverenskommelsen.

UTKAST