



Leanlink
Ao LKDATA

Datum
2010-05-19

Rev
1,5,0

Utredning om strategier kring vårt digitala kulturarv

Utredningsuppdrag beställt av Regionförbundet Östsam.
Utförd av LKDATA



Innehåll

Revisorer.....	3
1. Sammanfattning.....	4
2. Avgränsning.....	5
3. Behov - Inriktning.....	5
3.1. Summering Behov.....	6
4. Dataorganisation.....	7
4.1. Sammanslagen databas	7
4.2. Samsökning	8
4.3. Data integration (Heterogen sökning).....	8
5. Juridik.....	9
6. Säkerhet.....	10
7. Fjärråtkomst.....	10
8. Lagring.....	10
8.1. Långtidslagring.....	11
8.2. Dokumentformat.....	11
9. Lösningar i nationella projekt.....	13
9.1. Allmänt.....	13
9.2. Knowledge Management in Museums.....	14
9.3. Europeana.....	14
9.4. K- samsök.....	14
9.5. Sondera.....	16
10. Regionala projekt och Produkter.....	17
10.1. Västarvet.....	17
10.2. Carpe Cultura - Kalmar.....	18
10.3. Ödeshög kommun.....	19
10.4. Läns museet Östergötland.....	20
10.5. Sofie.....	20
10.6. Carlotta.....	21
10.7. Primus.....	21
10.8. Wunderkammer.....	22
10.9. Bygdeband.....	23
10.10. Tekniska lösningar för Bibliotek.....	23
10.11. CS-Library.....	23
10.12. Arena.....	24
11. Sammanfattad bedömning av olika verktyg.....	25
12. Rekommendation.....	26
Västarvet.....	26
Primus.....	26
Wunderkammer.....	27
Egenutveckling med opensource.....	27
CS-Library.....	27
13. Projektbeskrivning.....	28
13.1. Avtal.....	28
13.2. Centralt system.....	29
13.3. Konvertera och ansluta databaser.....	29



Konvertera databas.....	30
Ansluta databaser på distans (endast med CS-Library).....	31
13.4. Bygga en sökfunktion	31
13.5. K-samsök.....	32
14. Kalkyl.....	34
15. Systemförvaltning.....	35
16. Referenser.....	37
17. Bilaga 1 Kort beskrivning av begreppet OpenSource.....	38

Revisioner

Revision:	Datum:	Beskrivning:	Utförd av:
0.0.1	090731	Utkast	Jonas Wiman
0.0.3	090820	Disposition klar	Jonas Wiman
1.0.0	090923	Första delrapport	Jonas Wiman
1.3.0	090931	Andra delrapport	Jonas Wiman
1.4.0	100407	Utökad med projekt och drift	Jonas Wiman
1.5.0	100519	Slutlig version	Jonas Wiman



1. Sammanfattning

Meningen med en regional IT-satsning inom kulturarvsområdet är att knyta ihop det lokala med det nationella. Det handlar också om att sammanställa, bevara och tillgängliggöra information som idag finns utspridd hos många olika instanser och i många olika former. Den regionalt sammanställda informationsmängden kan länkas in i de nationella lösningar som finns med samma syfte och därmed öka i värde för alla.

Det finns två strategier för att sammanställa utspridd information. Antingen flyttar man in materialet och sammanställer det i ett enda system. Eller så kan man koppla ihop olika informationskällor och dynamiskt utföra sökningar i dem alla där de redan finns.

Att flytta information och inkorporera den i en regional databas har fördelar ur ett driftsäkerhets- och bevarandeperspektiv. Det kan också vara en hjälp för mindre informationsägare som får tillgång till professionella IT-resurser. Informationen kan fortfarande hanteras av den ursprunglige informationsägaren via distansuppkoppling. Att göra gemensamma sökningar i utspridda databaser är mer lämpligt när databaserna blir större och har egna krav på funktionalitet. Vid stora informationsvolymerna är det den enda möjliga vägen att gå.

Ett regionalt projekt i Östergötland föreslås kombinera dessa två strategier och utföra båda funktionerna i samma system. De organisationer som bidrar med data kommer i båda fallen att ha fullt ägarskap och tillgång till sitt eget material. Den regionala satsningen bör ur den enskilde informationsägarens synvinkel ses som hjälp till praktisk IT-hantering och att sprida det insamlade materialet i större sammanhang. Inte minst viktigt är också att det egna materialet kan användas tillsammans med andras material och därmed skapa ökad förståelse av vårt gemensamma kulturarv.

Utredningen har gått igenom ett antal projekt både nationellt och regionalt och kan konstatera att vi inte hittar några referenser som har skapat all den funktionalitet som vi efterfrågar. Några ligger nära, men ingen har alla delar. Det finns inte heller några kommersiella system som erbjuder allt vi vill ha. Däremot finns på marknaden alla olika delar som behövs och genom att kombinera olika lösningar så går det att bygga ett komplett system som motsvarar målbilden. Utredningen pekar ut vilka komponenter som är lämpliga att välja. Valen utgår från erfarenhet av praktisk IT-drift och pekar ut den mest funktionella och kostnadseffektiva lösningen. Det finns också en översiktlig beskrivning av hur ett konkret genomförande skulle kunna se ut och kosta.

Slutsatsen är att det är fullt möjligt att med rimliga medel bygga ett system som motsvarar alla behov vi har identifierat. Tidpunkten när systemet färdigställs ska dock inte ses som en slutpunkt utan en startpunkt i det stora arbete det innebär att fylla och utnyttja funktionen.

Utöver de IT-tekniska delarna finns många andra aspekter att ta hänsyn till. T.ex juridiska frågor, säkerhet och nomenklatur för dataorganisation över ämnesområden. Vissa kan utredas ytterligare men en del svar bör kanske få växa fram från erfarenheter som fås just i det praktiska genomförandet av projektet.



2. Avgränsning

Utredningen inriktar sig enbart på tekniska förutsättningar och möjligheter för att skapa en regional samverkan inom ABM-området. Verksamhetsmässiga behov, organisatoriska frågor eller ekonomi tas inte upp.

3. Behov - Inriktning

Inom området att samla in och digitalt publicera ABM-material finns i Sverige idag en mängd resurser och funktioner på många olika nivåer, från statliga initiativ till lokala ideella projekt. Den statliga utredningen *Kraftsamling* har gjort en grundlig utredning och rekommenderas för läsning.

Sammanfattningsvis finns ansatser på området på nationell, regional och lokal nivå. Det finns dock ingen tydlig samordning och ingen strategi för hur material från olika nivåer ska kunna länkas till varandra och bilda en helhet.

Störst och mest utvecklade är bibliotekens olika katalogsystem. Dom innehåller all relevant information och bra sökfunktioner finns. Bibliotekens system är dock inte kopplade till andra kulturområden utan fokuserar helt på de egna samlingarna och bokbestånden.

I andra änden av skalan hittar man privatpersoner eller små ideella föreningar som samlar material i egna arkiv och katalogiserar i en persondator med vanliga kontorsprogram.

Styrgruppen beslutar att arbetet inriktas mot att bygga en regional tjänst som kan fungera som en brygga mellan den nationella tjänsten K-Samsök och de olika lokala datasamlingar som finns. Det hade varit önskvärt att skicka data till den nationella söktjänsten Sondera också men möjligheten att göra det är mindre och prioriteras inte nu.

Den funktion som vi eftersträvar kallas fortsättningsvis för den *Regionala servern*, namnet betecknar en funktion oavsett teknisk implementation. Senare bör ett annat namn väljas.

Funktionen i den Regionala servern ska vara en säker lagringsplats för regionalt ABM-relaterat data. Den tar över och inkorporerar befintliga register och data från olika källor till en gemensam lagringsplats i en enda eller flera separata databaser. Detta kan göras som överföringar av engångskaraktär eller periodiskt från databaser som fortfarande utvecklas och uppdateras. Man kan också göra samsökningar i separata databaser och/eller bygga upp index för dessa. I de fall man tar över databaser permanent kan man ge informationssägare möjlighet att ansluta sig och uppdatera sitt eget data i det centrala systemet.

Huvudfokus för den regionala servern är att hantera mindre databaser från lokala aktörer. Större databaser som t.ex läns museets Sofia-system eller Bibliotekets samling av inscannat material kan generera för stor komplexitet och volym ifall de ska inkorporeras i den regionala servern. Det kan därför vara en fördel att koppla dessa direkt till K-samsök istället.

För publicering av innehåll bör man ha en regional portal som kan ha egna sökfunktioner och/eller använda externa söktjänster som K-Samsök, Sondera eller bibliotekens befintliga samsökning.



För bibliotekens del handlar arbetet i första hand om katalogsökning samt publicering av digitaliserat innehåll som kan utgöras av scannade texter, bilder eller ljud.

Projektarbetet inriktas på att definiera funktionen och föreslå en systemdesign för den regionala servern. Metoden är att inventera kommersiell marknad och vilka verktyg som används av andra regionala satsningar.

3.1. Summering Behov

Den regionala servern ska lösa följande uppgifter

- Ett i regionen centralt placerat IT-system med egen databas och systemfunktionalitet motsvarande ett huvudsystem för musei- eller arkiv-verksamhet. Det ska alltså inte bara bestå av en databas, utan även innehålla funktioner för administration och bearbetning av innehållet.
- Systemet ska kunna hantera och lagra text, bilder, ljud och film
- Innehållet i systemet ska kunna delas upp i samlingar. På samlingarna ska man kunna sätta behörigheter så att varje samling kan administreras och uppdateras av olika användare (informationsägare)
- Informationsägare ska kunna ansluta sig på ett säkert sätt över internet och administrera sin egen samling inom det regionala systemet.
- Till systemet ska man kunna importera databaser från en mängd olika system och databaser. Fältmatchning måste göras för varje typ. Importerna är av engångstyp och källsystemet förutsätts upphöra efter överföringen. Innehållet övertas av det regionala systemet.
- En funktion ska finnas för att kunna avsöka och skapa ett index av innehållet i databaser som finns placerade externt och inte importeras till det centrala systemet (funktion som en regional K-samsök). Det kan vara en integrerad eller separat teknisk lösning. Avsökningen görs periodiskt och automatiskt via säkra uppkopplingar till respektive system. Alla olika typer av databaser ska kunna avsökas, fältmatchning görs vid etablering för varje typ.
- Systemet ska erbjuda samsökningar inom alla egna samlingar kombinerat med externa databaser. Sökresultatet ska utgöras av en sammanhållen och sorterad informationsmängd oavsett informationens ursprung och typ
- Resultatet av sökningar ska kunna presenteras på en webb. Verktyg för webbpublicering kan ingå i systemet, men utveckling av komplett webb ingår inte i projektet
- Systemet ska ha en säker driftsmiljö och etablerad systemförvaltning. Information som har importerats ska anses ha en garanterad lagring
- Utöver det centrala systemet och sökfunktioner ska förvaltningen kunna erbjuda tjänster för säker långtidslagring av filbaserat material i stora volymer t.ex scannade böcker, bild- och film-arkiv. Rådgivning och konsulttjänster vid etablering av lagring bör kunna erbjudas.



4. Dataorganisation

System med olika användningsområden har olika uppsättningar av uppgifter (fält) som kan registreras i varje post. Även system med samma användningsområde men från olika leverantörer skiljer sig också åt i fråga om innehåll och benämningar på saker. Det kallas tekniskt för att man har olika "datamodeller". Den exakta datamodellen är inte synlig för användaren och betraktas ofta av systemleverantörer som en företagshemlighet.

Vid import och samsökning så uppstår problemet attfälten i olika system inte kan jämföras med varandra rakt av. Man behöver "översätta" innehållet för att få en användbar gemensam information. Vid arbetet skapar man en "översättningstabell" där man bestämmer sig för hur varje fält ska korrespondera med fält i andra system. Arbetet kallas fältmatchning. Många av de referenser vi talat med anger detta som ett tidsödande och svårt arbete med många fallgropar. Problemet är inte bara tekniskt utan kan också gälla skillnader i hur föremål och företeelser beskrivs inom olika fackområden. Man talar om olika "vokabulärer".

När man söker nationella referenser så återkommer detta som en svårighet flera gånger. Till exempel den statliga utredningen "Inför ett svenskt kulturnät" från 1996 adresserar samma fråga.

För att slippa lösa samma problem gång på gång har det funnits försök till att införa en gemensam standard som man kan hålla sig till. Nordiska Museet inrättade 1990 i detta syfte en arbetsgrupp med namnet "Insam" vars uppdrag var att ta fram en nationell standard och i förlängningen möjliggöra gemensam sökning i olika databaser. Arbetsgruppen las ner 1999 och dess resultat (Swesam) för en tynande tillvaro. Den enda tillämpning vi kunnat finna som använder Swesam är systemet Carlotta. Det finns också ett internationellt arbete inom ICOM som har skapat en internationell ISO-standard. Se vidare under punkt 9.2 Knowledge Management in Museums

Att uppnå en gemensam struktur och standardiserat sätt utbyta museidokumentation är oerhört lovvärt, men har inte slagit igenom för praktisk användning. När vi idag behöver etablera samarbete så måste man därför gå via fältmatchning i de existerande datamängderna. Fältmatchningen kan dock göras på olika sätt vilket beskrivs i nedanstående tekniska alternativ.

4.1. Sammanslagen databas

Den vanligaste metoden att lösa dataorganisationen är att slå ihop databaserna till en enda. Man bestämmer en datamodell för det centrala systemet och gör en fältmatchning för varje databas som ska importeras. Resultatet blir en enda stor databas som är organiserad på ett homogent sätt.

Det är ett stort arbete att importera många databaser och det växer med antalet varianter på system man vill konvertera data ifrån. Fältmatchningen måste gås igenom manuellt för varje typ av databas men därefter kan man bygga en automatisk funktion som gör volymarbetet. Det kräver att de ingående databaserna är kända och tekniskt dokumenterade. Viss handpåläggning och manuell hantering av undantag och fel behöver nästan alltid göras.



Ett alternativ som det också finns exempel på i vårt material är manuell omregistrering av alla registerposter i det gemensamma systemet. Man kan låta informationssägarna själva, kanske olika hembygdsföreningar, lägga in sina uppgifter manuellt. Metoden är hanterlig för mindre hembygdsföreningar men för stora databaser är det ett tidsödande och kanske oöverstigligt arbete. Det kommer också oundvikligen att ske registreringsfel under arbetet.

Det är stor skillnad på att importera en databas en gång eller att göra det flera gånger. Engångsimporter används när innehållet ska flyttas och det gamla systemet tas ur drift. Upprepade importers krävs om båda systemen ska finnas kvar och vara levande. Vid engångsimporter så kan en del saker lösas med manuell handpåläggning. Även om det kan vara mödosamt att matcha fält mot varandra är arbetet oftast överkomligt och tekniskt okomplicerat.

En upprepad import som ska köras automatiserat med en viss periodicitet ställer mycket högre krav. Främst på ”informationskvaliteten” i databasen, d.v.s att alla fält på alla poster verkligen innehåller precis det de ska och inget annat. Detta är normalt svårt att uppnå vid större informationsmängder. Därför krävs oftast löpande övervakning och felhantering av importen vid varje köringstillfälle. Vid vidareutveckling av endera av systemen så måste själva importfunktionen anpassas och det kräver en fortgående samordning av systemförvaltningen mellan samtliga inblandade system.

4.2. Samsökning

En annan metod att kunna söka i många databaser är att inte slå ihop dem utan istället låta dem finnas var för sig och utföra sökningar mot var och en - Samsökning. För att kunna göra det så krävs normalt att strukturen är identisk i varje databas eftersom själva sökfrågan som skickas ut är densamma för alla. Exempel på detta är t.ex samsökning i flera olika installationer av samma system, t.ex Sofia eller Primus. I fallet med Primus har man gjort detta till en del av produkten med tjänsten ”Digitalt Museum”.

En variant av samsökning som används av t.ex K-samsök och de flesta bibliotekssystem bygger på att ingående databaser kan presentera sitt innehåll efter en fastslagen mall oavsett underliggande struktur. Bibliotekskataloger använder för detta en internationell standard som heter z39.50. I K-samsök har man definierat en egen standard som alla måste följa.

Traditionell samsökning kräver alltså att fältmatchningen görs inom varje system i kopplingen till den gemensamma mallen. Genom anpassningen kan olika system besvara en standardiserad sökfråga som skickas likadant till varje databas. Att göra dessa anpassningar innebär förändringar i det egna systemet för alla som ska anslutas och det kan bli svårt eller omöjligt för mindre databasägare. Metoden är därför olämplig för den regionala databasen.

4.3. Data integration (Heterogen sökning)

Med modern teknik kan samsökning också göras mellan databaser som är sinsemellan olika. Istället för att anpassa alla databaser till en sökfråga som ställs på samma sätt till alla, så skraddarsyr man själva sökningsfunktionen med en anpassning för varje databas. Man översätter användarens sökning så att man skickar frågan på olika sätt för varje databas. Fältmatchningen är därmed inbyggd i själva sökverktyget så att varje databas får en fråga som

är anpassad efter dess uppbyggnad. Svaren sammanställs sedan centralt och kan presenteras på ett enhetligt sätt.

Metoden medför ett mindre ingrepp än att konvertera hela databaser, och det lämnar alla databaser intakta för uppdateringar och framtida utveckling. Tekniken har inte funnits särskilt länge och den ställer högre krav på servrar och programvara. Metoden kallas internationellt för ”data integration”. På svenska väljer vi att kalla det för ”heterogen sökning” som en motsats till sökningar i ett homogent data.

Den engelska Wikipedia har en utmärkt artikel om ämnet:

http://en.wikipedia.org/wiki/Data_integration

Det finns flera kommersiella verktyg som kan användas för data integration men ett av de mest intressanta alternativen är faktiskt gratis. Opensource-projektet **Talend** är designat just för att skapa sökningar i många olika databaser samtidigt. Det används t.ex i webbtjänster där man söker efter varor i många olika nätbutiker och jämför pris. Talend utgör inget komplett system utan en modul som man behöver komplettera med t.ex funktioner för presentation. Ett exempel på en webbtjänst som byggts med hjälp av Talend är: <http://www.allaplagg.se/>

Jag rekommenderar läsaren av den här rapporten att besöka Alla Plagg. Den utgör ett bra exempel på både heterogen samsökning och modern utvecklingsmetod. Gör man en sökning i Alla Plagg så sker just nu en sökning i 26 olika databaser som inte alls är anpassade till varandra eller till sökverktyget. Svaren kommer i en sammanställd lista på ett enhetligt sätt. Exakt samma funktion skulle kunna användas för att söka i olika ABM-databaser.

Alla Plagg är också ett exempel på modern och mycket effektiv systemutveckling där man använder moduler av färdig programvara som har satts ihop. Alla Plagg är helt och hållet byggt av moduler som hämtats gratis som opensource och sen anpassats. Därför har funktionen kunnat utvecklas av bara två personer som hobbyprojekt vid sidan av deras ordinarie arbeten. Grundsystemet byggdes på några dagar, att ansluta en ny databas tar i genomsnitt några timmar.

Exempel på kommersiella verktyg för samma typ av funktion är ”Stylus Studio” och ”Microsoft SISS”.

Vid heterogen sökning krävs ofta någon form av index som lagras centralt för att snabba upp sökningen. Ett index motsvarar en innehållsförteckning i en bok. Det skapas genom att man läser igenom boken (databasen) och utifrån det bygger upp indexet som en liten databas med bara rubriker i den centrala servern. Sökningar sker sen i första hand mot indexet och inte mot den verkliga databasen. Sökningen går därmed mycket snabbare. Utan index tar sökningen så lång tid som det tar för den långsammaste databasen att svara. Index används även vid modern samsökning i homogent data som t.ex K-samsök men inte i äldre varianter som Z39.50. För att skapa index använder man ett separat verktyg och även här finns gratisalternativ som opensource, t.ex Apache Lucene.

För en kort beskrivning av vad som avses med ”OpenSource”, se bilaga sist i dokumentet.

5. Juridik

Juridiska frågor runt framförallt Personuppgiftslagen och Lagen om upphovsrätt är central vid publicering av material via Internet.



Frågan har genomlysts som en del av arbetet med K-samsök och den statliga utredningen *Kraftsamlings*. Resultatet finns utförligt beskriven som Bilaga 5 till den utredningen. Sammanfattningsvis är läget osäkert och prejudikat saknas inom många områden. Det som de flesta referenter är överens om är att man vid sammanslagning eller samsökning av olika informationsmängder bör eftersträva att det juridiska ansvaret inte tas över utan ligger kvar hos den ursprungliga informationsägaren.

6. Säkerhet

Det finns en risk för en enskild hembygdsförening att publicera information om dyrbara föremål och därmed bjuda in till inbrott. Från Västarvet kommer tipset att för varje objekt kunna ange om dess fysiska placering ska döljas, och låta varje informationsägare själv avgöra vad som ska visas och inte. Ett objekt med dold fysisk placering kommer att falla bort från alla typer av geografiska söktjänster och information om placering skickas inte till K-samsök.

I tjänsten Carpe Cultura har man adresserat problemet med upphovsrätt för bilder. Man har där lagt till en funktion som kan göra en digital vattenstämpel i bilderna så att deras ursprung alltid kommer att synas även om de kopieras från internet.

7. Fjärråtkomst

Vissa databaser är statiska och uppdateras inte längre. Andra är levande och ägs och används av en informationsägare.

Att permanent flytta databaser till den regionala servern är inget problem om de är statiska. Men om de fortfarande behöver uppdateras måste de göras tillgängliga för sina ägare via distansåtkomst. Uppkopplingen bör gå via Internet och vara säker med avseende på kryptering och inloggning. Det är inte lämpligt att bygga denna funktionalitet ihop med den regionala servern eftersom det är ett stort projekt i sig.

Istället rekommenderas att man använder färdiga tjänster som kan kopplas på. T.ex har LKDATA en färdigpaketerad tjänst som med minimal anpassning kan användas för att detta. Tjänsten kan t.ex avropas av den regionala organisationen som sen självt sprider den till med de lokala informationssägarna.

8. Lagring

Att erbjuda en säker och stabil lösning för lagring av data är ett uttalat syfte med den regionala servern. Om man väljer alternativ med en regional server som utgår från en egen databas så kommer lagringen att ingå när man upphandlar en drift för hela den regionala servern.

Tjänsten ”lagring” är en normal del av en driftstjänst och det finns gott om leverantörer både i form av kommunala IT-enheter i länet eller hos privata företag. Priset kan antingen sättas som ett belopp per lagrad mängd och tidsperiod, eller ingå i en komplett tjänst som innefattar drift av servrar, löpande underhåll, felsökning m.m.

Lagring av speciella informationsmängder som t.ex filmarkiv, scannade böcker och annan utrymmeskrävande information kan lösas genom att man köper lagring till fasta priser som en



enskild tjänst. Det ger en säker och stabil lagring med god tillgänglighet. Extra säkerhet är linjärt kopplat till prislappen

Den regionala servern kan också byggas (i ett alternativ med enbart heterogen sökning) så att varje databas ligger kvar hos sin ursprunglige upphovsman. Sökningar görs via ett index som byggs centralt. Detta kräver ur ett lagringsperspektiv att varje informationsägare själv behöver adressera frågan. Det viktiga är att informationsägaren inser att lagring måste få kosta någonting och inte kan lösas gratis eller för lättvindigt.

8.1. Långtidslagring

Med undantag av runstenar och gravyrer i ädel metall så finns det idag inget medium som är bättre än text på papper för att bevara information under verkligt lång tid. Alla digitala medier har begränsad livslängd, tekniskt, fysiskt eller systemmässigt.

Tekniskt eftersom det inte längre finns utrustning som kan läsa den lagrade informationen.

Fysiskt genom att data förstörs av ren ålder på lagringsmediet.

Lösningen på båda dessa problem är relativt enkel. Man måste regelbundet kopiera allt data till nya format. Eftersom nya medier i regel har högre kapacitet än de gamla så är det inget stort problem. Sveriges Radios arkiv rekommenderar att all information oavsett typ av lagringsmedia skrivs om till nytt media en gång vart femte år.

De tjänster för lagring som ingår i driftstjänster för servrar eller som man hyr separat på marknaden, uppfyller oftast mycket väl kravet om periodisk omkopiering. Vid långa lagringstider får man inrikta sig på att ibland flytta materialet till nya leverantörer vartefter marknaden förändras. Det som spelar större roll än den enskilda lagringstjänsten är därför en aktiv huvudman som ställer krav på leverantören och bevakar att lagringen fungerar vid varje tidpunkt. I avtalen med olika leverantörer bör man se till att det finns goda möjligheter att få ut allt data på ett hanterbart sätt och säkert kunna flytta det till andra leverantörer.

8.2. Dokumentformat

Utöver de tekniska och fysiska begränsningarna finns också frågan om det finns system som kan läsa och tolka informationen i framtiden. När det handlar om lagring i datorer så pratar vi om olika filformat som ändras och utvecklas över tid. Av de idag vanliga formaten för datafiler är det få som har mer än 10 år på nacken, och majoriteten av de format som användes för 10 år sedan är inte möjliga att läsa idag även om man lyckats bevara de fysiska filerna och kopiera över dem till ny utrustning.

Samma problematik gäller också data som lagras i databaser. Dessa är ännu känsligare för att bli föråldrade eftersom de är beroende av programvara från en specifik leverantör. Oftast lämnar leverantören bara support för en eller två versioner bakåt i tiden. Eftersom företagen aldrig publicerar exakt hur databasen fungerar så finns mycket små möjligheter att återskapa data i en gammal databas där kanske inte ens företaget som skapade den finns kvar.

Utöver filformaten så har man också standarder på en lägre nivå. Det handlar om typsnitt och metoder för att skriva tecken för olika alfabet, t.ex våra ÅÄÖ. Den enda standard som inte ändrats sedan PC-datorn introducerades 1981 är att varje bokstav skrivs med åtta stycken ettor



eller nollor i datorernas minnen. Det är fortfarande standard och det finns inget som tyder på att det kommer att ändras.

Det är inte möjligt att inom denna utrednings ramar komma med en utförlig rekommendation angående lagring i alla olika situationer. Det finns ingen lösning som är rätt för alla situationer. Varje projekt måste hitta den lösning som passar ämnesområde och förutsättningar bäst. Man behöver också titta på vilka verktyg och metoder som i varje enskilt fall är lämpliga för att konvertera befintligt material till ett bra lagringsformat. Här följer istället en liten generell genomgång av format för lagring av olika typer av material.

Text och Bild

När det gäller lagring av text och bild så finns en generell lösning som används av många arkiv idag. Det är ett filformat som heter PDF/A. PDF/A är en variant av det välkända formatet PDF från företaget Adobe. För denna variant har Adobe släppt hela specifikationen fri och det innebär att man inte är hänvisad till programvara från Adobe för att läsa filerna. Många olika leverantörer kan skapa program som kan göra det. I en framtid där det teoretiskt inte längre finns några program som kan läsa PDF/A så kan man så länge den publicerade specifikationen finns kvar återskapa en programvara som kan läsa filerna.

PDF/A skiljer sig från det vanliga PDF-formatet genom att det bara kan innehålla text och bild, det kan inte finnas filmer, ljud eller inbyggd programkod (script). Formatet innehåller sin egen definition av typsnitt för texten, men är fortfarande beroende av att datorn som ska läsa filen använder 8-bitars teckenlängd.

PDF/A utgör ingen garanti för digital lagring i evighet, men det är det bästa alternativ vi känner till just nu. Riksarkivet rekommenderar också PDF/A för all digital långtidsbevaring. Läs mer om PDF/A på <http://en.wikipedia.org/wiki/PDF/A>

Ljud

Sveriges Radio håller just nu på att digitalisera hela sitt sändningsarkiv och utgör därför en bra referens för hur man lagrar digitalt ljud. Hos dem görs all lagring samma format som används på vanliga CD-skivor för musik. Det är ett filformat som heter CDDA. Filformatet är en internationell standard och väldokumenterat. Förutom själva ljudet innehåller filerna också en identitetsbeteckning i formatet ISRC som identifierar utgivare och inspelning, som ett slags ISBN-nummer för ljudinspelningar. ISRC-koder administreras i Sverige av IFPI (www.ifpi.se).

Sveriges Radios arkiv lagrar alla inspelningar som redan är digitala i sitt originalformat. Analogt material som man digitaliserar väljer man att avkoda med högre kvalitet än vad en CD skiva normalt har. Man använder därför 24 bitars upplösning och 96 kHz samplingsfrekvens. En vanlig musik-CD har 16 bitars upplösning och 44 kHz samplingsfrekvens.

Sveriges Radio tipsar också om att Visarkivet i Västervik finns nära oss och att där finns en hög kompetens runt lagring av ljud. Jag har dock inte varit i kontakt med dem.

Film

Jag har inte hittat någon referens för digital lagring av stora volymer film. Projektet Östgötafilm har dock sysslat med detta under lång tid och har samlat på sig mycket erfarenheter som inte behöver upprepas här.



Tekniskt använder Östgötafilm Adobe Flash Player för att visa filmer. Det är ett format som ägs av företaget Adobe men har mycket stor spridning på Internet.

Film lagras inte i fasta filformat som text och bild brukar göra. En datafil för film kallas för en "behållare" och det finns olika varianter av dem. Filmen som ligger i behållaren är kodad med en "Codec" och det är den som styr hur filmen ska spelas upp. Olika Codec's används separat för bild och ljud. Olika "Codec-format" kan finnas i samma typ av behållarfil, och olika behållarfiler kan innehålla filmer med samma Codec-format. Det går alltså inte att se via filnamnet vilket format en film är kodad med som man kan med nästan alla andra filtyper. Man måste veta det eller pröva sig fram.

För långtidslagring får detta följden att man behöver välja både ett behållarformat och ett codec-format för bild och ljud som man tror kan ha en lång livslängd och har möjlighet att hanteras av olika uppspelningsutrustningar långt in i framtiden.

Adobe Flash Player är speciellt genom att det alltid använder en fast uppsättning Codec-format som heter H.264 för bild och HE-AAC för ljud. H.264 är en ISO-standard och används även för TV-sändningar i HD.

Normalt undviker man format som ägs och kontrolleras av ett enda företag när man vill skapa stor tillgänglighet men eftersom man slipper variationer med olika Codec inom samma behållarformat och genom sin mycket stora spridning på Internet är Flash Video formatet generellt ett bra val för långtidslagring av film.

9. Lösningar i nationella projekt

9.1. Allmänt

På riksplanel finns två satsningar som utmärker sig. K-samsök och Sondera. Båda ger nationell sökning i underliggande databaser. K-samsök inom främst Museiområdet och Sondera inom Bibliotek och Arkiv. Tillsammans utgör de alltså en komplett nationell söktjänst för ABM-området.

K-samsök är en indexerings-tjänst som söker igenom anslutna databaser utan att kopiera dess data och bygger upp ett eget index över dess innehåll. Ambitionen är att ansluta så många databaser som möjligt och lösningen erbjuder en öppen anslutning både för informationsägare och mot webbplatser som vill göra sökningar.

För att göra en viss datamängd tillgänglig via K-samsök krävs en anslutningsprocess där man formaterar om sitt innehåll på ett sätt som är kompatibelt med K-samsök. Detta gör att det finns en tröskel för att kunna publicera ett material i K-samsök som bara lite större organisationer klarar av att ta sig över.

Sondera är en samsökningstjänst mellan tre databaser, Nationella Arkiv Databasen. LIBRIS och Svensk Media Databas. Tjänsten är mera statisk och har ingen uttalad ambition att ansluta fler databaser för sökning eller leverera sökresultat till andra webbtjänster. Däremot pågår en allt bredare anslutning av folkbibliotek och länsbibliotek till LIBRIS och därmed breddas innehållet alltmer.



9.2. Knowledge Management in Museums

Är ett forskningsprojekt som initierades 2005 av flera läns museer och universitet. Idag drivs det av sju läns museer bl.a Stockholms Läns museum, och Luleå universitet. Projektet är i huvudsak inriktat på att lösa problem med samsökning mellan olika databaser inom museiområdet. Man har även gjort flera andra delprojekt, bl.a ett arbete om långsiktigt digitalt bevarande som kan vara av intresse för oss.

De resultat som projektet publicerar handlar till största delen om att införa standardiserad terminologi och klassificering av information runt museiföremål. Man talar om att ”konceptualisera” och definiera vanliga begrepp för att alla ska uppfatta saker på samma sätt. Man jobbar också med en standard som heter ”Outline of Cultural Material” (OCM) som är ett vanligt klassifikationssystem på skandinaviska museer.

En annan standard som omnämns är CIDOC CRM Core som är en ISO-standard (ISO 21127). CIDOC ([http://cidoc.mediahost.org/home\(sv\)\(E1\).xml](http://cidoc.mediahost.org/home(sv)(E1).xml)) är en kommitté från det internationella samarbetsorganet för museer ICOM med syfte att verka för förbättrad museidokumentation. Deras standard CRM Core utgör enligt uppgift en ”modell för lagring och utbyte av museidokumentation”. Jag har inte satt mig närmare in i hur det fungerar. Standarden är dock kompatibel med ”Dublin-Core” som i sin tur är en metod för beskrivning av metadata som ibland används inom bibliotek.

Om man ska försöka sig på en försiktig bedömning av projektet som helhet så är detta ett forskningsprojekt på hög nivå. Mycket finns att lära och ta intryck av, men projektet ligger ganska långt från praktiska tillämpningar. Det är ett stort avstånd mellan den höga akademiska nivån i de här arbetena och de relativt traditionella verktyg som vi ser hos våra målgrupper och på marknaden.

Hemsida för Knowledge Management in Museums finns här: <http://www.kmmuseum.se/>

9.3. Europeana

På Europa-nivå finns ett projekt för att tillgängliggöra kulturarv:
<http://www.europeana.eu/portal/>

Inom Europeana finns även en variant som syftar till att bygga regionala portaler: ”Europeana Local”. Deras svenska presentation finns här:
<http://abm.murberget.se/europeanalocal>

Både Knowledge Management in Museums och K-samsök är deltagare i Europeana Local

9.4. K- samsök

K-samsök är den praktiska åtgärden som sprungit ur den statliga utredningen *Kraftsamling*. Det är en nationell söktjänst som länkar information från många olika databaser med kulturellt relaterat innehåll och ger en möjlighet till gemensam sökning.

Tjänsten har ett eget interface mot användare men är huvudsakligen avsedd att finnas som en osynlig underliggande funktion till olika e-tjänster. E-tjänster kan byggas av vem som helst och anslutas till K-samsök.



Projektet är under utveckling och har nyligen släppts i en första testversion.

Söktjänster som använder K-samsök finns idag hos ett fåtal organisationer

Riksantikvarieämbetet: <http://www.kringla.nu>

K-samsök demosida: <http://www.kulturarvsdata.se/demo/demostart.html>

Västarvet utnyttjar också K-samsök, mer om det nedan.

Organisation och projekt

Projektet drivs av Riksantikvarieämbetet som också ansvarar för att tjänsten får en långsiktig förvaltning. Centralmuseernas samarbetsråd stöttar projektet initialt med medel. Övriga medverkande bidrar med arbetstid. Arbetet påbörjades under vårvintern 2008. Tre arbetsgrupper har tittat på protokoll, metadata samt upphovsrättsliga och policymässiga frågor

Hemsida, dokumentarkiv:

http://www.raa.se/cms/extern/aktuellt/projekt/k_samsok_forsta_sidan.html

K-samsök togs i testdrift i februari 2009. Projektets tidsplan sträcker sig fram t.om 2010.

Planen var att släppa första skarpa version (1.0) i slutet av 2009. Där är man inte riktigt utan i skrivande stund, mars 2010, är det version 0.99 som används.

Teknik

K-samsök söker av och indexerar information från olika databaser och gör dem sökbara.

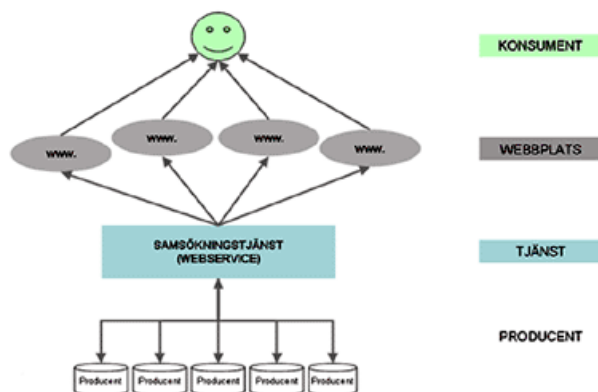
Funktionen kopierar inte data utan läser av databaser med ”harvesting-teknik” och skapar ett index som lagras i egen databas. Bibliotekskatalogen LIBRIS och EU:s kulturarvsportal Europeana är exempel på andra tjänster som fungerar med samma teknik.

För presentation fungerar K-Samsök som en webbtjänst (webservice) utan eget interface. De som vill utnyttja tjänsten gör egna webbplatser med sökfunktioner och använder K-samsök som underliggande söktjänst. Än så länge är det bara informationsägarna själva som börjat göra sådana egna webbplatser, men man lobbar för att få större spridning. Även kommersiella aktörer kan använda tjänsten på samma sätt, några intressenter finns redan.

K-samsök innehåller många olika datatyper, vilka som används väljs av födande system.

Information kan presenteras mot webbar i flera olika format, HTML, XML, RDF

I bilden nedan motsvarar den blå rutan K-samsök. De grå fälten är webbar som kan utvecklas av olika intressenter med olika inriktning





Roller

- Informationsproducenten:
Ansvarar för informationens kvalitet, förvaltning av informationen och copyright.
Ansvarar för driften av sina system och sin öppning mot K-samsök. Levererar valda data upp till K-samsök
- Samsökningstjänst (K-samsök)
Sammanställer information från producenterna. Lagrar ofta använd information temporärt eller permanent. Levererar information till webbplatser
- Webbplatsinnehavare:
Ansvarar för sin webbplats och paketering av informationen
- Konsument:
Använder kulturarvsinformation, interagerar och samskapar

Innehåll

Än så länge bidrar ett fåtal organisationer med sökbara data, t.ex. Riksantikvarieämbetet, Statens historiska museer och Västarvet. Fler organisationer ingår i samarbetet. Riksarkivet och Kungliga biblioteket är med som drivande i projektet. Sen finns också arbetsgrupper från: Etnografiska museet, Nationalmuseum, Tekniska museet, Arkitekturmuseet, Statens museer för världskultur, Statens historiska museer, Livrustkammaren Nordiska museet, Västerbottens museum.

K-samsök räknar främst med läns museer och hembygdsföreningar som ytterligare informationskällor

9.5. Sondera

Sondera är en söktjänst för att söka i Bibliotek och Arkiv. Adressen till söktjänsten är: <http://sondera.kb.se/>

Sonera är mer sluten än K-samsök och inriktar sig på samsök i tre stora databaser utan direkt ambition att utöka med flera i framtiden.

Baserna är:

- | | |
|--------|---|
| NAD | Nationella Arkiv Databasen. Riksarkivet. Lista över alla offentliga arkiv i Sverige. |
| LIBRIS | Katalog för samsök inom universitets och forskningsbibliotek, i viss omfattning vanliga folkbibliotek. KB |
| SMDB | Svensk Media Databas, KB. Arkiv med musik, film TV m.m. |

Sondera innehåller en egen söktjänst men inte något öppet interface för andra utvecklare. Däremot ges instruktioner om hur man länkar in deras sökruta på egna sidor. Tekniskt är jag inte riktigt säker på om Sonera gör en riktig samsökning eftersom resultatet inte presenteras sammanvägt utan i tre separata listor, en för varje ingående databas. Det kan istället vara frågan om tre parallella sökningar vilket ur teknisk synpunkt är en mycket enklare lösning.



Ägs och drivs av KB. Offentlig release den 090401

Pressrelease: <http://www.kb.se/aktuellt/nyheter/2009/Rakt-ner-i-nationalarkiven/>

Det finns idag inga planer på utökning av samarbeten eller integration med andra söktjänster. Men Sondera utesluter inte sådana möjligheter i framtiden.

LIBRIS utökas allteftersom med kataloger från folkbiblioteken, länsbibliotek m.m. och utvecklingen går emot att det kommer att bli en nationell databas över alla biblioteksbestånd. Sondera är inte ett tidsbegränsat projekt utan ska finnas långsiktigt. Utveckling kommer att ske efter behov.

10. Regionala projekt och Produkter

10.1. Västarvet

Västarvets sökportal finns här:

http://www.vastarvet.se/kulturvast_templates/Kultur_ArticlePage.aspx?id=7973

Det finns också andra varianter av sökportalen, t.ex en som anpassats för barn:

<http://www.vestrogothia.se>

Västarvets funktion liknar i flera delar det som vi tänkt för vår Regionala Server. Funktionen bygger på att man har en gemensam databas med SQL-server och till den importerar man data från alla databaser man vill söka i (sammanslagen databas). Originaldatabaserna har utgjorts av flera hembyggda Access-databaser samt tre stycken Sofie-system. Vid importen har man manuellt matchat de olika fälten i databaserna emot den gemensamma mallen vilket varit ett stort arbete. För att minska komplexiteten vid sammanslagningen av databaserna har man valt att göra separata databaser för föremål och bilder.

Importen har än så länge bara gjorts vid ett enda tillfälle och databasen är en statisk informationsmängd som med tiden åldras. Man arbetar nu på en funktion som löpande kan uppdatera informationen från de ingående databaserna när det förändras. Funktionen byggs med verktyget Microsoft SQL Server Integration Server (SSIS) och fungerar så att den via tidsstämplar (en märkning med tidpunkt för senaste ändring i varje post) i originaldatabasen känner av vilket data som har ändrats sedan senaste körning, och kopierar över bara ändringar och ny information. Funktionen är idag i drift för en av käll-databaserna.

Västarvets har inte någon funktion för att arbeta med informationen där den ligger i det centrala lagret (kallas mellanlagret).

Västarvet publicerar sina data till K-samsök och genom det får varje objekt med automatik en URL (internet-adress) som kan hittas av t.ex. Google. Det har gjort att 70% av trafiken idag kommer från personer som har fått träff på Google och inte gått via Västarvets egen söksida eller K-samsök. Man ser detta som ett oväntat men välkommet sätt att få ytterligare spridning och ökad lättillgänglighet.

Tidigare hade Västarvet en egen sökfunktion mot den egna databasen. Eftersom databasen är uppdelad i kategorierna: Foton och Föremål så fick man då nackdelen att sökningen gjordes per kategori och inte samlad. Eftersom ingen av de ingående käll-databaserna har med ljud eller film så fanns heller ingen funktion för det.



I den senaste versionen har man gått ifrån detta och har istället byggt en sökfunktion som går via K-samsök för att göra sökningarna. Man gör alltså sökningen direkt i K-samsök men förutom de parametrar som användaren söker på så skickar man med filter för att få träffar bara inom de egna samlingarna. Sökresultatet blir en sammanhållen lista inom alla kategorier (foton, föremål, m.m) och kan presenteras på tre olika grafiska sätt. Filtreringar kan göras inom sökresultatet m.m.

Själva sökfunktionen ligger på en egen server och kan anropas från flera olika håll, t.ex från olika museers egna hemsidor. Varje hemsida kan lägga på egna sökfilter och därmed anpassa sökresultaten till egna preferenser.

Den här nya sökfunktionen är fortfarande under utveckling. När den är klar avser Västarvet att släppa den fri för att användas och vidareutvecklas av vem som helst under villkor som motsvarar opensource .

Västarvet har också en funktion där besökare kan lämna kommentarer till varje objekt. Det ger värdefull ny information t.ex när folk känner igen personer och platser på bilder och meddelar detta. Informationen granskas och skickas sedan på datafiler tillbaka till resp. informationsägare som själv får avgöra om informationen ska läggas in i databasen eller ej. Under 2009 fick man in 700 upplysningar på detta sätt och trenden är ökande. Kvaliteten på kommentarerna är mycket hög.

Bedömning av tekniskt system:

Västarvet har valt att konvertera flera olika databaser till en enda och man har gjort det maskinellt utan manuell omregistrering. Det visar att det alternativet trots sin svårighetsgrad är tekniskt möjligt. Uppdelningen i olika kategorier var tidigare ett bekymmer vid samsökning men är inte det längre när man söker via K-Samsök. Man har inte köpt ett kommersiellt system utan utvecklat själva. Verktygen och metoderna är moderna och välkända. Export till K-samsök är välutvecklad.

Funktionen för periodisk uppdatering av databasen är fortfarande under arbete och måste betraktas som absolut nödvändigt för kontinuerlig drift..

Om vi fick tillgång till Västarvets system, inkl. funktion för periodisk uppdatering när den är färdigutvecklad, så skulle det kunna användas utan större förändringar för en regional server hos oss. Sökfunktionen via K-Samsök är också mycket bra. Det vi kommer att sakna jämfört med vår tänkta målbild är funktioner för att arbeta med informationen i den centrala databasen. Det blir inget centralt system, utan en låda som samlar in information och skickar den vidare till K-samsök. Därmed blir det svårt att lösa behoven för de små informationsägare där vi hade tänkt ta över deras databaser, skapa en säker lagring och driftsmiljö, och sedan låta dem arbeta vidare med sitt material över nätet.

Med en djupare analys finns möjlighet att vi kan hitta sätt att kombinera lösningar från Västarvet, t.ex den utmärkta sökfunktionen eller uppdateringsfunktioner via SSIS, med annan funktionalitet som väljer själva.

10.2. Carpe Cultura - Kalmar

<http://carpecultura.se/carpecultura.htm>

Carpe Cultura är en kommersiell tjänst från företaget Windh Technologies i Påskallavik. Den används för närvarande av Kalmar och Blekinge län och erbjuds öppet på marknaden.



Tjänsten produceras i en central server och distribueras över internet. Databasen stöder alla typer av objekt; föremål, bilder, ljud och film m.m. Men de hittills anslutna kunderna har än så länge bara bilder i sina databaser. Tjänsten har kopplingar till K-samsök.

Funktionen bygger på att olika databaser konverteras in till Carpe Cultura. Vid konverteringen slås de inte samman utan lagras var för sig vilket gör att varje informationsägare kan logga in och arbeta med sitt eget material. Sökning kan göras som traditionell samsökning eftersom allt data ligger lagrat på ett enhetligt sätt.

Efter importen skrotas normalt det gamla systemet och Carpe Cultura tar över som den lokala föreningens enda registersystem. Det finns ingen övergripande "regional funktion" utan varje informationsägare sköter och bestämmer själv om det egna materialet. T.ex kan man välja för varje objekt om det ska publiceras för allmänheten eller K-samsök eller inte alls.

För allmänheten byggs en regional webb där alla de ingående databaserna visas i en trädstruktur. Sökningar kan göras gemensamt i hela materialet.

Systemet håller också på att utveckla sökning på karta, presentation på tidslinje, samsökning med Libris och andra biblioteksdata via z39.50 men de nuvarande kunderna har inte tagit de funktionerna i drift än.

Import till Carpe Cultura görs maskinellt. Windh Technologies uppger man har stor erfarenhet av det och att det brukar gå bra. För små och hemgjorda system kan det ibland vara fördelaktigt att registrera om informationen manuellt.

Exempel på hur Carpe Cultura fungerar i Kalmar Län:

<http://daqapo.windh.net/daqapo/framework/framesetMain.jsp?PI=3.002834738372169>

Bedömning av tekniskt system:

Det är fullt möjligt för Östergötland att nappa på samma erbjudande från Windh som Kalmar och Blekinge har gjort. Det som krävs är att man skapar en affärsmodell för hur tjänsten ska finansieras och erbjudas till lokala föreningar. I drift kommer varje förening sköta sig självt och på regional krävs bara små arbetsinsatser – tyvärr också små möjligheter att påverka. Själva systemet Carpe Cultura har funnits i fem år och utvecklas fortfarande. Det är byggt med moderna metoder och följer standarder, förutom vad gäller databasen (NEO) som är helt egenutvecklad. Det är troligen en bra och kompetent databas, men den verkar vara helt okänd inom övriga IT-industrin.

Ett litet frågetecken också för hur pass färdigt systemet egentligen är, man kan inte undgå att se en del missar och småfel när man surfar runt hos Kalmar och Blekinge. Det är också förvånande att Kalmar som haft systemet i fyra år inte har hunnit lägga in fler databaser och fler objekttyper.

En egenskap med Carpe Cultura som måste övervägas är beroendet av ett relativt litet privat företag. Utvecklingen styrs efter de behov som företaget ser och om företaget skulle försvinna så gör sannolikt själva systemet det också. Särskilt den egenutvecklade databasmotorn NEO gör att det blir svårt för andra intressenter att ta över i ett sådant läge.

10.3. Ödeshög kommun

Egenutvecklat system i Lotus Notes. Databasen kan hantera alla typer av objekt (bild, ljud, film). Systemet används av alla kommuner inom IT-sam. Central teknisk drift men varje



kommun har en egen databas. Systemet har webbpublicering men ingen samsökning med andra system eller export till K-samsök. Det har länkar till externa källor även till GIS. Har länkar till externa textarkiv som öppnar upp filer med dokument.

Utvecklingen har prioriterat logisk struktur och grafisk presentation. Man har även jobbat med att få webben att hamna högt i träfflistor på sökmotorer.

Allt data har matats in i systemet manuellt av föreningar och privatpersoner. Mycket har skapats via studiecirkel. Längre fram siktar man på öppna inmatningsstationer på bibliotek eller via Internet. Överväger att släppa inmatning fri typ Wikipedia. Ansvar för upphovsrätt tas över av systemägaren.

Allmänna kommentarer från Ödehog är att man bör behålla lagring av data lokalt och inte flytta dem till en regional server. Syftet med det är att upprätthålla intresse och kompetens inom lokala föreningar. Materialet bör organiseras geografiskt med socken som huvudbegrepp.

Bedömning av tekniskt system:

Fungerar säkert väl men kommer i princip alltid att kräva manuell inmatning av data. Små möjligheter att integrera med andra system. Utveckling i Lotus Notes kräver specialiserad kompetens som är ovanlig på marknaden och system i Notes tenderar därför att bli isolerade öar. Inte ett alternativ för regional server men intressant att titta vidare på för struktur och presentation.

10.4. Läns museet Östergötland

Läns museet driver en publicering och sökfunktion för i första hand bilder. Den kan hittas under museets hemsida, rubrikerna *Utforska* och *Östgötabild*. Funktionen utnyttjar en webbpubliceringsmodul i systemet Sofie och allt som ska publiceras importerar först till Sofie's databas. Presentationsgränssnittet med sökfunktion ovanpå Sofie är egenutvecklat. I dagsläget använder samtliga bidragsgivare av innehåll till databasen egna Sofie-system och därför har man inga bekymmer med importer och konverteringar.

Bedömning av tekniskt system

Se punkt 10.5 nedan.

10.5. Sofie

<http://www.vbm.se/tjanster/sofie.html>

Kommersiellt system som ägs och utvecklas av Västerbottens Läns museum. Systemet finns i ca 400 installationer över landet.

Tekniskt bygger Sofie på en databas från "4-Dimension". Systemet är i princip ett register med inmatningsdialoger, rapporter och sökfunktioner. En modul för webbpublicering finns som tillval. Systemet kan bara hantera och söka i den egna databasen. En koppling till K-samsök är planerad. Systemet kan hantera konvertering från andra filformat i begränsad omfattning.



Bedömning av tekniskt system

Om man väljer en traditionell metod med konvertering av alla databaser till ett gemensamt system så är det möjligt att använda Sofie som regional server. Det kräver arbete med konverteringar vid import, utveckling av uppdateringsfunktioner för informationsägare och utveckling av ett nytt presentation- och sök-gränssnitt. Koppling till K-samsök kommer att finnas i systemet. Däremot saknas kopplingar till arkiv och biblioteksinformation. Som regional server bedöms Sofie ha begränsad funktionalitet och kräva ett omfattande import och utvecklingsarbete. Fördelen med systemet är att det redan finns hos flera regionala aktörer och därmed finns också en egen kompetens att hantera systemet. Även om Västerbottens museum står bakom systemet så sköts utvecklingen av en enda person, framtida utveckling måste därför betraktas som osäker.

10.6. Carlotta

<http://www.etnografiska.se/smvk/jsp/polopoly.jsp?d=1663&a=8054>

Carlotta är en föremålsdatabas som används av ett tiotal museer. Utvecklingen drivs gemensamt av museerna men genomförs praktiskt av ett konsultbolag. Carlotta bygger på standarden SWESAM och skulle därför teoretiskt kunna användas för samsökning mot andra databaser som följer samma standard. Carlotta använder Oracle databas och ett användargränssnitt skrivet i 4-dimension

Utvecklingen har prioriterat frågan om hur man konverterar olika begrepp från andra databaser och torde ha kommit långt på det området. Däremot har man inte lagt ner mycket arbete på utseende och användarvänlighet. Systemet ser helt enkelt inte särskilt bra ut och kan inte hantera andra objekttyper än just föremål med tillhörande bilder. Det finns inget stöd för ljud eller film.

Bedömning av tekniskt system

Kräver import och konvertering till central databas. Systemet har stora brister i presentationsdelen och saknar funktion för ljud och film. Det är också ensidigt konstruerat för museibehov och har inga funktioner för kopplingar arkiv eller bibliotek. Systemet bedöms som olämpligt för en regional server.

10.7. Primus

<http://primussverige.nordiskamuseet.se>

Ett kommersiellt system som är utvecklat i Norge speciellt för museers behov. Används av ett hundratal museer i Sverige och Norge. Det huvudsakliga systemet körs alltid i servrar hos leverantören men själva databasen kan placeras antingen i egen drift eller i leverantörens anläggning. Största svenska användare och även svensk kontakt för systemet är Nordiska Muséet. I Östergötland använder Flygvapenmuseum Primus.

Primus fungerar med en databas dit allt data importerar/konverteras. Inom databasen hanterar man samlingar som kan ha olika ägare. Upplägget motsvarar mycket bra de behov som uppstår med flera olika informationsägare i den regionala servern som har ansvar för sitt eget material. Webbpublicering görs genom en tjänst som heter Digitalt Museum (<http://www.digitalmuseum.se>). Vill man ha webbpublicering så kan den inte väljas bort. Därutöver kommer man i en version som släpps nästa år kunna bygga sin egen sökfunktion



som hämtar sina data från Digitalt Museum och utgör en regional sökportal. Samsökning sker mellan ingående samlingar i den egna databasen, samt andra museer som finns i Digitalt Museum.

Primus är enbart utvecklat för behov hos museer och saknar specialiserade funktioner för arkiv och bibliotek. Samsökning kan därför inte ske med de kategorierna av material.

Utveckling pågår för att skapa färdiga kopplingar från Primus till K-Samsök.

Primus kan hantera bilder, ljud och film och har även administrativa funktioner för museer som utlån, depositioner och accessionshantering.

Bakom Primus står ett utvecklingsbolag med 16 anställda, ägare är två museer och en statlig myndighet i Norge. Krav på utveckling och en långsiktig närvaro på marknaden är därmed väl tillgodosedda. Det finns en svensk fristående konsult med stor erfarenhet av att konvertera olika databaser till Primus.

Bedömning av tekniskt system

Primus är modernt och har en stabil bas. Det kräver import och konvertering till en central databas men stöder intern uppdelning i samlingar för olika informationsägare. Tillgång till systemet ges enkelt via webb och internet. Driftmodellen med drift av systemet hos leverantör, men möjlighet till egen utveckling av regional webb ger både avlastning från driftsfrågor och stora utvecklingsmöjligheter.

10.8. Wunderkammer

Kommersiellt system som utvecklas av företaget Mogul. Wunderkammer är inte ett färdigt och komplett system utan en verktygslåda som gör det möjligt att själv bygga ihop ett system för lagring, sökning och publicering av alla typer av ABM-information. Systemet bygger på flera olika opensource-komponenter (t.ex Fedora Commons och Apache Lucene) och delar också flera tekniklösningar med K-samsök. Det är samma företag som utvecklat K-samsök och Wunderkammer. Export till K-samsök från Wunderkammer är enkelt, och man kan även använda K-samsök som regional sökmotor och därmed slippa sätta upp sin egen.

Systemet är modernt och anpassningsbart till olika behov. Däremot kräver det antingen egna resurser med hög kompetens eller omfattande konsultinsatser från Mogul när det ska sättas upp för drift.

Wunderkammer bygger på en gemensam databas dit andra databaser importeras. Man uppger att man har övervägt att göra heterogena sökningar i flera databaser men avstått. Istället tänker man sig att märka varje post med information om ursprung och därmed lösa behoven av uppdateringar och äganderätt. Den ursprunglige informationsägaren behåller och kan vidareutveckla sin egen information trots att den lagras i samma databas som andra informationsmängder.

Det finns inga referenser på tjänster som använder Wunderkammer, förutom Riksantikvarieämbetet som använder delar av produkten. Mogul som företag har dock gedigna kunskaper inom området och jag känner ingen tvekan till att de kan få sitt system att fungera på ett bra sätt. Frågetecken dock för kostnaden.

En provversion av systemet som kan göra sökningar både i eget testdata och i K-samsök kan köras här: <http://wunderkammer.curalia.se/wunderkammer/sok>



Bedömning av tekniskt system

Kräver import och konvertering till central databas. Är byggt för att hantera alla typer av ABM-material. Systemet är inte färdigt utan levereras som en verktygslåda och är därmed krävande att installera och starta upp. Samtidigt är det mycket anpassningsbart och resultatet kan få hög kvalitet. Stor fördel med tät koppling till K-samsök, både i produkten och i kompetensen hos företaget som utvecklat båda systemen. Inga referenser.

10.9. Bygdeband

Kommersiellt system från Genline. Har avtal med Hembygdsförbundet. Hembygdsföreningar får systemet kostnadsfritt mot att man digitaliserar sitt material och lagrar informationen hos Genline. Externa sökningar i materialet kräver att man tecknar ett abonnemang med Genline som kostar pengar.

Hembygdsföreningarna äger fortfarande sitt eget material, de har det juridiska ansvaret för publiceringen och har också enligt avtalet rätt att använda materialet för andra syften än bara Genline. Avtalet innehåller särskilt tillstånd för hembygdsföreningarna att delge digitalt material till läns museer och kommuner, men samtidigt förbjuds de att publicera material på internet. En djupare juridisk analys av detta bör göras ifall sådant material ska användas i projektet.

10.10. Tekniska lösningar för Bibliotek

För bibliotekens behov av kataloger finns mycket väl utvecklade produkter. Största leverantör i Sverige är företaget Axiell med produkterna BookIT och Libra. Samsökningstjänster mellan olika bibliotekskataloger är väl definierade och finns redan på länsnivå. Tekniskt finns ett internationellt definierade protokollet Z39.50 som används för att samsöka mellan kataloger från olika huvudmän och olika tekniska system.

Bibliotekens samsökning är optimerad för att söka efter media i olika bestånd. Förutom den nationella söktjänsten Sondera har vi inte hittat några referenser på projekt där man samsöker innehåll i både bokbestånd och andra källor. Sondera gör dock ingen riktig samsökning utan har mer karaktären av tre parallella sökningar. Resultaten presenteras i varsina träfflistor. Varje utökning med en ny datakälla resulterar i ytterligare en lista med söksvar.

Biblioteket arbetar med producera digitalt material genom att fotografera och scanna äldre böcker. Tillsvidare lagras materialet bara som filer i en särskilt inköpt dator. Filerna lagras av bevarandeskäl i ett format som är mycket utrymmeskrävande. Det bedöms oralistiskt att använda befintligt bibliotekssystem för lagringen.

Uppgiften att lagra filerna permanent och publicera dem är inte löst ännu.

10.11. CS-Library

CS-Library är ett system som utvecklas av företaget Teknikhuset. Det är utvecklat med hjälp av opensource men säljs som en förpackad produkt. Funktionen är samsökning och presentation av information på webb. Den innehåller också publiceringsverktyg som används för att bygga en komplett webb och koppla material från databaser till redaktionella texter, nyheter m.m.

CS-library är utvecklat specifikt för bibliotek men det finns inga tekniska begränsningar som låser det till att bara användas där. Det är det enda verktyget av de undersökta som klarar en heterogen sökning mot många olika databaser. En anpassning av sökfunktionen görs för varje



databas, men i gengäld slipper databasen konverteras och kan ligga kvar i sitt originalformat. Systemet bygger upp en egen indexdatabas vilket ger snabba sökningar. Indexet ger också att behovet av kommunikation med originaldatabaserna kan inskränkas till periodisk uppdatering med korta eller långa intervall.

Systemet klarar flera typer av sökning, t.ex namn, geografisk plats och tidsintervall. Andra sökbegrepp kan utvecklas vartefter. Kartsökning finns inte men är möjlig att skapa. Sökresultaten kan sorteras efter flera olika begrepp, bl.a relevans till sökfrågan.

CS-Library har genom sin heterogena sökning möjlighet att fungera som en brygga mellan system för olika ändamål. Om man med andra system väljer att konvertera ihop alla föremålsdatabaser, fotoarkiv m.m. till en enda databas så får man normalt ändå inte in arkiv och biblioteksinformation. CS Library kan överbrygga detta genom att göra en samtidig sökning i t.ex ett museisystem, ett arkivsystem och ett bibliotekssystem, och presentera allt data sammanvägt och sorterat.

Bedömning av tekniskt system

Ett system med en annan utgångspunkt än de övriga. Det innehåller inget stöd för museets eller föreningens interna behov av systemstöd och kan därför inte ensamt lösa alla behov. Däremot utgör det en mycket kraftfull sökmotor och webbpublicering. I kombination med andra system ger det mycket intressanta funktioner.

CS-Library ger en frihetsgrad när det gäller vad som ska konverteras till en central databas och vad som ska ligga kvar i sitt originalformat. Man konverterar de databaser som är enkla att konvertera, medan man arrangerar samökning med dem som är svårare. T.ex kan man konvertera in små Accessdatabaser från hembygdsföreningar, men ordnar samsökning med museets Sofie-system och bibliotekets katalog.

10.12. Arena

Arena är en produkt från Axiell och marknadsförs som en sökmotor och webbpubliceringsverktyg med koppling till samma företags system för bibliotekskatalog, BookIT och Libra. Produkten har alltså ett liknande tillämpningsområde som CS-Library och är på ett liknande sätt nischad mot sökningar i bibliotekssystem. Det har också avancerade funktioner för att skapa en attraktiv webbplats och stöder alla typer av format som bilder, ljud och film.

Axiell säljer också sedan en kort tid tillbaka ett museisystem som heter Calm.

Utvecklingsplaner finns för att Arena ska kunna göra samsökningar i Axiells bibliotekssystem och Calm.

Bedömning av tekniskt system

Mitt intryck är att Arena är ett bra system men hårt knutet till leverantörens egna bibliotekssystem. Det är därför inte lika enkelt att anpassa till sökningar i andra system, och kanske särskilt inte till system som hanterar andra typer av objekt än böcker. Koppling till andra system måste utvecklas av leverantören och kan inte göras med konfiguration av standardfunktioner.



11. Sammanfattad bedömning av olika verktyg

- Västarvet Egenutvecklat system på en plattform av Microsoft SSIS. Kombination av central databas och periodisk insamling från externa databaser. Utmärkt sökfunktion via K-samsök
- Carpe Cultura Kommersiellt system med drift i leverantörens servrar. Uppges ha hög funktionalitet, men befintliga kunder har inte kommit så långt. Affärsmodell där leverantören tar över hela funktionen och vänder sig direkt till de enskilda databasägarna. Den regionala organisationen inte kan styra så mycket. Bygger tekniskt på en egen stängd databas som riskerar att ge inlåsnings effekter.
- Ödeshög Egenutvecklat i Lotus Notes vilket är ovanligt på marknaden. Små möjligheter att använda som regional server. Förespråkar manuell inmatning av allt data istället för maskinell konvertering. Kommit långt med struktur för innehåll.
- Sofie Kommersiell produkt av traditionell modell. Har webbmodul och planerar export till K-samsök. Anpassad för drift på enstaka museer och inte som regional server.
- Carlotta Samverkansprojekt mellan flera museer. Dyr databas (Oracle). Bygger på standardiserad objektbeskrivning SESAM. Dåligt grafiskt gränssnitt och bedöms inte lämpligt för regional server
- Primus Ger ett modernt och stabilt intryck. Bra uppdelning i central systemdrift och lokal webbutveckling. Troligen ett av de mest användbara systemen, men saknar samsökning med arkiv och bibliotek.
- Wunderkammer Kommersiell produkt som består av en verktygslåda för egen utveckling av funktioner. Mycket modernt och kompetent, men arbetskrävande. Inga referenser
- Bygdeband Kommersiellt system, låst för ett annat användningsområde. Ingen möjlighet att använda som regional server
- BookIT har viss funktionalitet men är inte avsett för tillämpningar som regionkatalogen. Dessutom pågår upphandling av nytt system och vi kan inte veta om BookIT kommer att vara kvar i biblioteken.
- CS-Library Ett renodlat sök och publiceringssystem. Modernt och flexibelt med heterogen sökning i flera databaser samtidigt.
- Arena Ett bra sök- och webbpubliceringssystem. Knutet till företagets egna bibliotekssystem och därmed svårt att anpassa för andra sökningar.



12. Rekommendation

Den här undersökningen är baserad på publicerad information och telefonintervjuer. Jag har inte sett något av systemen demonstreras. Av den information som framkommit har jag ändå tre tydliga spår som framstår som mest tilltalande. Detta kan givetvis ändras vid en mer ingående granskning.

De kommersiella system som har störst potential att användas som regional server är Primus och Wunderkammer. Västarvets egna lösning ligger också väldigt nära vår målbild. Alla tre kräver import och konvertering av ingående databaser till en central databas. I övrigt är det tre system med ganska olika egenskaper. Som ett fjärde alternativ står egen utveckling av funktionen med användning av opensource-verktyg.

För sökning och presentation så har Västarvet ett intressant spår där man utnyttjar K-samsöks funktioner. Detta ligger också exakt i linje med vad K-samsök själva eftersträvar. CS-Library är ett sök och webbpubliceringssystem med funktioner som ingen av de andra har. Den kan som enda system göra samsökningar i databaser som man inte anpassat eller kopierat in till den regionala databasen. Därmed kan den föra samman information från alla typer av databaser, Arkiv, Bibliotek och Museum, samt presentera detta på ett enhetligt sätt. Alternativet att kombinera CS-Library med ett grundsystem, som kan vara något av de andra tre, och dessutom göra sökningar i befintliga biblioteksystem är mycket tilltalande.

Det som särskiljer grundsystemen är inte bara deras funktioner utan också det sätt som systemen erbjuds till marknaden. Vilket engagemang har vi och vilka resurser vill vi lägga ner? Våra önskemål om hur vi vill arbeta med vår regionala server är kanske mer avgörande än de faktiska funktionerna i varje system.

Västarvet

Västarvet har byggt ett helt eget system som mycket väl motsvarar de funktioner vi har föresatt oss. Det enda som saknas är ett system för att arbeta med informationen centralt. Det får effekten att vi inte kan erbjuda lagring till små informationsägare på det sätt som vi tänkt oss. Västarvet är också villiga att dela med sig av sitt arbete och erbjuder just oss i Östergötland att delta i samarbetsprojekt för vidareutveckling.

Primus

Erbjuder ett traditionellt system och förvaltning med ett tydligt leverantör-kundförhållande. Systemdrift läggs hos leverantören, men den regionala webben byggs lokalt. Den regionala organisationen styr användning och webb, men inte systemutveckling. Ett projekt med Primus är betydligt mindre omfattande än Wunderkammer men inte lika flexibelt. Primus ger sammantaget färre osäkra faktorer och kanske mer avancerade funktioner än vad vi skulle klara att utveckla själva med andra verktyg.

Primus är huvudsakligen avsett för museer. Det är modernt och har en fördel i ett stabilt ägar- och utvecklingsförhållande och många gjorda installationer. Det finns bra sökfunktioner inom den egna databasen. Export till K-samsök och möjlighet till egen webbpublicering är under utveckling. Båda funktionerna måste finnas för att systemet ska kunna användas på det sätt vi vill.



Wunderkammer

Kräver stort eget engagemang, egen IT-kompetens, men ger stor frihet att skapa exakt den lösning man vill ha. Större uppstartskostnader än med färdiga system.

Produkten är öppnast och kan bäst anpassas till olika kravbilder. Den ger oss möjlighet att själva utforma vilken funktion vi vill. Vi kan åstadkomma sökningar inom den egna databasen och bra presentation. Sökningar över fackgränserna (museum, arkiv och bibliotek) är sannolikt möjlig. Export till K-Samsök är inbyggd.

Egenutveckling med opensource

Ett system kan byggas med hjälp av fria komponenter som utvecklas av breda intressegrupper över hela världen. Man kan få exakt de funktioner man vill ha, en svårighet blir snarare att definiera vilka funktioner det är man vill ha. Samsökning med arkiv och bibliotek samt webbpublicering kan byggas in (liknande CS-Library). Kräver stor egen kompetens eller konsultinsats. Utvecklingen har goda förutsättningar att fungera som samarbetsprojekt även med internationella kopplingar. Dyr utveckling men billigast i drift.

CS-Library

CS-Library är inte ett alternativ till de andra systemen utan ett möjligt komplement till samtliga. Det utgör en webbpubliceringsfunktion med en avancerad sökmotor som tillför möjligheten att göra sökningar utanför den egna databasen med stor flexibilitet. Dessutom ingår verktyg för att bygga en snygg och funktionsrik webb för allmänheten. Kombinationen av ett grundsystem och CS-Library ger den mest kompletta funktionen och är egentligen den enda lösning som uppfyller alla våra önskningar.

På sista raden är utredarens personliga rekommendation att bäst sammanvägt resultat fås av en kombination av Primus och CS-Library. Om man kan tänka sig att släppa på funktionalitet med lagring och drift av databaser gentemot små informationsägare så har Västarvet alla övriga funktioner redan klara. Vill man ha en större egen kontroll och en tät koppling till K-samsök samt är beredd att satsa mer på egen utveckling så är det närmaste alternativet att välja Wunderkammer.



13. Projektbeskrivning

Vi ska här försöka beskriva vilka olika aktiviteter som kan komma att behövas i ett verkligt projekt för att bygga en regional server för ABM enligt föreslagen modell. Det är inte meningen att skapa en komplett och exakt projektplan utan redovisningen ska ses som en översikt i syfte att ge en ungefärlig bild av projektets omfattning och innehåll.

Enligt diskussioner i styrgruppen utgår vi från två specifika scenarion

- Primus i kombination med CS-Library
- Wunderkammer

Ett projekt för att bygga en regional server kommer att bestå av fyra distinkta delar:

- Avtal, äganderätter, paketering av kostnader för olika aktörer
- Bygga upp det centrala systemet
- Konvertera och/eller ansluta databaser
- Bygga en sökfunktion

13.1. Avtal

En regional server kommer att ha många intressenter. Det är av avgörande betydelse att alla juridiska och ekonomiska förhållanden är helt klara från början. För att få så många informationsägare att ansluta sig som möjligt vill man också kunna erbjuda en bra och förutsägbar funktionalitet samt så låg kostnad som möjligt.

Det som absolut krävs är en huvudman som tar ansvar för hela funktionen. Denna måste upphandla och äga den regionala servern och erbjuda anslutning till olika informationsägare. Driften kan göras i egen regi eller upphandlas. För att ansluta databaser föreslås att man tar fram färdiga paketeringar för olika behov. Därmed slipper man förhandla om exakta villkor med varje avtalspart. Paketeringarna ska vara i form av avtal där äganderätt till material, rätt att publicera uppgifter samt kostnader regleras. Minst tre olika scenarion behöver hanteras:

- Regionala servern tar över materialet och allt ansvar i sin helhet. Ursprungliga informationsägaren kopplas bort
- Regionala servern tar över lagring och publicering men inte äganderätt och ansvar för vidareutveckling. Informationsägaren har distansåtkomst
- Regionala servern får rätt att söka i och publicera material som ägs och förvaras hos en annan huvudman (kräver CS-Library).

Man ska inte underskatta arbetet med att skapa dessa paketeringar och avtal. När systemet är i drift blir det också fråga om att uppsöka tänkbara informationsägare för att "sälja in" konceptet och skriva avtal. En viktig fråga att utreda är det juridiska ansvaret för publicering (PUL och upphovsrätt) i de alternativ där regionala servern inte tar över äganderätten.

De föreningar och organisationer som bidrar med information kommer i någon mening att kunna betraktas som kunder eller klienter. I drift kommer också löpande kontakter och support till dessa kunder också att ingå i arbetet för den regionala serverns personal.



Den uppskattade kalkyl för projektet som finns nedan tar inte upp den avtalsmässiga och ekonomiska delen av arbetet.

13.2. Centralt system

Det centrala systemet ska bestå av en eller flera fysiska eller virtuella servrar med de olika programvarorna installerade. Båda alternativen innehåller en central databas. Om CS-Library ska användas så kommer det också att finnas en separat databas med sökindex till både egna och externa databaser. I installationen ingår också paketering och distribution av eventuell klientprogramvara för administratörer.

Om databasägare ska kunna ansluta sig på distans och arbeta vidare med underhåll av sitt eget data så behöver man skapa en funktion för säker anslutning via Internet. Det är rimligt att anta att sådan funktion redan finns färdigutvecklad hos en driftleverantör, det kan vara ett krav vid val av driftleverantör. Funktionen för anslutning behöver i så fall endast konfigureras och anpassas för det aktuella behovet.

För Primus/CS-Library är installationen ett ganska enkelt arbete. Driften av Primus ligger hos leverantören i Norge och installationen handlar om att sätta upp lokal databas, lokal installation av CS-Library samt datakommunikation till leverantören.

En installation av Wunderkammer är en helt annan sak eftersom systemet inte är färdigt utan levereras som "byggklossar". Det kommer alltså att gå åt väsentlig tid som ren systemutveckling för att bygga ihop det system man vill ha.

Projektaktiviteter:

- Serverinstallation
 - Välja driftsplats och installera servrar
 - Distribuera eventuell klientinstallation
 - Installera grundsystem för egen drift (Primus och CS-Library)
 - Utveckla ett grundsystem för egen drift (Wunderkammer)
- Datakommunikation
 - Etablera datakommunikation till leverantör (Primus)
 - Anpassa distansåtkomst för interface till berörda informationsägare
- Drift
 - Sätta upp systemförvaltning och drift

13.3. Konvertera och ansluta databaser

Om man använder Primus/CS-Library är första steget när en ny databas ska anslutas är att avgöra om den ska konverteras och flyttas till Kulturarvsservern, eller om man ska ansluta till den på distans. För Wunderkammer finns bara det första alternativet. Dessa två metoder bör var för sig vara väl beskrivna och kunna utföras med en viss mån av automatik eller fastlagd procedur. I huvudprojektet föreslås att man tar med endast en databas av vardera typen i syfte



att utveckla själva metodiken. Därefter kan fler databaser anslutas antingen som fristående tilläggsprojekt, eller som standardiserade driftsåtgärder.

Projektaktiviteter:

- Beslut om konvertering eller anslutning till databas på distans
 - Beskrivning av befintligt system
 - Kontakt med leverantör av befintligt system
 - Undersöka möjlighet till datakommunikation med databasägaren (för sökning med CS-library)

Konvertera databas

Konvertering av databaser är en egen specialkompetens som man inte kan förutsätta finns inom en driftorganisation. I enkla fall eller om det är frågan om en typ av databas som har konverterats tidigare och finns dokumenterad så kan kanske driftoperatören klara det. Det vanligaste är dock att använda externa konsulter. Om man valt Wunderkammer som system så har leverantören egna konsulter som kan hjälpa till. Om man använder Primus så finns minst en tredjepartskonsult i Sverige som är specialiserad på att konvertera databaser till Primus.

Projektaktiviteter:

- Kartlägga aktuell databas
 - Tekniskt format
 - Datamässig struktur
 - Urval av poster
 - Detaljerad beskrivning av innehåll – datafält, nyckelbegrepp
 - Mappning av fält mot Kulturarvserverns databas
- Konvertering
 - Tvättning av data
 - Teknisk konvertering
 - Eventuell manuell hantering
 - Kvalitetskontroll
 - Import i Kulturarvsserver
- Om databasägaren ska ha distansåtkomst:
 - Etablera distansåtkomst för databasägaren
 - Säkerställa att gammalt system tas ur bruk och etablera rutiner för fortsatt underhåll och utveckling av innehållet i samarbete med databasägaren
- Om databasen ska tas över helt:



- Ingen mer aktivitet behövs. Vi behöver inte veta om det gamla systemet stängs av eller fortsätter att användas

Ansluta databaser på distans (endast med CS-Library)

Databaser som inte ska konverteras och lagras centralt utan endast finnas med i sökningar måste anslutas via fast datakommunikation. Här ställs lite högre krav på tillgänglighet och säkerhet i själva uppkopplingen eftersom det är maskin-till-maskin-anslutningar som behöver finnas tillgängliga dygnet runt. Man kan använda hyrda förbindelser, stadsnät inom vissa områden eller krypterade fasta förbindelser via Internet (kallas VPN-tunnlar). För att ansluta en databas som finns hos en kompetent IT-driftsorganisation så är bedömningen att det är ganska enkelt och kan göras med utrustning som redan finns i de flesta driftsmiljöer. Hos mindre aktörer kan det krävas att man hjälper till från centralt håll och kanske även investerar i viss utrustning som placeras ute vid driftsstället för databasen.

I tveksamma fall behöver man göra en avvägning om en databas ska kopplas upp på distans eller konverteras in och databasägaren erbjudas distansuppkoppling istället.

Att ansluta databaser på distans för att skapa gemensam regional sökning går till på ganska olika sätt beroende på om man använder Wunderkammer eller Primus/CS-Library.

Projektaktiviteter:

- Kartlägga aktuell databas
 - Tekniskt format
 - Datamässig struktur
 - Urval av poster
 - Detaljerad beskrivning av innehåll – datafält, nyckelbegrepp
 - Mappning av fält mot Kulturarvserverns databas
- Etablera datakommunikation
 - Välja och sätta upp teknisk datakommunikation
 - Identifiering och säkerhet för båda parter
 - Funktionstest och driftstart
- Skapa index
 - Avgöra vilka fält som ska vara sökbara
 - Sätta upp indexering
 - Kvalitetskontroll

13.4. Bygga en sökfunktion

Sökfunktionen kan byggas på många sätt och det är bara ambitionsnivån och medlen som sätter gränserna. Två huvudtyper kan identifieras.



Enklaste sättet är att bara bygga en sökmotor utan egen webbplats. Den kan sen användas av andra webbplatser både egna och andras för att tillgängliggöra materialet. Exempel på denna metod är K-samsök som inte har någon egen webb. K-samsök använder en avancerad teknik som kallas ”webservices”. Den kräver relativt mycket av den som ska sätta upp sök-webben. Vi kan välja en lite enklare funktion genom att göra en webbsida med bara en sökfunktion som andra sen kan länka till.

Ett annat alternativ är att bygga upp en egen webbplats runt sökfunktionen och inte låta andra använda sökningen på sina egna webbplatser. En sådan webb kan kompletteras med redaktionellt material och utvecklas hur mycket som helst. Diskussioner pågår om att webben Kulturarv Östergötland skulle behöva göras om med ett modernare publiceringsverktyg. CS-Library skulle mycket väl kunna vara ett sådant verktyg och man kan tänka sig att kombinera projekten så att den regionala serverns sökfunktion byggs direkt in i en ny Kulturarv Östergötland.

Att bygga en ny Kulturarv Östergötland är i sig ett stort projekt och därför begränsar vi oss i den här rapporten till att beskriva hur själva sökfunktionen kan tas fram.

Projektaktiviteter:

- Skapa sökfunktion
 - Sätta upp sökning mot intern databas
 - Skapa sökfunktion mot indexerade externa databaser (CS-library)
- Bygga webbsida med sökning och presentation av resultat
 - Byggs i CS-library eller annat publiceringsverktyg
 - Övrigt arbete med webb -> eget projekt

13.5. K-samsök

För att Kulturarvsserverns innehåll ska kunna visas via K-samsök krävs att man iordningsställer materialet efter en viss mall som K-samsök kan läsa av. När det gäller data som lagras i själva Kulturarvsservern så är det en relativt enkel åtgärd som man bara behöver göra en gång.

De databaser som inte lagras lokalt utan bara avsökts på distans för att vara sökbara är svårare att presentera för K-samsök. Funktionen kan säkert byggas men det är svårt att uppskatta hur stort arbete det kan vara. Det är möjligt att den behöver anpassas för varje ny databas som läggs till. Det finns inga referenser på området oavsett vilka system och produkter som väljs.

Primus har aviserat en modul för att kunna visa sitt innehåll gentemot K-samsök. I Wunderkammer finns denna med redan idag.

Projektaktiviteter:

- Från centralt system
 - Uppsättning av funktion i resp. system
 - Konfiguration av vilket data som ska presenteras
 - Genomgång av datakategorier för ev. anpassning mot K-samsök



- Mappning av fält mot K-samsök
- Från externa system
Utgår från att CS-library finns och är uppsatt för att skapa lokalt index
 - Skapa tekniskt interface som K-samsök kan ansluta emot
 - Urval av vilket data som ska presenteras
 - Mappning av fält mot K-samsök
 - Skapa sökinterface gentemot K-samsök



14. Kalkyl

Kalkylen inkluderar arbete av projektledare och systemutvecklare som kan arbeta i huvudsak heltid med projektet. Den inkluderar inte tid hos verksamhetspersonal hos olika informationssägare eller regional förvaltning.

Kalkylen utgår också från att alla förutsättningar och grundinformation för projektet finns klara vid projektstart. Arbetet med att upprätta leverantörskontakter, skaffa dokumentation för ingående produkter m.m. tillkommer.

Aktivitet	Arbetstid	Kalendertid
Avtal, tjänstepaketering	80 timmar	4 veckor
Sätta upp centralt system Primus och CS-Library	80 timmar	3 veckor
Sätta upp centralt system Wunderkammer	200 timmar	10 veckor
Konvertera en databas	30 timmar	1 vecka
Ansluta en databas på distans med CS-library	30 timmar	1 vecka
Bygga en sökfunktion exkl övriga webbfunktioner	40 timmar	2 veckor
K-samsök internt data	20 timmar	1 vecka
K-samsök externt data (enbart CS-Library)	20 timmar	1 vecka
Etablera drift och systemförvaltning	80 timmar	4 veckor
Summa Primus/CS-Library	380 timmar	17 veckor
Summa Wunderkammer	450 timmar	22 veckor



15. Systemförvaltning

För att driva en regional kulturarvsserver över tid behövs en kompetent systemförvaltning. I systemförvaltningens uppgifter ligger t.ex att sörja för driften av systemet, avtal och kontakter med leverantörer och samarbetspartners, budget och ekonomisk planering, planering av utveckling. Man kan också lägga in utbildning, riskanalyser, förvaltningsplaner, rapportering, beslutsunderlag och omvärldsbevakning.

För att göra det tydligt så är det utredarens fasta övertygelse att det måste finnas någon form av organisation för systemförvaltning med en årlig budget för att kulturarvsservern ska kunna fungera över tid.

Det finns många modeller för systemförvaltning, en bärande tanke i de flesta är en tydlig indelning av roller. Den modell som används inom Linköpings kommun har med vissa anpassningar den här indelningen:

Part Nivå	Nyttjande verksamhet	Systemägare	Teknisk drift
Strategisk nivå	<i>Politisk nämnd</i>		
Resursnivå (också styrgrupp)	<i>Chefer i berörd verksamhet</i>	<i>Systemägareombud med budget och delegation att ta beslut</i>	<i>Chef hos driftleverantör</i>
Uppdragsnivå	<i>Verksamhetsexperter</i>	<i>Systemförvaltare</i>	<i>Tekniskt systemansvarig</i>
Operativ nivå	<i>Kontaktpersoner</i>	<i>Systemexpert Leverantör Utbildare</i>	<i>Kundservice Teknisk personal</i>

Förvaltningen är alltså uppdelad i en matris där funktionerna kommunicerar både horisontellt på samma nivå och inom kolumnerna.

Detta är en avancerad modell och det finns självklart alla möjligheter att sätta upp alternativa och enklare organisationer. Det viktiga är att rollerna ska ha en definierad funktion och mandat.

Den allra enklaste indelning som kan rekommenderas är att dela förvaltningen i tre huvudområden. De motsvaras i princip av de tre kolumnerna i modellen ovan.

- Intressenter
- Systemägare
- Driftleverantör

Systemägaren är huvudman för hela funktionen och måste ha finansiella medel och beslutsmandat att driva och utveckla funktionen över tid. Systemägaren kan vara en institution eller grupp som har ansvar på en övergripande nivå. Det krävs i så fall också en utsedd systemförvaltare som har delegation från systemägaren att både ta operativa beslut och förvalta den löpande budgeten för systemförvaltningen.



Intressenter i systemförvaltningen är t.ex informationsägare som bidrar till innehållet och förvaltare av webbtjänster som publicerar information. Hit kan också räknas systemleverantörer, samarbetspartners och många andra personer och organisationer som på olika sätt har intresse av utvecklingen. Kanske Den allra största intressenten är självklart allmänheten.

Det är lämpligt att ha en tydlig åtskillnad på systemförvaltning och teknisk drift. Därför är teknisk drift lämpligen en upphandlad funktion. Funktionen upphandlas och styrs av systemägaren/systemförvaltaren. Avtal och teknik utformas så att systemet kan flyttas till annan driftsleverantör om behov skulle uppstå. Genom en upphandlad drift får man tillgång till kompetens på expertnivå utan att behöva bygga upp det själv inom förvaltningsorganisationen. Det finns också betydande stordriftsfördelar och därmed kostnadsbesparingar med att använda en befintlig serverhall.

I driften ska ingå all nödvändig teknisk utrustning, servrar, datorer, datakommunikationsutrustning m.m. Systemägaren äger alltså inte själva maskinparken utan bara programvara och ”funktion”. Driftleverantören fördelar själv systemet på lämplig hårdvara för att uppnå avtalad funktion och kan därmed arbeta aktivt för att sänka driftkostnaderna. Lagring med backup och säker förvaring av backupmedia ska ingå i drifttjänsten. Vidare kan man ta med en avtalad nivå på kapacitet mot internet (för publicering) och funktioner för att intressenter ska kunna anslutas säkert ska ingå.

I driftsleverantörens åtagande ligger normalt även att ta emot felanmälningar, åtgärda uppkomna fel, åtgärder vid planerat underhåll. Den totala servicenivån, systemets tillgänglighet gentemot intressenterna, specificeras i en särskild överenskommelse som kallas SLA – Service Level Agreement.



16. Referenser

Verksamhet	Namn	Kontaktinformation
Västarvet	Per Lekholm	Per.lekholm@vgregion.se 031-775 24 10
Ödehögs kommun	Magnus Lundberg	0390-21070
Länsmuseet Östergötland	Leif Valentinsson	013-230349
Carpe Cultura Windh Technologies	Susanna Windh	0491-97 000
Wunderkammer Mogul AB	David Haskia	08-506 661 00
Stadsbiblioteket	Mathias von Wachenfeldt	013-203105
Primus, Nordiska Museet	Kristina Berg	08-51 95 47 18 Kristina.Berg@nordiskamuseet.se
CS-Library, Teknikhuset	Lars Eriksson	090-70 87 16
Arena, Axiell	Peter Stiljestränd	046-270 04 72
K-samsök, RAÄ	Lars Lundqvist	lars.lundqvist@raa.se tel: 08-5191 8565
Sondera, KB	Birgitta Herdenberg	
Sveriges Radio	Claes Hollander Projektledare för digitalisering av sändningsarkiv	Via växel: 08-784 50 00

17. Bilaga 1 Kort beskrivning av begreppet OpenSource

Öppen källkod, engelska *Open Source*, är tekniskt sett datorprogram där källkoden är tillgänglig att använda, läsa, modifiera och vidare distribuera för den som vill. Det gör att användaren kan analysera och anpassa funktionen till sina behov. De modifieringar som användaren gör ska erbjudas tillbaka till den ursprungliga upphovsmannen och fritt till vem som helst. Dessa kan välja att använda ändringen eller göra ytterligare modifieringar

Öppen källkod är oftast men inte nödvändigtvis gratis. Det vanligaste är att man kan använda själva programvaran kostnadsfritt men vid behov får betala för utbildning, tryckta manualer och konsulttjänster i anslutning till produkten. Andra modeller förekommer också

Typiskt för öppen källkod är att utvecklingen sköts av ett antal personer, organisationer och företag som koordinerar sitt arbete via Internet. Detta till skillnad från klassisk proprietär källkod där det vanligen är ett enda företag som utvecklar och tillhandahåller produkten. Tidigare drevs utvecklingen främst av enskilda entusiaster, men har på senare år mer och mer övertagits av professionella företag som hittat andra affärsmodeller än att ta betalt för programlicenser.

Öppen källkod distribueras ofta som komponenter som inte är kompletta produkter utan tillhandahåller delfunktioner som kan byggas in i andra system. Därigenom har det uppstått en växande nisch för företag som använder opensource-komponenter och kombinerar dem med egen programvara till kompletta kommersiella system. Resultatet är proprietär programvara med en prislapp på licensen. Men utvecklingskostnaden och därmed licenspriset kan hållas på en mycket låg nivå jämfört med system som utvecklas med traditionella metoder.