



Digital informationshantering och infrastruktur för kulturarvet

Underlag för fortsatt arbete

Följande rapport presenterar Digisams arbete med digital informationshantering och infrastruktur för kulturarvet och utgör ett underlag för fortsatt arbete.

Rapporten är beslutad av Digisams styrgrupp den 17 oktober 2014.
Diarienummer: RA 06-2014/9438

Version 1.0. Licens CC-0. En senare version kan finnas på www.digisam.se.

Innehållsförteckning

INLEDNING	4
SAMMANFATTNING	5
BAKGRUND	7
DIGITAL KULTURARVSINFORMATION	7
INFORMATIONSHANTERING.....	8
PROBLEMFÖRMULERING	10
EN TENTATIV MÅLBILD	12
MÅLBILDENS KOMPONENTER	15
AGGREGERINGSTJÄNSTER.....	15
SEMANTISKA WEBBEN.....	18
STANDARDER	20
SYSTEM	25
DIGITALA INFRASTRUKTURER.....	28
GEMENSAMMA STÖDFUNKTIONER.....	30
SAMVERKAN MED FORSKNINGSSEKTORN	38
ETT EUROPEISKT PERSPEKTIV	40
DISKUSSION OCH ANALYS	43
DIGITAL KULTURARVSINFORMATION OCH SAMORDNAD INFORMATIONSHANTERING	43
DIGITALA INFRASTRUKTURER.....	46
ETT EUROPEISKT PERSPEKTIV	47
FÖRSLAG TILL VIDARE ARBETE	48
ORD- RESPEKTIVE AKRONYMLISTA	49

Inledning

Det övergripande målet i Digit@lt kulturarv¹, regeringens nationella strategi för arbetet med att digitalisera, digitalt bevara och digitalt tillgängliggöra kulturarvsmaterial och kulturarvsinformation, är att ”kulturella verksamheter, samlingar och arkiv i ökad utsträckning ska bevaras digitalt och tillgängliggöras elektroniskt för allmänheten”.

För att kunna öka tillgången till, användningen och bevarandet av digital kulturarvsinformation krävs en stabil och långsiktigt hållbar infrastruktur.

En digital infrastruktur innebär allt från internetuppkoppling till informationshantering med stöd av tjänster, verktyg och system, med mera. Tjänsterna kan till exempel vara anpassade för att strukturera metadata enligt vedertagna standarder samt använda sig av gemensamma arkivlösningar för lagring och åtkomst till distribuerade bevarandetjänster via den tekniska e-infrastrukturen.

Samordningssektariatet för digitalisering av kulturarvet, Digisam, ska enligt regeringsuppdrag (Ku2011/242/KA) till Riksarkivet och strategin Digit@lt kulturarv ta fram förslag till nationella riktlinjer för hur en samordnad digital informationshantering, samt ett samordnat och kostnadseffektivt digitalt långtidsbevarande av samlingar och arkiv, kan ske hos de statliga institutioner som samlar, bevarar och tillgängliggör kulturarvsmaterial och kulturarvsinformation. Förslaget ska även avse en roll- och ansvarsfördelning för arbetet med aggregering, digitalt tillgängliggörande och digitalt bevarande inom den statliga kulturarvssektorn. Vidare ska Digisam arbeta för att målen med regeringens nationella strategi Digit@lt kulturarv ska uppnås.

I december 2012² beslutade Digisams styrgrupp att samordningssektariatet under 2013, inom ramen för arbetet med infrastrukturfrågorna, skulle ta fram förslag till en modell över hur aggregatorerna kan samverka, vilken funktionalitet som behövs hos respektive aggregator och vilken gemensam domänspecifik nytta som kan hållas av en specifik aggregator. Modellen skulle också innehålla förslag på möjliga leveransvägar och aggregeringsformer samt vilken funktionalitet som är mest kostnadseffektiv att förvaltas centralt, utanför aggregatorerna.

Som ett led i detta har Digisam mellan februari 2013 och oktober 2014 arbetat med att sammanställa underlag och utreda förutsättningarna för att ta fram ett sådant förslag.

¹ http://digisam.se/images/docs/digitalt_kulturarv.pdf

² <http://digisam.se/index.php/docman/public-documents/omoss/styrgruppsprotokoll/21-protokoll-styrgrupp-digisam-2012-12-10>

Sammanfattning

Syftet med denna rapport är att beskriva hur situationen ser ut idag avseende behov, möjligheter och befintliga lösningar för hantering av digital kulturarvsinformation, inklusive tillgängliggörandet av kulturarvsmaterial inom det statliga kulturarvsområdet. Tillsammans med resultaten från Digisams förstudie om digitalt bevarande³ är avsikten också att denna rapport ska vara ett första steg på vägen till ett förslag om roll- och ansvarsfördelning för arbetet med aggregering, digitalt tillgängliggörande och digitalt bevarande av kulturarvet.

Rapporten tar avstamp i den tentativa målbild som har utarbetats i samband med ett antal workshoppar och modelleringar med representanter för de myndigheter och institutioner som är ansvariga för aggregeringstjänster (se s.15 för vidare information om aggregeringstjänster), eller med en annan koordinerande roll vad gäller digitalt kulturarv.

I rapportens första del presenteras bakgrunden och resultaten av de workshoppar och modelleringar som utförts med representanter från institutioner som förvaltar aggregeringstjänster. Modelleringarna genomfördes av Rolf Källman och Sanja Halling från Digisam med processtödd av Ambjörn Näve, Erik Isaksson och Fredrik Enoksson från Kungliga Tekniska Högskolan (KTH). Syftet med modelleringarna var dels att ta fram en övergripande modell över hur digital informationshantering, aggregering och digitalt tillgängliggörande inom kulturarvssektorn ser ut idag, dels att beskriva de möjligheter som finns och hur ett önskat läge ser ut med utgångspunkt i den tentativa målbild som de inledande gemensamma och övergripande workshopparna resulterat i (se exempel på modelleringskartor⁴).

Information om kulturarvssamlingar och bestånd hanteras idag (skapas, förvaltas, lagras) i tekniska system som är under ständig förändring, samtidigt som systemen inte alltid är kopplade till tillhörande datafiler (exempelvis högupplösta bilder på objekt, kontextuella metadata, med mera) och inte heller några långsiktiga bevarandelösningar. På så sätt splittras informationen både genom olika sätt att beskriva den på och genom olika system som hanterar respektive aggregerar den. Denna splittring gör att det idag läggs resurser på att uppdatera och berika information i system som inte alltid överför dessa ändringar till en långsiktig förvaltning och ett långsiktigt bevarande.

Vidare diskuteras i rapporten behov, möjligheter och utmaningar för en effektivare informationshantering, från det att informationen struktureras i institutionernas verksamhetssystem till att den aggregeras, tillgängliggörs, används och bevaras. För att kulturarvsinformation ska bli möjlig att använda och återanvända är standardisering, interoperabilitet och harmonisering centrala frågor, och för att standardisering och harmonisering av information mellan olika domäner ska kunna uppnås finns det ett övergripande behov av ett gemensamt tekniskt stöd vid registrering, aggregering och överföring av information samt en stödjande funktion för hur man ska gå tillväga. Den föreslagna modellen, som baseras på den tentativa målbilden, över en samordnad informationshantering presenteras i rapporten utifrån den funktionalitet som behövs gemensamt respektive hos varje enskild institution eller aggregeringstjänst. Den tentativa målbildens olika

³ <http://digisam.se/index.php/docman/public-documents/digitaltbevarande/40--8>

⁴ <http://digisam.se/index.php/docman/public-documents/infra/59-modelleringar-med-kth-1>

komponenter (aggregeringstjänster, semantiska webben, standarder, system, digitala infrastrukturer och gemensamma stödfunktioner) redovisas också i rapporten, del för del.

Rapportens andra del visar vilka beröringspunkter som finns mellan kulturarvs- och forskningssektorn, i synnerhet vad gäller e-infrastrukturer, samt hur det senaste arbetet kring de frågor som diskuteras i rapporten ser ut på europeisk nivå. Sammanfattningsvis behövs det idag en gemensam infrastruktur för kulturarvsinstitutionerna med stödjande funktioner för att länka information, bygga applikationer och göra information så kvalitativ, relevant och användbar som möjligt, för så många som möjligt. Genom ökad användning av de befintliga infrastrukturella resurser som kulturarvsinstitutionerna redan använder sig av, som exempelvis den nationella forskningsinfrastrukturen SUNET som sammankopplas med ett brett utbud av, i många fall, redan befintliga tjänster som är särskilt anpassade till kulturarvsdata, kan en samordnad och mer effektiv informationshantering uppnås.

Sammanfattningsvis redovisas ett förslag till fortsatt arbete för en samordnad digital informationshantering, utifrån den tentativa målbild som har tagits fram av arbetsgruppen och som redovisas i rapporten.

Bakgrund

Digital kulturarvsinformation

Kulturarvsinformation har skapats över lång tid och är under ständig förändring. Den digitala utvecklingen har gett oss verktyg för att visa institutionernas samlingar på ett sätt som eliminerar fysiska begränsningar och som därmed kan nå en bredare publik än någonsin tidigare. Tekniken har också gett kulturarvsinstitutionerna möjligheter att skapa digitala berättelser och engagera användare på den digitala kulturarvsarenan.

Metadata är strukturerad information, relaterad till olika slags material som används för identifiering och beskrivning, och underlättar åtkomst. Metadata som begrepp syftar på den information som beskriver ett objekt, det vill säga "data om data". Det finns flera olika typer av metadatastandarder som idag används för att hantera digital kulturarvsinformation, men det finns ingen metadatastandard som används för alla typer av behov och alla typer av samlingar. Detta är en central fråga för harmoniseringen av det digitala kulturarvet. Den konceptuella metadatastandarden CIDOC CRM har kommit längst i skapandet av en universell standard för kulturarvsdata, med ett definierat RDF-format.

Metadata skapas utifrån de traditioner (för till exempel katalogiseringar) som var gällande vid tiden då respektive samling skapades. Det innebär att objekt och dokument som från början kunde vara del av samma samling eller institution idag kan vara splittrade över flera typer av organisationer eller inom olika delar av samma organisation. Med andra ord kan till exempel en karta hanteras som ett föremål i ett museisystem, med tillhörande metadatastandarder och möjligheter till tillgängliggörande och bevarande, samtidigt som en karta av samma typ hanteras som arkivmaterial med helt andra beskrivningar och enligt andra rutiner och standarder vid ett bibliotek eller en annan institution (se bild 1 nedan).



Museiföremål



Arkivhandling



Kartografiskt
biblioteksmaterial

Bild 1. Samma typ av kulturarvsinformation hanteras på olika sätt vid institutioner.

Ett övergripande mål för den nationella digitaliseringsstrategin är att kulturarvsinstitutionerna ska öka mängden tillgänglig och långsiktigt bevarad digital kulturarvsinformation. Hur information tillgängliggörs och används har drastiskt ändrats i och med tillgången till nya digitala metoder, tjänster och verktyg. Idag kan man föra samman kulturarvsinformation som förvaras på olika institutioner och i olika domäner, men för att få ut den fulla potentialen är det fortfarande mycket som behöver göras.

För att synliggöra och öka användningen av samlingar är det nödvändigt att fokusera på kulturarvsinformationens struktur och kvalitet. Idag finns de tekniska förutsättningarna för detta, men tekniken är bara ett medel och det är viktigt att fokus ligger på vad som behöver uppnås med tekniken och inte på tekniken i sig. Grundläggande åtgärder för att öka potentialen i kulturarvsinformation är att se till att hela den digitala livscykeln är sammanhållande och att vedertagna internationella standarder används.

Informationshantering

För att kulturarvsinformation inte bara ska bli lätt för användarna att hitta, oavsett var den finns, utan också möjlig att använda och återanvända sömlöst är standardisering, interoperabilitet och harmonisering av data grundläggande frågor. En förutsättning för effektiv sammanlänkning av data från olika domäner är interoperabilitet mellan de digitala beskrivningarna av information från olika kulturarvsdomäner. Interoperabilitet kan åstadkommas genom användning av vedertagna internationella standarder och en harmonisering av domänspecifika standarder. Det finns olika strategier som institutionerna använder sig av när det gäller mappning mellan olika metadatamodeller. I de allra flesta fall görs detta mot den lägsta gemensamma nivån, vilket innebär att informationen försämras. Vad som istället vore önskvärt är att mappningen sker mot en gemensam nivå som tillåter beskrivning av högsta kvalitet, där all information blir interoperabel även om mängden information inte är densamma hos var och en av institutionerna.

Att tillgängliggöra digitalt kulturarvsmaterial handlar idag främst om olika former av publicering på internet. Att informationen finns på internet är ett första steg mot att informationen ska bli använd, men för att informationen ska kunna bli användbar i större skala är det helt avgörande att den tillgängliggörs på ett optimalt sätt. Det räcker inte att enbart publicera digitala filer på internet för att de ska kunna hittas av den intresserade. För en hög sökbarhet är en grundförutsättning att relevant metadata skapas för de digitala filerna. Detta gäller i synnerhet om informationen ska kunna nås genom aggregatorer eller andra tjänster som samsöker digitalt material som lagras hos olika förvaltare.

Aggregering, interoperabilitet och standarder är också kärnfrågor i de åtaganden som Sverige gjort genom Europeiska Unionens rådsslutsatser om Europeana⁵. Rådsslutsatserna ska bland annat uppfyllas genom den nationella digitaliseringsstrategin. I dessa framgår det att beständighet för Europeana och för de processer för digitalisering, tillgång och digitalt bevarande som inletts i medlemsstaterna bland annat förutsätter ”*aktivt och varaktigt stöd för digitalisering och elektronisk tillgång till kulturellt material och digitalt bevarande*” samt ”*vidareutveckling av praktiska samordnade lösningar i överensstämmelse med upphovsrätten och närstående rättigheter för att*

⁵ [http://eur-lex.europa.eu/legal-content/SV/TXT/PDF/?uri=CELEX:52008XG1213\(04\)&from=SV](http://eur-lex.europa.eu/legal-content/SV/TXT/PDF/?uri=CELEX:52008XG1213(04)&from=SV)

bidra till den elektroniska tillgången till ett så omfattande kulturellt material som möjligt”. Vidare uppmanas medlemsstaterna att inom ramen för sina nationella strategier fortsätta stödja integreringen av digitaliserat kulturellt material i Europeana, bland annat genom att främja standarder för interoperabilitet samt genom att inrätta eller förbättra portaler som sammanställer de nationella innehållen.

Idag splittras kulturarvsinformationen i Sverige både genom olika sätt att beskriva den på och genom olika system som hanterar respektive aggregerar data (se bild 2 nedan).

Exempel: Streiff, Gustav II Adolfs häst

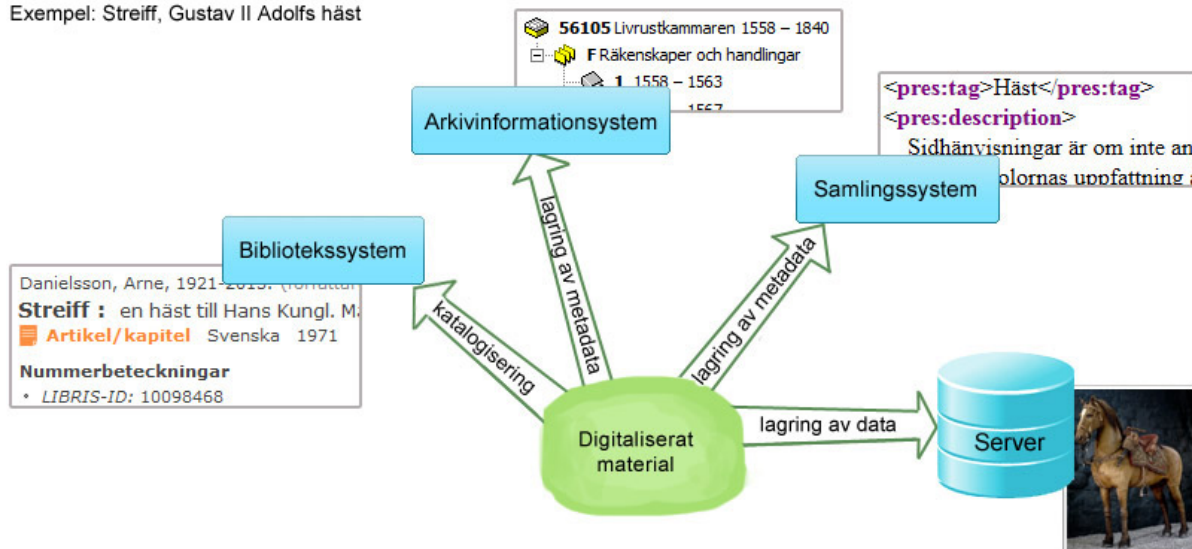


Bild 2. Olika domänspecifika beskrivningar av kulturarvsinformation vid institutioner.

Centrala roller i att tillgängliggöra informationen i dagens aggregeringsinfrastruktur spelar de nationella domänspecifika aggregeringstjänsterna NAD, Libris, SMDB, K-samsök och GBIF-Sweden. Dessa tjänster aggregerar metadata eller tar på annat sätt ett övergripande domänspecifikt ansvar. De myndigheter som förvaltar tjänsterna är Riksarkivet, Kungliga biblioteket, Riksantikvarieämbetet och Naturhistoriska riksmuseet.

En del av kartläggningen av den digitala informationshanteringen består i att undersöka hur integreringen av digitaliserat kulturarvsmaterial i Europeana kan stödjas. Digisam har därför även tittat på huruvida standarder som används av de aggregatorer som sammanställer de nationella innehållen är interoperabla på en nationell och internationell nivå, respektive på vilket sätt aggregering och interoperabilitet kan förbättras och/eller vidareutvecklas.

Problemformulering

Digisam har genomfört workshoppar och modelleringar med representanter från aggregeringstjänster samt andra kulturarvsinstitutioner för att ta fram ett underlag som beskriver nuläge, behov och förväntningar. Dessa workshoppar och modelleringar resulterade i att några områden ringades in som särskilt viktiga:

- Kvalitetsfrågor, datakvalitet.
- Semantiska webbens möjligheter.
- Gemensamma lösningar som byggs utifrån befintliga strukturer och kopplas till långsiktigt bevarande.

Ökad kvalitet av data, standardisering och framtida publicering av länkbar data är grundläggande för att all digital kulturarvsinformation i framtiden ska kunna bli meningsfull och användbar. Nedan redovisas de svårigheter som togs upp i relation till dessa områden.

Ökad kvalitet av data

Aggregeringstjänsternas kapacitet för att ta emot metadata och hantera vedertagna internationella standarder varierar kraftigt idag. En konsekvens av detta är att institutionerna ofta själva avsätter tid och resurser för att utveckla mappningsprocesser och strukturera olika typer av domänspecifik information för leveranser till olika aggregeringstjänster. Detta gäller både institutioner som förvaltar kulturarvsdata och de som förvaltar aggregeringstjänster. Metadata mappas under aggregeringsprocessen till format som i de flesta fall bygger på principen ”lägsta gemensamma nämnare”, vilket resulterar i en lägre metadatakvalitet.

Aggregering är idag i första hand ett sätt att göra informationen synlig och användbar, både på nationell och internationell nivå. Samtidigt ligger det en problematik i att det i de flesta fall saknas möjligheter att få tillbaka de kvalitetshöjningar som sker genom aggregeringstjänsterna till institutionens egna system för förvaltning och bevarande.

Semantiska webbens möjligheter

Den semantiska webbens möjligheter framhölls vid workshopparna, av representanter för aggregeringstjänsterna, som en teknik som kommer att effektivisera de processer som gör informationen användbar. En konsekvent implementering av länkad öppen data, verktyg m.m. ligger dock fortfarande långt fram i tiden.

Gemensamma lösningar som byggs utifrån befintliga strukturer och som har koppling till långsiktigt bevarande

Olika typer av standardiseringstjänster, konverteringsverktyg m.m., som i vissa fall skulle kunna användas gemensamt, utvecklas idag separat av respektive förvaltande organisation. Under workshopparna och modelleringarna rådde det därför konsensus kring behovet av gemensamma lösningar i form av stödfunktioner kopplade till de tekniska infrastrukturerna; dels i form av tekniska verktyg som aggregeringstjänster som institutioner skulle kunna nyttja gemensamt, dels i form av råd och stöd för att öka mognad, kunskap och deltagande i dessa frågor hos institutionerna.

När en övergripande bild av informationsflödena vid aggregering av metadata togs fram konstaterades även att en grundläggande förutsättning för interoperabilitet är kopplingar mellan de aggregeringstjänster som finns. Sådana kopplingar saknas idag. Det innebär att det idag saknas ett sätt för användare att på ett samordnat sätt söka/använda all informationen som finns i de olika aggregeringstjänsterna. Det konstaterades även att det inte finns behov av en enda gemensam aggregator utan att det finns fördelar med domänspecifik aggregering.

Det är viktigt att börja titta på hur det kan bli enklare och mer resurseffektivt för institutioner att publicera och/eller leverera data på flera sätt, då det redan idag finns flera parallella mappningsprocesser till olika aggregeringstjänster. Även i en helt distribuerad miljö skulle data behöva aggregeras för att förenkla för användarna att hitta relevant information.

En tentativ målbild

Workshopparna och modelleringarna ledde till konsensus kring att länkad öppen data är central i den framtida målbilden, samt att aggregering som sådan kommer att fortsätta vara viktig under en överskådlig framtid. Standardisering av data sågs som grundläggande i målbilden, och vikten av att nationella strukturer byggs i interoperabla och vedertagna standarder framhölls.

Inför workshop 2 beskrevs och presenterades tre olika alternativa framtidsscenarier⁶. Det rådde samstämmighet kring att den tentativa målbilden för det fortsatta arbetet bör utgå från scenario nr. 2 (se bild 3 nedan), som i sin tur innebär att gemensamma lösningar byggs utifrån befintliga aggregeringsstrukturer genom att de integreras och kopplas samman med varandra med stöd av olika tjänster samt data- och internetkapacitet.

Motivet för valet av den tentativa målbilden utgick ifrån att vi bör utgå från de system, de infrastrukturer, den funktionalitet m.m. som finns idag och fokusera på att knyta samman befintliga strukturer och aktörer med flexibla lösningar i stället för att förorda en från grunden helt ny infrastruktur. På så sätt kan befintliga infrastrukturella delar knytas samman inom den befintliga infrastrukturen med nya komponenter som kopplar samman olika delar och som kan förändras och/eller bytas ut i takt med att tekniken evolverar.

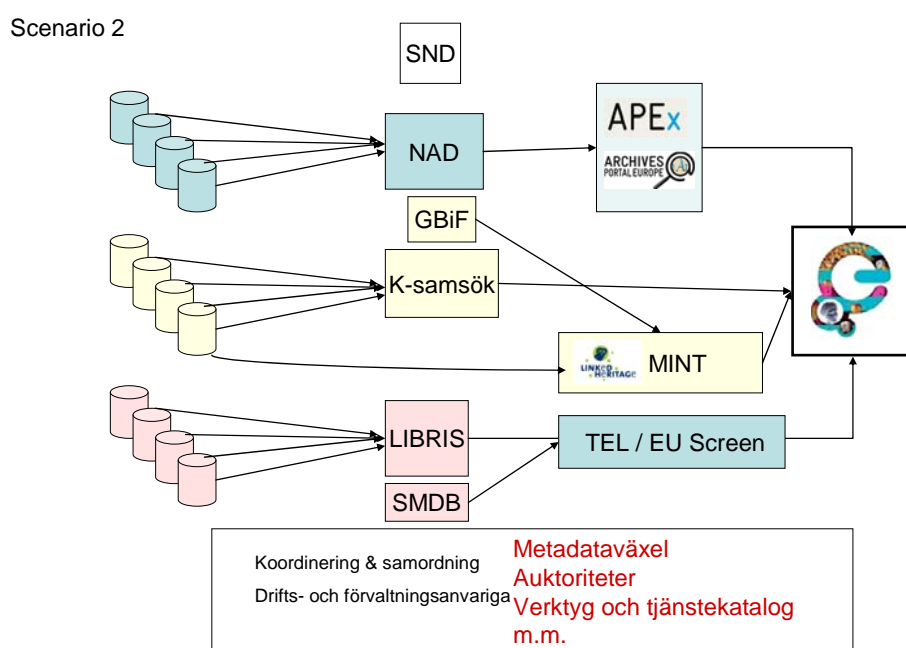


Bild 3. Tentativ målbild

Kulturarvsinstitutionerna skulle, enligt denna tentativa målbild, kunna effektivisera struktureringen av data till valfritt format genom olika tjänster, bland annat en gemensam metadataväxel med

⁶ <http://digisam.se/index.php/docman/public-documents/infra/10-workshop-aggregator-2013-04-05/file>

mappningsverktyg för strukturering respektive aggregering av information och andra funktioner som stödjer sammanlänkning av informationen.

Informationen skulle därefter kunna aggregeras vidare, antingen genom domänspecifika aggregatorer, eller direkt till europeiska/internationella projekt. En del av målbilden är att metadataväxeln fungerar som en gemensam tjänst för att standardisera och leverera olika typer av information. Eftersom kulturarvsinstitutioner bedöms att under en längre tid behöva domänspecifik aggregering på nationell, europeisk och internationell nivå skulle en sådan funktionalitet effektivisera processen, vad gäller både resurser och tid.

Målbilden har, sedan den togs fram, uppdaterats under fyra fördjupade modelleringar med nationella aggregatorer för arkiv, bibliotek, musei- och kulturmiljöinformation, samt naturhistorisk information.

Nedan en bild (bild 4) som sammanställer vilka verktyg som kan vara aktuella i framtiden. ”Verktygslådan” är ett ramkoncept som allt eftersom kan fyllas med de moduler det finns behov av för framtiden.

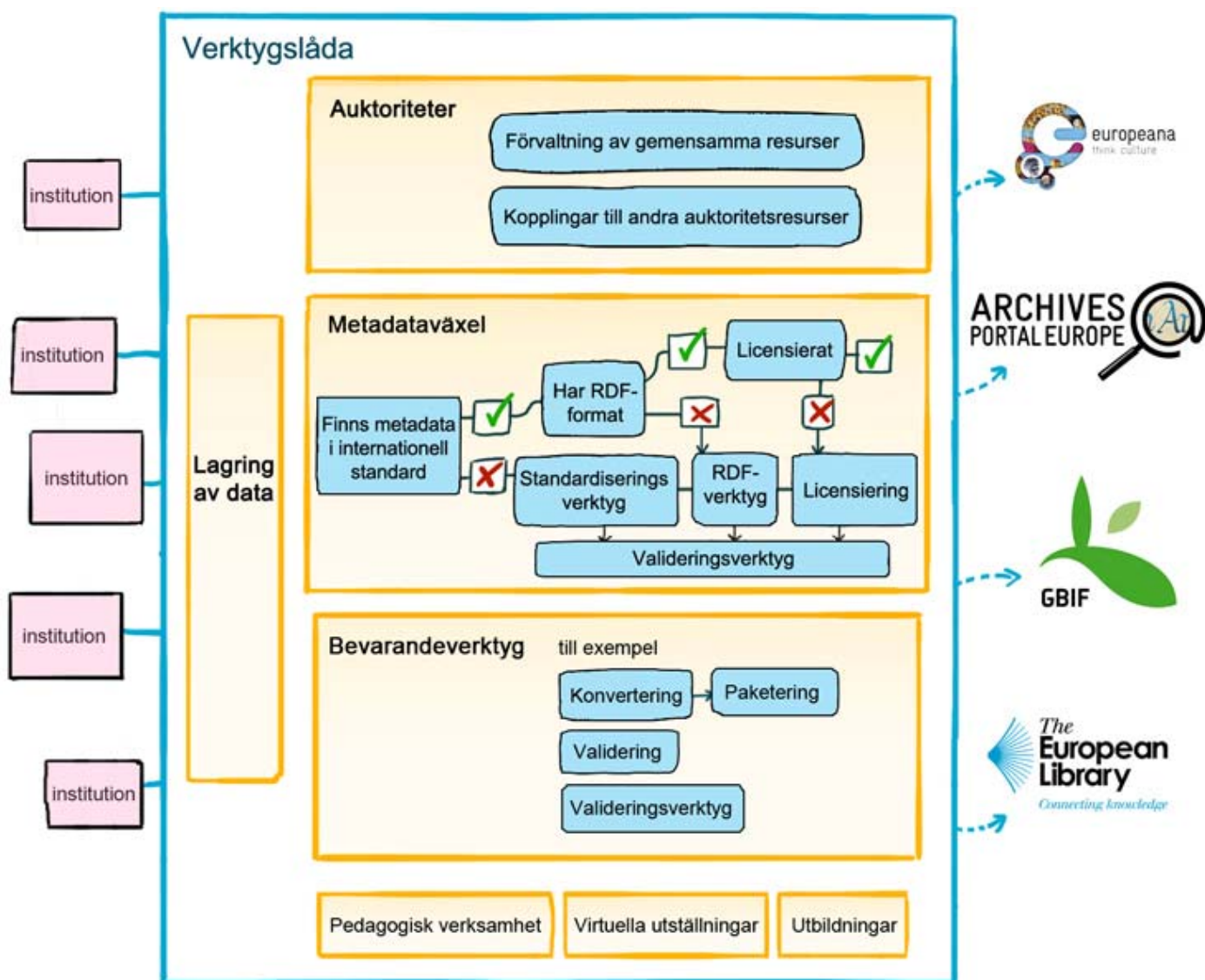


Bild 4. Verktygslåda.

På grundval av resultaten från alla workshoppar och modelleringar har den tentativa målbildens komponenter ställts upp på följande sätt:

- Aggregeringstjänster
- Semantiska webben
- Standarder
- System
- Digitala infrastrukturer
- Gemensamma stödfunktioner/verktyg och tjänster

Fördjupade beskrivningar av den tentativa målbildens komponenter som, till viss del, omfattar informationshantering på institutionsnivå, finns under avsnitt 6 i denna rapport. Till exempel handlar ökad kvalitet av data i stor utsträckning om vilka standarder och system som används vid institutionerna.

Vid framtagandet av rapporten har Digisam också tagit hänsyn till de dialogmöten som sekretariatet förde under 2012 med berörda myndigheter och institutioner, samt underlag från de samordningsprocesser som rör de myndighets- och institutionsvisa digitaliseringsplanerna. Underlaget består av Digisams 24 medverkande myndigheters och institutioners svar på bland annat ett frågeschema och en mängdmatrix som redovisades till Digisam i maj 2014.

Målbildens komponenter

I avsnitt 4 i rapporten presenterades den tentativa målbild som har tagits fram vid Digisams modelleringar och workshoppar med representanter från aggregeringstjänster samt andra kulturarvsinstitutioner. I detta avsnitt beskrivs målbildens komponenter närmare.

Aggregeringstjänster

Aggregatorer samlar in information från enskilda organisationer, samt tillser att organisationernas metadata överförs till Europeana eller andra internationella nätverk. Några av de viktigaste uppgifterna för aggregatorerna är att främja och implementera standarder för tillgängliggörande samt att tillhandahålla domänspecifik kunskap och kompetens till institutioner och Europeana. Det finns flera olika typer av aggregatorer, bland annat nationella, landspecifika, regionala eller projektaggregatorer.⁷ De sistnämnda är i många fall tematiskt inriktade.

Digital information skapas och tillgängliggörs idag utifrån de strukturer som har skapats för olika domäner, och det finns ett antal olika sökmöjligheter. De nationella, domänspecifika aggregatorerna spelar en central roll i infrastrukturen för det digitala tillgängliggörandet av samlingar och arkivbestånd. I Sverige finns idag fem stora nationella domänspecifika aggregatorer inom det statliga kulturarvsområdet: NAD (för arkiv), LIBRIS (för bibliotek), SMDB (för audiovisuellt material), GBIF (för naturhistoriska samlingar och observationer) och K-samsök (för kulturmiljö- och museidata). Dessa aggregatorer samlar in information från andra kulturarvsinstitutioner/myndigheter. Det är dock endast beskrivande metadata som levereras till aggregatorerna.

När det gäller tillgänglighet till arkivdata kan man exempelvis hitta information om arkiv i Sverige i Nationella Arkivdatabasen (NAD), och på europeisk nivå genom Europeiska arkivportalen/Archives Portal Europe. I båda fallen är informationen strukturerad enligt arkivstandarder. På bibliotekssidan är det Libris på nationell och TEL (The European Library) på europeisk nivå som är ingångarna för bibliografiskt material. För audiovisuella samlingar är det SMDB på nationell och EUROScreen på europeisk nivå. En samsökningsfunktion har utvecklats för biblioteks-, audiovisuell- och arkivdata, Sondera. Museisamlingar görs idag tillgängliga på olika sätt, exempelvis genom museernas hemsidor, museisystemens portaler med information från alla institutioner som använder samma samlingssystem, eller regionala portaler. Hur beskrivningar och metadata ser ut varierar beroende på vilka standarder som används vid institutionerna eller i museisystemen. De museer som levererar metadata till K-samsök gör det i K-samsöks format och informationen blir då sökbar genom applikationen Kringla. Det finns också flera museer i Sverige idag som levererar informationen direkt till Europeana genom olika EU-finansierade projekt. Deras metadata struktureras då enligt LIDO-format.

Den nationella aggregeringskartan (2013-04-05) illustreras i bilden nedan (bild 5).

⁷ <http://linkedheritage.cab.unipd.it/training/LO-03/sv/01.html>

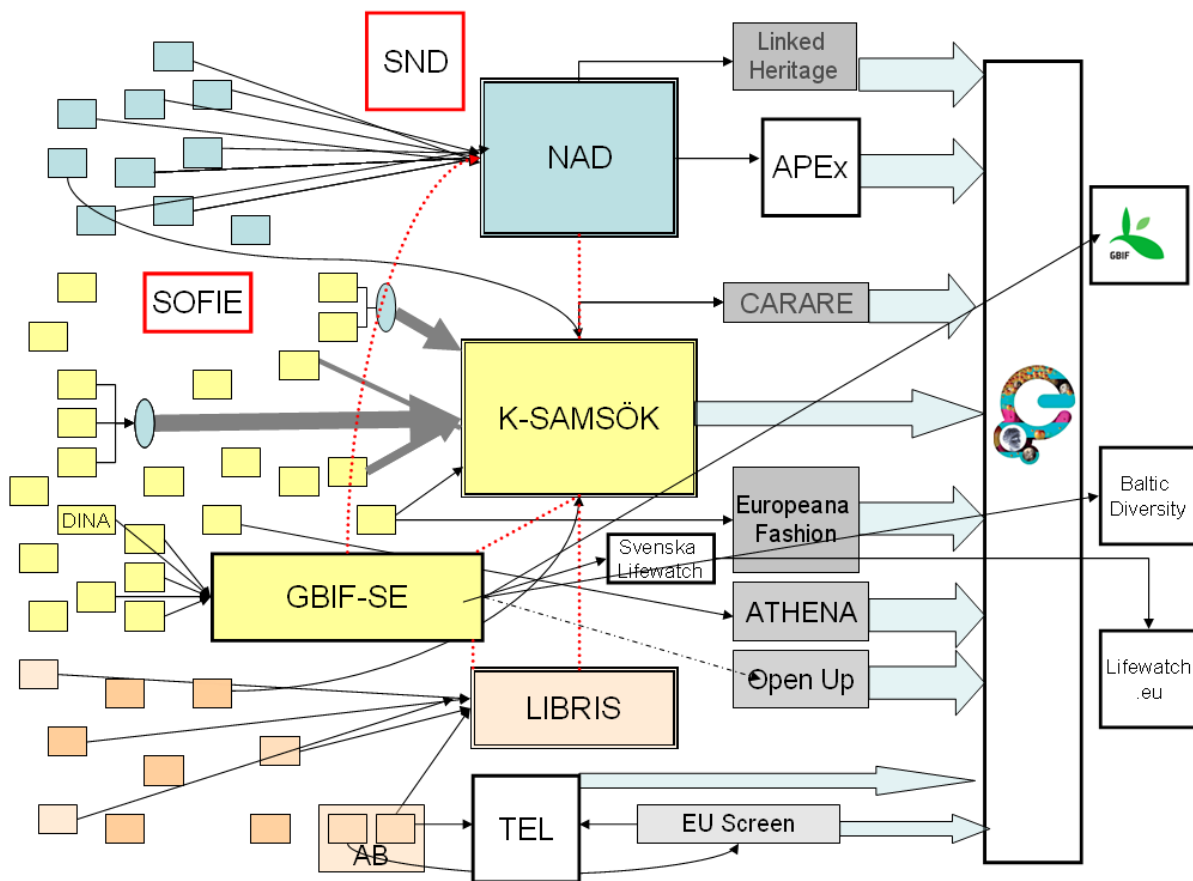


Bild 5. Befintliga aggregeringsflöden.

Viktigt att lyfta fram är att idag skapas interoperabilitet mellan olika domäner på en mer generell nivå först i Europeana. Däremot finns det betydligt mer information i metadata som tillgängliggörs på nationell nivå än vad som syns i Europeana.

De samsökningstjänster som finns på nationell nivå är inte sammankopplade. Försök har gjorts, exempelvis genom samsökningstjänsten Sondera där man kan samsöka i NAD, SMDB och Libris eller i musei- och kulturmiljöinformation genom Kringla. Utöver dessa sätt att tillgängliggöra kulturarvsinformationen idag finns det flera regionala portaler samt information som tillgängliggörs genom europeiska och internationella projekt och initiativ.

Aggregering av information från det statliga kulturarvsområdet sker också i form av ett europeiskt samarbete. En gemensam aggregeringsfunktion finns på europeisk nivå och går under namnet MINT (Metadata Interoperability Services). Denna webbaserade funktion för aggregering av metadata, fungerar som en europeisk aggregator som skickar informationen vidare till Europeana.

NAD

Den nationella arkivdatabasen (NAD) vid Riksarkivet är ett system för arkivsektorn som innehåller ett nationellt arkivbildarregister med information om arkiv och samlingar som finns i hela landet. NAD tillgängliggör register över arkivinstitutioner och hjälpdatabaser för topografisk indelning och förvaltningshistorik. Arkivdatabasen innehåller arkivförteckningar från arkiv förvarade hos Riksarkivet, vilka är sökbara i NAD. Övriga arkivinstitutioner kan länka till sina arkivförteckningar

i NAD. En förutsättning för publicering i NAD är att arkivstandarderna följs vid leverans av data.⁸ Det pågår för närvarande (september 2014) en satsning för att göra det möjligt för arkivinstitutioner att exportera sina arkivförteckningar till NAD. Om arkivförteckningarna exporteras till NAD i EAD-format kommer de också att kunna vidareexporteras till APEX. Cirka 120 institutioner är anslutna och har levererat uppgifter om drygt 125 000 arkivbildare. Antalet förtecknade arkivenheter (serier, volymer etc) uppgår totalt till 9.2 miljoner.

I det EU-finansierade projektet Archives Portal Europe network of excellence (APEX) samarbetar 28 länders nationalarkiv. Ett av de viktigaste målen inom projektet är aggregering av metadata och bildlänkar ifrån lokala och regionala arkivinstitutioner till den Europeiska arkivportalen (APE). NAD är idag den nationella samordningsfunktionen för APEX, och som så kallad Country Manager (CM) håller NAD i aggregeringsprocesserna gentemot den Europeiska arkivportalen. Inom ramen för det VINNOVA-finansierade projektet Culture Cloud har Riksarkivet bl.a. intresserat sig för hur lokala och regionala arkivförvarande institutioners aggregeringsprocesser skulle kunna underlättas.⁹

LIBRIS

LIBRIS är en nationell söktjänst med information om titlar på svenska universitets-, högskole- och forskningsbibliotek samt folkbibliotek. Här finns information om böcker, tidskrifter, artiklar, kartor, affischer, elektroniskt publicerat material, med mera. De registrerande biblioteken bidrar gemensamt till uppbyggnaden av katalogen. För närvarande finns sju miljoner titlar i databasen.¹⁰ Svensk nationalbibliografi och svenska auktoritetsposter är sedan sommaren 2011 fritt tillgängliga utan restriktioner, under CC0 licens och via API:er. Katalogposterna i Libris-databasen, härstammande från OCLC, kommer också att göras fritt tillgängliga under våren 2014.¹¹ De innehåller länkar till Wikipedia, DBPedia, LC Authorities (namn och ämnesord) samt VIAF.¹² Libris-databasen består av flera delmängder. Nationalbibliografier och auktoritetsposterna är delmängder som gjordes fritt tillgängliga under 2011. Katalogposterna i Libris har flera ursprung, de olika delmängderna har identifierats och arbete med att få tillåtelse att göra dem fritt tillgängliga pågår. Poster med ursprung i OCLC kommer att göras fritt tillgängliga under september 2014.

SMDB

Svensk mediedatabas, SMDB,¹³ är söktjänsten för Kungliga bibliotekets audiovisuella samlingar. Här finns uppgifter om tv, radio, video, biograffilm, skivor och multimedier. På det audiovisuella området har Kungliga biblioteket deltagit i projektet EUScreen¹⁴.

Riksarkivet och Kungliga biblioteket har tillsammans tagit fram en söktjänst, SONDERA, vari man kan samsöka material från NAD, LIBRIS och SMDB.

⁸ http://www.nad.riksarkivet.se/Information?infosida=_OmNADArkiv.cshtml&titel=Om NAD Arkiv

⁹ <http://www.vinnova.se/sv/Resultat/Projekt/Effekta/2012-01393/CultureCloud/>

¹⁰ http://librishelp.libris.kb.se/help/about_libris_swe.jsp?redirected=true&pref_is_set=&textsize=&contrast=&language=se

¹¹ http://librishelp.libris.kb.se/help/tech_swe.jsp?open=tech

¹² <http://www.kb.se/aktuellt/nyheter/2014/Sju-ars-forhandlingar-avslutade/>

¹³ <http://smdb.kb.se/>

¹⁴ <http://www.euscreen.eu/>

K-Samsök

Systemet K-samsök samlar och tillgängliggör kulturarv från svenska museer och andra minnesinstitutioner. Data om kulturarv som finns på en mängd olika institutioner och platser samlas in till K-samsöks databas och görs därifrån tillgänglig för användning, berikning och utveckling. K-samsök är nationell aggregator för leverans av data till Europeana. K-samsök har också till uppgift att göra det möjligt för externa applikationer att komma åt informationen genom ett API samt att tillhandahålla den skördade informationen genom URI:er och att förädla den skördade informationen med information från allmänheten och andra källor.

K-samsök har tagit fram flera söktjänster. I söktjänsten Kringla går det att samsöka i den musei- och kulturmiljöinformation som finns i K-samsök.

Utvecklare och beställare av tekniska tillämpningar, till exempel webbtjänster och mobilapplikationer, skapar spännande och intressanta lösningar utifrån K-samsöks innehåll. Detta bidrar till kunskap, lärande och upplevelser samt till näringslivsutveckling.

GBIF

GBIF (Global Diversity Information Facility) är ett internationellt nätverk som gör information om jordens arter fritt tillgänglig för alla via en sökbar internetportal.¹⁵ Naturhistoriska riksmuseet förvaltar en svensk GBIF-nod.

MINT

MINT är en webbaserad plattform som har utformats och utvecklats för att underlätta aggregering av metadata från europeiska institutioner. Flera kulturarvsinstitutioner i Sverige har deltagit i olika EU-finansierade projekt kring digitalisering och i en del av dessa, såsom Linked Heritage, Carare, Europeana Fashion, Partage Plus, Athena Plus, har metadata aggregerats till Europeana genom det europeiska aggregeringsverktyget MINT.¹⁶ MINT är baserad på öppen programvara som har utvecklats av det nationella tekniska universitetet i Aten (NTUA) inom ramen för Athena-projektet. Med MINT kan man i en och samma webbmiljö ladda upp, mappa, validera och leverera metadata som ska skickas till Europeana. MINT möjliggör också administration av användarroller, rättigheter och olika mappningsscheman.

Semantiska webben

Länkad data

Att informationen idag är strukturerad på olika sätt, med metadata som har skapats och hanterats i en rad olika standarder och system, gör att sökningarna oftast baseras på mappning till ”minsta gemensamma nämnare” och därmed inte resulterar i optimala svar. Komplexiteten i

¹⁵ data.gbif.org

¹⁶ <http://mint.image.ece.ntua.gr/redmine/projects/mint/wiki>

informationsstrukturerna är en av de bakomliggande orsakerna till det stora intresset för sammanlänkning av information som finns i olika domäner genom den semantiska webbens möjligheter, så kallad "länkad data".

Länkbar data

Det finns idag flera initiativ kring att lägga ut data i maskinläsbart format (exempelvis RDF), vilket gör data länkbara. Även om det finns en del publicerade länkbara data så finns det få exempel där sammanlänkningar med andra länkbara data har gjorts så att länkad data har skapats. Syftet med det digitala tillgängliggörande som sker genom länkad data är bland annat att göra det möjligt för utvecklare att bygga tjänster som baseras på dessa data; tjänster som bättre kan fånga upp användarnas behov.

Öppen data

Förutsättningar som behövs för att skapa sammanlänkade data som kan återanvändas är att dessa data är maskinläsbara och öppna för fri användning och återanvändning. Även öppen data i format som inte är läsbara kan skapa användbarhet, men själva maskinläsbarheten gör att den tekniska utvecklingen kan nyttjas till fullo.

Länkad öppen data

Öppna sammanlänkade datakällor kallas för länkad öppen data, "Linked Open Data." Länkad öppen data innebär att information kan lyftas från sitt ursprungliga sammanhang för att utan hinder användas och återanvändas i andra kontexter. "Open data handbook" som getts ut av Open Knowledge Foundation beskriver de legala, sociala och tekniska aspekterna av länkad öppen data och ger följande definition: "Open data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike".¹⁷ MetaSolutions har nyligen publicerat en Vitbok om länkade öppna data är skriven med svenska aktörer som målgrupp. Vitboken har som målsättning att skapa ökad kompetens och systematik kring processen för att tillgängliggöra länkade öppna data.¹⁸

Öppna standarder

Vidare ger användning av öppna standarder¹⁹ bättre förutsättningar för interoperabilitet samt bred användning och återanvändning av digitaliserat material. Med öppen standard avses här den definition som återfinns i European Interoperability Framework version 1.0, s. 9.²⁰

¹⁷ <http://opendatahandbook.org/>

¹⁸ <http://lankadedata.se/vitbok/>

¹⁹ <http://www.opengov.se/sidor/oppen-standard/>

²⁰ <http://ec.europa.eu/idabc/servlets/Docd552.pdf?id=19529>

Standarder

När man talar om det digitala kulturarvet nämns ofta olika metadatastandarder. Metadata som begrepp syftar på information som beskriver ett objekt, det vill säga ”data om data”. Metadata är viktiga för att dokumentera och hantera relationerna mellan informationsresurser. De används för att hitta, samla och bevara resurser över tid. Det finns idag flera olika typer av standarder som används vid kulturarvsinstitutioner för beskrivning av data, så kallade metadatastandarder. Dessa standarder kan vara tekniska, beskrivande, konceptuella, processuella, administrativa eller bevarandestandarder. Beskrivande standarder grundas ofta i Dublin Core,²¹ en metadatastandard ursprungligen skapad för att underlätta sökning och hantering av webbaserade dokumentresurser. Dublin Core består av femton beskrivande element och den används ofta som en slags minsta gemensamma nämnare. Tekniska standarder är oftast icke-beskrivande. Det finns även så kallade utbytesformat, som beskriver metadata som levereras för aggregering. Exempel på utbytesformat är LIDO eller EDM (Europeana Data Model).

Vanligtvis sker skapandet av metadata i institutionernas egna verksamhetssystem. Kulturarvsdata består då ofta av information som registreras och hanteras i verksamheten, händelser kring objekten, personer, organisationer och platser (ofta relaterade till händelser). Metadata i dessa system är ofta välbeskrivna. En svårighet vid sammankopplandet av information från olika domäner eller institutioner är att de standarder som används för dessa beskrivningar i många fall inte bygger på vedertagna internationella standarder utan är framtagna för att beskriva specifika egenskaper av kulturarvsmaterial inom den egna institutionen eller domänen. Det som ytterligare försvårar interoperabilitet mellan olika typer av digital kulturarvsinformation är att både metadata och data ofta finns i specifika system som kan vara egenutvecklade eller som endast har implementerat vedertagna standarder till en viss del. Metadata bidrar till förvaltningen av informationsresurser och hjälper till att säkerställa deras integritet både nu och i framtiden. Att använda sig av standardiserade beskrivningar är viktigt som ett första steg för att skapa interoperabilitet.

Beskrivande metadata finns i de flesta fall i levande system där informationen är under ständig förändring, samtidigt som systemen sällan har en koppling till varken data (bilder, kontextuella metadata, med mera) eller långsiktiga bevarandelösningar. Dessutom skördas en aggregerad del av dessa metadata av aggregatorer där skördad metadata modifieras utan att det alltid finns möjligheter att få denna information tillbaka till institutionens egna system respektive system för bevarande.

En så kallad ”metadataväxel” kan användas till att strukturera data enligt ett visst metadataschema. Den möjliggör även mappning av metadata fält från ett metadataschema till ett annat. Genom att det också finns möjlighet att uttrycka data i maskinläsbart format kan man också länka till olika datakällor, och därmed stödja semantisk interoperabilitet.²²

Det finns flera domänspecifika metadatastandarder och nedan följer några av de standarder som används inom arkiv, museum och biblioteksvärlden.

²¹ <http://dublincore.org/documents/dces/>

²² http://sv.wikipedia.org/wiki/Semantisk_interoperabilitet

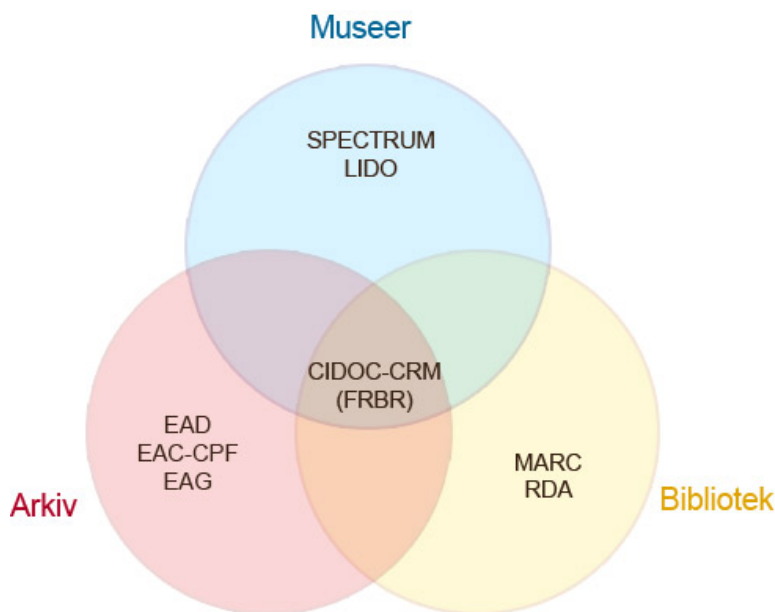


Bild 6. Domänspecifika metadatastandarder.

Museistandarder

När det gäller standarder som används inom ramen för samlingsbeskrivningar (och som därmed utgör kravspecifikationer för samlingsystem) finns det vissa rekommendationer för museer på både internationell och europeisk nivå i en handbok²³ som UNESCO och ICOM har tagit fram tillsammans. De standarder som rekommenderas där är bland annat SPECTRUM, CIDOC CRM och LIDO. Dessa lyfts även fram i de principer som ICOM har publicerat.²⁴ Fördelen med rekommendationerna är att de skapar grund för tvärspektoriell hantering av högkvalitativ metadata inom den digitala kulturarvsdomänen. Under Digisams arbete med kartläggningen av informationshanteringen kunde vi dock konstatera att standarderna oftast inte används alls, eller i väldigt liten utsträckning, vid svenska museer. Representanter för större samlingsystem uppger dock att deras datamodeller liknar dessa standarder samt att export till dessa standarder är möjlig som tillval. Ett undantag vad gäller standarder gäller naturhistoriska samlingar, där särskilda rekommendationer finns genom internationellt samarbete och där nätverket GBIF-Sweden stödjer institutionerna i frågor kring val, nedladdning och installation av samlingshanteringssystemen.²⁵

CIDOC - Conceptual Reference Model²⁶ är en konceptuell modell, ontologi och ISO-standard (ISO 21127) för semantisk integration av digital kulturarvsinformation. Den förvaltas av en intressegrupp inom ramen för ICOM (International Council of Museums) som går under namnet "The CIDOC Conceptual Reference Model Special Interest Group".²⁷ CIDOC CRM har en central roll när det kommer till insatser som handlar om tvärspektoriell hantering av metadata inom den digitala kulturarvsdomänen. Den övergripande inriktningen för intressegruppen är att göra CIDOC CRM till

²³ <http://www.unesco.org/new/en/culture/themes/museums/unescoicom-museum-studies-training-package/running-a-museum-a-practical-handbook>

²⁴ http://network.icom.museum/fileadmin/user_upload/minisites/cidoc/DocStandards/principles6_2.pdf

²⁵ <http://www.gbif.se/blog/article/samlingshanteringssystem>

²⁶ <http://www.cidoc-crm.org/index.html>

²⁷ <http://network.icom.museum/cidoc/working-groups/crm-special-interest-group/>

en grund för semantisk interoperabilitet för digital kulturarvsinformation, och därigenom skapa förutsättningar för semantiska länkar av hög kvalitet, med målet att omvandla domänspecifika informationskällor till tvärsektoriella kulturarvsresurser. För genomförandet på institutionell nivå finns det riktlinjer och bästa praxis. CIDOC CRM fokuserar på definitioner av relationer, snarare än klassifikation och händelsebaserad modellering, och använder sig till exempel av hierarkiska relationer för att göra sökningar enklare. Idag finns det ett stort fokus på länkad öppen data i kulturarvssektorn. Det finns dock ett behov av att förädla data på institutionell nivå och standardisera metadata på en tvärsektoriell nivå för att synliggöra data och göra den sökbar. CIDOC CRM är för närvarande huvudsakligen implementerad inom den europeiska museisektorn och är nu också harmoniserad med den konceptuella standarden för bibliotek - FRBR. Initiativ har också tagits för att harmonisera arkivstandarder med CIDOC CRM.

LIDO (Lightweight Information Describing Objects) är en tillämpning av CIDOC CRM. Oavsett vilken standard som finns i museernas databaser/system kan metadata oftast aggregeras i LIDO på ett sätt som behåller mycket mer strukturerad information från det ursprungliga formatet än vad andra format kan göra. LIDO är ett format som främjar interoperabilitet vad gäller aggregering av museidata på en europeisk nivå och används i de flesta EU-finansierade projekt som levererar data till Europeana. LIDO är resultat av ett samarbete mellan internationella aktörer inom museisektorn, som påbörjades 2008, med syfte att skapa en gemensam lösning för att tillföra kulturarvsinnehåll till webbapplikationer. Även om LIDO ursprungligen var framtagen för museisektorns behov används det idag i domänöverskridande sammanhang på grund av dess anpassningsförmåga och möjlighet till bevarande av integriteten hos rika metadata.

Europeana Data Model (EDM), är det nya formatet som har implementerats i Europeana och som bygger på Europeanas tidigare datamodell, ESE. Både ESE och EDM är baserade på Dublin Core med ytterligare särskilda Europeana-element. Innehållsleverantörer/aggregatorer måste anpassa sina metadata till EDM:s metadataprofil för att informationen ska kunna visas i Europeana.

SPECTRUM²⁸ är en öppen processuell standard för hantering och förvaltning av museisamlingar som används i stor utsträckning internationellt. SPECTRUM-standarderna förvaltas, utvecklas och publiceras av Collections Trust. Den infördes 1994 efter ett omfattande samarbetsprojekt. Idag används SPECTRUM i mer än 40 länder och i nästan 8000 museum. SPECTRUM består av: CORE SPECTRUM Standard (SPECTRUM 4.0 and Appendix 1), SPECTRUM Digital Asset Management, SPECTRUM Schema, samt ett arkiv över tidigare versioner av standarden. Under 2014 och 2015 kommer Riksantikvarieämbetet att tillsammans med Digisam och arbetsgrupper bestående av museirepresentanter översätta och bearbeta SPECTRUM.

Darwin Core är en metadatastandard för utbyte av biodiversitetsdata som utvecklas och tillhandahålls av TDWG-gruppen²⁹, som i sin tur arbetar med standarder inom biodiversitetsinformatik. Med Darwin Core kan man beskriva till exempel en biologisk museisamlingsdatabas eller en biologisk inventering. Vid Naturhistoriska riksmuseet används även ENSE. ENSE är en standard framtagen för svenska data, som tar hänsyn till internationella skillnader som till exempel sockenindelningar.

²⁸ <http://www.collectionslink.org.uk/programmes/spectrum>

²⁹ Biodiversity Information Standards <http://www.tdwg.org/>

ABCD (Access to Biological Collection Data) är en samlingsdatabasstandard för biodiversitetsdata som i likhet med Darwin Core utvecklas och tillhandahålls av TDWG.

ISCI³⁰, International Standard Collection Identifier (ISO 27730:2012), är en ISO-certifierad standard för beskrivning av samlingar.

På nationell nivå använder museerna i Sverige, inom ramen för aggregeringsaktiviteterna, K-samsöks egenutvecklade metadataschema som uttrycks i RDF-format, ett så kallat "K-samsök protokoll"³¹. Metadata som skördas överförs till K-samsöks format i de lokala så kallade "portarna" som utvecklas av en eller flera institutioner, där specifika mappningar görs som är anpassade till innehållsleverantörernas system och databaser. Fördelen med detta egenutvecklade system är att man har automatiska uppdateringar som gör att alla uppdateringar i databasen överförs till aggregatorn. Nackdelen är att uppdateringar och/eller ändringar i K-samsöks format innebär att portarna behöver anpassas, vilket av innehållsleverantörer ofta upplevs som tids- och resurskrävande.³² Detsamma gäller vid förändringar i system och/eller databasversioner.

Arkivstandarder

På en internationell nivå rekommenderas arkivstandarder som är utvecklade inom ramen för ICA.³³ Dessa arkivstandarder saknar dock ett utbytesformat och är renodlade beskrivningsstandarder. Det betyder att ingen XML är inblandad.

EAD (Encoded Archival Description) är en de-facto standard uttryckt i XML för arkivbeskrivningar. EAD är alltså ett utbytesformat baserat på XML, men används av många som en beskrivningsstandard. EADs "bakomliggande standard" från ICA är ISAD (G) (General International Standard Archival Description). ISAD (G) utvecklades först och därefter togs EAD fram. Utvecklingen av EAD drivs av en teknisk arbetsgrupp inom SAA (Society of American Archivists)³⁴.

EAC-CPF (Encoded Archival Context for Corporate Bodies, Persons, and Families) är även den en de-facto standard uttryckt i XML för att beskriva arkivbildare och upphovsmän. EAC-CPFs "bakomliggande standard" från ICA är ISAAR-CPF (General International Standard Archival Description). ISAAR-CPF utvecklades först, sedan togs EAC-CPF fram som utbytesformat. Utvecklingen av EAC-CPF drivs av en teknisk arbetsgrupp inom SAA (Society of American Archivists).

EAG (Encoded Archival Guide) är utbytesformatet i XML för att beskriva en institution med arkivhandlingar. EAGs "bakomliggande standard" från ICA är ISDIAH (International Standard for Describing Institutions with Archival Holdings). Utvecklingen av EAG har skett i Spanien och efter det har APEX-projektet tagit fram ett uppdaterat format influerat av arbetet med EAC-CPF. Det finns i dagsläget ingen klar ägare av formatet.

Användningen och implementeringen av utbytesformaten på en europeisk nivå har framförallt skett inom APEX-projektet där man bland annat tagit fram en så kallad apeEAD-beskrivning som

³⁰ http://www.iso.org/iso/catalogue_detail.htm?csnumber=44293.

³¹ <http://www.ksamsok.se/resurser/protokollversioner/>

³² <http://www.ksamsok.se/wp-content/uploads/2013/02/Teknisk-Review-Rapport-PA3.pdf>

³³ <http://www.ica.org/10206/standards/standards-list.html>

³⁴ <http://www2.archivists.org/>

används gemensamt av europeiska arkiv för leveranser till den europeiska Arkivportalen.³⁵ Inom Sverige gäller att större arkivinformati­onssystem som används vid framförallt kulturarvsinstitutionerna har implementerat exportfunktioner till Riksarkivets anpassning av dessa internationella standarder. Riksarkivet deltar också i det europeiska arkivsamarbetet inom ramen för APEX-projektet och är det nationella navet för leveranser från svenska arkiv till portalen.

EAD (Encoded Archival Description) är en XML-standard för arkivbeskrivningar. EAD är ett utbytesformat baserat på XML, men används av många som beskrivningsstandard. EADs "bakomliggande standard" är ISAD (G) (General International Standard Archival Description), som är en renodlad beskrivningsstandard. Alltså är ingen XML inblandad. ISAD (G) växte fram parallellt med EAD.

EAC-CPF (Encoded Archival Context for Corporate Bodies, Persons, and Families) är en XML-standard för att beskriva arkivbildare och upphovsmän. Den är liksom EAD ett utbytesformat baserat på XML, men används av många som beskrivningsstandard. EAC-CPFs "bakomliggande standard" är ISAAR-CPF (General International Standard Archival Description), som är en renodlad beskrivningsstandard. ISAAR utvecklades först, sedan togs EAC-CPF fram som utbytesformat.

EAG (Encoded Archival Guide) är utbytesformatet i XML för ISDIAH (International Standard for Describing Institutions with Archival Holdings).

Biblioteksstandarder

På en internationell nivå koordineras arbetet med biblioteksstandarder inom ramen för IFLAs Committee on Standards,³⁶ där det finns en lista över relevanta standarder publicerad.³⁷

Eftersom biblioteken samlar på publikationer som oftast finns på många andra bibliotek och i många olika länder så var de tidigt ute med att utveckla standardiserade katalogformat. Den huvudsakliga metadata­standarden för bibliotek är MARC (Machine-Readable Cataloguing)³⁸, vilken har används sedan 60-talet i olika vidareutvecklade versioner.

Vad gäller kulturarvsinstitutionernas bibliotek i Sverige används MARC-format vid de flesta. Kungliga biblioteket utvecklar dessutom en webbaserad katalogiseringsklient i projektet LIBRIS XL, vilket bland annat innebär att MARC21 kommer att mappas till det nya formatet JSON-LD (Java Script Object Notation for Linked Data). MARC-formatet kommer emellertid att finnas kvar som ett alternativt val. Som stöd inom Librisregistrering tillhandahåller Kungliga biblioteket en Formathandbok som är en bearbetad översättning av MARC21 Formatet (MARC Standards, Library of Congress, Network Development and MARC Standards Office).³⁹ MARC upplevs idag i vissa sammanhang som begränsande och är sedan flera år under omprövning.

RDA (Resource Description and Access) är en katalogiseringsstandard som innehåller anvisningar och riktlinjer för att formulera bibliografiska data som är anpassad för den digitala utvecklingen och

³⁵ <http://www.apex-project.eu/index.php/outcomes/standards>

³⁶ <http://www.ifla.org/standards-committee>

³⁷ <http://www.ifla.org/standards-committee/list-of-standards>

³⁸ <http://www.kb.se/katalogisering/Formathandboken/MARC-postens-uppbyggnad/>

³⁹ <http://www.kb.se/katalogisering/Formathandboken/>

för att hantera e-resurser, onlinekataloger och semantisk webb. Den innehåller anvisningar och riktlinjer för att formulera bibliografiska data. Nyligen påbörjades vid Umeå universitetsbibliotek ett samarbetsprojekt, ”Nya metadata i bibliotekets vaRDag” som bland annat handlar om införande av RDA, men också om huruvida RDA kan tillämpas för arkivmaterial som alternativ eller komplement till arkivstandarderna EAD och EAC-CPF. Projektet kommer att avslutas i december 2014.⁴⁰

FRBR (Functional Requirements for Bibliographic Records) är en konceptuell modell som handlar om att omvandla katalogstrukturer för att uppfylla användargränssnittet, i enlighet med definitioner av bibliografiska entiteter och relationer dem emellan, samt enligt en basnivå för vad bibliografiska poster bör innehålla. FBRoo är ett samarbete mellan arbetsgrupper för FRBR och CIDOC CRM som handlar om att skapa en formell ontologi som beskriver semantisk bibliografisk information.⁴¹

Standarder för audiovisuella format rekommenderas av bland andra IFLA och IASA⁴². Digisam har under våren inrättat en arbetsgrupp för ökad tillgänglighet till audiovisuellt kulturarvsmaterial som leds av Kungliga biblioteket och som kommer att fördjupa sig i frågor kring dessa typer av material.

Fornminnesinformation

Digital dokumentation av fornminnen, som exempelvis rapporter som lämnas in till Riksantikvarieämbetet efter avslutad arkeologisk undersökning av ett fornminne⁴³, hanteras idag framförallt i analog form, som pappersdokumentation. Större delen av dagens fältdokumentation (såsom rapporter, bilder, analysresultat och GIS-data) skapas digitalt, men då det saknas tekniska och formella förutsättningar (i form av föreskrifter) för samordnad informationsöverföring, sammanlänkning, lagring och bevarande av digital fornminnesinformation, överförs dessa digitala filer i många fall till pappersdokumentation.

För att effektivisera hanteringen av fornminnesinformation har Riksantikvarieämbetet initierat DAP-programmet (Digital Arkeologisk Process) som kommer att pågå mellan 2014 och 2018.

SND har också idag drygt 380 studier med viss digital arkeologisk fältdokumentation i form av framförallt GIS-data fritt tillgängliga, samt ytterligare 80 som är deponerade, under bearbetning och som kommer att tillgängliggöras.⁴⁴

System

Under de senaste tio åren har situationen vad gäller vilka system som finns på svenska museer förändrats drastiskt. För bara några år sedan var mängden databaser och system som användes av museerna stor och många var hembyggda register i till exempel Access, Excel, FileMaker eller Word. Denna bild har idag alltmer koncentrerat sig kring några få större system. Nya funktionella och tekniska krav ställs idag på systemen, bland annat stöd för vedertagna standarder.

⁴⁰ <http://www.kb.se/bibliotek/projektbidrag2/pagaende/>

⁴¹ <http://en.wikipedia.org/wiki/FRBROO>

⁴² <http://www.iasa-web.org/>

⁴³ Enligt Kulturmiljölagens 2:a kapitel

⁴⁴ <http://snd.gu.se/sv/catalogue/gis>

Museisystem

En jämförelse av de internationella museisystemen har gjorts av brittiska Collections Trust i en så kallad Software Survey⁴⁵, bland annat gällande vilka standarder som systemen stödjer samt vilka leverantörer som har implementerat SPECTRUM-rutiner i sina system.

De museisystem som används i Sverige idag är bland annat:

PRIMUS – ett norskt system som används av 40 institutioner i Sverige och fler än hundra i Norge.

MUSEUM Plus – ett tyskt system som används av åtta museer i Sverige och som även finns på institutioner i tjugo andra länder.

Carlotta – ett svenskt system som har utvecklats av Statens museer för världskultur och används av ett fåtal institutioner i Sverige.

TMS – ett amerikanskt system som används av Moderna museet i Sverige och cirka 6000 institutioner i andra länder.

Collective Access – ett open source-system som används av flera läns museer.

SOFIE - ett svenskt system som används av över 350 svenska museer och av mindre aktörer som hembygdsföreningar.

Naturhistoriska samlingssystem utgörs av en databas med ett användargränssnitt för både datahantering och funktioner. Ett samlingshanteringssystem för denna typ av material ska vara specialskrivet för de samlingar det handlar om, anpassat till standarderna ABCD och Darwin Core, och medge import och export av data genom öppna standarder.⁴⁶ Nätverket GBIF-Sweden stödjer institutionerna i frågor kring val, nedladdning och installation av samlingshanteringssystemen.

DINA (Digitalt Informationssystem för NATurhistoriska samlingar) är ett nationellt databassystem som ska hantera information om samlingar vid naturhistoriska museer i Sverige. Systemet är gjort för professionell samlingshantering och erbjuder stöd för registrering, hantering och presentation av de naturhistoriska samlingarna samt viktiga rutiner som exempelvis lånehantering. Programvaran utvecklas inom ett internationellt konsortium och hanterar även informationsflödet till andra portaler såsom GBIF. Ambitionen är att alla större naturhistoriska samlingar i Sverige i framtiden kommer att kunna använda och ha nytta av DINA-systemet.⁴⁷ Samlingssystemet kommer att byggas med programvaror i öppen källkod och bidra till gemensam, distribuerad utveckling.⁴⁸

Arkivinformationssystem

De arkivinformationssystem som används i Sverige idag är bland annat:

⁴⁵ <http://www.collectionslink.org.uk/software/survey>

⁴⁶ <http://www.gbif.se/blog/article/samlingshanteringssystem>

⁴⁷ <http://www.svenskalifewatch.se/sv/om-lifewatch/innehall/dina/>

⁴⁸ <http://www.nrm.se/forskningochsamlingar/forskningsavdelningen/biodiversitetsinformatik/pagaendeprojekt.13704.html>

ARKIS är ett arkivinformationssystem som används inom Riksarkivet. Standarder och format som används för leverans av data till Arkis är EAD (Encoded Archival Description) och EAC (Encoded Archival Context) i XML-scheman.

Visual Arkiv är ett arkivinformationssystem som hanterar olika processer, alltifrån registrering till att skriva in förteckningar, skapa etiketter samt söka bland arkivförteckningarna. Systemet ger möjlighet till publicering av arkivinformationen i en sökportal samt möjlighet för användare att länka till sin information från NAD.⁴⁹

Klara är ett arkivinformationssystem som stödjer bland annat registrering, klassificeringsstrukturer, arkivbeskrivningar, arkivförteckningar och utskrifter av etiketter. Systemet följer riktlinjerna i Riksarkivets författningar samt har sökfunktioner och exportfunktionalitet i formaten EAD, EAC och NADMarc.⁵⁰

AdLib (f.d. CALM)⁵¹ är ett system som används av framförallt arkiv, men som även kan användas till vissa andra samlingar, som hembygdssamlingar eller museisamlingar.

Bibliotekssystem

De bibliotekssystem som används i Sverige idag är bland annat följande:

LIBRIS är både en gemensam webbaserad katalog och en nationell söktjänst med information om titlar på svenska universitets-, högskole- och forskningsbibliotek samt folkbibliotek.

Micromarc är ett bibliotekssystem som används av flera bibliotek i Sverige och som finns i flera versioner.⁵²

KOHA är ett katalogiseringssystem som ännu inte har börjat användas i någon större utsträckning på svenska marknaden men som kan vara värt att nämna. Det är baserat på öppen programvara.⁵³

ALVIN (Archives & Libraries Virtual Image Network) är ett system som används gemensamt av flera universitetsbibliotek. Det utvecklas under 2010-2011. En utgångspunkt är att alla komponenter, om möjligt, skall vara open source-programvara.

Fornminnessystem

Inom ramen för nyligen initierade DAP-programmet (Digital Arkeologisk Process) kommer Riksantikvarieämbetet att bland annat samordna fornminnesinformationen tillsammans med länsstyrelserna, vidareutveckla infrastrukturen för fornminnesinformation nationellt och internationellt samt skapa lösningar för ett långsiktigt och säkert bevarande av digital fältdokumentation såsom rapporter, bilder, analysresultat och GIS-data. Syftet är även att verka för att all information ska vara fri att vidareutnyttja samt att utveckla tekniska gränssnitt tillsammans

⁴⁹ <http://www.visualarkiv.se/xtf/view?docId=om.html.xml>

⁵⁰ <http://www.knowit.se/klara> samt

http://www.knowit.se/Global/Knowit%20SE/Aff%c3%a4rsomr%c3%a5den/Prodblad_Knowit_Klara2.pdf

⁵¹ <http://www.axiell.se/adlib>

⁵² <http://www.bibliotekscentrum.se/1.0.1.0/2/1/>

⁵³ <https://joinup.ec.europa.eu/community/osor/news/more-libraries-switch-koha-catalogue-system>

med andra aktörer som gör att de olika verksamheternas system kan utbyta information och därefter eventuellt kunna registrera uppgifter i gemensamma system.

Digitala infrastrukturer

Begreppet digital infrastruktur betyder i de flesta sammanhang idag mer än internetuppkoppling och används även ofta när man talar om databassystem och/eller aggregeringsarkitektur.

Digital infrastructures – both physical- and service-based – are key enablers for the smart growth that Europe must achieve in the coming ten years in order to ensure its ability to compete internationally and generate wealth for its citizens. Indeed, this initiative is anchored in the Europe 2020 Strategy for smart, sustainable, and inclusive growth, which put digital infrastructures at the forefront as part of the flagship initiative, “Digital Agenda for Europe”.⁵⁴

Digitala infrastrukturer avser enligt denna definition både teknisk digital infrastruktur (internetuppkoppling, kapacitet) och tjänster kopplade till informationshantering, som till exempel grid-tjänster, datalagring och säkerhet. Utöver dessa tjänster finns det även andra tjänster som är av hög relevans för kulturarvssektorn, exempelvis bevarandetjänster eller en metadataväxel som skulle kunna användas på ett distribuerat sätt med hjälp av digitala infrastrukturer. Tjänster och verktyg kan finnas tillgängliga online via den tekniska infrastrukturen och underlätta arbetet med den digitala kulturarvsinformationens flöde genom att:

- implementera verktyg som kopplar samman databaser/system med olika tjänster (till exempel för registrering av auktoriteter, bevarandeverktyg, med mera).
- använda gemensamma tjänster för att strukturera eller aggregera information (till exempel en metadataväxel och semantifiering).
- använda lagringslösningar genom de digitala infrastrukturerna.
- använda gemensamma verktyg och tjänster för skapande av virtuella utställningar, appar, utbildningar, med mera.

Under modelleringsarbetet har det blivit tydligt att det idag inte existerar en sammanhållande infrastruktur för kulturarvsinformation som innehåller alla nödvändiga aktörer, delar och funktioner. Det har också blivit tydligt att det inte är aktuellt att sikta mot en enda sammanhängande infrastruktur för det digitala kulturarvet. I dagsläget arbetar vi i stället utifrån målet att med flexibla och skalbara lösningar koppla samman befintliga strukturer till en evolverande infrastruktur. Målet blir då att till exempel identifiera och implementera verktyg och tjänster som stödjer interoperabilitet och harmonisering av data. Hypotesen som resonemanget bygger på är att i en snabbt föränderlig värld leder ett sådant förhållningssätt till högre kostnadseffektivitet och kontinuerligt anpassningsbara lösningar.

⁵⁴ An extract from: European Commission Brussels, 19.10.2011 Com(2011) 676,

http://www.riksdagen.se/sv/Dokument-Lagar/EU/EU-forslag/Communication-from-the-Commiss_GZB6676/

SUNET är den självklara e-infrastrukturen till vilken flertalet statliga kulturarvsmyndigheter och kulturarvsinstitutioner redan är uppkopplade. Genom en djupare samverkan mellan kulturarvsmyndigheterna och SUNET skulle förutsättningarna för att utveckla och vidmakthålla en sådan infrastruktur kunna realiseras. Man borde även titta närmare på det nationella respektive europeiska Grid⁵⁵-samarbetet, SweGrid respektive EGI.

Detta skulle också kunna främja samordnade processer för hantering av digitalt kulturarv som förvaltas på regional och kommunal nivå, genom exempelvis universitetsbiblioteken eller landsarkiven som är naturliga anslutningspunkter. Landsarkiven är avdelningar vid myndigheten Riksarkivet och är som sådana redan uppkopplade mot SUNET. Landsarkiven spelar också en viktig roll i ABMK-samverkan (arkiv, bibliotek, museer och kulturmiljö) på regional nivå idag och skulle kunna ha ett utvecklat ansvar för utveckling, stöd och support i det regionala arbetet med digitalisering av kulturarvet. På så sätt skulle infrastruktur, rekommendationer, vägledningar, med mera kunna nå ut till alla som deltar i arbetet med digitalisering, användning och bevarande av vårt gemensamma digitala kulturarv.

Skalfördelarna i stora sammanlänkade digitala infrastrukturer skulle därigenom öka kostnadseffektivitet i åtkomst till digitalt kulturarv för forskning och allmänheten samt tillgång till gemensamma tjänster för kulturarvsinstitutioner. Under de senaste åren har en märkbart ökad dialog och samverkan mellan infrastrukturerna för forskning och för det digitala kulturarvet kunnat märkas. I Sverige har deltagandet i EU-projektet DC-NET lett till ett långt framskridet samtal om samverkan mellan SUNET och kulturarvsinstitutionerna representerade av Riksarkivet/Digisam. För närvarande pågår EU-projektet DCH-RP (Digital Cultural Heritage Roadmap for Preservation) där Riksarkivet/Digisam deltar med målet att på nationell nivå kunna implementera den vägkarta för digitalt bevarande som projektet kommer att resultera i. Inom ramen för DCH-RP utvärderas möjligheterna att kunna tillämpa bevarandetjänster tillgängliga via e-infrastrukturer och GRID-nätverk, bland annat genom användning av dataset från ett antal av Digisams medverkande myndigheter och institutioner.⁵⁶

Mycket skulle också kunna vinnas genom en samutnyttjad förvaltning av de stora forskningsinfrastrukturernas lager av forskningsinformation och kulturarvets lager av digitaliserad kulturarvsinformation. Det senare rymmer digitaliserade samlingar men även digital forskningsinformation som förvaltas av kulturarvsinstitutionerna. Samlat utgör denna information en betydande resurs för forskning men även för bred publik användning och återanvändning där åtkomst och användning av informationen kan styras av behörighet och licensiering. En gemensam infrastruktur skulle rimligtvis vara avsevärt mer kostnadseffektiv än två parallella och sannolikt även leda till ökad forskning.

I en samverkan mellan infrastrukturerna för forskningsinformation och digitalt kulturarv skulle också tjänster och verktyg utvecklade för respektive område till stor del kunna samutnyttjas.

Teknisk infrastruktur

Som redan nämnts, är de flesta statliga kulturarvsinstitutioner idag uppkopplade mot internet genom universitetens och högskolornas gemensamma organisation och infrastruktur för nationell och

⁵⁵ <http://www.egi.eu/about/faq/>

⁵⁶ <http://digisam.se/index.php/hem/entry/roadmap-foer-langsigtigt-bevarande-tas-fram>

internationell datakommunikation, SUNET. Inom ramen för det EU-finansierade projektet DC-net genomförde Riksarkivet i september 2010 en enkät om digitala infrastrukturer för kulturarvsinstitutioner i Sverige. Enkäten inriktades på de institutioner som var uppkopplade på andra sätt än genom SUNET. Resultatet från enkäten pekade på att kapaciteten i forskarvärldens infrastruktur (SUNET) för att hantera informationen i vissa fall är påtagligt mycket större än kulturarvsinstitutionernas kapacitet att tillhandahålla den. Det förekom stora variationer i fråga om bandbredd och mer än hälften av de tillfrågade institutionerna ansåg att den digitala infrastruktur som de har inte täcker deras behov. Bland annat lyfte många digital lagring (och ett nationellt grepp vad gäller lagring) som en funktion som saknas. Annat som togs upp är infrastrukturella tjänster som omfattar/möjliggör tillgängliggörande samt digitala arbetsytor.

Situationen kompliceras ytterligare av att vissa kulturarvsinstitutioner utgör en del av forskningsvärlden. Ansvaret för den forskning som kulturarvsinstitutionerna utför är också splittrat på ett flertal departement istället för samlat som oftast är fallet i våra grannländer. Resultatet blir att ingen egentligen får en tillräcklig informationsförsörjning.

Gemensamma stödfunktioner

Under modelleringsarbetet kring digitala infrastrukturer som Digisam genomförde uttrycktes ett starkt behov av stödfunktioner kopplade till de tekniska infrastrukturerna, dels i form av råd och stöd och dels i form av tekniska verktyg.

Flera olika verktygskataloger har skapats inom ramen för olika projekt och initiativ, ibland tematiska som till exempel kring bevarandet inom DCH-RP⁵⁷, textdigitalisering inom SUCCEED-projektet⁵⁸ eller som resultat av en nätverksgemenskaps intresse, som t.ex. COPTR.⁵⁹ Inom ramen för denna rapport tas de stödfunktioner upp som under de inledande samtal som Digisam har haft med medverkande institutioner har pekats ut som mest önskvärda.

De stödfunktioner som lyftes fram är fokuserade på:

- lagring och bevarande
- mappning (metadataväxel) inklusive semantiska stödfunktioner
- beständiga identifierare
- semantiska stödfunktioner för auktoritetsfiler

Utöver dessa funktioner har Digisam också tittat på:

- georeferenser
- UGC

⁵⁷ <http://www.dch-rp.eu/index.php?en/137/registry-of-services-tools>

⁵⁸ <http://www.digitisation.eu/tools-survey/index-coc>

⁵⁹ http://coptr.digipres.org/Main_Page

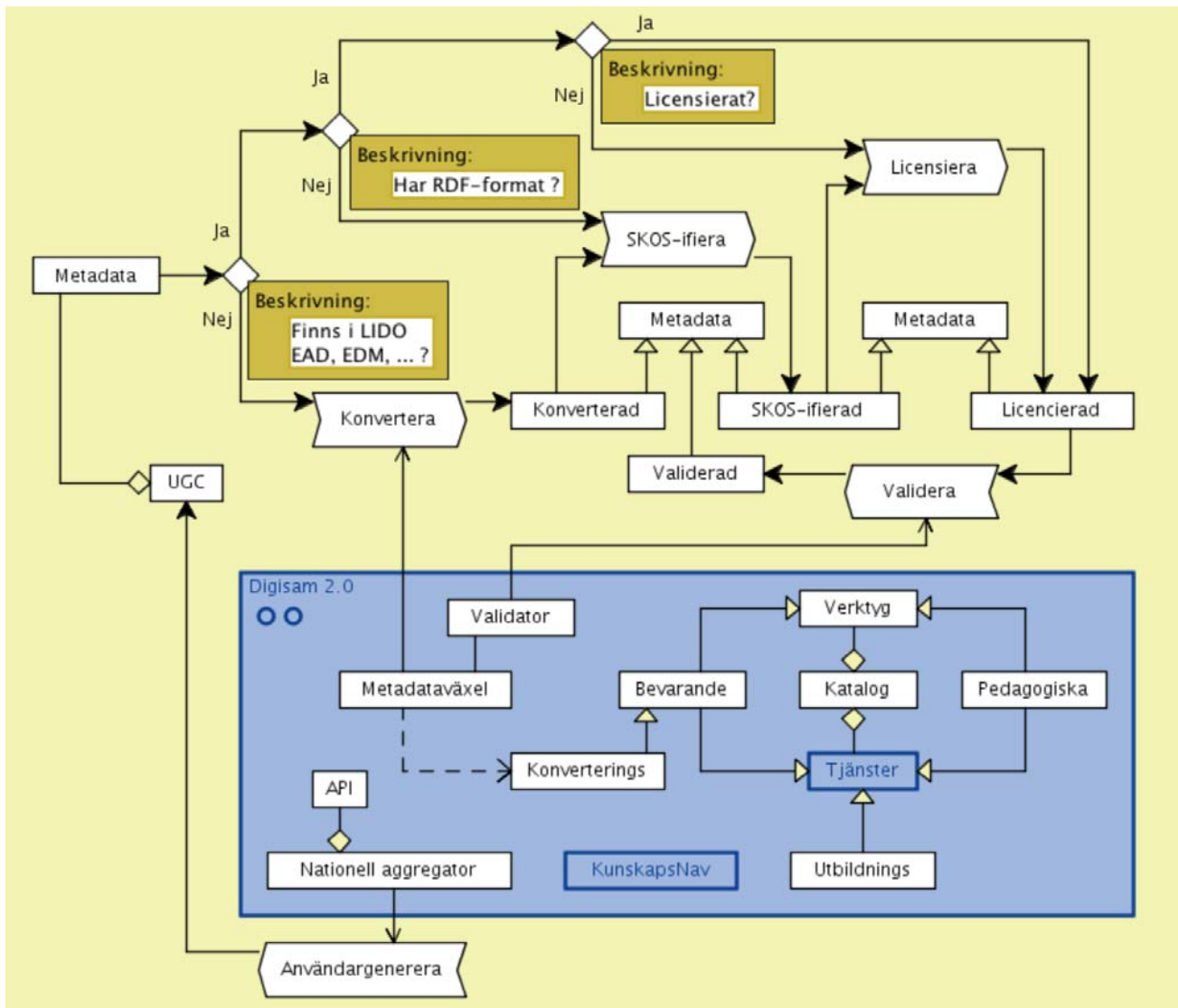


Bild 7. Stödfunktioner.

Ovan redovisas några exempel på pågående initiativ kring gemensamma tjänster och/eller stödfunktioner.

Stödfunktioner för lagring och bevarande

Lagring och bevarande är centrala tjänster inom ramen för infrastruktur för kulturarvsinformation. I Europeiska Unionens Rådslutsatser från 2008 uppmanas medlemsländerna att ”uppmuntra och samordna arbetet med att förbättra det långsiktiga digitala bevarandet och interoperabiliteten mellan systemen för tillgång till innehållet, bland annat genom att utbyta god praxis och sprida standarder, samt stödja forskning om och utveckling av innovativ teknik i syfte att särskilt få till stånd ett långsiktigt bevarande av digitaliserat kulturellt material, en flerspråkig navigering i och tillgänglighet till innehållet och en utveckling av nya användningar”.⁶⁰

Se vidare: Rapport Digitalt bevarande vid kulturarvsinstitutioner⁶¹.

⁶⁰ [http://eur-lex.europa.eu/legal-content/SV/TXT/PDF/?uri=CELEX:52008XG1213\(04\)&from=SV](http://eur-lex.europa.eu/legal-content/SV/TXT/PDF/?uri=CELEX:52008XG1213(04)&from=SV)

⁶¹ <http://digisam.se/index.php/docman/public-documents/digitaltbevarande/40--8/file>

Gemensamma mappningsfunktioner (metadataväxel)

I den europeiska MINT-aggregeringsplattformen finns en metadataväxel-funktion för att stödja skördning och mappning av metadata från innehållsleverantörernas modell till en referensmodell (LIDO) samt transformering av poster från LIDO till Europeana Data Modell (EDM).

CultureCloud är ett regionalt projekt som för närvarande testas inom ramen för ett regionalt pilotprojekt där Riksarkivet utvärderar möjligheterna att stödja innehållsleverantörer från arkivsektorn med gemensamma distribuerade funktioner för bland annat mappning av metadata på en regional, nationell och europeisk nivå. Arkitekturen kommer att vara baserad på CIDOC CRM, FRBRoo, APE EAD, med flera.

B2SHARE är en tjänst som har utvecklats inom ramen för EUDAT-projektet och som har som syfte att stödja synlighet och sökbarhet av informationen som lagras digitalt. I B2SHARE finns mappningsfunktioner och persistenta identifierare, anpassade till internationella standarder. Det finns möjlighet att hantera informationen under själva överföringen och efteråt. Man behöver inte ladda ner programvara utan tjänsten kan nås online. B2Share testas för närvarande av flera svenska kulturarvsinstitutioner. Testerna kommer att bli klara i augusti 2014.

Karma är ett verktyg med öppen källkod som gör det möjligt för användare att snabbt och enkelt integrera data från olika datakällor, inklusive databaser, kalkylblad, avgränsade textfiler, XML, JSON, KML och Web API. Användare kan integrera information genom att modellera den enligt en ontologi som de själva väljer, med hjälp av automatiserade processer. Verktøget kan automatiskt generera en ontologimodell som användare sedan kan justera och därpå publicera som RDF och/eller lagra i en databas.⁶²

Europeana Connection Kit (ECK) är ett verktyg som utvecklas inom EU-projektet där man identifierar befintliga arbetsflöden, standarder och verktyg för att kartlägga vilka funktioner som kan förenkla aggregering för institutioner för att kunna utveckla ett verktyg som stödjer dessa. Flera företag deltar i utvecklingen av verktyget, bland annat Museum Plus, där det bland annat ska finnas en funktion för metadatamappning.

Beständiga identifierare

I det europeiska samarbetet Athena-projektet togs det fram rekommendationer för beständiga identifierare inom kulturarvssektorn⁶³, enligt vilka kulturarvsinstitutionerna bör välja en infrastruktur för beständiga identifierare med hjälp av följande krav på systemet för beständiga identifierare:

- Tekniskt tillförlitligt (bland annat bör det finnas säkerhetskopiering med redundant teknik, registret över identifierare bör uppdateras automatiskt, säkerheten – trust/risk).

⁶² <http://www.isi.edu/integration/karma/>

⁶³ <http://www.athenaeurope.org/getFile.php?id=779>

- Auktoritativt (auktoritativ och trovärdig organisation som hanterar systemet, tilldelar identifierare och resolvers till resurser).
- Flexibelt (systemet bör vara tillräckligt flexibelt för att representera granularitet hos beskrivna metadata).
- Interoperabelt (systemet bör tillämpa öppna standarder).

Athena-projektets rekommendationer uppdaterades sedan i Linked Heritage-projektet i ”State of the art report on persistent identifier standards and management tools”.⁶⁴

Digital Preservation Europe (DPE)⁶⁵ rekommenderar att man följer dessa parametrar och att man utöver det också ställer krav på globalt unika identifierare, beständighet samt resolver-funktionalitet.

Vid tillfället då Athena-projektets rapport skrevs fanns det inga vedertagna internationella standarder för museiinformation, dock för biblioteksinformation, vilket påpekas i rapporten. Libris använder exempelvis sådana standarder för bibliografisk information. Det har funnits ett större europeiskt samarbete kring beständiga identifierare, PersID, där Kungliga biblioteket har deltagit som svensk representant⁶⁶. Syftet med projektet var att skapa en infrastruktur för beständiga identifierare. Projektet valde att inkludera de institutioner som redan hade använt URN:NBN "Uniform Resource Names for National Bibliography Numbers"⁶⁷ under en tid och därmed kunde bidra med erfarenheter och kunskap. Intentionen var att, utifrån de specifika erfarenheterna, arbeta fram något som skulle kunna gälla generellt och att hela tiden vara öppen för andra standarder. En prototyp för en meta-resolver tjänst togs fram inom ramen för projektet.⁶⁸

Inom ramen för ARIADNE-projektet kartläggs de olika formella standarderna för beständiga identifierare⁶⁹:

- the Uniform Resource Name (URN)
- the persistent URL (PURL); the Handle system;
- the digital object identifier (DOI);
- National Bibliography Numbers (NBNs);
- the Archival Resource Key (ARK);
- the Open URL.

Rekommendationer för utformning av URI:er från Europeiska kommissionen beskrivs i “Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC”, Europeiska kommissionen, 2012⁷⁰. Enligt denna rekommendation ska formen för

⁶⁴ <http://www.linkedheritage.eu/index.php?en/142/documents-and-deliverables>

⁶⁵ http://www.digitalpreservationeurope.eu/publications/briefs/persistent_identifiers.pdf

⁶⁶ <http://www.persid.org/>

⁶⁷ <http://www.kb.se/isbn-centralen/urnnbn/>

⁶⁸ http://www.persid.org/downloads/finalreports/PersID_Report_Part_4_final.pdf

⁶⁹ http://www.ariadne.ac.uk/issue56/tonkin#What_Is_a_Persistent_Identifier_and_Why

⁷⁰ <https://joinup.ec.europa.eu/sites/default/files/D7.1.3%20-%20Study%20on%20persistent%20URIs.pdf>

URI:er utformas enligt följande: "http://{domän}/{typ}/{koncept}/{referens}". Enligt samma rekommendation ska identifierare hos resurser som redan är unikt refererade återanvändas. Det finns flera rekommenderade resolver-tjänster, som till exempel Handle-systemet⁷¹. Kungliga biblioteket har en resolver-tjänst för URN:NBN som även är förberedd för DOI och Handle.

En sådan tjänst syftar till att dirigera webbläsaren till en lämplig eller aktuell kopia av objektet. Detta är en viktig strategi för att minska risken för felaktiga länkar genom att lägga ett lager mellan webbläsaren och målobjektet som tillåter webbläsaren att hitta rätt objekt och som vidarebefordrar till en aktuell kopia av objektet. Resolver-tjänsten kan vara en uppsättning omskrivningsregler på en server eller ett globalt nätverk av särskilda servrar.

Semantiska stödfunktioner för auktoritetsfiler

Beständiga identifierare är en förutsättning för långsiktighet i semantiska länkar. Det pågår för närvarande flera initiativ kring att skapa olika system som kan förvalta och hantera gemensamma auktoriteter. Några exempel på dessa är KulturNav och TMP.

KulturNav är en webbplats som förvaltas av KulturIT (för närvarande i Beta-version) där man ska kunna skapa, förvalta och distribuera gemensamma terminologier och auktoriteter med fokus på museer och andra kulturarvsinstitutioner.⁷²

Den Terminology Management Platform (TMP) som har börjat utvecklas inom ramen för Linked Heritage-projektet, och som numera vidareutvecklas i Athena Plus-projektet, har som syfte att tillhandahålla referensterminologi för kulturarvssektorn som lagras i SKOS RDF samt verktyg för semantisk koppling mellan terminologier och metadata. TMP är för närvarande i Beta-version och målet är att den ska kunna fungera som en webbtjänst för online-delning av terminologier, semantisk berikning och länkning av metadata till externa datakällor.

DISKA (digitala semantiska kulturarvsauktoriteter) är ett projekt som Digisam under 2013 bedrev i samarbete med KTH/ECE-skolan och med stöd från Vinnova. Projektet hade som mål att inventera listor/förteckningar av referenskaraktär på de centrala kulturarvsmyndigheterna, omarbeta ett strategiskt urval sådana "dataset" till RDF och publicera dem som länkbar öppen data⁷³.

Tjänster och verktyg för utbildningar

I undervisningen finns stort behov av nya innovativa former för användning av digitalt material. Den digitala utvecklingen har skapat en situation där det formella och informella lärandet kan samverka i nya former. Att möta dessa utmaningar förutsätter en ny form av strategisk kunskapsinfrastruktur, vilken stödjer den pågående utvecklingen där information i enskilda databaser länkas samman till en semantisk infrastruktur. Därigenom förstärks spridningen och användning av kulturarvsdata.

Kulturarvsinformation är tvärspektoriell samhällsinformation med både bredd och djup. Det ger rika kombinations- och användningsmöjligheter och lämpar sig väl för webbaserade utbildningar. Den

⁷¹ http://en.wikipedia.org/wiki/Handle_System

⁷² <http://kulturnav.org/>

⁷³ <http://diska.ece.kth.se/>

växande mängden digital kulturarvsinformation kan utnyttjas som en laborativ plattform vid utveckling av teknik och metoder för lärande och undervisning på alla nivåer samt i det livslånga lärandet. I europeiska sammanhang har flera projekt tittat på detta. Digikult-projektet har 2003 gjort en sammanställning över alla webbaserade utbildningar inom kulturarvs-/kulturvetenskapliga området, "Learning Objects from Cultural and Scientific Heritage Resources"⁷⁴ och idag har den listan blivit ännu längre. Inom ramen för bland annat Linked Heritage-projektet togs (år 2013) en del utbildningar fram om hanteringen av det digitala kulturarvet, speciellt riktade mot verksamma inom kulturarvsinstitutioner: "Linked Heritage Training Program/Learning Objects".⁷⁵ LEM-projektet "The Learning Museum Network Project"⁷⁶ är inriktat speciellt på museer. Cultura-projektet "Cultivating Understanding and Research Through Adaptivity"⁷⁷ tittar särskilt på användarnas upplevelse av det digitala kulturarvet och har utvecklat en stödfunktion för användare, en så kallad "Virtual Research Environment".

TElMap-projektet, där KTH deltar, är inriktat på den senaste tekniska utvecklingen och huruvida olika utbildningslösningar kan tas fram med nya tekniska möjligheter.⁷⁸

Tjänster och verktyg för digitala utställningar

En arbetsgrupp om virtuella utställningar har startat inom ramen för Linked Heritage-projektet, och numera även Athena Plus-projektet som kommer att undersöka hur man kan förbättra kvaliteten på metadata i virtuella utställningar samt hur dessa kan dokumenteras på lämpligt sätt. Movio-verktyget som är en öppen programvara kommer inom projektet att anpassas till kulturarvsinstitutionernas behov. När en utställning har tagits fram i Movio som webbapplikation går det till exempel att tillverka en app utifrån denna. Inom ramen för Athena Plus deltar Livrustkammaren och Nationalmuseum, med stöd av Digisam.

Inom ramen för EU-projektet Indicate⁷⁹ har en handbok tagits fram som beskriver hur digitala utställningar kan skapas av kulturarvsinstitutioner, "Handbook on virtual exhibitions and virtual performances".⁸⁰

Geografiska referenser

Geografiska referenser och tjänster kopplade till dessa är viktiga i beskrivningar av digital kulturarvsinformation.

Lantmäteriet ansvarar för de nationella geografiska referenssystemen. Dessa är SWEREF 99 (en anpassning till det europeiska ETRS 89) i tre dimensioner och i plan, RH 2000 i höjd samt RG 82 för tyngdkraft, samt den nationella geoidmodellen SWEN08_RH2000 för att säkerställa höjdmätning med GNSS.⁸¹

⁷⁴ http://www.digicult.info/downloads/digicult_thematicissue4_lres.pdf

⁷⁵ <http://linkedheritage.cab.unipd.it/training/LO-00/en/overview.html>

⁷⁶ <http://www.lemproject.eu/>

⁷⁷ <http://www.cultura-strep.eu/home>

⁷⁸ <http://learningfrontiers.eu/>

⁷⁹ <http://www.indicate-project.org/>

⁸⁰ <http://www.indicate-project.org/index.php?en/144/publications>

⁸¹ <http://www.lantmateriet.se/Kartor-och-geografisk-information/GPS-och-geodetisk-matning/Referenssystem/>

För närvarande etableras det en infrastruktur för "geospatial" information i Europa INSPIRE (EU-direktivet 2007/2/EC)⁸², med en Geo-portal⁸³ och ett register.⁸⁴

En kartläggning av GIS-system, koordinatsystem, etcetera, som är relevanta för kulturarvssektorn redovisas i publikationen "Geocoded Digital Cultural Content" som har tagits fram i EU-projektet Linked Heritage där Digisam deltog som partner⁸⁵. Inom ramen för denna kartläggning görs referenser till olika projekt och initiativ som utforskar dessa områden. Linked Heritage-projektets arbete grundas i det tidigare ATHENA-projektet där man publicerade "Guidelines for Geographic Information".⁸⁶ Vad gäller naturhistoriskt material har Gbif publicerat riktlinjer för hur man kan ange georeferenser för biologiska artdata.⁸⁷

Sammanlänknings av georeferenser genom länkad öppen data är också ett av huvudtemana i EU-projektet GeoKnow, "Making the web an exploratory place for geospatial data".⁸⁸

User Generated Content

Inom Europeana Inside finns det möjlighet att få tillbaka UGC-data till databaser, eventuellt genom delvis automatiserade processer, som en del av ECK verktyget.

Arbetsgrupper kring UGC finns både inom ramen för K-samsök och inom ramen för Digisams EU-nätverksgrupp. Digisams EU-nätverksgrupp handlar om samverkan och kunskapsutbyte för de institutioner som deltar i olika EU-projekt med bäring på digitalt kulturarv. Bakgrunden till nätverket är dels Digisams uppdrag, som bland annat syftar till att öka de svenska ABMK-institutionernas deltagande i relevanta EU-projekt, och dels den Europeana-krets som tidigare fanns samlad och där de svenska deltagarna i Europeana-relaterade projekt deltog. Syftet är att diskutera strategiska frågor både ur ett nationellt och europeiskt perspektiv och lägga grunden för framtida projekt.

I olika publika sökapplikationer som rör information från kulturarvsinstitutioner, exempelvis i Digitalt museum där man kan samsöka information från PRIMUS-an slutna museer, finns det också möjligheter för användare att lägga till information. För närvarande innebär sådan information i de flesta fall att det ställs krav på institutionerna att utveckla processer för att föra tillbaka utvalda delar av denna information till databaser. Detta sker oftast manuellt.

Bevarande

Digisam har genomfört en förstudie, bestående av intervjuer med sju av Digisams medverkande kulturarvsinstitutioner, arbete i arbetsgrupper och en jämförande analys av hur det digitala bevarandet ser ut på nationell och europeisk nivå. Resultaten från förstudien finns samlade i en

⁸² <http://inspire.jrc.ec.europa.eu/>

⁸³ <http://inspire-geoportal.ec.europa.eu/>

⁸⁴ <http://inspire.ec.europa.eu/registry/>

⁸⁵ <http://www.linkedheritage.org/getFile.php?id=482>

⁸⁶ <http://www.athenaeurope.org/getFile.php?id=787>

⁸⁷ <http://www.gbif.se/blog/article/guide-till-baesta-saettet-att-anvaenda-georeferenser>

⁸⁸ <http://geoknow.eu/Welcome.html>

nypublicerad rapport⁸⁹ som tar upp vilka behov som finns på området och hur 'state of the art' (dagens högsta nivå inom ett område) kan nås.

Resultaten visar bland annat att det saknas system anpassade för långtidsbevarande, samtidigt som datamängderna kommer att växa betydligt inom en snar framtid. De flesta intervjuade institutioner ser fördelar med en gemensam lagringslösning inom ABMK-sektorn och alla efterfrågar en gemensam stödfunktion. Idag saknas det även stöd för prioriteringar, rutiner och strategier kring lagring och bevarande; dock ses arbetet med de myndighets- och institutionsvisa planerna för digitalisering som ett steg på vägen.

I EU-projektet DCH-RP testade svenska kulturarvsinstitutioner gemensamma lagringstjänster inom det pågående arbetet med att ta fram en 'roadmap' för långsiktigt bevarande.

Digsam kommer framöver att ta fram:

- Ett förslag till en centralt förvaltd gemensam stödresurs bestående av stödfunktioner och samling av kompetenser och resurser kring bevarandefrågor.
- Ett förslag till en nationell roadmap för digitalt bevarande, baserat på bland annat resultaten av projektet DCH-RP och övrigt relevant underlagsmaterial.
- En utvärdering av för vilka verksamheter och situationer en centraliserad respektive distribuerad lagrings- och bevarandelösning för de statliga institutionernas kulturarvsinformation är att föredra och ett förslag till grundkrav för sådana lösningar.
- Förslag till former för utvecklad samverkan i lagrings- och bevarandefrågor mellan Riksarkivet, Kungliga biblioteket, Centralmuseerna och övriga statliga kulturarvsmyndigheter.

Projektet 4C (Collaboration to Clarify the Costs of Curation)⁹⁰ har i augusti 2014 publicerat sin Curation Cost Exchange⁹¹, som gör det enklare att dela med sig och få information om vad olika insatser kan kosta.

eARD

Projektet e-arkiv och e-diarium (eARD) har avslutats och det har inrättats en funktion vid Riksarkivet i juli 2014 för att förvalta och vidareutveckla FGS:erna.⁹²

eARK-projektet syftar till att ta fram en gemensam europeisk modell för e-arkivering. Flera nationalarkiv deltar, däribland Danmark, Norge och Estland. I projektet deltar också ett antal kommersiella företag och universitetsinstitutioner. Projektet är i högsta grad intressant för Sverige eftersom de lösningar som diskuteras har många beröringspunkter med de FGS:er som tagits fram

⁸⁹ <http://digsam.se/index.php/docman/public-documents/digitaltbevarande/40--8>

⁹⁰ <http://4cproject.eu/>

⁹¹ <http://www.curationexchange.org/>

⁹² <http://riksarkivet.se/fgs-earkiv>

inom eARD-projektet. Projektet har sedan förstudierapporten publicerat flera rapporter, bland annat kring best practices, format och tillgängliggörande.⁹³

Regeringen har i augusti 2014 gett Statens servicecenter, i samverkan med Riksarkivet, i uppdrag att dels utveckla en förvaltningsgemensam tjänst för e-arkiv, dels driva tjänsten samt erbjuda den till landets myndigheter. Den nya e-arkivtjänsten ska baseras på de FGS:er som Riksarkivet förvaltar. Denna nya tjänst kommer till en början att implementeras hos ett antal pilotmyndigheter och blir då ett första exempel på en förvaltningsgemensam lagringslösning.⁹⁴

Samverkan med forskningssektorn

Samverkan mellan KTH och kulturarvsområdet har ursprungligen påbörjats mellan Digisam och ECE-skolan. ECE-skolan har en ledande roll i den pedagogiska utvecklingen, forskningen och lärandet vid KTH. Skolan bedriver egen utbildning och forskning och ger stöd för såväl utbildning som forskning inom de andra skolorna vid KTH. Bakgrunden till samarbetet är idén om kulturarvssektorns samarbete med en tekniskt kvalificerad part. Samarbetet påbörjades med en avsiktsförklaring, ett "Letter of Intent", som skrevs under mellan Riksarkivet, Riksantikvarieämbetet, Kungliga biblioteket, Centralmuseernas samarbetsråd och KTH. Digisam fick i uppdrag att representera kulturarvinstitutionerna i det fortsatta planeringsarbetet. Vidare anordnades ett seminarium. Grundtanken med initiativet är att digitaliserat material som skapats av kulturarvsinstitutionerna inom ABM-sektorn omformas till pedagogiska instrument av forskare vid KTH för undervisning på olika nivåer. Utgångspunkten är att tillgängliggöra digitaliserat kulturarvsdata för undervisning men också för annan användning. Samarbetet har så småningom utökats till att sträcka sig längre än endast till lärandeperspektivet och har blivit en plattform för ömsesidig nytta. För KTH är nyttan forskning och utbildning och för kulturarvsinstitutionerna tillförs kompetens inom teknik och lärande. Särskilt kan forskning och utveckling kring kunskapskartografi, semantik och länkad öppen data nämnas som fokusområden där kulturarvsinstitutionerna, Digisam och KTH har gemensamma starka intressen och mycket att vinna på en utvecklad och fördjupad samverkan. Tillsammans ökar vi också attraktiviteten och relevansen för deltagande i EU- och internationella projekt vilket leder till kompetenshöjning och möjligheter till extern finansiering. KTH är också tillsammans med SUNET en viktig part i bryggan mellan nationell e-infrastruktur, det internationella GRID-nätverket och internet som helhet.

I det snabbt utvecklade och föränderliga webbaserade informationssamhället går frågorna horisontellt och vertikalt på samma gång och det finns inga skarpa gränser mellan teknik, humaniora och samhällsvetenskap. Ska vi nå såväl gemensamma som domänspecifika mål krävs samverkan. Inte minst gäller det gäller frågor om kulturarv och lärande inom kultur- och utbildningsdepartementens ansvarsområden.

På europeisk nivå finns det idag ett övergripande initiativ som handlar om att skapa en gemensam och tvärdisciplinär datainfrastruktur för forskningsdata, som med hjälp av olika tjänster underlättar

⁹³ <http://www.eark-project.com/resources/project-deliverables>

⁹⁴ <http://www.statenssc.se/Sidor/Nyheter/Pressmeddelanden/Gemensamt-e-arkiv-ger-samhället-miljonvinster.aspx>

hantering av stora datamängder. Projektet kallas EUDAT och dess konceptuella modell av datainfrastruktur integrerar flera infrastrukturer med stora mängder forskningsdata, tjänster för lagring och bevarande samt tillgängliggörande och användargränssnitt (se bild 8 nedan).

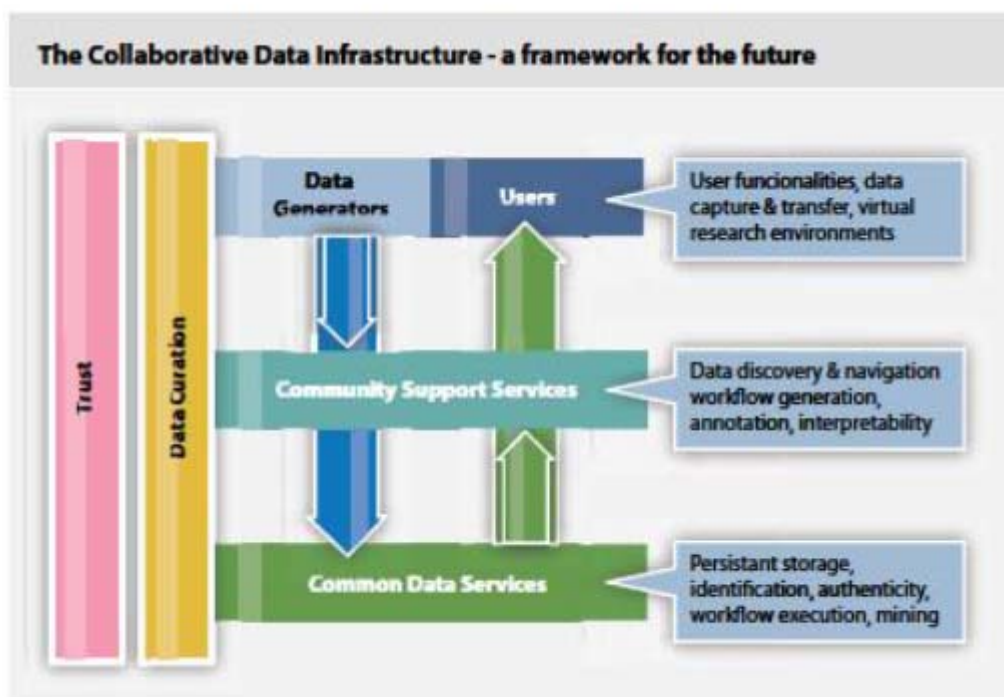


Bild 8. The collaborative data infrastructure - a framework for the future⁹⁵.

OpenAIRE (Open Access Infrastructure for Research in Europe)⁹⁶ är ett treårigt projekt som finansieras inom EU:s sjunde ramprogram för forskning och utveckling. Projektet ska underlätta för forskarna att uppfylla EU:s Open Access-krav genom ett nätverk av nationella projektkontor och en särskild OpenAIRE portal. Kungliga biblioteket är svensk nod i samarbetet.⁹⁷ För närvarande pågår OpenAIREplus-projektet, som handlar om tillgängliggörande av underliggande forskningsdata till publikationer.

Svensk Nationell Datatjänst är en nationell infrastruktur för samordning av forskningsdatabaser inom samhällsvetenskap, humaniora och medicin. SND ingår som den svenska noden i ett internationellt nätverk av dataarkiv som är en del av forskningsinfrastrukturen och arbetar med data inom humaniora, främst inom historia och arkeologi. SND har sedan 2011 arbetat mer aktivt för att kunna ta emot, beskriva, lagra och tillgängliggöra data inom humaniora. Under de senaste åren har en rad historiska och arkeologiska data beskrivits och tillgängliggjorts via SND, varav den största deponeringen består av drygt 400 digitala arkeologiska fältdokumentationer. Förändringar och anpassningar som SND infört på senare tid öppnar möjligheter för att kunna ta emot, tillgängliggöra och synliggöra mer digitalt arkeologiskt material. SND har förutsättningar för långtidslagring, uppdatering och återanvändning av materialet för både forskningsändamål och andra ändamål.

⁹⁵ Från rapporten "Riding the Wave", p. 31, <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>

⁹⁶ <https://www.openaire.eu/>

⁹⁷ <http://www.kb.se/openaccess/nyheter/2010/KB-svensk-nod-i-OpenAIRE/>

Ett europeiskt perspektiv

Det finns idag representanter från Sverige i både EU-expertgrupper och EU-finansierade projekt där man satsar betydande resurser på forskning kring digitalisering, digital tillgänglighet och bevarande. Dessa samarbeten har utvecklats under längre tid och inte sällan har det utifrån större sammanhang också uppstått flera mindre tematiska projekt. Därför kan det ibland förefalla svårt att navigera bland alla projekt och se ”röda trådar” i tillkomsten av nya projekt, om man inte tittar på hur dessa hänger ihop med tidigare och/eller större samarbeten.

MINERVA (Ministerial network for valorising activities in digitisation), 2002-2005, var ett EU-projekt som också fick uppföljningar genom Minerva Plus och Minerva EC. Minerva var ett nätverk av medlemsstaternas kulturministerier med fokus på samverkan kring goda exempel inom digitalisering av kulturell och vetenskaplig information. Inom ramen för projektet togs det fram rekommendationer och riktlinjer om digitalisering, tillgänglighet och långsiktigt bevarande av digitalt material som har använts i stor utsträckning i europeiska digitaliseringsprojekt. Minerva-projekten hade också kopplingar till MSEG- expertgruppen (Member States Expert Group on Digitisation and Digital preservation).

MICHAEL (Multilingual Inventory of Cultural heritage in Europe) var ett fortsatt samarbete, som uppstod genom nätverkande och resultaten från Minerva-projekten, med syfte att tillgängliggöra det europeiska digitala kulturarvet genom en flerspråkig kulturell portal. Michael-projektet följdes upp med Michael Plus och idag finns Michael Culture Association som förvaltar nätverket och är ett resurscentrum där man skapar synergieffekter med andra nätverk, initierar arbetsgrupper, anordnar seminarier, med mera.

ATHENA-projektet⁹⁸, 2008-2011, även det med kopplingar till MINERVA-nätverket, var en förebild för aggregering av en kritisk massa innehåll till Europeana. Projektet utvärderade befintliga standarder och verktyg som användes av kulturarvsinstitutioner och utvecklade och integrerade LIDO-standard⁹⁹ och Athena Ingestion Server and Mapping Tool – MINT¹⁰⁰.

DC-NET-projektet¹⁰¹ hade som syfte att främja den digitala infrastrukturen för tillgång till forskningsinformation för kulturarvssektorn inom EU.

Europeana är EU-kommissionens initiativ som arbetar med att göra kulturarvet digitalt tillgängligt, så att alla kan ta del av det och engagera sig i det. Europeana startade med att EU-projektet EDLnet presenterade en prototyp till en portal. Portalen fick namnet Europeana och därefter startade projektet som gick under samma namn. Projektet har bland annat tagit fram en webbportal där vem som helst kan söka bland olika poster som rör kulturarvet. Detta tack vare att arkiv, bibliotek och museer i 29 länder bidrar med data. Europeana arbetar också för kulturarvssektorn genom att bland annat sprida och överföra kunskap, ta fram tjänster och opinionsbilda. Kulturarvssektorn, tillsammans med experter inom bland annat teknik och juridik, ingår i projektets nätverk.

⁹⁸ <http://www.athenaeurope.org/>

⁹⁹ <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido/>

¹⁰⁰ <http://mint.image.ece.ntua.gr/redmine/projects/mint/wiki>

¹⁰¹ <http://www.dc-net.org/>

Linked Heritage¹⁰² (2011-2013) kommer att tillföra Europeana 3 miljoner nya objekt från både offentlig och privat sektor; målsättningen är att höja kvaliteten på metadata för att förbättra sökning, åtkomst och användning av innehållet i Europeana. I konsortiet som driver projektet ingår alla viktigare intressenter från 20 av EU:s medlemsländer, men även från Israel och Ryssland som har särskilda samarbetsavtal med EU. Projektet koordineras av Kulturministeriet i Italien, genom ICCU. Riksarkivet/Digisam deltar i projektet.

Athena Plus¹⁰³ (2013-2015), Access to cultural heritage networks for Europeana, är ett så kallat CIP best practice nätverk som samlar 40 partners från 21 medlemsstater. De centrala intressenterna utgörs av ministerier och myndigheter samt museer, bibliotek, arkiv, ledande forskningscentra och små och medelstora företag. Athena Plus bygger på de goda resultat som uppnåddes i det tidigare Athena-projektet där LIDO och Athena Ingestion Server and Mapping Tool (MINT) har utvecklats. Athena Plus har som mål att leverera stora mängder ny kulturarvsinformation till Europeana, förbättra sökningar, underlätta hämtning av material, terminologier, applikationer för turism och virtuella utställningar, experimentera med att förädla och återanvända metadata efter användarnas olika behov, med mera. Riksarkivet/Digisam deltar och kommer att aggregera data tillsammans med Nationalmuseum och Livrustkammaren och Skoklosters slott med Stiftelsen Hallwylska museet.

PREFORMA projektet,¹⁰⁴ där Riksarkivet och Kungliga biblioteket deltar, har fokus på problematiken kring implementering och validering av standarder för digitala filer. Projektet kommer att undersöka förutsättningar för kvalitet i syfte att etablera ett långsiktigt hållbart ekosystem runt utvecklade verktyg med olika intressentgrupper.

CIVIC Epistemologies är ett europeiskt projekt som kommer att utveckla och validera en roadmap för e-infrastrukturer för att stödja Europas medborgares deltagande i vetenskaplig forskning, med fokus på de möjligheter som öppen tillgång till information, öppen källkod och öppen innovation öppnar för digitalt kulturarv och digitala humaniora. CIVIC Epistemologies har som syfte att stärka praktiska sätt att öppna möjligheter som leder till ekonomiska och sociala fördelar. Projektet kommer också att göra en analys av de behov som forskare, medborgare och forsknings- och kulturarvsinstitutionerna har och ta fram en road map för hur man når dit.

ARIADNE (Advanced Research Infrastructure for Archaeological Dataset Networking in Europe) är ett EU-finansierat projekt. I detta projekt samarbetar SND med bl.a. Archaeology Data Service (ADS) vid University of York i Storbritannien. ARIADNE syftar till att föra samman befintliga forskningsdatainfrastrukturer så att forskare från olika länder kan använda europeiska arkeologiska digitala datamängder som tillsammans spänner över olika perioder, områden och regioner. Idag är dessa data svårtillgängliga och det finns en begränsad harmonisering av exempelvis metadatastandarder. Den integrerande verksamheten ska möjliggöra gränsöverskridande tillgång för forskare till datacenter, verktyg och vägledning, samt skapandet av nya webbaserade tjänster som bygger på gemensamma gränssnitt till databaser, tillgång till referensdatamängder och användning av ny teknik.

¹⁰² <http://www.linkedheritage.org/>

¹⁰³ <http://athenaplus.eu/>

¹⁰⁴ Preservation Formats for culture information/e-archives är ett projekt som är delfinansierat av Europeiska kommissionen, under FP7-ICT – programmet, och som är en s.k. förkommersiell upphandling. <http://www.preforma-project.eu/index.html>

Förutom ovanstående finns det ett antal organisationer/institutioner i Europa som har stort fokus på verksamhet som bidrar till state of the art, exempelvis DEN Foundation¹⁰⁵ (Digitaal Erfgoed Nederland/Digital Heritage Netherlands) som är ett nederländskt kunskapscentrum för digitalt kulturarv samt Collection Trust.

¹⁰⁵ <http://www.den.nl/english>

Diskussion och analys

Digital kulturarvsinformation och samordnad informationshantering

Samordnade insatser för storskalig digitalisering leder till att digital information som är sökbar, högkvalitativ, interoperabel och standardiserad skapas, vilket även lägger grunden för tillgängliggörande och bevarande av samma information. Digitalisering idag sker dock ofta i mindre skala och i projektform, vilket gör att det i många fall saknas kontinuitet i upparbetade processer för till exempel en materialtyp.

Under de senaste åren har digitaliseringsarbetet såväl i Sverige som internationellt handlat mycket om att uppnå en kritisk massa information på internet, bland annat för att kunna bygga innehållsrika portaler och tematiska applikationer som drar till sig användarnas intresse.

Aggregeringstjänster skördar idag en del av kulturarvinformationen (metadata) för att göra den synlig och användbar, men i de flesta fall saknas möjligheter att automatiskt få tillbaka kvalitetshöjningar som sker genom länknings och adderad metadata till institutionens egna system. Information från en institution finns då i de flesta fall i separata system, som exempelvis i institutionens eget samlingssystem, egen webbsida, systemets webbsida, de nationella/europeiska aggregeringstjänsternas plattformar, med mera. I många fall finns det då ingen koppling mellan olika digitala processer och dataset, från produktion till bevarande respektive tillgängliggörande. Idag finns det en tendens att se den semantiska webbens möjligheter och länkad öppen data som ett sätt att överbrygga dessa utmaningar, men det finns fortfarande vissa begränsningar för länkbar data inom kulturarvsområdet vilket i många fall leder till pragmatiska lösningar och ibland avsaknad av kvalitetskontroll. Sättet som kulturarvsinstitutionerna digitaliserar och hanterar sina digitala kulturarvsdata på är därför den viktigaste förutsättningen för att informationen från kulturarvsinstitutionerna ska kunna anpassas till att vara en del av länkad öppen data.

Ser man närmare på olika exempel på aggregering så visar det sig att det finns olika strategier för standardisering av data för överföring. I de allra flesta fall används ett så kallat exportformat. Exportformat för museiinformation är, i de flesta fall, något mer varierande och i högre grad egenutvecklade än vad som är fallet vid arkiv- och biblioteksinstitutioner.

Några kulturarvsinstitutioner använder olika typer av XML märkspråk¹⁰⁶ (till exempel EAD eller EDM) som exportformat som kan mappas till andra typer av XML/ RDF-format eller andra typer av maskinläsbara format. Andra institutioner använder också ett så kallat protokoll-format, vilket är en funktionalitet som tillåter sökningar genom en databas vid överföring, exempelvis Z39.50 eller OAI-PMH. I vissa fall kombineras dessa sätt genom att man kodar in mappningen i protokollet (exempelvis K-samsöksport). Fördelen med den sistnämnda metoden är att man kan göra inkrementella uppdateringar. Nackdelen är att det är resurskrävande och behöver göras om varje gång en uppdatering görs i antingen institutionernas system eller i mappningen. När länkbara maskinläsbara format som RDF används har man också behov av en databas för att lagra och söka

¹⁰⁶ <http://sv.wikipedia.org/wiki/XML>

igenom RDF-tripletter (subjekt-predikat-objekt). Ett vanligt sätt att hantera detta är att sätta upp en så kallad triplestore.¹⁰⁷ I YEAH-projektets handbok¹⁰⁸ konstaterar man att SQL-relationsdatabaser är problematiska då strukturen i databaserna gör dessa svårhanterbara i samband med triplestore.

Vikten av att använda sig av vedertagna standarder, metoder och format för det digitaliserade materialet framgår allt tydligare ju mer man fördjupar sig i frågan om förbättrad sökning och åtkomst. Man måste se den övergripande bilden; även om ett visst kulturarvsmaterial är säreget och en egen lösning för standarder eller format vid en första anblick verkar vara att föredra, så är det sällan att föredra i längden. Användning av vedertagna standarder gör det lättare att utveckla kopplingar till datamodeller och interoperabilitet samt ger möjlighet till att ta del av befintligt arbete med enhetligt standardiserat material som är i ständig utveckling. Allt detta främjar i längden största möjliga tillgång till och användningen av kulturarvsmaterial.

Som det redan har nämnts vore det önskvärt är att mappningen sker mot en interoperabel högsta gemensamma nivå. Naturligtvis är det grundläggande kravet att data ska vara interoperabel oavsett vilken kvalitet enskilda institutioner kan leverera. Högsta möjliga kvalitet är givetvis ett mål men inte en förutsättning. Därför är fördelen med CIDOC CRM att man redan från början kan (om man vill) få en enklare mappning som är kompatibel med eventuella senare ”djupare” mappningar som är den yttersta målsättningen. Det innebär också att man inte behöver göra om några mappningar, bara lägga till mer information. Det fungerar inte på samma sätt med andra modeller som är framtagna med minsta gemensamma nämnare som utgångspunkt.

Det finns exempel där man skapat standardiserat rdf-format direkt från databasen. Ett nederländskt baserat företag, Delving, har exempelvis tagit fram en beskrivning av hur en sådan mappning kan gå till, (s.k. X3ML).¹⁰⁹ I denna modell blir resultatet CRM RDF.

Idag finns det vissa begränsningar och avsaknad av vedertagna standarder för länkbar data inom kulturarvsområdet, vilket leder till pragmatiska lösningar och ibland avsaknad av kvalitetskontroll. En mer kontextuell syn på öppna standarder utifrån en önskan att utnyttja de potentiella fördelarna skulle därför kunna ge en viss flexibilitet.¹¹⁰ Sättet som kulturarvsinstitutionerna digitaliserar och hanterar sina digitala kulturarvsdata på är därför, tillsammans med en stödjande teknisk infrastruktur, den viktigaste förutsättningen för att informationen från kulturarvsinstitutionerna ska kunna anpassas till att vara en del av länkad öppen data. Det finns idag flera initiativ kring hantering av kulturarvsterminologier i maskinläsbart format.

Ser man på den nationella nivån så arbetar man idag med att omvandla Libris till ett system som i helhet bygger på länkbar data.¹¹¹ Vad gäller arkivsektorn finns det idag ett pågående arbete inom ramen för det europeiska samarbetet i APEX-projektet med att integrera semantiska beskrivningar för arkivmaterial.¹¹² I K-samsök uttrycks metadata i en RDF-baserad modell, vilket skapar möjligheter till länkning.

¹⁰⁷ <http://en.wikipedia.org/wiki/Triplestore>

¹⁰⁸ <http://www.ltu.se/centres/Centrum-for-langsiktigt-digitalt-bevarande-LDB/Vara-projekt/2.40019/YEAH-You-Enhance-Access-to-History-1.86175>

¹⁰⁹ <https://github.com/delving/x3ml/blob/master/docs/x3ml-language.md>

¹¹⁰ <http://www.archimuse.com/mw2007/papers/kelly/kelly.html>

¹¹¹ <http://librisbloggen.se>

¹¹² http://www.apex-project.eu/images/docs/D61_Web20_In_Archival_Applications.pdf

EU-projektet Linked Heritage¹¹³ (Coordination of Standards and Technologies for the enrichment of Europeana), där Digisam deltog, var det första europeiska projekt som testade att länka metadata i LIDO-standarden från europeiska kulturarvsinstitutioner (framförallt museer men även bibliotek och arkiv). Arbetet i Linked Heritage grundade sig i stor utsträckning på resultaten från ett tidigare EU-projekt kallat Athena, som i sin tur utvecklade LIDO. Linked Heritage har gett ut publikationen "Your terminology as a part of the semantic web recommendations for design and management" med rekommendationer för kulturarvsinstitutioner (framförallt museer) kring terminologier i kulturarvsdata i RDF/SKOS-format¹¹⁴ och flerspråkighet. Den handlar om hur man kan hantera bland annat flerspråkig kulturarvsdata för att göra det möjligt att länka informationen till andra datakällor. Vissa domänspecifika initiativ finns också idag kring att skapa modeller för länkad öppen data inom kulturarvssektorn, bland annat Linked Data for Libraries (LD4L)¹¹⁵ och LiAM (Linked Archival Metadata).¹¹⁶

Auktoritetsfiler som tillgängliggörs online och länkas samman med andra datakällor är ett exempel på länkad öppen data. För närvarande pågår uppbyggnad av olika plattformar för hantering av terminologier (auktoritetslistor, tesaurier, med mera), till exempel Kulturnav¹¹⁷ och "Terminology Management Platform"¹¹⁸ inom ramen för Athena Plus-projektet.¹¹⁹ Stöd för länkning till sådana externa öppna resurser kan även byggas in i registreringsprocessen. Andra exempel på olika typer av samarbeten kring att sammanlänka stora informationsmängder genom länkad öppen data är bland annat OCLCs samarbete med Europeana¹²⁰ samt samarbetet mellan CHAIN-Reds Knowledge Base, Europeana, Cultura Italia och Isidore där man har utvecklat en sökfunktion som bygger på länkbara datamängder.¹²¹

Länkad data innebär en optimering av sökmöjligheterna och ett första steg i sammanlänkningen av idag spridda kulturarvsdata. Maskinläsbarhet ger också möjligheter till praktiska tillämpningar, till exempel i form av virtuella utställningar eller applikationer för utbildning. Sökmöjligheterna och sammanlänkningarna är dock fortfarande beroende av hur beskrivningarna av data ser ut i grunden.

Det är också viktigt att uppmärksamma att den dynamiska karaktären hos länkade öppna dataobjekt och frånvaron av en central administration innebär risker för framtida tillgänglighet och användbarhet.¹²² Det finns pågående teknisk utveckling och forskning idag som handskas med denna problematik. Att bevara länkade data i form av databaser eller webbsidor är inte så problematiskt då det redan idag finns processer för detta. Det som däremot är svårt att bevara är länkar som skapats mellan olika dataobjekt, på grund av avsaknad av rutiner för att hantera och bevara dessa objekt och länkar långsiktigt. Idag finns initiativ för att skapa verktyg för detta; exempelvis har man tittat på metoder för att skapa arkivinformativpaket med länkad data inom PRELIDA-projektet.¹²³ Det finns också flera verktyg som testar dessa funktionaliteter idag,

¹¹³ <http://www.linkedheritage.org/>

¹¹⁴ <http://www.w3.org/RDF/>

¹¹⁵ <https://wiki.duraspace.org/pages/viewpage.action?pageId=41354028>

¹¹⁶ <http://sites.tufts.edu/liam/2014/04/24/version-099/>

¹¹⁷ <http://kulturnav.org/>

¹¹⁸ <http://www.culture-terminology.org/>

¹¹⁹ <http://www.athenaplus.eu/>

¹²⁰ <http://oclc.org/research/news/2012/08-20.html>

¹²¹ <http://ecsg.dch-rp.eu/semantic-search>

¹²² [http://prelida.eu/results/deliverables\(D\)](http://prelida.eu/results/deliverables(D))

¹²³ [http://prelida.eu/results/deliverables\(D2.5\)](http://prelida.eu/results/deliverables(D2.5))

exempelvis Hyberlink.¹²⁴ En annan problematik är att de identifierare som pekar mot olika resurser inte alltid är beständiga.

Sammanfattningsvis finns det ett starkt behov av både konceptuell harmonisering av information mellan olika domäner och tillgängliggörande av samlingar som maskinläsbar och länkbar data. En infrastruktur för detta samt stödjande verktyg vid registrering och/eller aggregering behövs för att skapa så enkla och automatiserade processer som möjligt. Länkbar öppen data i kombination med implementering av vedertagna internationella standarder och stödfunktioner är ett första steg för att öka användningen av kvalitativa kulturarvsdata.

Digitala infrastrukturer

Idag, när mängden digitalt kulturarv ökar kraftigt, framgår det tydligt att kulturarvsinformationen måste exponeras på ett sätt som optimerar potentialen i informationen. Detta innebär bland annat att kulturarvsinformation på nätet måste anpassas och struktureras för interaktivitet. Det är också viktigt att effekten av den ökande mängden kulturarvsinformation inte ger en mättnadseffekt så att man slutar leta om det man söker efter inte finns med bland de första sökträffarna. Det fordras alltså fungerande system för att länka information, bygga applikationer och göra kulturarvsinformation så kvalitativ, relevant och användbar som möjligt. Det finns idag en ökande tendens till användning av gemensamma tekniska stödfunktioner och infrastrukturer för att skapa ökad användning av vedertagna internationella standarder. En stödjande övergripande digital infrastruktur där befintliga infrastrukturer kopplas samman med olika tjänster samt datakapacitet kan byggas upp successivt.

En målbild där olika infrastrukturer integreras med varandra med stöd av olika tjänster och data- och internetkapacitet finns det även behov för på nationell nivå. Som ett resultat av workshopparna rådde samstämmighet kring ett framtidsscenario med en preliminär modell för informationshantering som kan byggas upp successivt utifrån den nuvarande aggregeringsstrukturen.

Vad gäller de tekniska förutsättningarna så är tillgången till en kraftfull infrastruktur en förutsättning för att kulturarvsinstitutionerna ska kunna synas digitalt och att göra sitt material lättillgängligt och användbart. Genom ökad samordning och användning av en gemensam digital infrastruktur med ett brett utbud av tjänster som är särskilt anpassade till kulturarvsdata, skulle kulturarvsinstitutionernas digitala samlingar och bestånd på ett bättre sätt kunna bli öppna och tillgängliga för både forskningen och allmänheten. Det innebär inte att en ny gemensam digital infrastruktur behöver sättas upp, utan att man i så stor utsträckning som möjligt bör använda sig av de befintliga infrastrukturella resurserna som kulturarvsinstitutionerna redan använder sig av, som exempelvis den nationella forskningsinfrastrukturen, SUNET.

Förutsättningar för stödfunktioner, som till exempel metadatastandarder, obrutna länkar, digitalt bevarande på kort och lång sikt, säkerhet och licensiering/upphovsrätt grundas på informationsförvaltningsprocesser hos institutionerna. Det finns ett framtidsperspektiv där en del av dessa stödjande processer flyttar, i högre utsträckning eller till och med helt och hållet, ut från institutionsnivå till de digitala infrastrukturerna. Till exempel beskriver Collections Trust ett

¹²⁴ <http://hiberlink.org/developments.html>

framtidscenario där museisystem blir webbaserade och öppna i sin helhet (COPE – Create Once Publish Everywhere).¹²⁵ Under modelleringsarbetet kring digitala infrastrukturer som Digisam har genomfört uttrycktes ett starkt behov av stödfunktioner kopplade till de tekniska infrastrukturerna, dels i form av råd och stöd och dels i form av tekniska verktyg.

Allt som händer med den digitala informationen, från att den skapas och under livscykeln av ett digitalt objekt, kan beskrivas genom processer i en samordnad digital informationshantering. På så vis kan man säga att samordnad digital informationshantering består av:

- processer på institutionsnivå när digital information skapas (registreras, katalogiseras, förtecknas) och under tiden den finns i institutionernas databaser och system,
- processer som sker genom användning av digitala infrastrukturer (inklusive aggregering och tillgängliggörande) samt
- lagrings- och bevarandeprocesser som kan ske på flera nivåer.

Implementering av vedertagna internationella standarder och stödfunktioner behövs idag som ett första steg i processen att skapa interoperabilitet och harmonisering av data för att kulturarvsinformation ska långsiktigt bli möjlig att hitta, använda och återanvända. Användning av digitala infrastrukturer och möjligheter som där finns för att nå datakapacitet, bredbanduppkoppling och användning av olika tjänster och verktyg ligger till grund för detta.

Ett europeiskt perspektiv

Behovet av att skapa en stödjande datainfrastruktur för att hantera digital information i olika former, från olika aktörer och på olika nivåer har på en generell nivå beskrivits i Europeiska Kommissionens rapport från 2010, ”Riding the Wave”¹²⁶. I rapporten föreslås uppbyggnaden av en gemensam och tvärdisciplinär datainfrastruktur, som med hjälp av olika tjänster underlättar hantering av stora datamängder och integrerar flera infrastrukturer. Inom ramen för projektet EUDAT¹²⁷ arbetar man för närvarande med denna fråga.

EU-finansierade projekt med bäring på digitalt kulturarv, där svenska representanter ofta deltar, har skapat starka nätverk och en kontinuerligt ökad samsyn bland medlemsländerna i frågor om bland annat infrastruktur.

¹²⁵ <http://www.collectionstrust.org.uk/digital/create-once-publish-everywhere-cope>

¹²⁶ <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>

¹²⁷ <http://eudat.eu/>

Förslag till vidare arbete

Grunden i digital informationshantering av samlingar och arkivbestånd består i att information skapas, hanteras och tillgängliggörs på ett standardiserat sätt som stödjer interoperabilitet och säkrar långsiktigt digitalt bevarande. Användning av vedertagna internationella standarder är en förutsättning både för att skapa interoperabilitet mellan data från olika kulturarvsdomäner och för att skapa tillgång för utvecklare till kvalitativa datakällor.

De nationella, domänspecifika aggregeringstjänsterna spelar en viktig roll i infrastrukturen eftersom aggregering idag fortfarande är ett nav för tillgängliggörande och användning av samlingar och arkivbestånd. De bedöms också som viktiga ur ett framtidsperspektiv. Samtidigt ses den semantiska webben som en framtidsbild, där digitalt tillgängliggörande sker genom att kulturarvsinstitutionerna lägger ut öppen länkbar data för att göra det möjligt för utvecklare att bygga tjänster som fångar upp användarnas behov. Stödfunktioner som stödjer standardisering och strukturering av digital information, tillgängliggörande och långsiktigt bevarande av länkade data behöver tittas närmare på och olika rekommendationer behöver knytas till olika steg i dessa processer.

För att målbilden ska kunna nås identifierades följande behov:

1. Samverkan med SUNET, och andra e-infrastrukturer, för att undersöka möjligheterna till en digital e-infrastruktur anpassad till kulturarvsinstitutionernas behov och som i huvudsak handlar om:

- Tillgången till en digital e-infrastruktur med kraftfull bandbredd, datakapacitet, digital lagring, tillgängliggörande, gemensamma digitala arbetsytor, support m.m.
- Utveckling av en gemensam handlingslinje för kulturarvsmyndigheternas nyttjande av den digitala e-infrastrukturen såväl kulturarvsinstitutionerna sinsemellan som mellan kulturarvsinstitutionerna och forskarvärlden.

2. Utveckling och validering av en roadmap för samordnad digital informationshantering, inklusive digitalt tillgängliggörande och bevarande.

Förslag till en modell över en samordnad digital informationshantering samt en vägledning för en handlingsplan för att nå dit.

3. Utveckling och validering av ett antal stödfunktioner samt utvärdering av befintliga stödfunktioner (tekniska, rekommendationer och handledningar) för digitalisering, tillgängliggörande och bevarande av det digitala kulturarvet.

I målbilden beskrivs en så kallad ”verktygslåda” som ett ramkoncept som successivt kan fyllas med moduler för lagring, metadataväxling, valideringsverktyg, m.m.

Resultaten av dessa arbetsområden kommer att vara grundläggande för framtagandet av nationella riktlinjer.

Ordlista

ABMK – Arkiv, bibliotek, museer, kulturmiljö.

Aggregering - Aktiviteten att samla eller knyta ihop information som hör ihop eller har gemensamma egenskaper.¹²⁸ Mer generellt innebär aggregering en process där data summeras (eller mer generellt kombineras) till ny data.¹²⁹ Inom digitalt kulturarv står ofta uttrycket för insamling av metadata från olika institutioner för överföring till olika typer av webbportaler, alternativt för tillgängliggörande via API:er.

Aggregator - I kulturarvsdata och i synnerhet i Europeanas kontext är aggregatorer organisationer, projekt eller institutioner som samlar in metadata från enskilda organisationer och ser till att deras metadata och format standardiseras och överförs till Europeana enligt Europeanas riktlinjer och rutiner.¹³⁰ Det finns olika typer av aggregatorer; landspecifika (nationella eller regionala, domänöverskridande eller domänspecifika), projektaggregatorer och oberoende organisationer.

API (Application Programming Interface) - En regeluppsättning för hur en viss programvara kan kommunicera med en annan programvara.¹³¹

Arkivsystem - System som hanterar digitala arkivhandlingar från myndighetens/ institutionens verksamhet. Enligt eARD-projektets definition är det ett system med sådan funktionalitet och uppmärkning av information så att långsiktigt säkerställande kan uppnås.¹³²

Auktoritet - Källa som enligt överenskommelse utgör standard för stavning, termer eller policy vid registrering.¹³³

Auktoritetspost - Post bestående av en auktoriserad form för en term eller namn kombinerad med andra informationselement som bidrar till att identifiera och särskilja det som den auktoriserade formen betecknar.¹³⁴

Auktoritetsfil – En auktoriserad form för att identifiera en term eller ett namn.

Beständiga identifierare/PID (persistent identifiers) - Unika kodsträngar kopplade till varje objekt/post. PID:ar är en förutsättning för att man sedan med fasta länkar/URI (uniform resource identifiers) ska kunna skapa referenser mellan olika objekt/poster i ett nätverk som t.ex. Internet.

Bevarandesystem - System som används för att säkerställa bevarande av elektroniska handlingar.¹³⁵

Datafil - En systematisk strukturerad mängd digital information.

¹²⁸ eARD-projektets definition, <http://riksarkivet.se/publicerade-rapporter-fran-eard>

¹²⁹ <http://www.beslutsstod.nu/ordbok-beslutsstod-och-business-intelligence/vad-ar-data-aggregering/>

¹³⁰ <http://linkedheritage.cab.unipd.it/training/LO-03/sv/01.html>

¹³¹ http://sv.wikipedia.org/wiki/Application_Programming_Interface

¹³² http://riksarkivet.se/Media/pdf-filer/Projekt/eARD_Begreppslista_DP1_v2.xls

¹³³ <http://digisam-ra.blogspot.se/p/begreppslista-bilddatabaser-och.html>

¹³⁴ Ibid.

¹³⁵ eARD-projektets definition: http://riksarkivet.se/Media/pdf-filer/Projekt/eARD_Begreppslista_DP1_v2.xls

Digitalt bevarande - Säkrande av innehåll så att framtida användare kan nå det, förstå och ta det till sig.

Kortsiktigt bevarande – IT-lösningar som används kortsiktigt, max 5 år.

Mellanlångt bevarande – IT-lösningar som används så länge ett system håller, max 10 år.

Långsiktigt bevarande – IT-lösningar som används för att kunna hålla längre än livstiden för systemet.

Digital infrastruktur - Allt från internetuppkoppling till informationshantering med en hel rad tjänster kopplade till det, som till exempel grid-tjänster, datalagring, säkerhet, interaktivitet, med mera.

Digitalisering - Operativt arbete med överföring av information, data och metadata till digitala format.

Digital kulturarvsinformation - Begreppet har använts för att beteckna både digital information om samlingar (metadata) och de digitala samlingarna själva (digitala objekt).

Digital lagring - Teknisk lagring av data.

Domän - I kulturarvssammanhang används ordet domän ofta för att specificera ett visst ABMK-område, exempelvis arkiv.

e-Infrastruktur - Verktyg för elektronisk kommunikation, lagring, bearbetning och visualisering av data samt datorer för storskaliga beräkningar, visualisering, modellering och lagring, nätverk för dataöverföring och avancerat användarstöd som hör ihop med detta.¹³⁶

Georeferenser - Inom ramen för rapporten avses att uttrycka beskrivning av en lokalitet i en mappningsbar representation.

Grid (Grid computing) - Är ett sätt att dela datakraft och datalagringskapacitet via Internet och genom särskilda höghastighetsnät. Skillnaden mellan grid och moln är att grid är en open source-teknik som utvecklats speciellt för den akademiska forskningen.¹³⁷

Interoperabilitet - Förmågan hos olika system att fungera tillsammans och kommunicera med varandra. Exempel på detta kan vara att systemen kan använda samma protokoll eller läsa och skriva samma filformat. Interoperabilitet uppnås genom samordning och (internationella) standardiseringsprocesser.

Interstruktur - Inom ramen för rapporten avses flera sammankopplade infrastrukturer.

Kulturarvsinstitutioner - Institutioner som samlar, bevarar och tillgängliggör digital kulturarvsinformation.

Licensiering - En markering av materialets upphovsrättsliga status.

Länkbar öppen data - Avser resurser som är fritt tillgängliga på nätet i maskinläsbara, icke-proprietära format och i enlighet med datamodellen RDF¹³⁸ (resource description framework) som möjliggör meningsbärande länkar och därmed den framväxande semantiska webben.

¹³⁶ <http://www.vr.se/download/18.1d68133213a3f73f1e93e4/1349857069368/Vetenskapsradets+guide+till+infrastruktur+2012.pdf> och <http://www.snic.vr.se/news-events/events/vaelkommen-till-workshop-om-e-infrastruktur>

¹³⁷ <http://www.egi.eu/about/faq/>

¹³⁸ <http://www.w3.org/RDF/>

Länkad data - Sammanlänkning av information som finns i olika domäner genom den semantiska webbens möjligheter.

Länkad öppen data - Avser resurser som är fritt tillgängliga på nätet i maskinläsbara, icke-proprietära format och i enlighet med datamodellen RDF¹³⁹ (resource description framework) är sammanlänkade och därigenom utgör en del av den semantiska webben.

Mappning - Inom ramen för rapporten avses en jämförelse mellan två metadata scheman där kompatibla kopplingar mellan olika metadata fält definieras.

Metadata - Data om data (Anmärkning: Metadata är knutna endera till ett informationssystem eller till ett informationsobjekt med syftet att beskriva, administrera eller använda detta. Termen används vanligen för strukturerande uppgifter om digitala (och analoga) system. Metadata delas ofta upp i olika typer, exempelvis teknisk metadata som skapas för eller genereras av ett datasystem, administrativa metadata som beskriver hur ett system fungerar eller hur ett objekt skapats och metadata som rör objekten i systemet och deras egenskaper).¹⁴⁰

Metadata standarder - Olika standarder för metadata används inom kulturarvssektorn, t.ex. CIDOC-CRM.

Metadata växel - Inom ramen för rapporten avses tjänster och/eller verktyg som kan skapa automatiserade eller halvautomatiserade mappningar mellan olika metadata scheman.

Metadata standarder - Olika standarder används inom olika kulturarvssektorer, t ex MARC (bibliotek) och LIDO/CIDOC-CRM (museer).

Myndighetsarkiv/ institutionsarkiv - De arkiv som bildas av allmänna handlingar från myndighetens/ institutionens verksamhet, och som omfattas av arkivlagen (1990:782).¹⁴¹

Ontologi - I arbetet med elektronisk informationshantering används ofta ett formaliserat sätt att beskriva en ontologi för en domän. Genom att beskriva ontologin på detta sätt möjliggörs härledning till definitioner av klasser, egenskaper och begrepp.¹⁴²

Resolver-tjänst - En sådan tjänst syftar till att dirigera webbläsaren till en lämplig eller aktuell kopia av objektet, som till exempel Handle-systemet¹⁴³. Detta är en viktig strategi för att minska risken för felaktiga länkar genom att ett lager läggs mellan webbläsaren och målobjektet som tillåter webbläsaren att hitta rätt objekt och som vidarebefordrar till en aktuell kopia av objektet. Resolver-tjänsten kan vara en uppsättning omskrivningsregler på en server eller ett globalt nätverk av särskilda servrar.

Roadmap - En vägledning i att skapa en aktivitetsplan.

Semantisk beskrivning - Medan världswebben är en standard för att presentera (sammanlänkade) dokument så går den semantiska webben längre och standardiserar även hur man uttrycker data och kunskap. Man har bl.a. standardiserat hur man refererar till objekt som man vill beskriva. De identifieras med unika identifierare, så kallade URI:er, som tekniskt sett är webbadresser. Data och

¹³⁹ <http://www.w3.org/RDF/>

¹⁴⁰ <http://digisam-ra.blogspot.se/p/begreppslista-bilddatabaser-och.html>

¹⁴¹ <http://www.notisum.se/rnp/sls/lag/19900782.htm>

¹⁴² http://sv.wikipedia.org/wiki/Ontologi#Ontologibegreppet_i_informationsarbete

¹⁴³ http://en.wikipedia.org/wiki/Handle_System

kunskap uttrycks sedan genom så kallade tripletter sammansatta av subjekt, predikat och objekt som var för sig identifieras på ett standardiserat sätt. Att beskriva semantik på ett standardiserat sätt innebär att redogöra för på vilket sätt man semantifierat sitt datainnehåll.¹⁴⁴

Semantiska webben - Ett begrepp myntat av World Wide Web Consortiums (W3Cs) chef Tim Berners-Lee, som också är skapare av världswebben (WWW). Begreppet beskriver metoder och teknik för att möjliggöra för maskiner att förstå innebörden eller "semantiken" i informationen på webben.¹⁴⁵

Semantisk interoperabilitet - Förmågan att utväxla information mellan datorsystem på ett sådant sätt att mottagande system utan manuell intervention kan tolka informationens innebörd och producera för slutanvändaren användbara resultat. Detta förutsätter en för sändare och mottagare gemensam referensmodell över kunskapsområdet.¹⁴⁶

Samlingar - De mängder av objekt som en kulturarvsinstitution förvaltar, t ex böcker, konst, museiföremål, enskilda arkiv och andra forskningsmaterial.

Samlingssystem - Content Management System, system som hanterar digitala objekt som en kulturarvsinstitution förvaltar, ofta vid museer.

State of the art - Dagens högsta nivå inom ett specifikt område.

Teknisk infrastruktur - Nätverk för dataöverföring.

Term - Benämning för begrepp inom visst fackområde¹⁴⁷

Terminologier - En mängd ord, uttryck och termer som är specifika för ett visst fackområde auktoritetslistor, tesaurier, m.m.¹⁴⁸

Tesaurus/thesaurus - Kontrollerad ordlista med de semantiska relationerna mellan angivna termer (Anmärkning: En tesaurus är ofta hierarkisk och omfattar i regel ett särskilt ämnesområde).¹⁴⁹

Utbytesformat - Inom rapporten definieras utbytesformat som format som gör att informationen kan överföras mellan olika IT-system. Utbytesformat har i de flesta fall utvecklats i samband med överföring av digital information.

Öppen data - Avser resurser som är fritt tillgängliga på nätet för att användas, återanvändas och publiceras i maskinläsbara, icke-proprietära format.

Öppen standard - En standard som, i motsats till en proprietär teknisk specifikation, tillåter vem som helst att implementera den.¹⁵⁰

¹⁴⁴ <http://digisam.se/index.php/docman/public-documents/principer/37--7/file>

¹⁴⁵ http://sv.wikipedia.org/wiki/Semantiska_webben

¹⁴⁶ http://sv.wikipedia.org/wiki/Semantisk_interoperabilitet

¹⁴⁷ Ibid.

¹⁴⁸ <http://sv.wikipedia.org/wiki/Terminologi>

¹⁴⁹ <http://digisam-ra.blogspot.se/p/begreppslista-bilddatabaser-och.html>

¹⁵⁰ http://sv.wikipedia.org/wiki/%C3%96ppen_standard

Akronymlista

ABCD - Access to Biological Collection Data

ABMK - Arkiv, bibliotek, museer och kulturmiljövård

ATHENA - Access to Cultural Heritage Networks Across Europe

CIDOC CRM - CIDOC Conceptual Reference Model (CRM)

COPE - Create Once Publish Everywhere

DAP - Digital Arkeologisk Process

DCH-RP - Digital Cultural Heritage – Roadmap for Preservation

DC-NET - Digital Cultural Heritage Network

DISKA - Digitala Semantiska Kulturarvsauktoriteter

EAC-CPF - Encoded Archival Context for Corporate Bodies, Persons, and Families

EAD - Encoded Archival Description

EAG - Encoded Archival Guide

ECE-skolan - Skolan för teknikvetenskaplig kommunikation och lärande vid KTH

ECK - Europeana Connection Kit

EDM - Europeana Data Model

ENSE - Nationell Standard för Naturhistoriska Samlingsdatabaser

ESE - Europeana Semantic Elements

EUDAT - European Data Infrastructure

FGS - Förvaltningsgemensamma specifikationer

FRBR - Functional Requirements for Bibliographic Records

GBIF - Global Biodiversity Information Facility

ISAAR-CPF - General International Standard Archival Description

ISCI - International Standard Collection Identifier

ISDIAH - International Standard for Describing Institutions with Archival Holdings

KTH - Kungliga Tekniska Högskolan

LIDO - Lightweight Information Describing Objects

LOD - Linked Open Data

MARC - MACHine-Readable Cataloging

MICHAEL - Multilingual Inventory of Cultural heritage in Europe

MINERVA - Ministerial network for valorising activities in digitisation

MINT - Metadata Interoperability Services

MOU - Memorandum of Understanding

NREN - National Research and Education Network

OCLC - Online Computer Library Center

OpenAIRE - Open Access Infrastructure for Research in Europe

RDA - Resource Description and Access

RDF - Resource Description Framework

SUNET - Swedish University Computer Network

SKOS - Simple Knowledge Organisation System

TDWG - Taxonomic Databases Working Group/ Biodiversity Information Standards

TMP - Terminology Management Platform

UGC - User Generated Content

URI - Uniform Resource Identifiers

XML - Extensible Markup Language

Digisam är ett samordningssekretariat för digitalisering, digitalt bevarande och digitalt tillgängliggörande av kulturarvet.

Samordningssekretariatet är en av regeringen beslutad verksamhet vid Riksarkivet.

www.digisam.se