

# Ediffah — En digital infrastruktur för forskningsbibliotekens arkiv- och handskriftssamlingar

## 1. Övergripande mål

Projektet syftar till att göra det möjligt att skapa skapa allmänt tillgängliga kataloger och arkivbeskrivningar över bibliotekens handskriftssamlingar och personarkiv. Systemet skall vara enkelt, och innehålla ett arbetsflöde sådant att posterna som genereras också med automatik kan återfinnas i Nationella arkivdatabasen [NAD], vårt gemensamma bibliotekssystem [LIBRIS], såväl som i lokala databaser.

Vi föreslår ett antal åtgärder, som resultat av ett samarbetsprojekt mellan institutioner inom biblioteks- och arkivväsendet. Resultatet avses att bli en kombination av arbetsflöden och förbättrade programvaror som kommer att göra forskningsbibliotekens samlingar mera synliga och lättillgängliga. Genom att förenkla produktionen av poster kommer biblioteken att snabbare kunna registrera sina samlingar. Den framtida driften av en gemensam databas för dessa typer av arkiv skall också utredas.

### 1.1. Produkter

Efter projektets slut skall följande produkter finnas:

- 1 Kortfattade arkivbeskrivningar i det XML-baserade formatet Encoded Archival Description [EAD] för arkiven vid de största svenska forskningsbiblioteken. I samband med att detta görs skall också auktoritetsfiler för namn i databaserna finnas tillgängliga i formatet Encoded Archival Context [EAC]. Möjligheterna för konvertering av befintliga databaser undersöks. Alla NAD-poster konverteras, och för andra tjänster görs konvertering i experimentell skala.
- 2 Minst en webbplats som implementerar sökning och navigering i sådana beskrivningar. Webbplatsen skall vara baserad på open-source programvara och erfarenheterna av detta arbete sammanfattas i en rapport.
- 3 Användningen av EAD för katalogisering av arkivalier vid forskningsbibliotek och dess specialsamlingar skall ha studerats, och några verktyg för sökning, navigering och produktion av EAD-poster skall ha testats och jämförts med katalogisering direkt i NAD. Ett av de testade verktygen kommer att användas i DP2. Erfarenheter från detta skall redovisas i en rapport.
- 4 Etablerade kommunikationsvägar som tillåter kommunikation mellan det utvecklade systemet och Nationella Arkivdatabasen [NAD].
- 5 Etablerade kommunikationsvägar som tillåter kommunikation mellan det utvecklade systemet och samkatalogen LIBRIS. Syftet med detta är att användare som söker material av och om särskilda personer skall kunna återfinna även deras arkiv.
- 6 Utredning av förutsättningarna för driften av en gemensam databas för EAD-beskrivningar av arkivsamlingar vid forskningsbiblioteken. Webbplatsen som utvecklas inom DP2 är en experimentell prototyp för en sådan söktjänst.

Varje produkt utgör ett delprojekt i projektplanen. Vissa produkter kommer att finnas kvar efter projektets slut, och kommer vara tillgängliga för nedladdning och återanvändning, dock utan underhåll. Webbplatsen kommer att vara i drift ett år efter att projektet avslutats.

## **2. Deltagande organisationer och kontaktpersoner**

Projektgruppen består av organisationer med samlingar, erfarenheter och teknologi som är viktiga för att uppnå projektets mål. Personerna har alla erfarenheter som ligger i linje med projektets innehåll.

- **BD Lund**, Biblioteksdirektionen Lunds universitet (Sigfrid Lundberg, forskningsingenjör, och Jörgen Eriksson, bibliotekarie)
- **GUB**, Göteborgs Universitetsbibliotek (Pia Schmidt de Graaf, bibliotekarie med ansvar för arkiv)
- **KB**, Kungliga Biblioteket (Anders Burius, chef för handskriftsavdelningen)
- **LIBRIS**, KB-LIBRIS det nationella bibliotekssystemet (Mats Herder)
- **RA**, Riksarkivet (Göran Kristiansson, leder utvecklingen av NAD)
- **UB Lund**, Universitetsbiblioteket, Lunds universitet (Birgitta Lindholm, föreståndare för handskriftsavdelningen)
- **UUB**, Uppsala Universitetsbibliotek (Inga Johansson, tf chef för handskriftsavdelningen)

## **3. Projektets innehåll och tidsplanering**

De olika delprojektens innehåll och projektets tidsplanering sammanfattas i Tabell 1. Här följer en mer detaljerad beskrivning.

### **3.1. Skapande av kortfattade arkivbeskrivningar i det XML-baserade formatet EAD (Dp. 1)**

Som framgår av bakgrundsteckningen nedan (se Bakgrund) finns redan idag mycket kortfattade arkivbeskrivningar för forskningsbibliotekens arkiv i den gamla NAD-databasen som distribuerades i form av en CD-skiva.

Dp 1 syftar till export av dessa data i EAD för återföring till respektive bibliotek. Samtidigt skall även kortfattade auktoritetsposter i EAC-format skapas. EAC fyller två funktioner: Det används för distribution av normaliserade namnformer i arkivdatabaser, men även för "öppen" länkning mellan arkivdatabaser och för identifiering av databaser som är lämpliga för samsökning. Riksarkivet och arkivinstitutioner (främst inom biblioteksvärlden) använder EAC inom LEAF-projektet (Linking and Exploring Authority Files) för auktoritetsutbyte [LEAF]

Dp1 kommer också att inventera de befintliga databaser som finns vid bl. a GUB och UU och göra en uppskattning av det arbete som kommer att krävas för att konvertera dessa.

### **3.2. Skapande av minst en webb-plats som implementerar sökning och navigering i EAD-beskrivningar (Dp. 2)**

EAD är ett format som tillåter katalogisering av komplexa resurser vilka är hierarkiskt strukturerade. En ytlig arkivbeskrivning ger inte mycket mer än en Dublin Core beskrivning av en samling (jämför DCMI Collection Description WG's arbete [DC])

Ett arkiv är nästan alltid hierarkiskt organiserat och man kan ha valt att lägga registreringen på objektsnivå, på volymnivå eller - vid en rudimentär beskrivning - på en ännu högre nivå. Typiskt sett kan ett arkiv bestå av kapslar. EAD är anpassat för katalogisering av hierarkiskt organiserade samlingar, och kan även innehålla löpande prosa i beskrivningarna. Namn på platser och organisationer och andra viktiga detaljer kan markeras för automatisk indexering och sökning.

EAD används numera vid många arkivinstitutioner, eller konsortier av sådana [EAD-sites]. Archives Hub, "A national gateway to descriptions of archives in UK universities and colleges" [AHUB] är en sådan, som dessutom har gjort sina verktyg till öppen källkod. Inspiration kommer också att hämtas från andra implementeringar, t ex från Library of Congress fi nding aid's project[LOC]

Dp. 2 syftar till att med minsta möjliga ansträngning implementera en webbplats för sökning och navigering av arkivbeskrivningar. Webbplatsen skall byggas med fri programvara och eventuella anpassningar för svenska förhållanden skall göras tillgängliga för intresserade. Härigenom skall det enskilda bibliotekets samlingar kunna synliggöras.

Webbplatsen skall potentiellt kunna fungera som en gemensam resurs för samtliga bibliotek eller kunna replikeras vid de olika institutionerna. Vidare skall den metod för katalogisering som väljs eller utvecklas i Dp. 3 kunna integreras i webbplatsen.

Den starkaste kandidaten när det gäller programvara för sökning är sökmaskinen Cheshire II (som huvudsakligen har sitt huvudsakliga hem i Berkeley) [CHESHIRE] som delvis har utvecklats av personal vid Archives Hub och Manchester Information & Associated Services [MIMAS]. Båda tjänsterna är utvecklade med stöd av Joint Information Systems Committee [JISC] och är delar av dess utbud av samlingar.

Varje arkiv eller bibliotek som deltar i ArchivesHub har en egen server och den centrala tjänsten samsöker alla 40 eller en del av dem beroende på frågans innehåll. Programvaran levereras färdig bränd på CD från den centrala HUBen i Manchester och kommer komplett med operativsystem, redigeringsverktyg och databashanterare. Mer information om hur ArchivesHub fungerar fi nns tillgängligt på Internet.[AHS] Det fi nns inget i Archives Hub som avslöjar för användaren att tjänstens hela arkitektur är baserad på databaser spridda över de brittiska öarna.

Om Ediffah byggs distribuerat som Archives Hub eller centraliserat som en centraliserad lösning i likhet med LIBRIS eller NAD är en fråga som skall utredas under projektets gång. Ett självklart krav är att programvaran skall kunna länka till digitaliserade versioner av beskrivna objekt.

Beslutet att de svenska forskningsbibliotekens arkiv skulle kopplas till den Nationella arkivdatabasen [NAD] är gammalt. Denna koppling skall fi nnas kvar. Här ställer vi i istället frågorna: Hur skall vi använda EAD, och fi nns det användningsområden som går utöver det självklara, att beskriva arkiv.

Syftet med Dp. 3 är tvåfaldigt:

- att utreda användningen av EAD. Om användningen kommer att bli begränsad till sökning i arkivalier och kommunikation med NAD kan de funktionella kraven bli enklare än om det finns vidare användningsområden. Jämför användningen av EAD för beskrivning av digitala samlingar inom Open Archives Initiative [OAI]. Likaså påverkas användningen av EAD av vilken fokus katalogiseringen ges och vilka katalogiseringsregler som används.
- att ge de svenska forskningsbiblioteken en infrastruktur för att katalogisering i EAD. För det ändamålet skall vi testa några produkter och välja ut en kandidat. Kandidaten skall jämföras med katalogisering direkt i NAD's produktionsserver.

### **3.4 Etablering av kommunikationsvägar mellan arkivdatabaser vid bibliotek och andra databaser (Dp. 4 och 5)**

Tre funktionella krav skall tillgodoses: arkivalier skall synas i LIBRIS, i NAD och på det enskilda bibliotekets webbplats.

Dp. 4 och 5. Etablering av kommunikationsvägar mellan bibliotekens framtida arkivdatabaser och de två nationella katalogerna NAD och LIBRIS.

LIBRIS kommer inom projektet att producera en specifikation och katalogiseringsregler för en sammanfattande MARC-post vilken skall peka på en arkivbeskrivning. Specifikationen kommer att vara en anpassning av en redan etablerad konkordans mellan EAD och MARC21 [Cookbook]. Procedurer för uppdatering av exporterade MARC-poster i Libris kommer att etableras.

Samarbetet med NAD kommer att vara något mer komplicerat, eftersom det kan bli tal om tvåvägskommunikation, dvs poster skall kunna exporteras till bibliotekens lokala databaser, och även kunna importeras i NAD. Å andra sidan existerar redan de nödvändiga verktygen. En särskild procedur för export av EAC-poster måste också etableras om katalogisering skall ske i andra databaser än i NAD.

### **3.5 Utredning av etablerandet av en samsökningstjänst för personarkiv (Dp. 6)**

Det kan vara av intresse att etablera en EAD-baserad samsökningstjänst för bibliotekens arkivsamlingar. Syftet med Dp. 6 är att finna former för driften av en sådan. En sådan tjänst skall fungera med de arbetsföden som är produkten av Dp. 3, 4, och 5, ha långsiktig finansiering samt en organisatorisk hemvist.

## **4. Kostnader**

Kostnaderna sammanfattas i Tabell 2.

## **5. Bakgrund**

De svenska forskningsbiblioteken har ansevärd samlingar av arkivalier, främst i form av brev, manuskript och andra efterlämnade handlingar efter personer vars livsverk legat inom estetiska, politiska eller vetenskapliga områden.

Projektet avser att väsentligt underlätta tillgängligheten till detta material - ovärderligt inte bara för forskning och högre utbildning utan också för en bred allmänhet.

Tabell 1. Ediffah-projektets olika delprojekt och kortfattad beskrivning av deras innehåll

<b>Dp</b>	<b>Startvecka i projektet och Varaktighet (veckor)</b>	<b>Beskrivning och Deltagare (ansvarig partner i fet stil)</b>
1	Vecka 2. 10 veckor	Omvandlingen av posterna i CD-versionen av NAD till EAD och EAC. Konvertering av existerande databaser Deltagare: <b>RA</b> , UUB, GUB
2	Vecka 12. 10 veckor	Etablering av webbplats med sök- och navigeringsfunktioner för EAD. Deltagare: <b>BD</b>
3	Vecka 2. 20 veckor	Användning av EAD för katalogisering av arkivalier vid forskningsbibliotek och dess specialsamlingar. "Best practices", katalogiseringsregler och verktyg. Deltagare: <b>KB</b> , alla
4	Vecka 24, 8 veckor	Etablering av kommunikationsvägar till, och från, NAD. Deltagare: <b>RA</b> , BD
5	Vecka 24, 8 veckor	Etablering av kommunikationsvägar till LIBRIS. Deltagare: <b>LIBRIS</b> , BD
6	Vecka 32, 20 veckor	Utredning av förutsättningarna för drift av söktjänst för personarkiv. Deltagare: <b>KB?</b> , alla

Tabell 2. Kostnader och personalinsatser inom projektet

Deltagare	Beräknad tidsåtgång	Kostnad
RA	2 pm	100,177 SEK
BD	3 pm + 0.5 pm (administration)	175,310 SEK
LIBRIS	1 pm	
GUB	1,5 pm	87,658 SEK
KB	1 pm	-
LUB	1 pm	50,089 SEK
UUB	1,5 pm	75,133 SEK
Resor	5 möten: 2 allmänna och 3 inom DP 3.	150,000 SEK
Summa	638,367 SEK	

Hur hittar en förhoppningsfull forskare fram till de källor som finns vid svenska

forskningsbibliotek. Svaret på frågan är att han eller hon sannolikt knappast hittar källmaterial, åtminstone inte på egen hand. Det finns ingen strukturerad, sök- eller navigerbar information om dessa samlingar på de aktuella bibliotekens webb. Det existerar inga länkar från bibliotekens kataloger till arkivbeskrivningar. I den mån strukturerad information om samlingarna finns, ligger den i äldre versioner av Riksarkivets databaser och saknar där länkar till annan nätburen information. För material vid forskningsbiblioteken är dessa beskrivningar ytterst översiktliga, inkompleta och tyvärr inaktuella.

Metoden som återstår för låntagare är informationsåtervinning per telefon och tågresor.

1987 påbörjades datoriseringen av de svenska arkivregistren, vilka samlades under en hatt i Nationella Arkivdatabasen, NAD. Från år 1995 utgavs denna på CD-ROM, ett distributionsmedium som övergavs från år 1998 till förmån för ett webbaserat system [NAD].

En betydande del av arkivalierna vid de svenska forskningsbiblioteken katalogiserades ytligt för användning i NAD. Efter 1998 har dessa katalogposter inte underhållits och inga nya lagts till. Katalogisering såväl som sökning var baserat på en äldre MARC-liknande standard som övergivits av Riksarkivet till förmån för ett nytt arkivsystem Arkis2, vars arkitektur är baserad på EAD, ett XML/SGML baserat detaljerat metadatasystem för beskrivning av samlingar [EAD].

De beskrivningarna av arkivalier som skapades vid registreringsenheten REA i Sandslån[ABM] har kommit med i NAD, via den nya Arkis2 baserade arkivdatabasen. Detta gäller för vissa av de större forskningsbiblioteken (KB, UU och GU), och kommer inom en nära framtid finnas med i NAD på Internet, medan de samlingar som katalogiserades via andra initiativ (t ex de vid LUB) inte finns tillgängliga via moderna gränssnitt.

Att snabbt drastiskt förbättra situationen låter sig inte göras utan mycket stora och kostsamma investeringar. Även om pekuniära resurser funnes i överflöd, kräver arbetet insatser från kompetent personal med kännedom om samlingarna. Kompetens kan nyrekryteras, men även ett omfattande åtgärdsprogram är en kamp emot tiden; kunskapen om samlingarna har i stora stycken traderats oralt mellan generationerna. Dagens traditionsbärare tillhör inte yrkeskårens yngre medlemmar.

Att förenkla framställningen av poster och hur de görs maximalt synliga så långt som möjligt ser vi som viktigt i denna situation, eftersom den stora kostnaden är för personal till postframställning. Arkis2 t. ex., som är RAs nya verktyg för att skapa arkivposter, är ett komplext verktyg, byggt för att kunna hantera alla typer av arkiv och materialtyper, med en relativt hög inlärningströskel. Ett verktyg, särskilt anpassat till universitetsbibliotekens typ av arkiv, ser vi som en viktig komponent i arbetet med att skapa ökade möjligheter för registrering av dessa arkiv.

## 6. Referenser

[AHUB]	<a href="http://www.archiveshub.ac.uk/">http://www.archiveshub.ac.uk/</a>
[AHS]	<a href="http://www.archiveshub.ac.uk/arch/index.html">http://www.archiveshub.ac.uk/arch/index.html</a>
[ABM]	<a href="http://www.kb.se/bibsam/bidrag/abmrappport.htm">http://www.kb.se/bibsam/bidrag/abmrappport.htm</a>
[CHESHIRE]	<a href="http://cheshire.lib.berkeley.edu/">http://cheshire.lib.berkeley.edu/</a>
[Cookbook]	<a href="ftp://jefferson.village.virginia.edu/pub/ead/EADCookbook.pdf">ftp://jefferson.village.virginia.edu/pub/ead/EADCookbook.pdf</a>
[DC]	<a href="http://dublincore.org/groups/collections/">http://dublincore.org/groups/collections/</a>

[EAC]	<a href="http://www.crxnet.com/leaf/news_online.html#eac2">http://www.crxnet.com/leaf/news_online.html#eac2</a>
[EAD]	<a href="http://www.loc.gov/ead/">http://www.loc.gov/ead/</a> , se också <a href="http://jefferson.village.virginia.edu/ead/">http://jefferson.village.virginia.edu/ead/</a>
[EADsites]	<a href="http://jefferson.village.virginia.edu/ead/sitesann.html">http://jefferson.village.virginia.edu/ead/sitesann.html</a>
[JISC]	<a href="http://www.jisc.ac.uk/">http://www.jisc.ac.uk/</a>
[LEAF]	<a href="http://www.leaf-eu.org/">http://www.leaf-eu.org/</a>
[LIBRIS]	<a href="http://www.libris.kb.se/">http://www.libris.kb.se/</a>
[LOC]	<a href="http://www.loc.gov/rr/ead/">http://www.loc.gov/rr/ead/</a>
[MIMAS]	<a href="http://www.mimas.ac.uk/">http://www.mimas.ac.uk/</a>
[NAD]	<a href="http://www.nad.ra.se/">http://www.nad.ra.se/</a>
[OAI]	<a href="http://www.openarchives.org/OAI/2.0/guidelines-repository.htm#setDescription">http://www.openarchives.org/OAI/2.0/guidelines-repository.htm#setDescription</a>

### **Revision**

**Version:** \$Revision: 1.12 \$

**Datum:** \$Date: 2004/02/25 14:03:53 \$

**Dokumentansvarig:** \$Author: sigge \$