

Samordning av ansvarsbibliotekens ämnesportaler Grov teknisk kravspecifikation

Henrik Åslund
Tomas Friberg

KUNGL. BIBLIOTEKET

BIBSAM, avdelningen för nationell samordning och utveckling
Box 5039, 102 41 Stockholm. Humlegården. Tel: 08-463 4272
www.kb.se/bibsam/ E-post: bibsam@kb.se
2002-05-13



1. Inledning	3
2. Lägesbeskrivning	3
2.1 Svenska ämnesportaler – kort historik	3
2.2 Internationell utblick	4
2.3 Kostnader.....	5
3. Existerande systemlösningar	6
3.1 Kommersiella lösningar.....	6
3.1.1 ENCompass	6
3.1.2 MetaLib/SFX.....	6
3.1.3 Portalmaker.....	7
3.1.4 Z-portal.....	7
3.1.5 Slutsats.....	7
3.2 Icke kommersiella lösningar.....	7
3.2.1 Projekt:	8
3.2.2 Applikationer:.....	8
3.2.2 Protokoll och standarder:.....	9
3.2.3 Z39.50-applikationer:.....	9
3.2.4 Harvesting:	10
3.2.5 Sammanfattning.....	10
3.3 Ansvarsbiblioteken nuvarande lösningar	10
3.3.1 Sammanfattning.....	11
4. Innehåll	12
4.1 Innehållsstruktur	12
4.1.1 Ämnesrubriker och klassning	12
4.1.2 Metadata	13
4.2 Ambitionsnivå	13
4.3 Sammanfattning.....	14
5. Gränssnitt	15
5.1 Producenternas gränssnitt.....	15
5.2 Användarnas gränssnitt	15
5.3 Sammanfattning.....	16
6. Teknik	17
6.1 Tekniska förutsättningar	17
6.1.1 Databasmodellen	17
6.1.2 Samsökningsmodellen.....	18
6.2 Teknisk plattform.....	19
6.3 Sammanfattning.....	19
7. Förslag till fortsatt arbete	20
7.1 Vidareutveckling av någon av de befintliga portalerna	20
7.2 Användning av befintlig portalprogramvara	20
7.3 Utveckling av egen programvara och teknisk arkitektur.....	20
7.4 Sammanfattning och förslag.....	21
8. Bilagor	23
Bilaga 1 Projektbeskrivning	23
Bilaga 2 Projektorganisation	24

1. Inledning

Projektets syfte, bakgrund, specifikationer och avgränsningar beskrivs i bilaga 1. I bilaga 2 återfinns projektorganisationen.

2. Lägesbeskrivning

2.1 Svenska ämnesportaler – kort historik

I Sverige har ämnesportaler byggts upp inom eller i nära anknytning till ansvarsbibliotekssystemet, och uppbyggnaden av ämnesportalerna har under flera år varit en av BIBSAM prioriterad ansvarsbiblioteksuppgift. De olika portaler som finns idag har dock olika tillkomsthistoria. En del av förklaringen till detta ligger i ansvarsbibliotekssystemets uppbyggnad och tillväxt. Uppbyggnaden av ansvarsbibliotekssystemet har pågått under många år, och fem av de idag 11 huvudansvariga ansvarsbiblioteken har utnämnts så sent som under åren 1996-1999 (ansvarsbibliotek för samhällsvetenskap, kvinno-, mans- och genusforskning, biologi, juridik och geovetenskap). Först i och med utnämningen av ett geovetenskapligt ansvarsbibliotek 1999 blev systemet ämnesmässigt heltäckande.

EELS, numera förkortning för Engineering E-library Sweden, började byggas upp redan 1994, och det fanns då ytterst begränsad internationell erfarenhet att bygga på. EELS har utvecklats i samarbete mellan de svenska teknologibiblioteken, som har ett gemensamt ansvar för innehållet. Ansvaret för den tekniska uppbyggnaden och utvecklingen ligger hos NetLab, Lunds universitets bibliotek. BIBSAM har genom ansvarsbiblioteksmedel stått för en stor del av uppbyggnads-, samordnings- och utvecklingskostnaderna.

ESSEL, Earth Science Sweden Electronic Library, började byggas upp inom ramen för ett särskilt projekt som startade i slutet av 1995. Ambitionen var att skapa ett nationellt elektroniskt bibliotek för geovetenskaperna. Första delen av projektet redovisades till BIBSAM våren 1997. Något ansvarsbibliotek för geovetenskap fanns då ännu inte utsett. I rapporten till BIBSAM begärdes fortsatt stöd för uppbyggnad och stabilisering av den databas som hade skapats, vilket också beviljades. Arbetet pågick under 1997 men gick sedan i stå. I sin behovsanalys hösten 1999 föreslog det nya ansvarsbiblioteket att arbetet med ESSEL skulle återupptas. Detta skedde också under år 2000. För närvarande arbetar ansvarsbiblioteket med att i samarbete med andra geobibliotek fylla ESSEL med innehåll, men arbetet har hittills gått trögt.

NOVAGate (lantbruk) startade som ett samnordiskt projekt med finansiering från Nordiska Ministerrådet 1997. BIBSAM har finansierat vissa delar av SLUB:s arbete med ansvarsbiblioteksmedel sedan starten.

AGORA (humaniora), Rättskällan (juridik) och SamWebb (samhällsvetenskap) har alla tillkommit inom ramen för ansvarsbibliotekssystemet under de senaste åren, och arbetet med dessa portaler har i viss mån samordnats från början. BIOGATE (biologi) byggdes upp som ett särskilt utvecklingsprojekt, finansierat av BIBSAM, i samarbete mellan Ekologiska biblioteket och NetLab i Lund.

Samtliga ovanstående portaler bygger på kooperativt länksamlande. Den stora länksamling som tidigt utvecklades inom medicinområdet har dock byggts upp helt och hållet inom

Karolinska institutets bibliotek, och kan till stora delar sägas vara en mans verk. Med tiden har länksamlingen förvandlats till en databas, och försök har gjorts att få de övriga medicinbiblioteken att bidra till länksamlandet. Intresset för detta har dock hittills varit ytterst måttligt.

För ekonomi, psykologi och pedagogik samt för kvinno-, mans- och genusforskning saknas ännu så länge en portal. Inom samtliga dessa ämnesområden finns dock omfattande samlingar med kvalitetsbedömda länkar. Handelshögskolans bibliotek, ansvarsbibliotek för ekonomi, har byggt upp en organisation för kooperativ insamling och bedömning av ekonomiska länkar och anordnar också möten för ”länkansvariga”. För psykologi och pedagogik har arbetet med att bygga upp en portal pågått under en längre tid, och en portal för ämnesområdena bör kunna sjasättas under de närmaste månaderna. Inom ansvarsbiblioteket för kvinno-, mans- och genusforskning har diskuterats huruvida man för detta ämne skall bygga upp en separat portal eller om ambitionen i stället bör vara att arbeta för att integrera KMG-länkar i andra befintliga ämnesingångar, eftersom ämnet till sin natur är gränsöverskridande. Ansvarsbiblioteket har nu beslutat sig att bygga upp sin portal i nära anslutning till SamWebb.

Sommaren 2001 presenterade INETMEDIA:s nyhetsbrev vad man kallade sju ”absoluta favoriter” bland virtuella bibliotek och kataloger. Av dessa sju favoriter var tre stycken ämnesportaler uppbyggda inom ramen för ansvarsbibliotekssystemet de senaste åren, nämligen AGORA, Rättskällan och SamWebb. Ett mycket gott betyg, om man betänker mängden ”virtuella bibliotek och kataloger” som idag finns tillgänglig på Internet.

2.2 Internationell utblick

I olika länder pågår för närvarande arbete med att bygga upp nationella elektroniska forskningsbibliotek. Som exempel kan nämnas våra grannländer Finland och Danmark. Där har relativt stora ekonomiska resurser satsas på uppbyggnaden av FinELib respektive DEF. Inom ramen för DEF har man i Danmark byggt upp en företagsekonomisk portal, en musikhistorisk och en portal för ”klinisk information”. Uppbyggnad av fler portaler pågår. I Storbritannien arbetar JISC, the Joint Information Systems Committee, intensivt med uppbyggnaden av DNER, the Distributed National Electronic Resource, ett brittiskt elektroniskt forskningsbibliotek, där uppbyggnad och samordning av ämnesportaler ingår som ett väsentligt inslag i arbetet

I Norge arbetar man för närvarande med uppbyggnad av ämnesportaler i anslutning till BIBSYS-samarbetet.

Försök att samordna ämnesportaler över nationsgränser har gjorts de senaste åren. Ett mindre nordiskt projekt, NISBIG (Nordic Interconnected Subject-Based Information Gateways), finansierades av NORDINFO och pågick under åtta månader 1999-2000. Projektet initierades av DTV, Danmarks Tekniske Videncenter & bibliotek, NetLab och Oslo universitetsbibliotek. I projektet gjordes ett försök att samordna EELS (Sverige) med IPG (Internet Pointer Guide, Danmark) och Juridisk nettviser (Norge) I projektet testades en mjukvara som hade utvecklats inom ett amerikanskt projekt för samordning av ämnesportaler m m, Internet Scout Project (ISP).

RENARDUS är ett pågående EU-projekt, koordinerat av det nederländska nationalbiblioteket, som syftar till att samordna existerande europeiska ämnesportaler. EELS är en av dessa portaler, och NetLab är partner i projektet. Även NOVAGate ingår i RENARDUS genom den

finska partnern, Det vetenskapliga biblioteket i Vik, som är en del av Helsingfors universitet. Företrädare för RENARDUS-projektet är intresserade av att få till stånd ett närmare samarbete med fler svenska ämnesportaler.

2.3 Kostnader

BIBSAM:s totalkostnader för uppbyggnad av portaler inom ansvarsbibliotekssystemet under de senaste fem åren, 1997-2001, kan inte ges med absolut tillförlitlighet, eftersom kostnader för portal-arbete i vissa fall ligger inbakade i andra, större poster, t ex i samordnings- och administrationskostnader för ansvarsbiblioteket. Kostnaden under dessa år har dock uppgått till minst 8 miljoner kronor, dvs minst ett helt årsanslag för ansvarsbiblioteken har den senaste femårsperioden använts för ämnesportal-uppbyggnad. De senaste två åren har drygt två miljoner årligen av de tillgängliga ca åtta miljonerna avsatts för att täcka särskilda kostnader i samband med uppbyggnaden och utvecklingen av ämnesportaler. De senaste åren har varit en uppbyggnadsperiod, och man skulle kunna förmoda att utvecklingskostnaderna går ner när samtliga portaler har kommit in i en reguljär driftsfas. Detta motsägs dock av erfarenheterna från äldre portaler i systemet, som visar att det ständigt finns ett behov av kostsam teknisk vidareutveckling. De senaste två åren har t ex vidareutvecklingen av EELS kostat 860 000 kronor.

Att inom ramen för ansvarsbibliotekssystemet fortsätta att finansiera vidareutvecklingen av sinsemellan olika tekniska plattformar för ämnesportalerna medför merkostnader som är svåra att motivera. På sikt finns det stor risk att systemet utarmas genom att portal-utbyggnaden tar allt större andel av de tillgängliga resurserna för ansvarsbiblioteken. Det innebär att andra angelägna utvecklingsprojekt inte kan finansieras.

3. Existerande systemlösningar

3.1 Kommersiella lösningar

För att undersöka om det finns några på marknaden redan existerande lösningar som skulle kunna användas för att samordna ansvarsbibliotekens portaler så har vi valt ut fyra stycken produkter som kortfattat kommer att presenteras lite närmare. Det är:

- **ENCompass** från Endeavor Information Systems
- **Metalib/SFX** från ExLibris
- **Portalmaker** från Sunstone.
- **Z-portal** från Fretwell-Downing

3.1.1 ENCompass

ENCompass är en ”tilläggsmodul” till bibliotekssystemet Voyager som hanterar elektroniskt material och elektroniska resurser. Den är utvecklad i samarbete med Cornell University och består av flera delar, exempelvis LinkFinderPlus, vilka fyller olika funktioner.

Med hjälp av ENCompass kan man knyta samman ett antal olika databaser. Dessa kan vara i olika format t.ex. Dublin Core, MARC21 eller andra egendefinierade format. Det finns till dags dato inte någon fungerande demo av ENCompass, man är istället hänvisad till beskrivningar och skärmdumpar vilket gör att det är svårt att sätta sig in i dess funktioner.

Kortfattat kan man säga att ENCompass är utvecklat för tre huvudfunktioner:

- administrationsverktyg för databaslicenser och användare
- verktyg för att bygga upp ”lokala digitala samlingar” - där man med detta menar bilddatabaser (inskannade textdokument m.m.)
- verktyg för att länka till elektroniska resurser på artikelnivå – bl.a. med hjälp av standarden OpenURL. Allt detta blir så knutet till en Voyager OPAC som fungerar som en ”portal” till alla resurserna

Mer information finns på <http://encompass.endinfosys.com>

Att välja ENCompass för en gemensam portal är definitivt att beteckna som ”overkill” då mycket funktionalitet i produkten inte är tänkt för webbresurser utan för andra typer av digitala dokument.

3.1.2 MetaLib/SFX

MetaLib/SFX är två produkter varav SFX även är en del av MetaLib. SFX finns även som en stand-alone produkt och är enkelt uttryckt en länk-server som administrerar vidarelänkning till elektroniska dokument, webbtjänster eller andra Internetresurser. MetaLib är baserat på en Aleph-databas men är i övrigt helt frikopplad från bibliotekssystemet med samma namn.

MetaLib består också av flera andra delar: Universal Gateway, som är en ren samsökningsfunktion via HTTP, Z3950 m.m., samt ResourceStore som med hjälp av ett webbgränssnitt är ett sätt att bygga upp länksamlingar för användaren. Dessa kan annoteras, beskrivas och fördes med Dublin Core eller andra Metadata. MetaLib stödjer också OpenURL.

Se mer av SFX på <http://www.sfxit.com/>
En demo av MetaLib finns på <http://www.metalib.com/V>

Inte heller MetaLib ter sig som ett användbart alternativ för ansvarsbiblioteken. Visserligen innehåller själva portalen de samsökningsmöjligheter som önskas men den så kallade ResourceStore ligger långt från den funktionalitet man behöver.

3.1.3 Portalmaker

PortalMaker är en ren samsökningsprodukt utvecklad av Sunstone tillsammans med produkten IndexMaker. Den innehåller i sig ingen databas i egentlig bemärkelse utan tillhandahåller istället ett gränssnitt för samsökning av olika former av Internetresuser. Den innehåller vissa intressanta funktioner såsom en bevakningsmodul där man kan ställa regelbundna sökfrågor mot vissa intressanta söktjänster (sökprofiler), samordnad resultatlista, olika presentationsformat (såsom e-post, fax) samt konfigurera inställningar för operatörer, tecken m.m.

IndexMakers sökrobot letar upp och hämtar hem alla angivna webbsidor för att därefter automatiskt fritextindexera dessa och göra dem sökbara. För en fungerade demo:

<http://portalmaker.sunstone.se/demo/logindemo.jsp>

Portalmaker löser inte, i sig självt, några av de problemställningar som finns rörande utvecklingen av en samlad ämnesportal, men man kan se att produkten har vissa önskvärda tilläggfunktioner.

3.1.4 Z-portal

Detta är en programvara som utvecklats av det engelska företaget Fretwell-Downing Informatics. Det är en portalprogramvara som kan användas för att söka i z39.50-databaser. Denna kan utökas med andra produkter –man föreslår Z' mbol från danska Indexdata för att även inkorporera databaser som inte har en Z39.50-ingång. Läs mer om Z-portal på <http://www.fdggroup.com/fdi/zportal/>

3.1.5 Slutsats

Efter en översiktlig genomgång av dessa produkter kan vi se att ingen av dem motsvarar de specifika behov som behövs för utvecklandet av en gemensam portal för ansvarsbiblioteken. Både Metalib och ENCompass är dock produkter som innehåller mycket intressanta funktionaliteter inom närliggande områden - som ett parantetiskt tillägg så är man i Finland i slutfasen av sin upphandling för en nationell ämnesportal och i motsats till andra nordiska länder (se vidare 3.2 nedan) anser man att deras nationella ”ämnesportalsproblem” ska lösas genom upphandling av en kommersiell produkt.

3.2 Icke kommersiella lösningar

Under den här rubriken presenterar vi olika initiativ som har tagits inom ämnesportalernas område. Det kan vara såväl olika projekt och/eller programvaror som protokoll och framtagna standarder. Det har tagits en stor mängd initiativ vilka har resulterat i såväl intressanta projekt som användbara programvaror. Några av dem kommer vi att gå igenom utförligare medan vissa av dem enbart kommer att presenteras med en URL.

Inledningsvis så finns det också några äldre men bra översikter av programvaror för att skapa ämnesportaler:

Netlab	http://netlab.lub.lu.se/SBIG_Survey/Survey.html
Ukoln	http://www.ukoln.ac.uk/metadata/roads/product-comparison/
Imesh	http://www.ilrt.bris.ac.uk/discovery/2000/09/imesh/

Vi är medvetna att det kan råda en viss oklarhet när vi presenterar de olika delarna eftersom de begreppsmässigt hakar in i varandra men vi ska dock försöka vara tydliga när vi menar en programvara, ett projekt eller ett protokoll.

3.2.1 Projekt:

Renardus	http://renardus.lub.lu.se/
DEFF 1	http://www.deff.dk/admin/
DEFF 2	http://www.deflink.dk/arkiv/dokumenter2.asp?id=879
Bibsys	http://roads.bibsys.no/ROADS/cgi-bin/ep/
Desire	http://www.desire.org

Om vi börjar med att titta på våra nordiska grannar så har Danmark och Norge valt två olika lösningar. I Norge har man exempelvis valt att använda ROADS för att bygga en nationell portal trots att programvaran inte vidareutvecklas. En trolig förklaring är att BIBSYS har kompetensen att själva anpassa och vidareutveckla programvaran för sina behov. Lanseringen av ämnesportalen har dock skjutits upp ett flertal gånger framför allt beroende på att man har svårigheter att få till stånd ett avtal med OCLC om användningen av Dewey-systemet.

I Danmark har man valt motsatt modell – man har decentraliserade portaler som man numera skapar och underhåller med hjälp av en egenutvecklad programvara. Man använde tidigare sig av harvesting-funktionalitet i den så kallade DEF-Vejviser och med tonvikten lagd på att beskriva de danska bibliotekens resurser, men nu med DEF-portalen övergår man istället till att fokusera sig på rena ämnesgångar.

Ett annat exempel på en distribuerad lösning med ett gemensamt gränssnitt är RENARDUS som nu även innehåller de danska ämnesportalerna. Projektets målsättning är att förbättra tillgången till existerande webbaserade samlingar av kulturella och vetenskapliga resurser i Europa och det valda tillvägagångssättet är att fungera som “mäklare” genom ett gemensamt gränssnitt för sökning och bläddring i distribuerade nationella ämnesportaler.

Renardus är tekniskt komplicerat och här är inte rätt plats att försöka beskriva uppbyggnaden. De nationella ämnesportalerna extraherar normaliserad data enligt en Renardus profil till en lokal Z39.50 kompatibel “Renardusserver”. För att kunna bläddra hierarkiskt i systemet gör man stora insatser för att mappa olika klassifikationssystem till ett gemensamt nämligen Dewey. Vidare har man valt att från de nationella ämnesportalerna mappa till ett antal metadatafält med Country som enda undantaget.

3.2.2 Applikationer:

Imesh	http://www.dli2.nsf.gov/internationalprojects/JISC/IMESHtoolkit.pdf
Scout Portal toolkit	http://scout.cs.wisc.edu/research/SPT/index.html

Def-toolkit 1 http://www.deff.dk/admin/fagportaler/konfighjalp1_5/intro.html
Def-toolkit 2 http://www.deflink.dk/upload/doc_filer/doc_alle/817_Rapport%20DEF-Toolkit.pdf
ROADS <http://www.roads.lut.ac.uk/index.html>

Målsättningen med ROADS var att utveckla en programvara som kan användas för att sätta upp ämnesportaler och utveckla standarder för indexering, katalogisering och sökning samt att undersöka möjligheterna för samsökning mellan olika ämnesportaler. Programvaran är mycket konfigurerbar och tillåter ett flexibelt angreppssätt vad beträffar systemadministration, men den är ganska komplicerad att underhålla och driva.

ROADS är framtaget av ett antal utvecklare i ett konsortium, och källkoden är fritt tillgänglig vilket gör det möjligt att utveckla systemet vidare. Dock utvecklas inte systemet längre inom JISC utan all utveckling sker av användarna själva. I Sverige har flera ämnesportaler valt ROADS, och utveckling och anpassning till svenska förhållanden har gjorts av NetLab vid Lunds universitets bibliotek och av Sveriges lantbruksuniversitetets bibliotek.

The Internet Scout Project har tagit fram en programvara (Scout portal toolkit) som är så pass ny att den ännu bara finns i beta-version och den är inte allmänt tillgänglig ännu. Av de få dokument som finns tillgängliga så kan man dock se att programvaran bygger på fria resurser såsom Linux, Apache och MySQL.

En mycket intressant programvara är DEF-toolkit som är framtagen i samarbete med det danska företaget Indexdata. Programvaran har använts till att utveckla de sju ämnesportaler som finns i Danmark. Enligt DEF's årsberättelse har det också utarbetats en ny kravspecifikation för nästa version av programvaran. Sedan årsberättelsen skrevs har en version 1.5 utvecklats och i september kommer version 2.0 att vara färdig. I motsats till ROADS så sker det en stabil utveckling av program, sökfunktioner och användargränssnitt

3.2.2 Protokoll och standarder:

Z3950 <http://lcweb.loc.gov/z3950/agency/>
OAI <http://www.openarchives.org>
Dublin Core <http://dublincore.org/index.shtml>
URN <http://www.kb.se/bus/dc/dcstart.htm>

Oavsett vilken lösning man använder för en ämnesportal så är anpassningen till standarder mycket viktig och bland dessa intar naturligtvis Dublin Core en särställning. Vi behöver kanske inte säga så mycket mer om detta än att de beskrivna resurserna naturligtvis ska märkas upp med en lämplig metadatamodell som kan kompletteras exempelvis med URN.

OAI, The Open Archives Initiative, utvecklar och marknadsför standarder för interoperabilitet som befrämjar en effektiv spridning av "innehåll". Det ursprungliga syftet med OAI var att tillhandahålla en infrastruktur för interoperabilitet bland webbplatser med stöd för egenarkivering och därigenom lansera dem brett. "OAI" är en kombination av organisatoriska principer och tekniska specifikationer som förser innehållsleverantörer med relativt enkla mekanismer för att göra deras metadata externt tillgängliga. Denna externa tillgänglighet gör det sedan möjligt för tjänsteleverantörer att utöka funktionaliteten med hjälp av den information som gjorts tillgänglig från andra som implementerar OAI.

3.2.3 Z39.50-applikationer:

Zebra/ZAP/Yaz-php <http://www.indexdata.dk>

Gate-Z	http://www.rdn.ac.uk/projects/gate-z/
ZBig-server	http://tsadev.speciesanalyst.net/Z.X/
Cheshire II	http://cheshire.lib.berkeley.edu/index.html
Rudolf	http://rudolf.opensource.ac.uk/about/contents.html

På detta område finns det en nästan oöverskådlig mängd material och det har under de senaste åren skapats en mängd applikationer för Z39.50-sökningar. Nästan alla ämnesportalsprojekt har på ett eller annat sätt använt eller undersökt protokollets möjligheter.

De flesta applikationerna bygger på ett eller annat sätt på programvara utvecklad av det danska företaget Indexdata som själva tillhandahåller ett antal olika programvaror som man kan ladda ner från deras hemsida. Exempelvis ZAP som man kan installera på sin webbservare och skapa en egen Z39.50-gateway inom loppet av ett par minuter och ZEBRA som är en programvara för att indexera och söka i fritext med en Z39.50-server.

Programvaran ZBig-server såg intressant ut ett tag eftersom den är speciellt inriktad på Z39.50-sökning i ämnesdatabaser, men det verkar som om utvecklingen av programvaran har avstannat. Liknande funktionalitet kan också fås i "freeware-produkter" från Indexdata.

3.2.4 Harvesting:

Pavuk	http://www.idata.sk/~ondrej/pavuk/
Combine	http://www.lub.lu.se/combine/

För att automatiskt indexera webbsidor på i förväg utpekade webbservrar använder man sig av en harvester. Vi har tittat på två stycken: Pavuk som används i ASF (Advanced Search Facility) och Combine som utvecklades inom DESIRE-projektet och används som del i DEF-toolkit och inom SAFARI. Om man ska använda harvesting som ett komplement till beskrivna ämnesresurser, exempelvis för att hämta hem OAI-märkta dokument, så är båda dessa programvaror fullgoda val.

3.2.5 Sammanfattning

Med en återkoppling till de svenska ansvarsbibliotekens behov av en gemensam teknisk plattform så finner vi att ur denna översiktliga genomgång av diverse icke-kommersiella projekt och programvaror så är det mest intressanta det som sker i Danmark. Av tillgänglig dokumentation kan man se att man lagt ner stor möda på den "katalogiseringsmodul" som man använder till sina portaler samt att man arbetar tillsammans med Indexdata för distribuerade samsökningar. DEF-toolkit är en programvara som är under utveckling och den verkar täcka in en hel del av de svenska behoven - dock ser man att den även är mycket komplicerad att sköta. Vi ser också med intresse att Novagate överger ROADS för att satsa på DEF-toolkit som man planerar att byta till vid årsskiftet 2002/03.

Vi anser vidare att man vid utformningen av kravspecifikationen bör försöka att harmonisera en gemensam plattform med Renardus bland annat genom användningen av Z39.50.

3.3 Ansvarsbiblioteken nuvarande lösningar

Detta material bygger på den enkät som skickades ut under 2001 där man ville försöka få en samlad bild av ansvarsbibliotekens utbud av ämnesingångar. Under förarbetet till denna rapport ställdes en del tekniska frågor angående ansvarsbibliotekens utveckling av ämnesportaler. Nedan följer en kort beskrivning av vilka tekniska lösningar som har valts.

EELS, NOVAGATE, SLU, Handelshögskolan, KIB, ESSEL (Geovetenskap), Samwebb Genusforskning, AGORA, Rättskällan, Pop-webb, BIOGATE, Arbetslivsinstitutet

Handelshögskolan har ännu ingen strukturerad länksamling utan väntar på arbetsgruppens rekommendationer så deras svar har mer karaktären av deras önskemål. Psykologisk-Pedagogiska biblioteket kommer att använda samma databas som SUB använder för Rättskällan för att skapa POP-webb och samma sak gäller med Genusforskning som kommer att använda sig av Samwebb för sina länkar.

SLU har lämnat två svar: ett för NOVAGATE och ett för sitt elektroniska referensbibliotek. Arbetslivsinstitutet har lämnats utanför i denna sammanställning då man har sin databas-lösning under uppbyggnad.

När enkäten gjordes så hade biblioteken följande databaslösningar:

3	ROADS (ASCII text)
3	Microsoft SQL Server
1	Oracle
1	Glimpse (ASCII text ?)
1	Postgresql

Samtliga av dessa kan också tillhandahålla poster i något vanligt text-format. Biblioteken har anpassat sina länksamlingar så att de ska passa just sina användares behov.

Biogate använder sig också av harvesting - en databas är uppbyggd med hjälp av en robot som hämtar hem intressanta webbplatser. Dessa sidor som ligger i en Zebra-databas (Indexdata) kan sedan samsökas med hjälp av ZAP (också en produkt från Indexdata) med de manuellt inlagda resurserna.

3.3.1 Sammanfattning

Utifrån enkätsvaren och från studier av ansvarsbibliotekens befintliga lösningar så har vi en god bild ämnesingångarnas uppbyggnad och struktur. Inmatningsformulär, metadataprofiler och databasarkitektur är några av de områden där man bör ta tillvara redan utfört arbete för att på så sätt inte uppfinna hjulet på nytt eller att göra om samma misstag. Det har framkommit synpunkter på att man skulle kunna undersöka de befintliga lösningarna närmare och se om man kan ta någon av dem "rakt av" för den gemensamma databasen. I detta fall skulle en i sammanhanget oberoende part titta närmare på och välja ut någon av exempelvis Agora, Samwebb eller Rättskällan.

Hur man än går vidare så är det mycket viktigt att den lösning man väljer för en gemensam plattform tar tillvara de erfarenheter och det arbete som redan utförts av ansvarsbiblioteken.

4. Innehåll

I denna grova tekniska kravspecifikation kommer vi inte att ha några synpunkter på innehållet i ämnesportalerna. Vilket innehåll som portalerna fylls med är upp till ansvarsbiblioteken. Vi kommer däremot att kommentera innehållsfrågor där vi ser att det har implikationer för en gemensam ämnesportal. Detta avsnitt kommer att handla om hur innehållet i ämnesportalerna beskrivs och struktureras.

4.1 Innehållsstruktur

4.1.1 Ämnesrubriker och klassning

Innehållet i en multidisciplinär ämnesportal måste kläs i en övergripande struktur som täcker in ansvarsbibliotekens olika ämnesområden och som kan användas för att skapa en allmän ingång till portalen. Under denna övergripande struktur skapar man sedan den hierarkiska finfördelningen utifrån ansvarsbibliotekens befintliga ämnesresurser.

Det finns några olika sätt som man kan göra detta på: antingen så definierar man ett övergripande skikt av rubriker som delar in och avgränsar de olika portalerna – exempelvis att man under rubriken *Samhällsvetenskap* återfinner hela Samwebb. Man kan också tänka sig att man skapar en större finfördelning av dessa övergripande rubriker så att man under – exempelvis - rubriken *Bok- och biblioteksväsen* finner rubriker hemmahörande under *Bokväsen och skrift* i Agora samt *Biblioteks- och informationsvetenskap* i Samwebb.

Biblioteken har redan definierat sina egna ämnesrubriker och hierarkier och detta är naturligtvis strukturer som måste bevaras i den gemensamma databasen. För ämnesöverskridande rubriker (ämnen) måste man skapa mappningar mellan dessa olika platser i hierarkierna.

I ovanstående exempel med *Bok och biblioteksväsen* så kan ett sådant val av rubrik endast avspegla den allmänna ingången till databasen medan i ”Agora”-ingången så finns enbart *Bokväsen och skrift*.

Vidare bör man även ha klart för sig hur den hierarkiska finfördelningen ser ut i de olika portalerna – hur många nivåer det finns?

En annan fråga är den språkliga: önskemål har kommit från flera bibliotek om att man vill ha sin ingång både på svenska och på engelska. Några av biblioteken har redan arbetat fram mappningar mellan sina svenska och engelska rubriker så dessa är det bara att ta tillvara eftersom flerspråkighet måste stödjas i databasen från första början. Vi kan dock se ett problem är med de bibliotek som saknar svenska rubriker (Essel, Biogate, Novagate och Diseases and Disorders). Om man väljer att inte mappa mellan rubriker utan att exempelvis lägga hela Biogate under Biologi så skulle man enklare kunna använda sig av den ämnesspecifika ingången.

Här finns också mycket arbete utfört av ansvarsbiblioteken som man absolut bör ta tillvara och gruppens arbetsuppgift blir mer att sammanställa detta och att se vilka behov av översättningar och mappningar som finns. Detta är dock ett arbete som vi bedömer inte behöver vara klart innan den slutgiltiga kravspecifikationen påbörjas utan kan löpa parallellt med detta arbete.

4.1.2 Metadata

När det gäller mappningar mellan metadata visar sammanställningen av enkätsvaren att det enbart är Titel och URL som är täckta genom alla 10 ämnesportaler och därefter följer Beskrivning, Resurstyp och Upphovsman som täcks in av ca 70-80 % av portalerna. Inom ämnesområdena humaniora, samhällsvetenskap, juridik finns en större täckning av metadata – och förmodligen även inom ekonomi, psykologi-pedagogik och genusforskning.

Vi föreslår att man tar fram gemensamma riktlinjer för metadatamärkning av posterna i databasen. Här finns mycket arbete redan utfört av biblioteken och det blir mer frågan om en sammanställning av dessa. Man bör dock ha ett flexibelt angreppssätt och utgå både från antalet obligatoriska fält samt att de olika ämnesingångarna befinner sig på så skilda nivåer vad gäller implementering av metadata.

Man bör hitta ett antal ”minsta gemensamma nämnare” i metadatauppsättningen men som inte är begränsande utan som sedan kan byggas på av de bibliotek som önskar en djupare metadatamärkning.

4.2 Ambitionsnivå

När det gäller ambitionsnivån för den gemensamma ämnesportalen finns det åtskilliga frågor att ställa även om vi även här ser att de flesta övervägningar kommer att göras av ansvarsbiblioteken.

- Vilken typ av resurser skall beskrivas?

Till att börja med kan ställa sig frågan om det är webbplatsen eller enskilda resurser som skall beskrivas. En webbplats är naturligtvis en resurs i sig men kan ju bestå, och gör det ofta, av många olika resurser. En enda bild på en webbsida kan vara en viktig e-resurs. Väljer man att bara beskriva resurser på webbplatsnivå kan man komma i ett läge där man finner att en webbplats inte fyller de kvalitetskrav man har satt upp men att det på webbplatsen finns en resurs som mycket väl stämmer överens med kvalitetskriterierna. Men eftersom webbplatsen i sig inte gör det kommer det enskilda dokumentet inte med i portalen.

- Svenska eller utländska resurser?

Skall ämnesportalen innehålla både svenska och utländska resurser? Internationellt byggs det upp portaler inom de flesta ämnen och konkurrensen om användarna ökar. Stora forskarinstitutioner så som MIT och Harvard satsar förmodligen mycket mer i sina portaler än vad vi i Sverige kan göra. I diskussioner som har förts har det påpekats att vi borde använda dessa istället för att lägga ner resurser på att bygga upp egna. De svenska portalerna bör satsa på svenska resurser. Frågan är då hur användarna av de svenska ämnesportalerna kan tillgodogöra sig de ”utländska resurserna”. Väljer man att enbart beskriva svenska resurser i den gemensamma ämnesportalen måste man säkerställa samsökning med de utländska portalerna i fråga – om inte är man tillbaka på ruta ett. En målsättning med den gemensamma portalen är att erbjuda användarna ett integrerat sökverktyg, att kunna göra en sökning i mer än en ämnesportal samtidigt. Att söka efter svenska resurser i en portal för att därefter gå in i

en utländsk portal och göra om sökningen är inte en godtagbar lösning. Lösningar på problemet kan vara Z39.50 eller Open Archive Initiative.

- Tvåspråkig portal?

Samsökning med utländska portaler. Vilket språk skall den nationella ha? Två språk kräver ytterligare mappning

- Hur många obligatoriska fält?

Ju färre obligatoriska fält ju sämre sökprecision. Avvägning mellan hur precisa sökfrågor som skall kunna ställas. Minsta gemensamma nämnare. Var skall ambitionsnivån läggas?

- Skall det finnas gemensamma kvalitetskriterier när det gäller vilka resurser som skall med i portalen?

Vi ser det som en tillgång om det finns gemensamma kvalitetskriterier men den frågan måste i så fall diskuteras och lösas av ansvarsbiblioteken tillsammans.

4.3 Sammanfattning

Under avsnitt 4.1 så framkommer det en del frågor kring den ämnesstruktur som en gemensam portal måste ha. Hur definierar man den övergripande strukturen? Hur mappar vi engelska och svenska rubriker? För att få svar på dessa frågor är vårt förslag är att man tillsätter en mindre arbetsgrupp som får till uppdrag att:

- Definiera den övergripande ämnesstrukturen för den gemensamma databasen
- Samla in och sammanställa ansvarsbibliotekens ämnesrubriker
- Om så behövs – definiera mappningar mellan rubriker
- Samla in och sammanställa redan befintliga mappningar mellan engelska och svenska rubriker
- Klarlägga möjligheterna och behoven av översättning och mappning av ämnesingångar som enbart har engelska rubriker

Det finns också några andra delar som skulle ingå här: exempelvis klassningar eller harmonisering av namn på resurstyper. Att man vid insamling och sammanställning av ämnesrubriker även tar med detta - kanske kan man för detta ändamål utnyttja delar av databasen GEN-praxis som har mappningar mellan klassning och SAB-rubriker <http://appl.libris.kb.se/SABrubriker/SABrubriker.html>

Under avsnitt 4.2 har vi redogjort för en del av de komplikationer som kan uppstå när det gäller ambitionsnivå för de olika ingående ämnesportalerna. Vi ger inga rekommendationer när det gäller dessa frågor utan lämnar det till ansvarsbiblioteken själva. Vi tror dock att arbetet med en gemensam ämnesportal vinner på en viss harmonisering i dessa frågor. Möjligtvis kan ovanstående föreslagna grupp även adressera dessa frågeställningar.

5. Gränssnitt

När det gäller gränssnitt finns det ett antal olika frågeställningar att reda ut. Vi har delat upp problemområdena i två, dels producenternas gränssnitt, dvs. de gränssnitt som skall finnas för inmatning av information om resurserna i databasen, och dels användarnas gränssnitt.

5.1 Producenternas gränssnitt

Producenternas gränssnitt, det vill säga gränssnittet för dem som lägger in poster (beskrivningar av webbresurser) i systemet, bör kunna hantera en mängd olika administrativa funktioner.

- Skapa och redigera rubriker
- Skapa, redigera och makulera poster
- Verktyg för redigeringshjälp
- Redigera visningsformatet
- E-post-hantering för administrativa funktioner (länkkontroll mm.)
- E-post-hantering för slutanvändare
- Gränssnitt för statistik – antal nyinlagda poster, ämnesområde, redaktör, bibliotek
- Administration av nyhetsservice
- Ämnesanpassade inmatningsformulär

Vidare bör gränssnittet vara anpassat för att kunna hantera statistik, nyhetsservice, kalendarium etc. Man behöver också kontrollmekanismer vid inläggning och vid uppföljning såsom verifiering av URL:er.

Här finns det dock mycket kunskaper som kan tas tillvara från ansvarsbibliotekens nuvarande lösningar där man har tagit fram bra formulär och hjälpfunktioner. Med tanke på bibliotekens olika nivåer så är det en bra idé att man har olika formulär för olika bibliotek – där så behövs.

5.2 Användarnas gränssnitt

När det gäller användarnas gränssnitt till systemet har arbetsgruppen diskuterat en lösning som består av två olika gränssnitt.

Det första gränssnittet, **ansvarsbiblioteksgränssnittet**, är det gränssnitt som helt och hållet formges av ansvarsbiblioteket. I dag finns det 11 ansvarsbibliotek så om alla utformar sina egna kommer det att finnas 11 ansvarsbiblioteksgränssnitt. Varje gränssnitt får en egen URL (exempelvis: <http://ansvars.bibliotek.se/eko/> för den ekonomiska ämnesportalen) som ansvarsbiblioteken använder för att lotsa sina användare till ”sin” del i ämnesportalen.

Färgval, logotyper, teckensnitt etc. kommer att överensstämma med ansvarsbibliotekets eller moderorganisationens övriga webbsidor.

Det är också upp till ansvarsbiblioteket att lyfta fram de ämnesportaler som man anser bäst passa in i ansvarsbibliotekets informationsstrategi. Exempelvis genom att välja att sökningar i systemet i första hand görs i den egna ämnesportalen samt att erbjuda samsökning i ett urval av de andra ämnesportalerna, och slutligen låta användarna själva välja om ytterligare ämnesportaler skall samsökas.

Bibliotek som inte är ansvarsbibliotek men som har ett intresse av att deras användare kommer direkt till ett av ansvarsbiblioteksgränssnitten kan enkelt lotsa sina användare till önskad ämnesportal

Det andra gränssnittet, **allmänhetsgränssnittet**, är det gränssnitt som användare som befinner sig utanför något av ansvarsbibliotekens nätverk kommer att hamna på. Detta gränssnitt kommer att formges av BIBSAM/LIBRIS i samarbete med ansvarsbiblioteken. Allmänhetsgränssnittet presenterar alla ämnesresurserna på ett neutralt sätt. Varje ämnesresurs presenteras med länkar till det ansvarsbibliotek som ansvarar för ämnesresursen. De olika typer av sökningar som kan göras presenteras.

5.3 Sammanfattning

För att kunna katalogisera resurser i ämnesportalen behövs, naturligtvis, ett inmatningsgränssnitt. Ett absolut krav är att man kommer överens om en miniminivå och därefter utökar gränssnittet beroende på det egna ansvarsbibliotekets krav. De existerande ämnesportalerna har redan lagt ner mycket arbete kring dessa frågor och det är naturligtvis något som man bör ta tillvara i det fortsatta arbetet.

Man bör sträva efter att kunna formge olika användargränssnitt utifrån de olika ansvarsbibliotekens krav. Det finns även behov av ett användargränssnitt som formges för de användare som inte kommer in i portalen via ansvarsbibliotekens gränssnitt.

6. Teknik

Detta avsnitt går igenom tekniska förutsättningar och möjliga lösningar. De förslag som förs fram innehåller olika grader av komplexitet så de tekniska förutsättningarna varierar med vilken modell man väljer, men utgångspunkten för resonemangen i avsnittet tas från två olika vägar: en databasmodell och en samsökningsmodell.

Databasmodellen presenteras mest ingående och finns i två olika ”versioner” – en central och en distribuerad. Det är dels den versionen där man arbetar mot en gemensam databas för registrering av poster och som innehåller ett allmänt gränssnitt mot slutanvändaren samt en specifik ingång med ett egendefinerat gränssnitt för vart och ett av ansvarsbiblioteken. Dels en modell med en gemensam databas där man utbyter poster med import och export av poster mellan den enskilda ämnesportalen och den gemensamma databasen.

6.1 Tekniska förutsättningar

De tekniska grundförutsättningar är att databasen ska kunna lagra poster som beskriver Internetresurser och som sedan ska kunna återvinnas genom strukturerade ämnesingångar. Sökning ska kunna göra på ett ämne, flera ämnen eller alla ämnesresurser.

6.1.1 Databasmodellen

En gemensam databas måste först och främst kunna ta emot ansvarsbibliotekens poster vid en initialladdning och detta innebär att databasen måste skapas på ett sådant sätt att så mycket data som möjligt går att föra över.

Detta betyder också att den gemensamma databasens arkitektur måste definieras utifrån de befintliga portalernas strukturer och att den som har den mest utvidgade ”formatet” måste ligga till grund för denna arkitektur. Ansvarsbiblioteken har ju själva utvecklat sina ämnesportaler men man har kommit olika långt – och på olika vägar. Utifrån svaren på den av BIBSAM utskickade enkäten så finns dokumentation över omfattningen av de olika portalernas tabeller och fält. Där kan man se att det är Agora som har den djupaste beskrivningsnivån och som således får stå modell för den övergripande arkitekturen.

När den övergripande arkitekturen specificerats så den kan lagra alla data så måste man definiera de ”konverteringstabeller” som behövs inför laddningen. Dessa tabeller mappar de olika fälten från de befintliga portalerna till den gemensamma och det finns följande krav på tabeller/fält och funktioner:

- Primärnyckel - länknummer
- Titel – DC.Title
- Undertitel och Alternativtitel
- Filing
- Länk (URL) – DC.Identifier
- Beskrivning – DC.Description
- Beskrivning på annat språk – DC.Description
- Språk/geografisk indelning
- Datumhantering och kontroll
- Standardnummer (URN:NBN)
- Upphovsman/Medupphovsman – DC.Publisher

- Institution – DC.Publisher
- Organisation
- Resurstyp – DC.Type
- Ämnesord – DC.Subject
- Ämnesrubriker/Nivåer/Språk
- Klassning - DC.Subject
- Access/Copyright - DC.Rights
- Felhantering/Länkkontroll
- Redaktör
- Identifikator (Biblioteks/portal-identifikation)
- Statistik – (allmän, antal besök mm.)
- Områdesbevakning (ämnesområde/låntagare/epostadress)

Vidare tillkommer naturligtvis tabeller för de administrativa funktioner som beskrevs i avsnitt 5.1. Några saker kan dock ytterligare beskrivas: Dubblettkontroll – sådan måste ske på unika nycklar såsom URL eller URN, Redigeringshjälp – kopplingar till olika hjälpsystem som kan utvecklas – exempelvis Dewey/SAB, GEN-praxis, Svenska ämnesord men även mer generella hjälpfunktioner.

Till detta kommer ett internt sökgränssnitt för att söka fram redan befintliga poster i portalen. Här kan man tänka sig att man använder enklare sökfunktionalitet eller att man använder delar av det publika gränssnittet. För slutanvändaren måste det dock finnas ett gränssnitt som innehåller sökfunktionalitet såsom booleska operatörer, högertrunkering och frassökning. Vidare måste det finnas möjligheter för slutanvändaren att skapa sig en ”profil” inom ett område eller att på annat sätt koppla sig till en nyhetsservice eller liknande.

Vidare måste gränssnittet vara mycket flexibelt och kunna skapas med lokal profilering - det vill säga olika layouter, logotyper m.m. och där finns också de skript/mallar till webbsidor som behövs för att generera webbsidor med träfflistor, strukturer, poster m.m. Beroende på hur komplexa dessa mallar blir så kan det behövas någon form av utbildning för biblioteken i hur man arbetar med dessa.

Om man ämnar använda sig av den distribuerade databasmodellen så behövs några specifika funktioner för att hantera logistiken vid import och export.

- Uppmärkning av poster
- Skript/program som tar ut de poster som ska exporteras
- Konvertering/mappning
- Laddningsprogram/skript med dubblettkontroll

6.1.2 Samsökningsmodellen

Det kan komma att visa sig att det finns bibliotek som inte kan ingå i databasmodellen av olika skäl och som ett alternativ presenterar vi en samsökningsmodell med hjälp av Z39.50-protokollet. Det finns önskemål om att man ska kunna samsöka den gemensamma portalen med andra resurser så någon form av Z-sökningar bör stödjas.

Exempelvis om det visar sig att den intellektuella och manuella insatsen för att integrera någon eller några av de befintliga ämnesresurserna inte är realistisk så finns denna modell som en tänkbar lösning. Det betyder att man från det gemensamma webbgränssnittet ska kunna komma åt distribuerade resurser utan att det märks att de inte är lagrade lokalt.

Detta är en modell som följer det danska mönstret och innebär att man sätter upp en Z39.50-server lokalt på detta/dessa bibliotek och som knyts till den lokala ämnesportalen. Detta medför också att den lokala portalen kan samsökas med andra resurser såsom den egna bibliotekskatalogen, andra ämnesspecifika databaser mm.

Man kan också använda sig av harvesting för att skapa egna sökbara databaser av andra viktiga webbresurser och göra dem tillgängliga via det gemensamma gränssnittet. Därmed så närmar man sig frågan om automatisk selektion av "kvalitetsgranskade" resurser.

6.2 Teknisk plattform

När det gäller frågan om teknisk plattform så lämnar vi sist i detta avsnitt en rekommendation men vi ändå att frågan hålls öppen för den eventuella framtida kravspecifikationen och till det projekt som senare ska bestämma den slutgiltiga lösningen.

Några viktiga punkter som inte bör glömmas bort är:

- Inköp av hårdvara – maskin/er
- Inköp av programvara – befintlig eller vidareutvecklad
- Konsulter – specialutveckling
- Underhåll av programvara – löpande kostnad
- Central driftsorganisation – konfiguration, backup, ftp, webbserver, Z-server och övrig systemadministration (exempelvis förändringar i exportrutiner)
- Central supportorganisation
- Samarbete med ansvarsbibliotekens egna tekniska organisationer – lokala anpassningar, kompetensnivåer, utbildning
- Kontakt och resurspersoner i ansvarsbiblioteks organisationer

Många har betonat vikten av en central supportorganisation som kan hantera de problem som uppkommer. Man kan också tänka sig att man utbildar och försöker höja kompetensnivån i ansvarsbibliotekens tekniska organisationer – speciellt vad gäller Z39.50 – något som kan vara till gagn för svenskt biblioteksväsende i framtiden.

6.3 Sammanfattning

Vi är av den uppfattningen att den gemensamma databasen måste definieras utifrån de befintliga ämnesportalernas strukturer och att den ämnesportal som har den djupaste beskrivningsnivån får ligga till grund för den slutliga strukturen.

Vi tror att en distribuerad modell med import och export mellan en databas och de befintliga ämnesportalerna är en allt för komplicerad lösning och vi tror att en helt gemensam och central databas, tekniskt sett, är en mer framkomlig väg.

7. Förslag till fortsatt arbete

Som vi ser det så finns det tre olika vägar man kan ta i det fortsatta arbetet med en gemensam databas för ämnesportaler:

1. Att vidareutveckla någon av de befintliga portalerna
2. Att använda en befintlig portalprogramvara
3. Att utveckla programvaran och den tekniska arkitekturen från grunden

7.1 Vidareutveckling av någon av de befintliga portalerna

Det första alternativet innebär att man låter en oberoende part och/eller representanter för ansvarsbiblioteken tillsammans gå igenom de befintliga egenutvecklade lösningarna utifrån de behov som är uppställda i detta dokument med målsättningen att välja en av de tre lösningarna (Samwebb, Agora eller Rättskällan) för det fortsatta arbetet. I detta alternativ ingår också att man ser på vilka utvecklingsmöjligheter som systemen kan ha exempelvis för att hantera harvesting, samsökningar eller andra funktionaliteter som kan anses viktiga. För att detta skall kunna vara ett alternativ måste programvaran stödja möjlighet till individuella anpassningar av användargränssnittet.

7.2 Användning av befintlig portalprogramvara

När det gäller det andra alternativet, där man använder sig av en befintlig programvara, så ser vi att den enda portalprogramvaran som är under utveckling och som är harmoniserbar med Renardus är DEF-toolkit. Det som vi ser som ett problem är att det finns för lite dokumentation om programvaran, dess användning och vilka problem och begränsningar som den kan tänkas ha för att man ska kunna skapa sig en uppfattning om dess användbarhet. Om man ska vidare på denna linje är vårt förslag att man kontaktar de danska utvecklarna och att man begär att få tillgång till programvaran för utvärdering. Denna utvärdering innebär att man installerar programvaran och sätter upp en test-portal utifrån de kriterier som är uppställda i detta dokument. Användargränssnittets flexibilitet måste också undersökas så att målsättningen med särskilda ingångar för vart och ett av ansvarsbiblioteken kan uppfyllas eftersom vi till skillnad från den danska modellen förespråkar en gemensam databas uppbyggd kring DEF-toolkit. I DEF-toolkit finns redan harvesting inbyggt i form av programvaran Combine, som utvecklats inom DESIRE-projektet och används av SAFARI.

7.3 Utveckling av egen programvara och teknisk arkitektur

Det tredje alternativet har sina för- och nackdelar. Det kan ta längre tid att ta fram en sådan lösning än med de föregående alternativen. Det kan också vara en fördel så man bygger in önskad funktionalitet från grunden. En möjlig negativ effekt av denna lösning kan vara att man inte kan ta tillvara lika mycket av ansvarsbibliotekens redan gjorda arbete.

Om man väljer detta alternativ rekommenderar vi att man i den framtida kravspecifikationen utgår från den portal som har den ”djupaste beskrivningsnivån” och att denna får ligga till grund för den övergripande databasarkitekturen. Den valda modellen för den övergripande ämnesstrukturen ska sedan infogas i det gemensamma lagringsformatet. Vi rekommenderar vidare att man väljer den centraliserade modellen utan import- och exportfunktionalitet och att man till portalen implementerar en Z39.50-server.

När det gäller den tekniska plattformen och dess lösning så rekommenderas att man anskaffar en Intel-baserad server på vilken man installerar Linux som operativsystem. Detta är visserligen ett operativsystem som kräver en viss erfarenhet men det är snabbt, billigt och det finns funktionalitet och programvara som är intressanta för våra behov.

Till databas förslår vi att man använder MySQL för Linux – det är en väl utvecklad databashanterare som är gratis tillgänglig. Den kan utan problem hantera de datamängder som det rör sig om i detta fall. Vi rekommenderar vidare att webbsidorna utvecklas i skriptspråket php och att man använder Apache som webbservar. Fördelarna med att använda php är att det finns inbyggd funktionalitet mot MySQL som skulle kunna användas och att man med hjälp av Indexdatas produkt Yaz/PHP kan bygga in Z39.50-funktionalitet från webbsidorna mot andra resurser. Vi föreslår också att man implementerar Indexdatas Z39.50-server Zebra mot portalen för att göra den åtkomlig via Z39.50. Funktion för harvesting av utvalda webbresurser bör också utvecklas.

7.4 Sammanfattning och förslag

Vi föreslår att man i första hand väljer förslag 2, det vill säga att man närmare utvärderar DEF-toolkit som grund för en gemensam ämnesportal. Detta med anledning av att programvaran är utvecklad och testad, katalogiseringsmodul och kopplingar till RENARDUS existerar. Skulle det efter analys visa sig att programvaran motsvarar de svenska ansvarsbibliotekens behov finns det möjlighet till snabb implementering av systemet och uppbyggnad av en gemensam ämnesportal. Vi tror att det finns möjligheter till support av programvaran från det danska utvecklarna, vi tror också att det finns stora möjligheten att få tillgång till källkoden för egen utveckling.

För att inte förlora tid om det skulle visa sig att DEF-toolkit inte håller måttet är det viktigt att BIBSAM och LIBRIS fortsätter hålla sig uppdaterade om vad andra aktörer på marknaden (t.ex Endeavor och ExLibris) gör när det gäller utveckling av verktyg för uppbyggnad av portaler och sökning i elektroniska resurser.

Kontakt har tagits med Biblioteksstyrelsen i Köpenhamn och en första förfrågan angående villkor för användning av DEF-toolkit är ställd. De första svaret var positivt men mer ingående förhandling måste naturligtvis göras i laga ordning.

Som det visat sig i detta dokument så finns det en hel del frågor kvar när det gäller en eventuell samordning av ansvarsbibliotekens ämnesportaler. Kvarvarande frågor är av två slag, dels frågor som rör de tekniska aspekterna och dels frågor har att göra med portalernas innehåll och administration.

För det fortsatta arbetet så förslår vi att två arbetsgrupper utses

Arbetsgrupp 1 har som sin uppgift sätta upp, testa och utvärdera DEF-toolkit och relaterade tekniska frågeställningar så som harvesting och samsökning. Förutom representanter för ansvarsbiblioteken och KB bör kontakter tas med NetLab vid Lunds universitets bibliotek som har stor erfarenhet och kompetens när det gäller ämnesportaler. Flera av dem som var med vid utvecklingen av DEF-toolkit arbetar idag på Netlab. NetLab har också en stor roll i utvecklingen av RENARDUS.

Arbetsgrupp 2 tittar på frågor som inte rör den tekniska delen såsom:

- Samordning av kvalitetskriterier, metadataprofiler
- Definition av den övergripande ämnesstrukturen
- Insamling och sammanställning av ansvarsbibliotekens ämnesrubriker
- Definiera mappningar mellan rubriker
- Insamling och sammanställning av existerande mappningar mellan engelska och svenska
- Rubriker
- Klarlägga vilka möjligheter och vilka behov som finns av att översätta och mappa
- ämnesingångar som enbart har engelska rubriker

Skulle det däremot visa sig att DEF-toolkit inte håller för de svenska behoven så föreslår vi att man då går vidare med förslag 3, en egen utveckling av programvara och teknisk arkitektur utifrån en genomarbetad teknisk kravspecifikation.

8. Bilagor

Bilaga 1 Projektbeskrivning

Kungl biblioteket
Kjell Nilsson
2001-11-09

Projektbeskrivning för förstudie angående samordning av ansvarsbibliotekens ämnesportaler.

Bakgrund

BIBSAM har under de senaste fem åren lämnat ekonomiskt stöd till flertalet ansvarsbibliotek för uppbyggnad av s.k. ämnesportaler. Totalt har drygt 8 miljoner kronor betalats ut i detta syfte. Arbetet har mestadels skett på försöksbasis, och på olika tekniska plattformar.

I september 2000 tillsattes en arbetsgrupp med deltagande från flertalet ansvarsbibliotek och med uppgift att se över möjligheterna till samordning av arbetet med portalerna. I maj 2001 utfärdade gruppen enhälligt en rekommendation att detta arbete i fortsättningen skulle samordnas i en gemensam databas, något som ansågs ge klara fördelar i form av lägre kostnader för såväl registrering som utveckling, samt förenklad samsökning av portalerna.

Vid ett möte den 25 september i år med företrädare för BIBSAM, LIBRIS och samtliga ansvarsbibliotek inbjöds dessa senare att anmäla intresse för deltagande i ett projekt med målet att åstadkomma en sådan samordning. Tio ansvarsbiblioteksområden (av sammanlagt elva) inkluderas i projektet. KTHB (Teknologi med grundvetenskaper) har valt att stå utanför.

Syfte

Att framställa en grov teknisk kravspecifikation för en gemensam databas för ämnesportaler samt ett förslag till plan för fortsatt arbete.

Specifikationer och avgränsningar

En översiktlig värdering skall göras av möjligheterna att använda befintlig programvara. Möjligheter till samordning med den planerade e-tidskriftskatalogen i LIBRIS skall beaktas.

Tidplan

Arbetet skall vara slutfört senast den 15 mars 2002.

Bilaga 2 Projektorganisation

Styrgrupp: Mats Herder (LIBRIS), Gunilla Jonsson (KB/IDA), Kjell Nilsson (BIBSAM)

Uppgifter: att på ett övergripande plan leda och följa upp arbetet samt se till att erforderliga resurser ställs till förfogande.

Projektgrupp: Henrik Åslund (BIBSAM; projektledare), Tomas Friberg (LIBRIS).

Uppgifter: utföra det dagliga arbetet med förstudien; projektledaren leder arbetet och rapporterar fortlöpande till styr- och referensgrupperna.

Referensgrupp:

Gullbritt Åhman - Göteborgs universitetsbibliotek (*Samhällsvetenskap samt Kvinno-, mans- och genusforskning*)

Peter Lindgren - Arbetslivsbiblioteket (associerat ansvarsbibliotek för *Arbetsmiljö*)

Erik Stattin - Karolinska institutets bibliotek (*Medicin*)

Per Åkerlund - Sveriges lantbruksuniversitetets bibliotek (*Lantbruk, skogsbruk, veterinärmedicin och miljövard*)

Marjatta Hauska - Handelshögskolans i Stockholm bibliotek (*Ekonomi*)

Åse Paulsson - Lunds universitets bibliotek, Ekologiska biblioteket (*Biologi*)

Eva Müller - Uppsala universitetsbibliotek (*Geovetenskap samt Humaniora*)

Catarina Ericson-Roos - Stockholms universitetsbibliotek (*Juridik samt Psykologi, pedagogik och folkbildningsforskning*)

Uppgifter: att fortlöpande lämna synpunkter på projektgruppens arbete.