

Teknisk samverkan mellan svenska ämnesportaler – Test och utvärdering av portalprogramvaran TKL

*Colm Doyle, Jörgen Eriksson,
Sigfrid Lundberg och Tomas Schönthal*

Biblioteksdirektionen
Lunds universitets bibliotek

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 ansvarsbiblioteken 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.

Kungl biblioteket/BIBSAM presenterade i maj 2002 förstudierapporten "Samordning ansvarsbibliotekens ämnesportaler. Grov teknisk kravspecifikation", skriven av Tomas Friberg, LIBRIS-avdelningen, och Henrik Åslund, BIBSAM. <<http://www.kb.se/bibsam/utredn/amnesportaler.pdf>> I rapporten föreslås bl.a. en utvärdering av programvaran DEF Toolkit som används för hantering av ämnesportaler inom ramen för Danmarks Elektroniska Forskningsbibliotek.

Uppdraget

Den 4 oktober 2002 beviljades Biblioteksdirektionen vid Lunds universitet medel för installation och testdrift av nämnda programvara samt deltagande i utvärderingen.

Syfte med projektet

Att fastställa programvaran DEF Toolkits lämplighet för uppbyggnad av de svenska ansvarsbibliotekens gemensamma ämnesportal.

Genomförande

Arbetet har delats in i följande moment:

- utarbetande av ett kriterieunderlag som utgör grunden för testen. Kriterieunderlaget översattes därefter till ett testschema med en översiktlig lista av punkter som skulle testas av antingen Lund eller referensgruppen under testfasen.

- tekniskt förarbete; installation, konfiguration och förberedelser för test av programvaran. Dessutom har en arkiverad mailinglista satts upp <<http://www.lub.lu.se/def-ref/>> och en projektwebbplats skapats <<http://www.lub.lu.se/knowtech/projects/toolkit/>>. Webbplatsen ger tillgång till all dokumentation och testinstallationer.
- test; (bedömning av möjligheter till förändring av gränssnitt, hantering av flera ämnesområden i samma tjänst, olika registreringsformulär, import av befintliga data från existerande ämnesportaler m.m.)
- rapport

Projektet att testa och utvärdera programvaran DEF Toolkit Lite (TKL) har pågått sedan slutet av januari 2003. Avsikten var från början att utvärdera programvaran Toolkit Classic (TKC). I januari meddelades från DEF att utvecklingen och anpassningen av TKC avbrutits och inte skulle gå vidare på grund av att leverantören av TKC bedömdes att inte kunna uppfylla sina åtaganden. I stället valde DEF att ytterligare vidareutveckla TKL. Projektgruppen för DEF Toolkit-projektet beslöt då att gå vidare med en utvärdering av enbart TKL.

Projekt- och referensgrupper

Projektgruppen

Teknisk förberedelse och test av programvaran genomfördes i Lund av en grupp bestående av Colm Doyle, Jörgen Eriksson, Sigfrid Lundberg och Tomas Schönthal. Arbetet har följts av regelbundna möten med projektgruppen som består av följande representanter från Lund, Bibsam och Libris.

- Colm Doyle (LUB)
- Karl Isaksson (Bibsam)
- Ulla-Britt Nordin Siebolds (Bibsam)
- Kristin Olofsson (Libris)

Referensgrupp 1

Förutom projektgruppen beslöt man att knyta en referensgrupp bestående av representanter för ansvarsbiblioteken till projektet. Syftet har varit att ge deltagarna möjlighet till att testa programvaran och att ge synpunkter utifrån "sitt" ansvarsbiblioteks synvinkel. Huvudparten av deltagarna i referensgruppen har anknytning till de ämnesportaler som har byggts inom det svenska ansvarsbibliotekssystemet. De har finansierats av Bibsam inom ramen för uppgiften att sammanställa systematiska ingångar till kvalitetsbedömda informationsresurser på Internet/WWW. Samtliga ansvarsbibliotek erbjöds att delta. Undantaget i referensgruppen är Malmö högskola som har fått medel från Bibsam för att utveckla sin "Länksamling för pedagoger" som i nuläget är uppbyggd med statiska html-sidor till en ämnesportal.

I referensgrupp 1 deltar följande:

- Maria Cederblad: (Malmö Högskola) - Länksamling för pedagoger

- Ann-Sofi Cullhed (Uppsala UB) - Agora < <http://agora.ub.uu.se/> >
- Ylva Gavel, (Karolinska institutets universitetsbibliotek), Diseases, Disorders and Related Topics, Biomedical Information Resources <<http://www.mic.ki.se/Diseases/index.html>> <<http://www.mic.ki.se/Other.htm>>
- Kjell Jansson (Kungl Tekniska högskolans bibliotek), - fd. EELS
- Åse Paulsson (Biologibiblioteket, Lund) - Biogate < <http://biogate.lub.lu.se/> >
- Alexandra Machol Ravnskov (Stockholms UB), - Rättskällan <<http://www.sub.su.se/rattskallan/home.asp>>
- Gullbritt Åhman (Göteborgs UB) – Samweb <<http://www.ub.gu.se/samwebb/>>
- Per Åkerlund (Sveriges lantbruksuniversitets bibliotek) - NOVAgate <<http://novagate.nova-university.org/nova.html>>

Referensgrupp 2 - policy, innehåll och kvalitet

Projektgruppen samt referensgrupp 1, som redan bildats för utvärderingen av TKC och TKL, fick i uppgift att sammanställa en teknisk kravspecifikation. Det finns en rad frågor som det var viktigt att särskilja från den tekniska utvärderingen, men som likväl måste undersökas närmare. I Friberg och Åslunds rapport föreslogs det att titta på frågor 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

Därför tillsattes Referensgrupp 2 med uppdraget att föreslå lösningar på de "mjuka" frågorna.

Det övergripande syftet är samma för båda grupperna, dvs att fastställa programvaran DEF Toolkits lämplighet vad avser en gemensam databas för de svenska ansvarsbibliotekens ämnesportaler. Syftet för grupp 2 är att undersöka de generella koordineringsbehov för en gemensam svensk ämnesportal som finns och att lämna förslag på hur de bäst löses. Uppdraget utgår från BIBSAM. Gruppen består av Gullbritt Åhman, GUB (sammankallande), Inga Nyman, SUB och Lena Åkerbäck Ågren, UUB.

Referensgrupp 1 skall lämna en rapport om den tekniska utvärderingen den 31 maj 2003. Det är tänkt att Referensgrupp 2 ska kunna lämna ett policydokument innan midsommar. Styrgruppen för DEF Toolkit-projektet har sedan att besluta om DEF Toolkit Lite kan användas för en gemensam svensk ämnesportal för de ingående

ansvarsbiblioteken. Beslutar styrgruppen att gå vidare med TKL fortsätter Referensgrupp 2 sitt arbete under hösten 2003 med målet att TKL kan implementeras och tas i drift i början av 2004.

Allmän beskrivning av testförloppet

Testet delades upp i två områden.

a) LUB skulle testa och bedöma generella krav som t ex serverarkitektur, möjligheter till förändring av gränssnitt, hantering av flera ämnesområden i samma tjänst, olika registreringsformulär, import av befintlig data från existerande ämnesportaler m.m.

b) Referensgruppen testade verktyget utifrån ett testschema baserat på kriterieunderlagets sammanställning.

Testfasen började med ett möte för referensgruppen i Lund 05/03 där vi gick igenom programvaran och testschemat. Under mötet fick alla i referensgruppen tillgång till var sin "portal" att arbeta med. Testet pågick till vecka 15. Kompletterande tester fortsatte 14-25 april. Under testfasen fördes det en aktiv diskussion om verktyget dess begränsningar och möjligheter på DEF-REF mailinglistan. Diskussionen är arkiverad och finns tillgänglig på < <http://www.lub.lu.se/def-ref/> >

Efter referengruppsmötet kom det fram att det fanns behov av en kortfattad manual eller lathund för att underlätta för referensgruppen att sätta sig in i systemet. I detta syfte skapades följande dokument som är tillgängligt i dokumentarkivet <<http://www.lub.lu.se/knowtech/projects/toolkit/dokument.html>>

- "Administration av TKL"
- "Att skapa, editera och makulera poster"

Själva testschemat var indelat i fyra områden

- A. Innehåll i tjänsten
- B. Gränssnitt
- C. Användargränssnitt
- D. Central drift och utveckling

där specifika krav och önskmål var numererade. Anledning till detta var att få ett överskådligt sätt att arbeta med och presentera svaren till listan över krav och önskemål för ett gemensam svensk ämnesportal. Målsättningen var att testa praktiskt eller göra en bedömning av hur TKL lever upp till var och en av punkterna. Kraven och önskemålen skulle testas på olika sätt beroende på deras natur. I vissa fall handlade det om att testa programvarans egenskaper eller bakomliggande strukturer medan det i andra fall handlade om funktionalitetstester som genomfördes av referensgruppen.

Samtliga representanter fick en TKL med administrativt och användargränssnitt att arbeta med. Testportalerna är tillgängliga från följande URL. <<http://www.lub.lu.se/knowtech/projects/toolkit/testportaler.html>>. Vi delade upp oss i tre grupper

- Testgrupp 1 = A=AGORA, B=Biogate,
- Testgrupp 2 = S=SAMWEBB, SU=SUB, K=KIB, E=EELS; N=Novagate, MAH,
- Testgrupp 3 = Lund

Testgrupp 1 fick ett administrativt gränssnitt som innehöll poster från deras egna ämnesportaler. Anledning till detta var att bedöma möjligheten att importera poster från befintliga system, i detta fall från SQL och ROADS-baserade tjänster. Vi tittade även på möjligheten att importera resurser från statiska html länksamlingar. Målet var även att testa import av länkar från KIBs länksamling. Efter diskussion och undersökning kom vi fram till att det pga av tjänstens uppbyggnad vilket tillsammans med kopplingen till MESH skulle vara för resurskrävande att få fram en motsvarande TKL portal under testfasen. Resultatet av importen var att Grupp 1 fick en förbättrad möjlighet att jämföra TKL med sin egen befintliga programvara, med egna poster i TKL. De fick därmed testa andra uppgifter än grupp 2.

Testgrupp 2 fick ett mer eller mindre tomt skal nära TKLs default att arbeta med och hade möjlighet att testa portalens grundfunktionalitet.

Under testförloppet höll vi regelbunden kontakt med Indexdata, det danska företaget bakom programvaran. Detta gjordes bla genom att bevaka TKL-utvecklingsmailinglista men även genom specifika diskussioner och frågor när något dök upp under ämnesportalens eller våra egna tester. Vi blev tidigt medvetna om att TKL-programvaran var under snabb utveckling pga de krav och önskemål som de danska ämnesportalerna och den centrala Def-redaktionen ställde på systemet. En ny version skulle släppas under våren men datum för släppet var oklart. Under hela testförloppet och en tid därefter var vi beredda att ersätta vår testversion med den nya för att se om en del av de brister som vi upptäckte blivit lösta. Den nya versionen av det administrativa gränssnittet släpptes till sist mitt i rapportskrivandet, men vi har försökt att undersöka om något har hänt som påverkar slutsatserna från testen.

Toolkit Lite

Bakgrund

Det ursprungliga förslaget från rapporten "Samordning av ansvarsbibliotekens ämnesportaler. Grov teknisk kravspecifikation" var att testa Def Toolkit. I början av uppdraget uppmärksammades att det fanns två parallella utvecklingar som med viss rätt kunde kallas för Def Toolkit, nämligen Toolkit Classic (TC) och Toolkit Lite (TKL). Toolkit Classic var tänkt att ersätta och fullfölja utvecklingen av det verktyg som används i DEFs fagportaler. TKL skulle bli ett mera basalt alternativ för den som inte behövde alla de finesser som TC var tänkt att innehålla. Utvecklingen av Toolkit Classic avbröts på ett mycket sent stadium (i januari 2003) och då beslutade DEF att i stället finansiera en vidareutveckling av Toolkit Lite (TKL) så pass att det skulle kunna användas av de existerande fagportalerna. Den utvecklingen pågår nu under våren 2003. Tanken är att de danska ämnesportalerna i DEF samarbetet skall kunna gå över TKL som portalverktyg under 2003.

Om programvaran

TKL ("Toolkit Lite") < <http://www.indexdata.dk/tkl/> > är ett mjukvarusystem för att driva informationsportaler och andra webbplatser. Enligt Indexdatas beskrivning kan det användas (förutom ämnesportaler) även för organisations- eller personliga hemsidor, dvs ett slags Content Management System. Med TKL leveras en "grundportal", där kan det rymmas en hierarkisk organiserad samling metadata och länkar till externa resurser, men som även kan presentera annat eget innehåll; artiklar, nyheter, mm. TKL är baserat på öppna standarder. TKL-dokument/resurser lagras i XML och presenteras via XSLT-baserade stylesheets. Mjukvaran ska stödja XML schema, SOAP och Z39.50 inom kort.

Alla de programvaror som används i TKL är **portabla** i den mening att de fungerar under flera olika operativsystem. Det vill säga programmeringsspråket PHP, XSLT-processorn Sablotron, XML-programvarorna expat och XML-Lib och sökmaskinen Zebra fungerar alla under olika smakriktningar av Windows NT, Unix och Linux och tillsammans med en webbservrar av olika märken. Däremot är TKL självt bäst testad under olika varianter av Linux och alla av oss kända installationer kör under webbservern Apache.

Vad TKL gör, egentligen

TKL bygger på en lika enkel som god idé. En webbsajt, ämnesportal eller vad som, har en viss uppsättning med dokument av olika typ att leverera till dess användare. Alla filer som skall levereras av TKL ges fil-suffixet .tkl. Webbservern, normalt Apache (se ovan) konfigureras så att filer med detta suffix passerar genom ett program som heter shell.php. Detta program har en rad olika uppgifter, men den viktigaste är att välja en XML-stilmall (skriven i XSLT) som passar ihop med dokumentet. Ett axplock av dokumenttyper som är standard i TKL är

Dokumenttyp	funktion
link	metadata för en resurs
news	nyhetsblänkare
article	längre text (på prosa)
subject	en ämneskategori
search	En sökning i portalen
soap	En "XML-baserad" sökning i en soap-databas
portal	En portals toppsida

Var och en av dessa dokumenttyper knyts till XML-stilmallar, och all väsentlig konfigurering och programmering görs följaktligen i XSLT. Även sökningar. För att, till exempel, möjliggöra access till en Z39.50-databas finns ett antal nya funktioner i XSLT definierade i TKL i form av särskilda länkar (URLar) vilka sedan hanteras av shell.php.

Stilmallen som hör samman med ett dokument av typen 'subject' utnyttjar en funktion som ger alla länkar till underordnade ämneskategorier, och en annan funktion för att få tillgång till alla 'link'-dokument som hör samman med den

aktuella ämneskategorin.

Genom att man kan skapa nya dokumenttyper med en särskild schema-editor blir systemet mycket flexibelt. När planerade förbättringar som uppladdning av filer och ett system för redigering av stilmallar blivit förverkligade är TKL i allt väsentligt ett Content Management System.

Implementeringar

Under senaste halvåret har flera implementeringar av TKL igångsatts. Listan kommer med all sannolikhet att öka under det närmaste halvåret när flera av de danska ämnesportalerna i DEF samarbetet ersätter sina befintliga portalverktyg med TKL. Följande är en lista av befintliga tjänster som använder eller för tillfället utvecklar tjänsten mha TKL:

- Def Portalen < <http://www.deff.dk/> >
- Aalborg Universitet - Tidskriftvejviser < <http://portaler.aub.auc.dk/journals/> >
- DOAJ (LUB) < <http://www.doaj.org/> >
- University of North Texas (Portal om Texas historia, bara påbörjad)
- Material.dk < <http://materials.dk/> >
- Svenskt Resurscentrum för Vetenskaplig Kommunikation ScieCom.org <<http://www.sciecom.org/>>. Betaversionen ligger på <<http://sciecom.lub.lu.se/>> (Släpps inom kort)

Sammanfattning

TKL är under snabb utveckling och därmed har det varit vanskligt att göra en utvärdering av en ganska tidig version av programvaran. Det finns tydliga nackdelar och buggar i produkten som vi återkommer till i följande kapitel men vi vill passa på att nämna några generella kommentarer utöver de tekniska.

Redan nu kan vi konstatera att TKL ser ut att ha en relativt trygg framtid. Den är nära kopplad till den nationella utvecklingen av ämnesportaler i Danmark. Programvaran har skapats i synnerhet för ämnesportaler och används redan idag av flera ämnesportaler eller liknande tjänster. Indexdata är ett känt företag som genom åren levererat flera bra produkter. Ett samarbete mellan danska och svenska nationella lösningar omkring koordination av utveckling av TKL är väl inte en omöjlighet att tänka sig. Sett ur en svensk ämnesportals perspektiv krävs dock en hel del utveckling och överenskommelser för att komma dit hän.

Omvärlden - en orientering

Trenden vad gäller ämnesportaler är som vi uppfattar den på väg mot förenkling vad gäller det manuella arbetet och integration i den meningen att länksamlingen är en mer eller mindre integrerad del i en portal som också erbjuder andra tjänster. En av orsakerna är de förbättringar som de allmänna robotgenererade söktjänsterna på Internet har genomgått vad gäller sökgränssnitt och rankningsalgoritmer (Google!) vilket gör att behovet inte känns lika angeläget. En annan är den relativt höga kostnaden för att manuellt bygga upp en ämnesportal, vilket tillsammans

med bristen på användarundersökningar som visar på att nyttan verkligen uppväger kostnaden har gjort att den manuellt skapade ämnesportalen med rik metadata allt oftare ifrågasätts. Ett löfte som inte heller har förverkligats är de möjligheter till arbetsfördelning och samordning på internationell nivå som de tidiga visionärerna drömde om. Renardus, som är det störst anlagda försöket i den riktningen har nu efter projektfinansieringens slut mycket svårt att hitta någon huvudman som är beredd att finansiera det gemensamma goda. Försök görs också med att skapa hybrider mellan manuellt uppbyggda och robotgenererade ämnesportaler för att få lägre kostnader som tillåter långsiktig drift men ändå samtidigt behåller delar av kvalitetsurvalsmodellen. Den nya versionen av EELS är ett sådant exempel liksom den danska DEF-portalen Material.dk.

Portalverktyg - vad har hänt. Nya verktyg.

Fram till år 2002 var ROADS den enda tillgängliga programvaran specifikt designad för att bygga ämnesportaler. Detta har ändrat sig från 2002 och nu finns det några alternativ vid sidan av TKL tillgängliga som open source. Vi kommer att lista dem här. Vi har inte haft tid inom ramen för den här utvärderingen att testa dem.

- Scout Portal Toolkit < <http://scout.wisc.edu/research/SPT/download.html> > SPT har utvecklats av Internet Scout Project med stöd från Andrew Mellon Foundation och sägs vara en nyckelfärdig lösning för att bygga internetportaler. En presentation av programvaran finns på <<http://www.dlib.org/dlib/november02/almasyl11almasyl.html>>
- iVia < <http://infomine.ucr.edu/iVia/> > iVia har utvecklats i anslutning till University of Californias Infomine och understödjer både manuellt skapande av metadata och robotgenererade index. En presentation av iVia finns på <<http://www.dlib.org/dlib/january03/mitchell/01mitchell.html>>

Kommersiella verktyg

Vad gäller kommersiella verktyg hänvisar vi till genomgången som gjordes i Friberg och Åslunds rapport Samordning av ansvarsbibliotekens ämnesportaler. <<http://www.kb.se/bibsam/utredn/amnesportaler.pdf>>

Det stämmer att produkter som Metalib mm. innehåller intressanta funktioner i närliggande områden men vi kan konstatera att slutsatsen gäller även idag att ingen av de kommersiella lösningar motsvarar de specifika behov som behövs för utvecklandet av en gemensam portal för ansvarsbiblioteken.

Ämnesportaler internationellt - exempel på uppbyggnad av broker-lösningar

Här följer en mycket kortfattad genomgång av hur olika länder/tjänster har hanterat samordning av ämnesportaler.

Storbritannien

Storbritannien satsade tidigt på att bygga upp ämnesportaler som ett av projektområdena inom elib-programmet. Inom ramen för samma projektområde utvecklade man också portalprogramvaran ROADS som länge var det enda fritt tillgängliga programmet för ämnesportaler. I en andra fas har man utvecklat en

samsöknings- och koordinationstjänst, Resource Discovery Network (RDN). Utvecklingen av ämnesportalerna och den centrala tjänsten har finansierats av Joint Information Systems Committee (JISC). RDN < <http://www.rdn.ac.uk/> >

Målgupp: primärt internetanvändare inom högre utbildning och forskning. Bläddring: RDN består av en mycket grund bläddringsstruktur som länkar till de delagande ämnesportalernas bläddringshierarkier. Om man klickar på länken Business i RDN så hamnar man i den avdelningen inom Social Sciences Information Gateway (SOSIG), om man väljer Engineering så hamnar man i EEVL. Man har alltså valt att inte skapa ett gemensamt klassifikationssystem eller ett gemensamt användargränssnitt för bläddring. Sökning: De enskilda ämnesportalerna kan samsökas i en enkel sökruta från RDN. RDN använde ursprungligen Z39.50 för att åstadkomma samsökningen, men har nu övergått till att använda OAI-MHP för att skörda data från ämnesportalerna till en gemensam databas. Databasen är i sin tur möjlig att skörda med OAI-MHP eller samsöka med z39.50. För mera information om teknologin se artikel av Peter Cliff: Building Resource Finder, Ariadne issue 30 <<http://www.ariadne.ac.uk/issue30/rdn-oai/>>.

Portalprogramvara: De enskilda ämnesportalerna använder ROADS eller olika hemmayoutvecklade program för sina portaler. ROADS slutade utvecklas och underhållas centralt i slutet av 90-talet så även de ROADS-baserade tjänster har gjort ganska mycket egenutveckling.

Danmark

I Danmark har det byggts fagportaler inom ramen för DEF-projektet 1999-2002. De enskilda fagportalerna är självständiga i förhållande till huvudportalen men måste uppfylla vissa interoperabilitetskrav vad gäller metadata och teknik.

Den gemensamma ingången finns på < <http://www.deff.dk/> >. Målgupp: primärt internetanvändare inom högre utbildning och forskning. Bläddring: Man har ingen gemensam bläddring utan fagportalerna hittas bland andra resurser under repektive ämne under länken "Faglige links". Sökning: Det finns samsökning som bygger på Z39.50. Ett sökgränssnitt där man kan välja att samsöka mellan grupper av tjänster; DEF fagportaler, Andre Fagportaler och Generelle faglige links (den gemensamma portalens länksamling). Möjligheten "Avancerad sökning" < <http://soeg.deff.dk/?page=search&grp=all> > innebär att de enskilda portalerna är grupperade efter ämne och man har möjlighet att göra en egen kombination av ämnesportaler, både DEF-portaler och andra (I dagsläget främst RDN-portaler).

Portalprogramvara: Inom ramen för projektet använde samtliga fagportaler en vidareutveckling av DTVs Internet Pointer Guide (IPG). Allteftersom nya krav på funktionalitet ställdes under projektets gång bedömde man att gränsen för att lappa och bygga på IPG var nådd och beslöt att under 2002 utveckla en ny programvara (Toolkit Classic) baserat på de samlade funktionalitetskraven som kommit fram under projektets gång. Samtidigt initierade man en utveckling av en enklare variant, Toolkit Lite (TKL). Utvecklingen av Toolkit Classic försenades och i januari 2003 beslöt DEF-projektets ledning att lägga ner utvecklingen. I stället satsar DEF på en utökning av TKLs funktionaliteter bl. a. med stöd för Z39.50 och harvesting i första omgången. Status är att alla fagportalerna utom Materials.dk

fortfarande använder den gamla programvaran i väntan på den nya versionen av TKL. Materials.dk som bygger på robotinsamling av resurser med automatklassificering använde sig av TKL från början. < <http://www.deflink.dk/> >

Norge

Bibsys emneportal < <http://sgate.bibsys.no/cgi-bin/ep> > Målgrupp: studenter och anställda inom universitet och högskolor. Norge valde från början en centraliserad lösning. År 2000 sattes under BIBSYS ledning upp en ämnesportal där universitets- och högskolebiblioteken gemensamt registrerar resurser. Det finns 27 "fagredaktioner" med medlemmar från de flesta BIBSYS-bibliotek och en redaktionsgrupp med överordnat ansvar för tjänsten. BIBSYS står för driften. Portalen täcker alla ämnen. Programvara som används är en modifierad variant av ROADS. Mera om BIBSYS emneportal < <http://www.bibsys.no/roads/> >

Finland

Finnish Virtual Library (FVL) < <http://www.jyu.fi/library/virtuaalikirjasto/engvirli.htm> > Målgrupp: Finska forskare FVL består av något över 60 ämnesingångar som underhålls av 18 finska universitets- och högskolebibliotek. Man har gemensamma urvalskriterier. I övrigt består samordningen av en ingångssida med länkar till de enskilda ämnessamlingarna, en enkel samsökning som inte täcker riktigt alla ämnesingångarna, gemensam hjälp och användarfeedback. Programvara som används är framför allt ROADS. Ett bibliotek använder sig av TRIP Highway och några samlingar är uppbyggda som länklister i HTML.

Tyskland

Die Virtuelle Fachbibliothek < <http://www.virtuellefachbibliothek.de/> > Der Deutschen Forschungsgemeinschaft (DFG) finansierar uppbyggnaden av en gemensam portal och de ämnesspecifika tjänsterna. De sistnämnda byggs upp av forskningsbibliotek runt om i landet. UB/TiB ? koordinerar projektet och man samarbetar bland annat med META-LIB, ett metadata-projekt för samordning av metadata inom tyska bibliotek. Den gemensamma portalen länkar till de enskilda ämnesportalerna och man planerar att införa någon form av samsökning i framtiden.

Nederländerna

Dutchess < <http://www.kb.nl/dutchess/index.html> > Den nederländska ämnesportalen är uppbyggd efter en centralistisk modell med en gemensam portal som täcker alla ämnen. Dutchess utvecklades av Koninklijke Bibliotheek. Efter utvecklingsfasen har flera universitetsbibliotek engagerat sig i uppbyggnaden och man har formaliserat samarbetet. Dutchess använder sig av ROADS som portalprogramvara.

Renardus

Renardus < <http://www.renardus.org/> > Inom EU-projektet Renardus utvecklades en broker-lösning för ett större antal europeiska ämnesportaler. Samsökningen

bygger på Z39.50. Mera intressant är att man har utvecklat sambrowsing genom att mappa de olika portalernas klassifikationssystem mot DDC. Inom ramen för den delen av projektet har metodologi för konsistent mappning tagits fram liksom verktyg för att understödja mappningsarbetet. Renardus framtid efter projektiden är osäker, PICA/OCLC har visat intresse för att finansiera tjänsten men ingenting är klart.

Utbyte av poster mellan tjänster

Eftersom det finns önskemål om utbyte av poster mellan ämnesportalerna har vi också som hastigast tittat på den frågan i förhållande till de ovan beskrivna tjänsterna. Ingen av dem har någon modell/system för utbyte av poster. Frågan var föremål för en WP i Renardus och SUB-Göttingen tog fram en rudimentär programvara och det skrevs en rapport i ämnet. Behovet av utbyte av poster visade sig vara lågt liksom överlappningen mellan tjänsterna.

Sammanfattning

En slutsats man kan dra av denna (ytliga) genomgång är att de länder/tjänster som har utvecklat decentraliserade system har begränsat samsökning i de gemensamma portalerna till i de flesta fall enkel sökning. Där använder man sig av oftast av Z39.50 och i ett fall av OAI-MHP. Egentlig sambläddring förekommer inte i någon av tjänsterna. I stället sorterar man de deltagande tjänsterna efter en grov ämnesindelning och länkar sedan till de enskilda ämnesportalerna.

En annan slutsats är att trots flera initiativ till koordinering, samsökning mm. finns ingen given lösning varken vad gäller programvara för den enskilda portalen eller för nationell/ämneskoordinering av tjänster. Sverige kan mycket väl delta i ett Renardus samarbete och enskilda svenska portaler kan öka samarbete med liknande tjänster i andra länder men det finns dock ingen "killer application" som enkelt löser de "mjuka" samt tekniska problemen på vägen dit. Oberoende på vilken teknisk lösning som väljs måste mål och överenskommelser vara på plats och teknisk utveckling ske. Budskapet är dock tydligt från en analys av förutsättningar i andra europeiska länder - utan en medveteten långsiktig och strategisk koordinering, integrering och nära samarbete är förutsättningen för att driva ämnesportaler på lång sikt minimal.

Testresultat

I det följande presenterar vi slutsatser från de tester som har utförts. Uppdelningen var som följer

- LUB testade generella krav som t ex severarkitektur, möjligheter till förändring av gränssnitt, hantering av flera ämnesområden i samma tjänst, olika registreringsformulär, import av befintlig data från existerande ämnesportaler m.m.
- Referensgruppen testade verktyget utifrån ett testschema baserat på kriterieunderlagets sammanställning.

Schemat är indelat i fyra områden

- A Innehåll i tjänsten
- B Gränssnitt
- C Användargränssnitt
- D Central drift och utveckling

För samtliga områden förs en diskussion och ges sammanfattande slutsatser som bygger på resultaten som presenterats i testschemat. Dessutom har vi lagt korta kommentarer i Bilaga 1 - Testschema samt svarat ja eller nej till hurvida TKL tillfredställer de krav och önskemål som formulerades i testschemat.

A. Innehåll i tjänsten

Grundläggande förutsättningar

Vi utgår från några av de befintliga ämnesportalerna med existerande metadata och klassifikationssystem/ämnesstrukturer.

Testförlopp

Mål för testet var att identifiera de begränsningar och klargöra de möjligheter som TKL-programvaran tillför vad gäller hanteringen av innehållet i de svenska ämnesportalerna. Behandling av import av resurser från befintliga tjänster, möjligheten till export av resurser och stöd för relevanta standarder är frågor som belyses här. Vi undersökte även hur pass väl TKL hanterar komplexa metadatarelationer.

- Import av data
- Export av data
- Ämnesrubriker och klassifikation
- Metadata - Datamodellen

Import av Data och automatisk konvertering av ämnesportaler till TKL

Automatisk konvertering är en procedur för att helst utan manuellt ingripande översätta en ämnesportals samtliga poster från ett format till ett annat, i detta fall till TKL. Det är frågan om en skraddarsydd rutin anpassad till portalformatet (ROADS t ex) och till den aktuella portalens fältuppsättning och benämningar. Det bör betonas att konverteringen i detta fall är experimentell; det är alltså inte meningen att replikera någon ämnesportal fullt ut. Konverteringen översätter naturligtvis heller inte portalens användargränssnitt och layout; det måste göras manuellt.

En sak måste däremot alltid göras manuellt i samband med att en portals data konverteras, nämligen att upprätta ett metadata-schema för posterna.

I det aktuella projektet har två olika portaler konverterats: Agora, <<http://agora.ub.uu.se/>> (MS SQL-baserad). Biogate, <<http://biogate.lub.lu.se/>> (ROADS-baserad). Dessutom övervägde vi att konvertera någon del av KIBs samling av statiska HTML-sidor men avstod från det, när vi efter samtal med KIBs representant fått klart för oss vidden av ett sådant åtagande; det fanns helt enkelt inte någon rimlig möjlighet att klara av något sådant inom ramen för detta projekt.

I båda fallen har ett färdigt användargränssnitt för den dansk/engelska Bibliothecia-portalen utnyttjats, samma som för de experimentportaler, som inte tillhandahåller egna dataposter och bläddringsstrukturer.

I följande avsnitt beskrivs konverteringarna av Agora resp Biogate var för sig. Avslutningsvis sammanfattas konverteringsresultaten.

Konvertering av Agora

Underlaget till Agora består av tabelldumpar, inalles 20 st filer samt en detaljerad dokumentationsbilaga, som bortsett från ett par uppenbara sakfel betr fältens inbördes ordning har varit till stor hjälp för att förstå materialet.

Agoras poster beskriver humanioresurser över hela världen. Användargränssnittet är svenskt/engelskt, medan posterna ofta följer de beskrivna resursernas eget språk. Det bör noteras att posterna är klassade på två sätt: Dels med fullständiga Sabkoder och dels med positioner i en särskild bläddringsstruktur i fyra nivåer. Det är den senare som har realiserats i Unix filsystem med hjälp av underkataloger och s k symboliska länkar.

Bland de konverterade fälten märks identifier (URL), description, MailAddress (till utgivare), title, StandardNumber, (t ex ISBN), SabCode, ResourceType, SubjectWord (fria ämnesord), Institution (utgivande) samt Person (ansvarig utgivare).

Några fältpar har representerats av kombinationen gemensamt XML-element med attribut. Fältparet CreatedDate/EditDate har inte, som sig borde, representerats enligt TKL-standard, som attribut till postens huvudelement, link. Att åstadkomma detta innebär dessbättre endast en trivial ändring av konverteringsprogrammet.

Följande auktoritetsfiler har automatiskt härletts ur materialet: PersonInstitutionType (utgivande institutions/persons roll), ResourceType, Sab (klassifikationsskoder/benämningar), StandardNumberType (kontrollerad vokabulär för fältet StandardNumber) samt SubjectWordType (huruvida ett ämnesord avser ämne, geografisk omfattning eller tidsperiod).

För att åstadkomma detta har 15 av Agoras tabeller utnyttjats. Det har ibland varit omständligt att återge kopplingarna mellan de olika tabellerna, men det har alltid löst sig tack vare Agoras utmärkta fältdokumentation.

Av inalles ca 2.300 poster har drygt 1.700 framgångsrikt kunnat konverteras. Att en så stor andel poster har fallit ifrån, tycks bero på att felaktiga radbrytningar i tabelldumparna; konverteringsprogrammet har slaviskt följt Agoras dokumenterade fältsyntax.

En fungerande, första version av ett schema för Agora har också upprättats, vilket var en nyttig övning, eftersom det blottlade en del svagheter hos TKL's schemaeditor.

Konvertering av Biogate

Underlaget till Biogate bestod av ca 1.700 s k IAFA-templates, som samtliga lät sig konverteras. Biogate var enklare att konvertera, eftersom all information om en resurs är samlad i en egen fil. Biogate, som beskriver biologiresurser över hela

världen, är strikt engelskspråkigt både till innehåll och gränssnitt. Klassifikationssystemet och bläddringsträdet är ett och samma och består av tolv huvudingångar samt är fyra nivåer djupt. Det har realiserats i TKL på samma sätt som Agoras bläddringsstruktur.

Bland de konverteradefälten märks title, identifier (URL), description, Author-Name, Category, Comments (opublicerade redaktörskommentarer), Geography (geografisk täckning), Keyword (fria ämnesord), Language (resursens språk, som övervägande är engelska) samt PublisherName.

Följande auktoritetsfil har automatiskt härletts ur materialet: Geography.

En fungerande, första version av ett schema för Biogate har också upprättats.

Sammanfattning och slutsatser

Konverteringsprogrammet, som både tar hand om Agora och Biogate utgörs av ett Perl-script om knappt 800 rader. Det har framgångsrikt använts för att överföra de båda portalernas väsentligaste datafält för experiment i TKL.

KIBs samling av statiska HTML-sidor bedömdes som alltför svåra att konvertera inom ramen för detta projekt.

Konverteringsprogrammet bör relativt enkelt kunna anpassas för att åstadkomma en produktionsmässig konvertering av Agora och Biogate, förutsatt att underlag för det upprättas av portalernas utgivare. Även om exempelvis inte alla Agoras fält hittills har hanterats på ett sätt som överensstämmer med äkta Agora, har åtminstone något av de svårare fallen framgångsrikt kunnat emuleras i TKL. Att förfina programmet, så att produktionsmässighet uppnås för dessa båda portaler, bör därför kunna uppnås med en betydligt mindre insats än det hittills nedlagda arbetet.

För att konvertera andra portalers data krävs mer arbete, men vissa grundläggande delar som bildandet av bläddringsstrukturen bör kunna återanvändas.

Export av data

Det är vår bedömning att återanvändbarheten av URLografiska poster i de svenska ämnesportalerna är en synnerligen viktig fråga. Dessa poster skall inte bara kunna återanvändas inom ämnesportalsväsendet utan även inom andra tjänster av helt andra slag. Det skulle kunna röra sig om tjänster liknande My Library < <http://mylibrary.lub.lu.se/> > och Mitt Kursbibliotek < <http://mittkursbibl.lub.lu.se/> >.

Tre vägar ut från en TKL-baserad portal existerar eller kan utvecklas.

- OAI-PMH < <http://www.openarchives.org/> >. TKL kommer med en nästan körfärdig OAI-server. Vi ger exempel på hur det fungerar i DOAJ (jmf <http://www.doaj.org/articles/questions/#metadata> >)
- RSS < <http://purl.org/RSS/> >. Indexdata har rapporterat en implementering av detta
- Z39.50

Av dessa förutser vi att den första och sistnämnda har störst potential. OAI-PMH har blivit något av en Swiss-Army-knife för att lösa problem som rör överföring av

bibliografiska poster från en databas till en annan.

Det är jämförelsevis enkelt att implementera en OAI-server med TKL. I skrivande stund kräver det en stunds programmering i php och XSLT, men detta torde vara en engångsinsats för varje portal. Att utöka en servers syntaxrepertoar är också relativt enkelt. Vad som krävs är en syntax som är möjlig att validera mot en XML Schema. Tre sådana existerar idag, vilka bör täcka de flesta behov

- oai_dc < <http://www.openarchives.org/OAI/openarchivesprotocol.htm> >
- MODS < <http://www.loc.gov/standards/mods/> >
- Marc21 < <http://www.loc.gov/standards/marcxml/> >
- oai_def_portal

TKL kommer med stöd oai_dc och ett speciellt oai_def_portal som används för samarbetet mellan ämnesportaler inom DEF. Ett beslut att stödja TKL skulle öppna möjligheten för större nordisk samarbete i fråga om vidareutveckling på detta området.

TKL kommer med en sökmaskin, idZebra, som talar Z39.50 med omvärlden. Att vidga detta till att täcka olika former av samsökning är en enkel uppgift. Det kräver dock en viss utvecklingsinsats, eftersom implementeringen av Z39.50 i TKL inte är helt ortodox, databasen fungerar mer som en native XML search engine än som Z39.50-server. Att utveckla en traditionell Z39.50 target är närmast trivialt.

Som alltid kräver framgång när det gäller delning av metadata också samordning och överenskommelser när det gäller val av metadamodeller.

Ämnesrubriker och klassifikation

Allt material i en TKL-baserad tjänst är dokument. De resurser som skall presenteras i en tjänst matas in som katalogposter. För varje typ av katalogpost finns ett inmatningsformulär. Så långt känner man igen sig i många andra ämnesportalverkyg som har byggts under åren. Det stora skillnaden är dock att all navigering baseras på att man bygger och beskriver resurser i en katalogstruktur. Alltså katalogposterna inordnas i ett hierarkiskt klassifikations- eller ämnesordssystem. Resultatet syns i varje sida i administrationsverktyget och vi återkommer till detta under administrationsgränssnittsdiskussionen.

Förutom den överordnade hierarkiska ämnessystem finns även möjlighet att lägga till, importera och anpassa andra standardiserade hierarkier, datum, tid, andra klassifikationsträd mm.

Att använda bläddringssystemet på detta viset har vissa fördelar och bygger på en tydlig logik. Det ställer dock krav på val av ämnessystem samt på funktioner och gränssnittet i portalsystemet själv. På det praktiska planet innebär det ett annat sätt för redaktörer att arbeta. Det ställer krav på enkla och smidiga redskap för att få en översikt av bläddringssystem och de inlagda resurser. Något som inte fungerar tillfredställande idag. Valet av ett allt för stort och komplext bläddringssystem skulle också ställa till problem för den dagliga skötseln av ämnesportalen.

Alltså var det inte oväntat att detta sätt att strukturera portalen har varit svårast för testpersonerna att acceptera. Vår egna erfarenheter från att bygga nya portaler från

grunden mha TKL visar dock på att det går att bygga tjänster på detta viset utan stora problem. Flexibiliteten i systemet innebar t ex att vi kunde i DOAJ < <http://www.doaj.org/> > använda LCC som klassifikationssystem samt en förenklad visualisering av detta system. Vi har upplevt att de största svårigheterna ligger i administrationsgränssnittets nuvarande upplägg och de buggar i testversionen som upptäckts efterhand. Vi återkommer till detta i kapiteln om administrationsgränssnittet.

Sammanfattningsvis kan vi konstatera att det går att anpassa presentationen av en ämnesstrukturerad databas men att det kräver en viss utveckling.

Metadatamodellen

Som nämns ovan kan allt innehåll i en TKL-baserad tjänst ses som dokument. De resurser som skall presenteras i en tjänst matas in som katalogposter. För varje typ av katalogpost finns ett inmatningsformulär. Likaså är all navigering baserad på att man bygger en katalogstruktur. I varje sida i administrationsverktyget finns några stående rubriker som reflekterar detta.

När man lägger in material, dvs katalogposter till en TKL-baserad tjänst, redigerar man sin text genom att fylla i formulär. När texten är färdig är det ett dokument.

Alla dokumenttyper går att redigera med hjälp av ett formulär. Definitionen av dokumenttyper görs genom att redigera ett schema. I vårt arbete med import av resurser från Agora och Biogate innebar det anpassning av relativt komplex metadata till en TKL-schema. Resultaten visade att det var möjligt att definiera och anpassa metadataformuläret så att standardiserade element, delement och kontrollerade hierarkier kunde användas. Alltså stöd finns för alla metadatafält som nu finns i portalen eller som finns i DCMES (Dublin Core). Samtidigt är det fullt möjligt att lägga till "lokalmetadata" som är till för den enskilda portalens behov. Det som saknas i dag var en "bäst-före" funktion, dock det skulle gå att utveckla utan större problem. De största svårigheterna ligger igen i administrationsgränssnittet.

B. Administrativt gränssnitt

Grundläggande förutsättningar

Vi förutsätter att det kommer att utvecklas ett centralt samt individuella gränssnitt till delmängder av data.

Testförlopp

Mål för testet var att belysa TKLs administrativa gränssnittet. Samtliga representanter fick ett administrativt gränssnitt att arbeta med. Testportalerna är tillgängliga från < <http://www.lub.lu.se/knowtech/projects/toolkit/testportaler.html> >.

På detta viset fick referensgruppen möjlighet att bedöma hur det är att arbeta praktiskt med TKL. Vid slutet av testfasen lämnade alla in en sammanfattning av sina intryck av verktyget och redogjorde för hur pass väl TKL skulle fylla deras ämnesportalens funktionella behov. (Se Bilaga 2 - Dokumentarkiv - Referensgruppens rapporter)

Ämnesportalernas synpunkter på administrationsgränssnittet

Som vi har konstaterat tidigare bygger den administrativa systemet på annorlunda principer än det som är vanligt i portalverktygsimplementationer. Allt material i en TKL-baserad tjänst är dokument. De resurser som skall presenteras i en tjänst matas in som katalogposter men man bygger och beskriver resurser i en katalogstruktur. För att skapa en ny post i systemet bläddrar man sig ner till den relevanta nivån och lägger in sin beskrivning där.

Att skapa innehåll i en ämnesportal är ett tidskrävande intellektuellt arbete där en av de allra viktigaste redskapen är själva gränssnittet där innehållet matas in. Ämnesportalerna har olika organisationella modeller för hur inmatningen sker. Agora är t ex ett kollaborativt arbete där över trettio personer är verksamma som redaktörer. Å andra sida finns portaler där en eller två personer sköter inmatning och skötsel av tjänsten. I alla tjänster finns ett visst antal redaktörer som arbetar kontinuerligt och för dessa blir det enklare att lära sig ett nytt system. För andra där det kan dröja mellan gångerna man använder systemet blir det betydligt svårare att sätta sig in i ett nytt system. Det ställer höga krav på lätthanterliga och lättöverskådliga funktioner - något som inte testinstallationen uppfyllde.

Ett exempel av konsekvensen av den speciella uppbyggnaden av bläddringssystemet är hur man dubbelklassar ett dokument. För att dubbelklassa en post måste man upprätta en referens till posten från en annan katalog. Att upphäva dubbelklassningar eller att klassa om en post från början kan därmed inte göras via det vanliga registeringsformuläret, utan man måste använda ett särskilt formulär för det, där man f n tvingas orientera sig via klassifikationskoderna istället för klartextsbenämningar. Förutom att det är långsamt, inbjuder det till oavsiktliga fel som att man kan råka placera en post i fel i hierarkin eller rentav utanför den, radera en post eller t o m åstadkomma bestående fel i bläddringsstrukturen, vilket testpersonerna faktiskt har råkat ut för. Testpersonerna har inte lyckats länka till poster utanför den egna portalen.

Liknande erfarenheter har upplevts av redaktörerna för de nya portalerna i Lund. Det är "onödigt krångligt" att redigera resurser och ändra dubbelklassningar. Överskådliget innebär att resurser tar alldeles för lång tid att hitta.

Bläddringssystemet är den del av TKL, som testpersonerna har haft svårast för att acceptera.

Formulären

TKL är ett mycket flexibelt system. I och med att allt material i en TKL-baserad tjänst kan ses som dokument innebär det att för varje typ av katalogpost finns ett inmatningsformulär. När man lägger i material, dvs katalogposter till en TKL-baserad tjänst, redigerar man sin text genom att fylla i formulär. När texten är färdig är det ett dokument.

Testpersonerna har uppskattat att man förutom dataposterna kan redigera auktoritetsfiler, metadataschemat och filen med medarbetarna. Tyvärr framstår formulären som klumpiga och inte allt för användarvänliga; man tvingas scrolla fram och tillbaka alldeles för mycket. Det speciella formuläret, som används för att redigera bläddringsstrukturen hade testpersonerna genomgående haft svårt för att

tillägna sig. Auktoritetsfilseditorn uppfattades som lite väl trög för stora vokabulärer. Vidare är det självklart olämpligt att man tvingas upprepa klassifikationsskoden varje gång man definierar en ny översättning av en viss term.

En del av de andra funktioner som erbjuds inom samma administrativa gränssnittet upplevs som positiv, nyhets- och artikelpublicering tex. En av testpersonerna har uttryckt sig uppskattande om nyhetspubliceringssystemet men efterlyser automatiskt hanterade första- och sistadatum.

Stabilitet

Testpersonerna har genomgående klagat på att TKL ännu inte är färdigutvecklat och därför behäftat med en hel del allvarliga fel. De allvarligaste klagomålen gäller redigeringen av bladdringsstrukturen. Andra fel gäller schemaeditorn där man skapar och anpassar metadata, dvs tjänstens datamodell. Naturligtvis är detta inget som man ändrar varje dag, men helt klart upplevs det som otympligt. Efter att ha definierat ett fält kan man inte gå tillbaka och ändra vissa av dess parametrar i efterhand, t ex inmatningsformatet för att ta ett exempel. Schemaeditorn har stor potential, men i praktiken fungerar inte underfält så bra. Det är heller inte möjligt för närvarande att definiera mer än ett attribut per fält. Det var möjligt att mata in text (t ex HTML-taggar för att länka till externa hjälptexter), som fick formuläret att krascha med fullkomliga felkoder. Länkkontrollfunktionen rönste uppskattning, medan dublettkontrollfunktionen inte tycks fungera tillfredställande.

Effektivitet

Det är livsviktigt att registreringen kan ske så effektivt som möjligt: Dels är den avsatta tiden begränsad och dels löper redaktörerna annars risken att förlora motivationen med följd att portalen självdör. Därför är det viktigt att stabilitetsproblemen och problemen med dubbelklassning blir lösta på ett bra sätt, så att man kan dra fördel av exempelvis en distribuerad redaktion.

Pågående utveckling

Det är både glädjande och en aning frustrerande att utvärdera en programvara stadd i snabb utveckling. Våra slutsatser angående det administrativa gränssnittet är baserade på version 1.1 av TKL, den senaste versionen 1.2 av programvaran släpptes nyligen. De flesta synbara förändringarna gäller just det administrativa gränssnittet.

I skrivande stund har vi inte hunnit portera våra egna portaler till version 1.2. De mest påtagliga förändringarna är visuella; gränssnittet är inte bara visuellt mera tilltalande utan där finns även påtagliga förbättringar när det gäller säkerheten vid inloggning och framför allt är gränssnittet flerspråkigt. De flesta av formulären genereras nu med XSLT från portalens XML-filer, och alla förklarande texter finns samlade på i en fil för enkelt underhåll.

En av de komponenter i systemet som fått lägst betyg av testgruppen har tyvärr inte förändrats. Detta gäller den så kallade schema-editorn, varmed definitionen av portalens olika dokumenttyper definieras.

En komponent som fått lågt betyg och som faktiskt blivit bättre är editorn för auktoritetsfiler. Den består fortfarande av en "lång lista" av formulär, men den nya utformningen av sidorna gör att man kan leva med detta. En del av deltagarnas frustration när det gäller detta verktyg måste nog ses som ett pedagogiskt problem. TKLs auktoritetsfiler är inte tänkt att fungera som en tabell i en relationsdatabas (RDBMS). Om man försöker lägga en stor del av SABs ämnesord i en TKL-auktoritetsfil gör man något den aldrig var tänkt för.

Det går utmärkt att implementera en funktionalitet som liknar separat lagring av ämnesord i en tabell i en RDBMS. Det kräver en del utveckling, men den behöver inte vara omfattande eftersom XML har lämpliga verktyg för ändamålet. Till exempel XPointer och XLink.

Även det faktum att TKL 1.2 tillåter bindningar även till SQL utöver de existerande till SOAP och Z39.50 ges det nya möjligheter när det gäller förfinade data-modeller.

C. Användargränssnitt

Testförlopp

Mål för testet var att identifiera de begränsningar och klargöra de möjligheter som TKL-programvaran tillför vad gäller användargränssnittet. Samtliga representanter fick testa en mer eller mindre default version av användargränssnittet. Möjligheten att ge skraddarsydd gränssnitt under testfasen var begränsad.

Sammanfattning av referensgruppens synpunkter

Eftersom det krävs kunskaper i XSLT för att man ska kunna påverka användargränssnittet alltifrån layout till funktionalitet, har det funnits stora begränsningar i vad referensgruppen kunde testa vilket flera av testpersonerna har beklagat.

Som nämnts ovan fanns i den version av TKL vi testat inga möjligheter att via det administrativa gränssnittet redigera stilmallarna. Ett sådant är nödvändigt för att de enskilda portaladministratörerna på ett enkelt sätt själva skall kunna utforma sina gränssnitt.

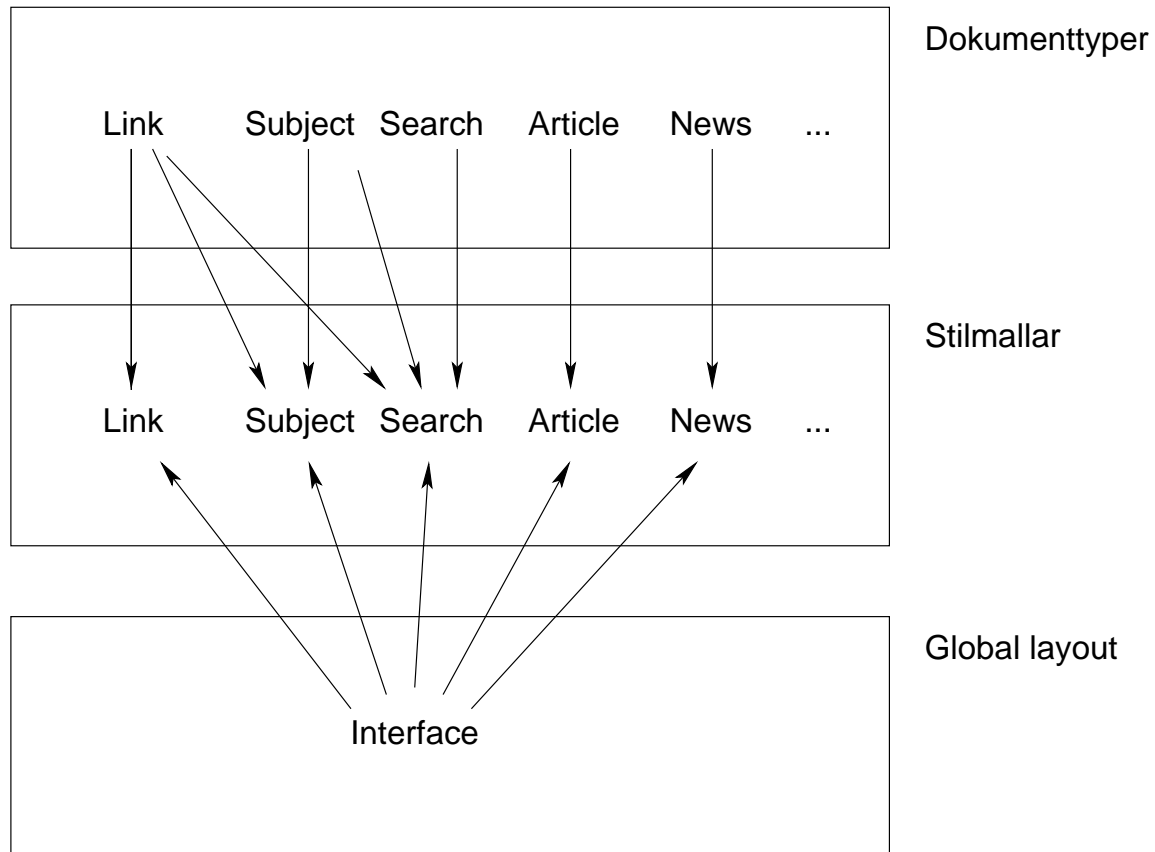
XSLT är kanske med sina fyra år på nacken en relativt ny teknologi och därför är inlärningströskeln högre än för traditionella system vars användargränssnitt byggs med mallar, templates eller vars programmering utförs som inbäddad i html-sidor (php, perl, asp m fl populära system erbjuder sådana faciliteter). Vidare kräver XML-baserade system renare taggning än för traditionell webbredigering. Ett gränssnitt tillverkat i programvaror som MS Frontpage kan kräva en hel del arbete att överföra till TKL.

Det är därför sannolikt så att många portaler kommer att behöva XML/XSLT-stöd av ett slag som många biblioteks IT-personal inte kan bistå med. Behovet är naturligtvis störst under portalernas uppbyggnad, men på många håll kan det finnas behov av även stöd för smärre justeringar även under drift.

Å andra sidan är vi bara i början av en utveckling mot att kunskap i XSLT är ett lika naturligt krav på en webbdesigner som kunskap i bildbehandling (text behandling är viktigt det med - det finns trots allt mer text än bild på Internet). Vi ser detta som ett övergående problem.

Hur ett användargränssnitt i TKL byggs

Elegansen i TKLs användning av XSLT blir uppenbar av följande schema:



•

Varje dokumenttyp som ligger i en *.tkl-fil kan levereras från servern givet att det existerar ett tillhörande XSLT-skript. All global formgivning är samlad i en stilmall som anropas från de enskilda stilmallarna. Som nämnts ovan finns det också ett bibliotek av funktioner tillgängliga inifrån XSLT som är implementerade som URLar, vilka returnerar XML. Detta används till exempel för att inkludera dokument av typen link i dokument av typen subject. Detsamma gäller träffar i filer av typen subject vid sökning, som visas i träfflistan. Klickar man på en sådan får man bläddringsidan för den aktuella kategorin.

Ingen kedja är starkare än den svagaste länken. Den svagaste länken i TKL är redigeringen; det existerar en hel rad XML-dokumenttyper som inte går att redigera enkelt i ett formulär, men som likväl går att presentera för användare med hjälp av TKL. Enda kravet är att det går att skriva en stilmall.

Metoden är tillräckligt generell för att tillåta oss att ersätta de enkla subject-dokumenterna med något mera komplicerat. För att ta något närliggande så skulle det gå att använda Mesh i XML i samma plats <<http://www.nlm.nih.gov/mesh/filelist.html>>. Inom varje sådan fil kan TKL inducera en Z39.50-sökning på aktuella Mesh-termer för att åstadkomma listor med länkar, och likaledes generera länkar till andra termfiler från NLM.

D. Central drift och utveckling

Testförlopp

Mål för testet är att identifiera de begränsningar och klargöra möjligheter som TKL-programvaran ger vad gäller uppbyggande och driftsrutiner kring ett gemensamt system.

Grundläggande förutsättningar

Gemensam databas definierar vi som

- portalerna går att underhålla och utveckla gemensamt som ett centralt system och att samordningsvinster därvid görs.
- det finns, eller går att utveckla, ett system som gör att de olika portalerna kan återanvända varandras metadata för resurser som är tvärvetenskapliga eller inom gränsområdena mellan deltagande tjänster och att underhållet av katalogposterna därvid förbilligas.

En gemensam databaslösning har flera fördelar. Användningen av gemensam programvara förbilligar underhåll och utveckling, givetvis under förutsättning att alla portalers krav är rimligt lika eller om de kan jämkas samman. Den uppenbara fördelen skulle naturligtvis vara att genom ett utvecklingsamarbete kommer en viss facilitet bara att behöva köpas en gång.

Även rent redaktionellt arbete kan göras billigare, en post som tillverkats inom en portal skall kunna återanvändas i en annan med delvis överlappande målgrupper och inriktning. Sådan återanvändning kan ske på tre olika sätt

- Samsökning - vilket ger en virtuell samkatalog av självständiga portaler
- Direkt kopiering av poster mellan portaler.
- En stor central portal, vilken ämnesmässigt "stratifieras" på ett sådant sätt att den kan fås att framstå som kanske ett dussin självständiga portaler

Självklart ställer dessa tre olika vägar olika krav på hård- och mjukvaror. Därutöver tillkommer den olika grader av sammanjämkning som krävs för framgång. Vi har avsiktligt ovan sorterat de tre olika sätten efter den grad av samarbetsvillighet som skulle avkrävas av varje portal för att nå till målet.

Priset för att nå fram till samordning är olika högt beroende på hur nära man vill samarbeta, och det är naturligtvis högst för nivå tre. Priset skulle vara högt för alla, och vägen till målet kan vara lång och därtill snårig och beströdd av törnen.

Det är i form av omkatalogisering av de befintliga portalernas resurser de vassaste taggarna visar sig. Ämnesklassningar i enlighet med nya ämnesvokabulärer måste

tillföras. Nya fält tillkommer och måste befolkas av beskrivningar. Om man väljer den högsta graden av integrering kan **dessutom** tidigare investeringar i form av fält som inte passar in i den nya gemensamma strukturen och som mödosamt fyllts av hårt katalogiserande bibliotekarier komma att behöva skatta åt förgängelsen.

Oavsett vilken nivå av samordning det svenska akademiska portalväsendet kommer att kunna enas om, kommer denna inte att kunna nås utan utveckling av en åtminstone delvis gemensam metadatamodell för samsökning av resurser, samt att portalerna kan förmås att samsas om en överordnad ämneshierarki som kan länkas eller "mappas" mot de som används i de befintliga portalerna.

Den test av TKL-programvaran som vi har genomfört behandlar inte dessa frågor, och dessa problem ligger utanför denna rapports område. Här följer en genomgång av vad som är möjligt och/eller rimligt att göra med TKL.

Val av teknisk arkitektur

Oavsett vilken nivå man väljer för integrering av finns det pengar och resurser att spara enbart på samordnat mjukvaruunderhåll. Det finns ytterligare vinster om tjänsterna samlokaliseras på en central server eller ett centralt kluster. Samlokalisering av tjänsterna till en central server eller ett centralt serverrum innebär också att de kommer kunna underhållas av personal som kommer att lära sig att sköta dem.

Hur kan man tänka sig att organisera ett svenskt nätverk av ämnesportaler? Vi har slängt ihop en skiss över ett bläddringsträd.

I figuren återfinns olika ämnen A,B,... Q samt U. Ämnet U är unionen av det mänskliga vetandet. Inringade ämnen tänker vi oss som självständiga ämnesportaler, med delvis egna kvalitets- och selektionskriterier och dessutom sitt eget gränssnitt. Varje portal har sitt eget ämnesråd - portalen U har ett träd som direkt importerar de övriga portalerna och slutligen har vi ett litet specialarrangemang mellan B och C vilka delar på ämnet J (det återfinns alltså i båda portalerna). Dessamma gäller ämnet L som finns i C och D. Eftersom ämnet K i sig är en underordnad term till ämnet C finns hela K-portalens bestånd även i C.

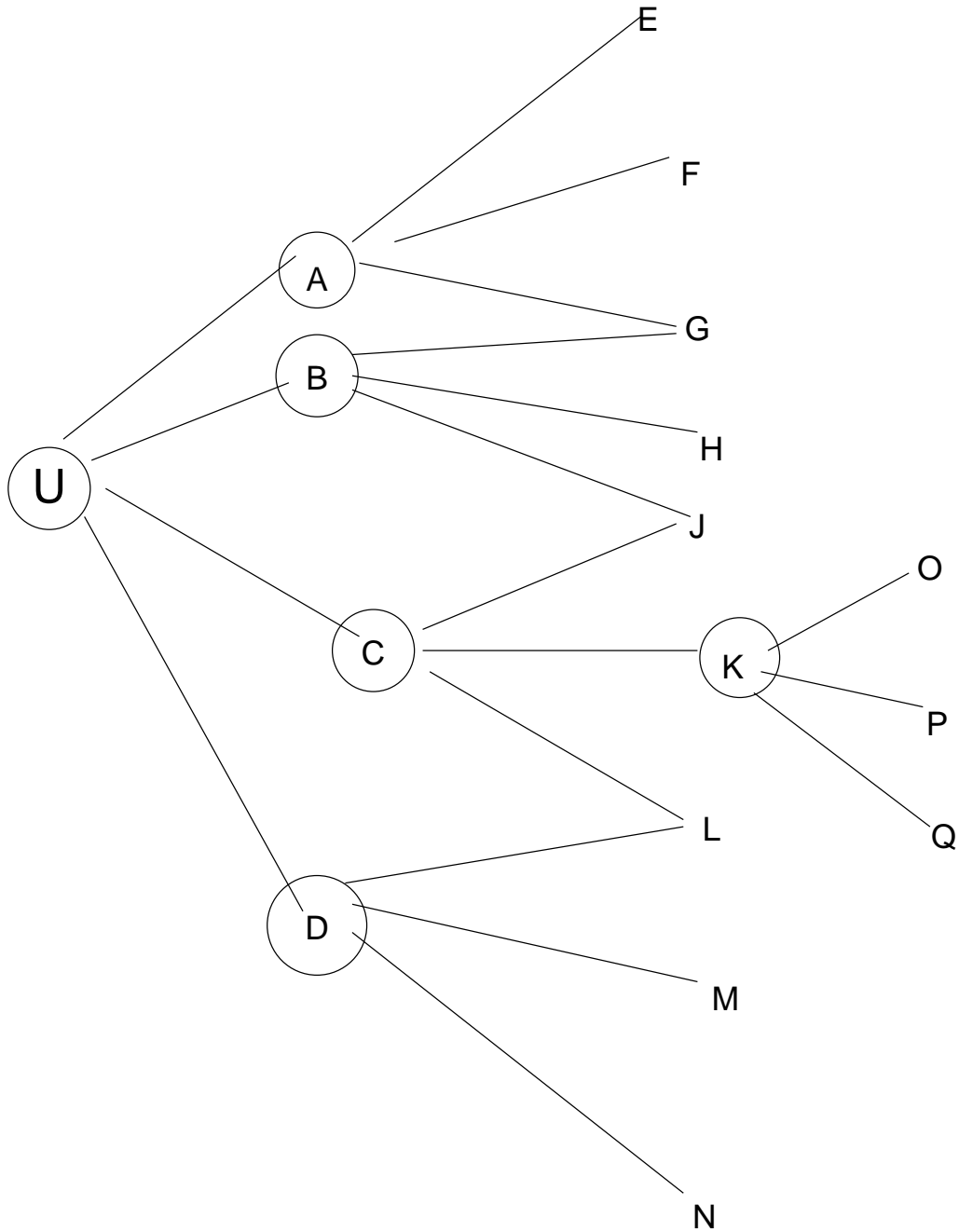
Tack vare TKLs öppna arkitektur är det med en modest teknisk arbetsinsats (men långt gående samordning när det gäller metadata och vokabulärer är detta en möjlig arkitektur). Det är inget förslag från vår sida, utan snarare ett exempel som är tänkt att väcka till eftertanke.

Ett mindre ambitiöst sätt att åstadkomma samordning är samsökning. Exakt hur det skall göras beror av valet av integrationsnivå. Den enklaste metoden är att helt enkelt harvestera metadata från samtliga portaler och lägga posterna i ett sökindex.

Det finns många sätt att lösa integrationsproblemet. Många, kanske de flesta men alls inte alla, är förenliga med TKL. Vad vi vet är att **alla** har det gemensamt att de kräver samordning när det gäller de "mjuka frågorna".

Slutsatser

Efter den genomförda testen av TKL anser vi att programvaran är ett kraftfullt och flexibelt verktyg för skötsel och drift av ämnesportaler. Utvärderingen har



•

försvårats av att programvaran är en relativt ofärdig produkt. Den skulle kräva en del vidareutveckling för att planteras i ett svenskt nätverk av ämnesportaler. Vi kan inte heller utesluta att vissa av våra slutsatser var obseleta redan medan vi skrev rapporten. Utvecklingen av TKL har fortsatt efter att vi avslutat vårt arbete.

Som referensgruppen har konstaterat var den version som testades en buggig produkt. Den bygger på ett ovanligt koncept och därför kan man förvänta sig att det blir en ganska stor omställning för svenska ämnesportalsredaktörer att vänja sig vid

den.

Version 1.2 har på klara förbättringar vad gäller administrativt gränssnitt och användarvänlighet men det behövs fortfarande utveckling för att tillfredsställa referensgruppens synpunkter och önskemål. Schema-editorn t ex förblir oförändrad i den nya versionen. Om vi tittar på listan av krav och önskemål som ställdes i testschemat klarar TKL av en ganska stor andel redan nu och ännu fler med en som vi bedömer det rimlig vidareutvecklingsinsats.

Som vi har nämnt tidigare anser vi att "better is the enemy of good enough". Dock måste "good enough" innebära att det inte försvårar eller förhindrar det redaktionella arbetet. Det är avgörande att det finns ett buggfritt och användarvänligt registreringsformulär.

TKL är skapat som ett verktyg för att bygga enskilda portaler, inte för brokerlösningar. Som vi har kommenterat i vår internationella överblick finns det ingen nyckelfärdig lösning för att bygga en nationell svensk ämnesportaltjänst.

Utifrån våra erfarenheter under testfasen kan vi konstatera att TKL mycket väl skulle kunna användas som grundstomme att bygga den nationella lösningen på. För att det skall fungera måste flera förutsättningar vara uppfyllda:

- Klara direktiv och dimensioneringen för arbetet nationella portalen är på plats
- Det måste råda konsensus mellan portalerna om de "mjuka" frågorna; överenskommelser som bildar underlaget till den nationella tjänsten.

Att användargränssnittet kräver XSLT-kunskaper för att kunna förändras är en nackdel idag. Kanske inte om några år. Oavsett vilket, balanseras denna nackdel av att tjänsten är tänkt att drivas centralt. Problemet behöver inte vara avgörande eftersom den kompetensen då bara behöver finnas på ett ställe.

Av vad vi kan bedöma har TKL om man jämför med andra freeware produkter, en ovanlig trygg bas och tydligt vidareutvecklingsfokus tack vare samordning och samarbete med en nationell organisation. En jämförelse med ROADS för att se de tidigare bristerna i ett område som myllrar av egenutvecklingar och återvändsgränder. Företaget bakom TKL har bevisligen levererat flera bra och användbara produkter tidigare. En annan klar fördel med ett val av TKL är att det kan öppna vägen till samarbete om samma verktyg med en annan nationell nordisk ämnesportaltjänst.

Bilaga 1 - Testschema

A. Innehåll i tjänsten

Krav	Angelägenhet	Uppfyllt	Kommentar
A.1			Import av data.

Krav	Angelägenhet	Uppfyllt	Kommentar
A.1.1	Skall	Ja	Möjlighet till import av data från existerande ämnesportaler. Gjort för Biogate (ROADS-baserad) och Agora (MS-SQL-baserad). I och för sig utanför TKL som sådant. Se vidare "Import av data".
A.2	Skall	Ja	Export av data. Filernas format (XML) innebär inte något problem. Däremot innebär den omständighet att klassningarna inte kan avläsas ur själva posten utan istället avgörs av i vilken katalog posten är förlagd samt av eventuellt ytterligare referenser till poster från andra kataloger en försvårande om än inte oöverstiglig omständighet.
A.2.1	Bör	Nej	Renarduskompatibilitet. TKL är inte Renarduskompatibelt. Tekniskt är detta relativt enkelt att åstadkomma, däremot kräver detta att det deltagande tjänster etablerar en detaljerad omvandlingstabell mellan tjänsternas klassifikationssystem och DDC.
A.2.2	Bör	Ja	OAI-kompatibilitet. TKL stödjer en delmängd av OAI, och det är enkelt att konfigurera exporten av metadata.
A.2.3	Bör	Ja	Z39.50-kompatibilitet. Med en liten ansträngning åstadkoms ett fullt utbyggt stöd för Marc 21 och BIB-1 sökattribut. Kräver samordning av tjänsternas metadatamodeller.
A.2.4	Skall	Ja	Kompatibilitet med centrala standarder, främst XML, Dublin Core, Z39.50. Se ovan. Dessutom erbjuder TKL SOAP-funktionsanrop.

Krav	Angelägenhet	Uppfyllt	Kommentar
A.2.5	Skall	Ja	Sidor, där länkarna presenteras, ska vara synliga för sökrobotar (dvs det insamlade materialet ska inte vara "gömt" i databasen). I en TKL-portal kan alla bläddringsupp-slag nås via länkföljning från ingångssi-dan, vilket medför att de också kan sam-las in automatiskt av en robot. Det går naturligtvis inte för någon söktjänst att förutsätta för fritextsökresultat, efter-som det inte går att förutsätta att några länkar leder till dessa.
A.2.6	Skall	Ja	Inlagda poster skall kunna återexporteras till ansvarsbiblioteken (t ex för användning i applikationer, som ligger utanför ramen för det gemensamma sys-temet. Se A.2.
A.3			Ämnesrubriker och klassifikation.
A.3.1	Skall	Nej	Hantering av flera ämnesområden i samma tjänst/på samma server. Det finns stora möjligheter att påverka presentationen av en ämnesstrukturerad databas. Kräver viss utveckling.
A.3.2	Skall	Ja	Olika klassifikationssystem/kontrollerade vokabulärer. TKL kan användas för att presentera samma data på olika sätt. I DOAJ används både LCC och en visualisering av detta system. Ämnesredaktörerna och tillgång till en separat ämnesportal som vars bläddringsstruktur strikt följer LCC.
A.4			Metadata/datamodellen
A.4.1	Skall	Ja	Inmatningsgränssnitt - olika registrerings-formulär. (Se B.1)
A.4.2	Skall	Ja	Fördefinierade metadatafält - obliga-toriska fält. Se B.1.
A.4.3	Skall	Ja	Det skall finnas stöd för alla metadatafält som nu används i portalen och/eller som förekommer i Dublin Core Element Set. Se B.1.

Krav	Angelägenhet	Uppfyllt	Kommentar
A.4.4	Skall	Nej	Fält för "bäst-före"-datum bör finnas i formuläret, så att temporära resurser kan döljas i användargränssnittet, då det angivna datumet överskridits. Datum för senaste modifiering av en post lagras med sekundnogrannhet i varje post. Det går att utveckla en sökning som väljer ut poster som inte kontrollerats under någon given tidsperiod.

B. Administrativa gränssnitt

Krav	Angelägenhet	Uppfyllt	Kommentar
B.1	Skall	Ja	Skräddarsydda registreringsformulär. Metadatamodeller kan definieras med hjälp av TKLs schema-editor.
B.1.1	Skall	Ja	Inmatningsformulär ska finnas och ska skall kunna hantera nyregistrering, uppdatering och makulering av poster. Alla elementära databasoperationer stöds av TKL.
B.1.2	Skall	Ja	Möjlighet till anpassning av registreringsformulären skall finnas. Klaras av med hjälp av schema-editorn, där man inte bara kan definiera vilka fält man ska ha utan också hur de ska matas in, huruvida de ska vara upprepningsbara och/eller obligatoriska. Det går också att mata in kortare hjälptexter, men tyvärr går det inte att länka till externa hjälptexter. Möjligheterna att påverka det redaktionella gränssnittet har ökat i TKL version 1.2, i vilket formulär skapas med hjälp av XSLT. Därför är det numera möjligt att konstruera länkar till annan dokumentation inifrån verktyget.
B.1.3	Skall	Ja	Registreringsformuläret skall tillhandahålla möjlighet till och god hantering av upprepningsbara fält utan principiell begränsning av antalet fält (t ex titlar, upphovsmän).

Krav	Angelägenhet	Uppfyllt	Kommentar
B.1.4	Skall	Ja?	Registreringsformuläret skall tillhandahålla möjlighet till och god hantering av förvalsmenyer/flervalmenyer. Klarar i skrivande stund bara av envalsmenyer.
B.1.5	Skall	Ja?	Registreringsformuläret skall ha fileringsfunktion för titlar. Ett elementärt system för avlägsnande av danska bestämda artiklar (den, det m fl) från titlar vid sortering finns.
B.1.6	Skall	Ja	Registreringsformuläret skall kunna kopplas till kontrollerade vokabulärer. (Schemaeditor: Inmatningstyp = dropdown + koppling till auktoritetsfil.)
B.1.7	Skall	Ja	Möjlighet för redaktörerna att vid behov själv lägga till nya ämnesord och klassningar. Redaktörerna kan tillåtas att uppdatera auktoritetsfiler via ett särskilt formulär (genom att det går att lägga individuella auktoritetsfiler i egna kataloger om så önskas, kan man undvika "befogenhetsinflation". Däremot kan TKL inte tillåta redaktören att för ett och samma fält välja mellan rullgardinsmeny och textinmatningsruta som i Agora, där det sistnämnda valet medför utvidgning av auktoritetsfilen.
B.1.8	Skall	Ja?	Geografiska och kronologiska ämnesord skall kunna särskiljas från andra ämnesordstyper. TKL klarar detta, vilket dock är en fråga för om val av metadatamodell. Lösningen skulle vara att skapa två fält, låt oss kalla dem temporal coverage respektive spacial coverage (efter var de hör hemma i Dublin Core). Vanliga ämnesord skulle hamna i subject.
B.1.9	Skall	Ja	Dublett kontroll skall finnas för länkadress, ämnesord och klassningar. Finns för URL men troligtvis inte i övrigt, alltså att exempelvis att ett visst ämnesord upprepas i en post.

Krav	Angelägenhet	Uppfyllt	Kommentar
B.1.10	Skall	Ja	Hjälptexter skall finnas kopplade till formuläret. TKL understöder hjälptexter, som lagras lokalt i schemafilen. (jfr B.1.2)
B.1.11	Skall	Ja	Förhandsgranskning av alla uppgifter som registrerats och/eller som i formuläret legat som standardval skall finnas. Förhandsgranskning finns så tillvida att en post kan markeras såsom icke publicerad. Den kan i så fall granskas av andra - fast inte i tjänstens gränssnitt utan i metadataformuläret.
B.1.12	Skall	Ja	Om obligatoriska fält lämnats oifyllda skall posten ej accepteras i databasen och förklarande felmeddelande visas.
B.1.12.1	Skall	Ja	Vissa fält skall vara obligatoriska = registrering skall ej kunna slutföras utan att dessa har ifyllts. Se B.1.12.
B.1.12.2	Bör	Ja	Fördefinierade metadatafält - obligatoriska fält. Se B.1.
B.1.13	Skall	Ja	Vid makulering av poster skall inte anknutna ämnesord eller klassningar makuleras ens om ingen annan post har knutits till dessa.
B.2			Postadministration.
B.2.1	Skall	Ja	Datum för skapande av ny post och datum för ändring av befintliga poster skall skapas som dolda värden i tabellen och vara separerade från varandra. Klaras av, fast f n buggigt.
B.2.2	Skall	Nej	Varje post ska kunna knytas till ett ID-nr för registrerande redaktör. Varje post knyts till redaktörens användaridentitet. Man får i princip denna funktionalitet genom att man byter lösenord på en identitet när en redaktör pensioneras men återanvänder identiteten.

Krav	Angelägenhet	Uppfyllt	Kommentar
B.3	Skall	Ja	Stöd för flerspråkighet. TKL klarar flerspråkighet på såväl redaktörs-, användargränssnittsnivå och fältinnehållsnivå. Version 1.2 har dessutom möjlighet till språkval för redaktörer. Det administrativa gränssnittets texter lagras i en XML-fil.
B.4	Skall	Ja	Sökformulär för internt behov skall finnas (redaktörmässig sökning). Redaktörerna får använda tjänstens sökfunktionen, och de kan, om de är inloggade, ladda poster i inmatningsverktyget.
B.5	Skall	Ja	Accesskontroll. Se B.8.1 och B.8.2. Accesskontrollen för redaktörer har i vissa avseenden förbättrats i version 1.2. Det är dock fortfarande omöjligt att logga ut.
B.6	Skall	Ja	Hantering av presentation av kontrollerade vokabulärer. Finns en enkel auktoritetsfilseditor samt möjlighet i schemaeditorn att knyta inmatningsfält i form av rullgardinsmenyer till auktoritetsfiler. Presentationen av träffar görs med XSLT, och det går att hämta termer via deras ID i auktoritetsfilerna.
B.6.1	Bör	Ja	Det skall finnas stöd för automatisk kontroll av att termer hämtade ur en thesaurus är korrekta. Se B.6
B.6.2	Bör	Ja	Det skall vara möjligt att vid indexering representera en indexeringsterm ur en thesaurus med ett unikt ID istället för förklartexten för ämnesordet. Klaras av med TKLs auktoritetsfiler. Ämnesord som hanteras via auktoritetsfilerna blir tyvärr inte sökbara vid sökning.

Krav	Angelägenhet	Uppfyllt	Kommentar
B.6.3	Skall	Ja	Thesaurusen skall kunna underhållas oberoende av länkarna, lämpligen genom att man importerar unika ID med tillhörande klartexter som en separat tabell. Att automatiskt bilda auktoritetsfiler enligt TKL är enkelt. Se vidare "Import av data".
B.6.4	Skall	Ja	Flerspråkiga thesaurer skall kunna importeras och hanteras av bläddringsstrukturen. Klaras av, under förutsättning av att tesarens struktur är oberoende av språkval. (Två givna termer får alltså inte stå i olika relation till varandra i olika språk).
B.6.5	Skall	Nej	Ändring/makulering av inlagda ämnesord och klassifikationer (spärrad uppdateringsbehörighet). Notera att i TKL ligger posterna ligger i klassifikationen, dvs som en fil i bläddringsstrukturen. Det är alltså inte möjligt att makulera en klassifikation utan att först flytta eller makulera posterna i den. Att förändra auktoriteter kräver normalt administratörsbehörighet.
B.6.6	Skall	Nej	Stöd för mappning av av ämneshierarkier, resurstyper (och ämnesord?) mellan de portaler som integreras i tjänsten skall finnas så att sökning och visning i databasen kan ske. Understöds inte av TKL internt. Ett sådant system har implementerats i DOAJ separat utanför TKL.
B.6.7	Bör	Nej	Att kunna utnyttja databasen "Svenska ämnesord" för registrering och sökning med ämnesord. Inte utan ytterligare utveckling. Notera att vad vi vet stödjer inte "Svenska ämnesord" några standardiserade format för export. Det skulle gå definiera en nätverksburen form av vokabulärsauktoritet om så varit fallet.
B.7	Skall	Ja	Länkkontroll.

Krav	Angelägenhet	Uppfyllt	Kommentar
B.7.1	Bör	Nej	Inmatningssystemet skall automatiskt kunna kontrollera om innehållet på en sida har ändrats sedan länken som pekar ut den lades in. Bör gå att implementera med viss utveckling. Vad det skulle kosta beror naturligtvis på hur förändring skall mätas.
B.7.2	Skall	Ja	Central, automatisk länkkontroll måste kunna integreras. Baserad på HTTP statuskoder.
B.7.3	Skall	Nej	Länkkontrollen skall kunna koplas och styras till ansvarig ämnesredaktör. (Förantlett av att användarklick på trasiga länkar.) Förefaller onödigt med tanke på B.7.4. Detta förutsätter loggning av användning av utgående länkar, och automatisk kontroll av just dem. En normal ämnesportal är inte större än att alla länkar kan kontrolleras regelbundet.
B.7.4	Skall	Ja	Länkkontrollen skall dessutom kunna göras centralt med jämna mellanrum. Se B.7.2
B.7.5	Skall	Ja	Statusrapporterna skall kunna identifiera olika typer av länkfel. Klaras troligen av. För närvarande dock baserat på HTTP-statuskoder.
B.8	Skall	Ja	Administration av redaktörer.
B.8.1	Skall	Ja	Access till registrering skall kunna begränsas och vara lösenordsbelagd. Klaras av men buggig i nuvarande version. Förbättrat i version 1.2.
B.8.2	Bör	Nej	Redaktörerna skall själva kunna ändra sina lösenord och namnuppgifter. Möjligt om redaktörerna har administratörsstatus
B.9	Skall	Nej	E-postsystem. TKL har inget eget e-postsystem, men kan skicka e-mail från tips-formulär och dylikt.

Krav	Angelägenhet	Uppfyllt	Kommentar
B.9.1	Skall	Nej	Deltagande redaktörer skall ha egna (men opersonligt utformade för att underlätta vid redaktörsbyte) e-postadresser, dit användarkontakter och länkfelsrapporter skall kunna styras direkt utan central förmedling. Skulle förmoda att ingen existerande portal-programvara har detta inbyggt i själva portalen. Det är å andra sidan trivialt att implementera utanför genom mail-alias.
B.9.2	Bör	Nej	Skräppostfilter bör finnas på servern. Kan med fördel läggas utanför TKL.
B.10	Bör	Ja	Systemet skall vara skalbart; antalet redaktörer skall därför inte behöva vara intressant.

C. Användargränssnitt

Krav	Angelägenhet	Uppfyllt	Kommentar
C.1	Bör	Ja	Usability-aspekter bör beaktas, även för användare med funktionshinder. Det finns inget som hindrar implementatörer att bygga såväl utmärkta som förfärliga gränssnitt med TKL.
C.2			Skräddarsydda sök- och bläddringsgränssnitt.
C.2.1	Skall	Ja	Möjlighet till skräddarsydda sök- och bläddringsgränssnitt skall finnas för både för varje deltagande tjänst och för användare. TKL erbjuder enkel anpassning av portalers metadatamodeller. Det råder dock en omvänd relation mellan möjligheter till samordnad sökning och navigering mellan portalerna och antalet implementerade egenheter inom var och en av dem.
C.2.2	Skall	Ja	Flera gränssnitt (även ett gemensamt, icke ämnesspecifikt). Ett sådan gränssnitt måste dock utvecklas separat. Det följer inte med TKL (som man ju får gratis).
C.3			Presentation av resultat.

Krav	Angelägenhet	Uppfyllt	Kommentar
C.3.1	Skall	Ja	Det skall vara möjligt att konstruera en URL som adresserar sig till en viss rubrik i bläddringsstrukturen (vanligtvis svarande mot en viss term ur thesaurusen) Se A.2.5.
C.3.2	Skall	Ja	Länk till ämnesrubrikerna i bläddringsstrukturen skall på alla nivåer vara stabil, så att det går att i andra sammanhang direktlänka till olika ämnesrubriker oavsett hierarkisk nivå. Se A.2.5.
C.3.3	Skall	Ja	Rubrikerna i ämnesstrukturen skall gå att presentera som länkar till sidor där innebörden av termen förklaras. Klaras av (relativt enkel anpassning av XSLT script).
C.3.4	Bör	Ja	Viktiga grenar i hierarkin skall kunna från rotnivån i bläddringsstrukturen (dvs man skall inte behöva bläddra sig ner till rätt nivå i hierarkin för att hitta länkar indexerade med termer ur den grenen av hierarkin. Finns i princip redan. Går att åstadkomma även med specialfallsprogrammering av XSLT-script (givetvis tillrådligt bara om bläddringsstrukturen är stabil).
C.3.5	Skall	Nej	Polyhierarkiska thesaurer (där samma term kan återfinnas på flera platser i hierarkin) skall hanteras korrekt i bläddringsstrukturen. TKL klarar inte detta eftersom bläddringsstrukturen är hårt knuten till filsystemet. Å andra sidan kan TKL leverera allt som går att formulera i XML och kan omvandlas till hypertext med XSLT. MESHs struktur är av detta slag. Efter granskning av XML-kodningen av MESH som den levereras av NLM, uppfattar vi denna uppgift som görbar men dessutom sannolikt karaktärsdanande och nyttig uppgift för någon stöddig XSLT-programmerare.

Krav	Angelägenhet	Uppfyllt	Kommentar
C.3.6	Bör	Nej?	Bläddringsstrukturen skall kunna presentera alla länkar indexerade med termer ur ett visst delträd av hierarkin på samma sida. (Detta kan vara önskvärt om antalet utnyttjade termer från thesauren är stort i förhållande till antalet indexerade länkar.) TKL klarar inte detta som standard. Vi ställer oss dessutom frågan om det verkligen är önskvärt att presentera de delar av (en omfattande) vokabulär som inte används?
C.3.7	Skall	Ja	Om vissa nivåer i en stor, hierarkisk thesaurus inte utnyttjas vid indexering, skall luckorna kunna hanteras av bläddringsstrukturen. Kan realiseras som tom vidarenavigeringsnod, som utelämnad nod eller som nod med upplysningen att inget material finns förlagt hit. Man kan även rekursivt slå samman poster till närmast liggande bredare termer.
C.3.7.1	Skall	Ja	Ämnesrubriker, ämnesord och klassningar till vilket för tillfället inga poster finns knutna skall döljas i sökformuläret och resultatlistor i det externa gränssnittet. Se C.3.7.
C.3.8	Skall	Ja	Rubrikerna i bläddringsstrukturen inom ett ämnesområde skall kunna bestå både av termer ur en thesaurus och egna termer, som kompletterar thesauren.
C.3.9	Skall	Ja	Användaren skall vid bläddring kunna begränsa till resurser på ett visst språk. Kommentar: Detta är det klassiska "filterproblemet", som går igen i många söktjänster. Lösningen, som i och för sig inte har med TKL att göra, utnyttjar att XSLT-script har förmågan att testa huruvida en CGI-variabel är definierad eller inte. I fallet att den (t ex en språkvals-menyn) är det tar XSLT-scriptet istället för alla poster i aktuellt bläddringsuppslag exempelvis bara de med language = 'se'.

Krav	Angelägenhet	Uppfyllt	Kommentar
C.3.10	Skall	Ja	Gränssnitt för ämnesdefinierad bläddring i databasen skall finnas, med möjlighet till vidare navigering på minst 4 hierarkiska nivåer.
C.3.11	Skall	Ja	Under varje ämne i blädderstrukturen skall en sekundär sortering efter resurstyp vara möjlig, och i de fall en resurs knutits till flera resurstyper skall den visas under varje resurstypsubrik. Sekundär resurstypsortering kan realiseras med XSLT. Dock osäkert om en träff kan upprepas för alla de resurstyper den klassats som. Standard är sortering efter titel.
C.3.12	Bör	Ja	Under varje ämne i blädderstrukturen bör dessutom alfabetiska listor kunna väljas. Om önskemålet gäller sortering alfabetiskt map titel finns det redan. Andra sorteringar går direkt att åstadkomma med hjälp av XSLT. Om alfabetiska ämnesuppslagslistor ska kunna väljas, ligger detta troligen utanför TKL som sådant. En framkomlig väg skulle kunna vara att ämnesträdet traverseras t ex en gång per dygn och en alfabetisk länksamling framställs. Den görs tillgänglig på varje bläddringsnivå via det XSLT-script, som framställer bläddringsuppslagen.
C.3.13	Skall	Ja	I den bläddringsbara ämnesstrukturen skall kontaktuppgifter till ämnesansvarig redaktör vara lätt tillgänglig. Är faktiskt så enkelt som en lay-outfråga.
C.4			Sökfunktion.
C.4.1	Skall	Ja	Unicode. I skrivande stund är inmatningen fortfarande i ISO-8859-1, men detta bör enkelt gå att ändra. Sökindex är däremot i UTF-8.

Krav	Angelägenhet	Uppfyllt	Kommentar
C.4.2	Skall	Ja	"Specialtecken" (å, ä, ö, etc) skall kunna anges med eller utan diakritiska tecken och ändå ge samma sökresultat. Tja, det bor ga att astadkomma detta. Fragan ar om man vill att ål (fisk) och al (trad) skall ge samma traffar?
C.4.3	Skall	Ja	Frassökning. Finns inte som standard. Enkelt och billigt.
C.4.4	Skall	Ja	Högertrunkering. Standard.
C.4.5	Skall	Nej	Maskering. Bör gå att implementera genom sökning med regularutryck. Vem vill ha det?
C.4.6	Skall	Ja	Sökning med booleska operatorer. Standard. Försökspersonerna har rapporterat att bara enordssökning i testportalerna klaras av. Detta är en känd bugg, en fix kom i början av februari. Prova t ex sökning på "chemistry and molecules" och "chemistry or molecules" i DOAJ. Sökspråket är CCL, men fältsökning enligt denna standard stöds ej.
C.4.7	Skall	Nej	Fritextssökning. Fritextsökning i hela posten är standard. Däremot finns inte fritextsökning i dokumentent. Kräver harvesting. Kan utvecklas, och inte särskilt tidskrävande.
C.4.8	Skall	Ja	Strukturerad sökning (kombinerade sökfält). Se C.4.6. Boolesk fältsökning går att implementera, och är relativt enkelt att åstadkomma.
C.4.9	Bör	Ja	Ämnesord ur thesaurerna skall vara sökbara både som fraser och fritext. TKL ger träffar i såväl kategorier som i poster, ungefär som Yahoo brukade göra förr i tiden.
C.4.10	Bör	Nej	Ett sökbart index över det innehåll som pekats ut av länkarna skall kunna byggas. Kräver en insamlingsrobot, som producerar TKL/link-poster enligt uppgjort schema. Se C.4.7.

Krav	Angelägenhet	Uppfyllt	Kommentar
C.4.11	Bör	Ja	Ämnesord ur thesauren skall vara sökbara både som fraser och fritext. Som standard gäller fritext. Se C.4.9.
C.4.12	Skall	Ja	Fritextssökning samt strukturerad sökning skall finnas tillgänglig från alla sidor. En enkel sökruta finns redan tillgänglig från alla portalens sidor. Att därifrån länka till ett mera avancerat sökformulär kräver bara en enkel XSLT-ändring. (Se C.4.6.)
C.4.13	Skall	Ja	Den detaljerade katalogposten skall ge möjlighet till vidare sökning på relaterade poster i databasen. Att generera länkar för sökning efter relaterade poster i databasen ska kunna realiseras enkelt. Se C.5.2.
C.4.14	Bör	Ja	Den detaljerade katalogposten skall ge möjlighet till vidare sökning på relaterade poster i LIBRIS (och gärna andra tjänster), åtminstone på ämnesord. Z39.50- och SOAP-gränssnitt är integrerade i TKL.
C.5			Svarspresentation (vid sökning och bläddring).
C.5.1	Bör	Ja	Användaren skall kunna välja utförlighetsnivå på presentationen. Se C.5.2.
C.5.2	Skall	Ja	Det skall kunna finnas en lättillgänglig möjlighet att gå in i den detaljerade katalogposten. Fullpostvisning i TKL har framgångsrikt implementerats i DOAJ.
C.6			Nya poster.
C.6.1	Skall	Ja?	Flagga för nyinlagda poster skall finnas. Se C.6.3.
C.6.2	Bör	Nej	Möjlighet att flagga för poster som använder andra alfabet än det latinska. Om alla poster är i UTF-8 är det inte trivialt att avgöra vilka alfabet de kräver.

Krav	Angelägenhet	Uppfyllt	Kommentar
C.6.3	Skall	Ja?	Separat utmatning av nyinlagda poster skall finnas på särskild sida. Borde kunna gå att realiseras med ett XSLT-script, som utnyttjar det datum posten skapades.
C.7	Skall	Ja	Formulär för länkförslag och synpunkter från användaren skall finnas. Standard
C.8	Bör	Nej	Personalisering. Populärt krav. Går att åstadkomma här. Session-cookies finns implementerade redan och kan utnyttjas för ändamålet.

D. Teknisk arkitektur. Vad klarar TKL?

lb lb lb lb . Krav	Angelägen	Uppfyllt	Kommentar
D.1.1			Samsökningsmodellen
D.1.2	Bör	Ja	Möjlighet till samsökning med andra ämnesportaler och t ex LIBRIS Samsökning mellan TKL-portaler går lätt att implementera, antingen via OAI-harvesting av dem, eller genom att var och en av dem öppnar en Z39.50 server (vilket i princip ingår i TKL).
D.1.3	Bör	Nej	Central databas Även om samtliga svenska ämnesportaler samlades på en server är det tveksamt om man skulle kalla denna samling av tjänster för en central databas. Det beror på vad man menar med 'central' och med 'databas'.

lb lb lb lb . Krav	Angelägen	Uppfyllt	Kommentar
D.1.4	Bör	Ja	Portalen och databasen skall kunna användas som ett helt integrerat, gemensamt och centralt system Det är fullt möjligt att, utifrån ett nätverk av TKL-baserade tjänster (oavsett om de lever sitt liv i en central databas eller ej), åstadkomma en central tjänst som för användaren framstår som helt integrerad och gemensam. Det enda som begränsar möjligheterna att uppnå detta är mjuka frågor som gemensamma metadatamodeller och katalogiseringsregler.
D.1.5	Bör	Nej	Central databas för integrerade tjänster/portaler skall finnas, eller en motsvarande arkitektur som till alla delar och lika enkelt erbjuder samma möjligheter till administrativ och teknisk samordning/rationalisering som en central databaslösning. (Se diskussionen i D. Central drift och utveckling)
D.1.6	Bör	Nej	Samsökningsmodell för icke integrerade tjänster/portaler skall skapas. Måste utvecklas. (Se diskussionen i A. Innehåll i tjänsten)
D.2			Teknisk plattform
D.2.1	Bör	Ja	Plattformsoberoende De i TKL använda komponenterna fungerar i både Windows och Unix/Linux miljöer. Verktyget är dock bara testat under Unix.
D.2.2	Bör	Ja	Stöd av XML (export/import av data via XML)
D.2.3	Bör	Ja	Skall kunna hantera Unicode UTF-8 stöds.
D.2.3	Bör	Ja	Språkform. Portalen måste stödja möjlighet till flerspråkighet. Se även ovan B.3.

lb lb lb lb . Krav	Angelägen	Uppfyllt	Kommentar
D.2.4	Bör	Ja	Skalbarhet TKL skalerar för alla vanliga storlekar på akademiska svenska ämnesportaler, dvs 2000 -- 5000 poster.
D.3	Skall	Ja	Autentificering
D. 4	Skall	Nej	Hantering av statistik för den gemensamma och lokala tjänster Kan utvecklas
D.4.1	Skall	Nej	Hantering av statistik: Det skall finnas statistiktjänster, både användarstatistik och "katalogiseringsstatistik". Användarstatistiken bör kunna skilja på katalogisatörers och slutanvändares besök. Kan utvecklas
D.5	Bör	Ja	Nyhetservice TKL kommer med en nyhetservice integrerad. Den kan även leverera material med RSS.
D.5.1	Bra att ha	Nej	Möjlighet till e-postkopplad nyhetservice skall finnas.
D.6			Kalendarium
D.6.1	Bra att ha	Nej	Möjlighet till databasregistrering av nyhetsblänkare till "Aktuellt"-spalter/motsvarande skall finnas.
D.6.2	Bra att ha	Nej	Kalendarium bör finnas
D.7	Bör	Nej	Insamling av resurser med robot Kan utvecklas.
D.7.1	Bör	Nej	En "harvester"-funktion för att automatindexera utvalda webbplatser. Kan utvecklas

Teknisk samverkan mellan svenska ämnesportaler - dokumentarkiv

Bakgrund

- DEF-utvärdering - projektorganisation (.doc)
- Projektbeskrivning. DEF Toolkit - test och utvärdering (.doc)
- Materials.dk: vejledning til administrationsgrænsefladen (Hjälptext för en av de danska fagportaler som använder TKL)
- TKL Dokumentation: IndexData

Kriterieunderlag

- Kriterieunderlag V0.01 Skickat till referensgruppen 2002-12-20 (.doc)
- Kriterier för utvärdering av DEF Toolkit: Sammanställning av svaren från ansvarsbiblioteken 2003-01-28 (.doc)
- Kriterier för utvärdering av DEF Toolkit: Sammanställning av svaren från ansvarsbiblioteken 2003-02-11 (.doc)

Testschema

- Testschema 1.0 Presenterad på referensgruppsmötet i Lund 2003-03-05 (.doc)
- Administration av TKL ger tips om användaradministration, att ändra meta-dataformuläret och att skapa en kontrollerad vokabulär (.doc)
- Att skapa, editera och makulera poster även med information om att "dubbelklassa" en post och att flytta en post i ämneshierarkin (.doc)

Referensgruppens rapporter

- Maria Cederblad: (Malmö Högskola) - Länksamling för pedagoger (Rapport)
- Ann-Sofi Cullhed (Uppsala UB) - Agora (Rapport)
- Ylva Gavel, (Karolinska institutets universitetsbibliotek) (Rapport)
- Kjell Jansson (Kungl Tekniska högskolans bibliotek), - fd. EELS (Rapport)
- Åse Paulsson (Biologibiblioteket, Lund) - Biogate (Rapport)
- Alexandra Machol Ravnskov (Stockholms UB), - Rättskällan (Rapport)
- Gullbritt Åhman (Göteborgs UB)- Samweb (Rapport)
- Per Åkerlund (Sveriges lantbruksuniversitets bibliotek) - NOVAgate (Rapport)