

Open Repositories 2007

23-26 januari 2007, San Antonio, Texas USA

Sofia Arvidsson, Bibliotek & länderesurser, Högskolan i Borås

Bakgrund

September 2006 installerade Bibliotek & länderesurser ett system för digital publicering för högskolans institutioner; Borås akademiska digitala arkiv (BADA). BADA används för digital publicering av forskningspublikationer, studentuppsatser och andra digitala objekt. <http://bada.hb.se>

Som teknisk plattform används programvaran DSpace. Denna plattform är utvecklad vid Massachusetts Institute of Technology och lanserades 2002 som ett så kallat Open Source-program (öppen källkod). Lärosäten och andra organisationer kan använda plattformen utan någon licensavgift. DSpace används idag vid över 100 lärosäten över hela världen, och i Sverige av bland andra Högskolan i Borås, Göteborgs universitet, Malmö högskola och Högskolan i Halmstad.

2006 hölls den första konferensen för öppna digitala arkiv; Open repositories 2006 i Sydney, Australien http://www.apsr.edu.au/Open_Repositories_2006/. 23-26 januari 2007 hölls den andra konferensen, Open Repositories 2007 i San Antonio, Texas USA <http://openrepositories.org/> och Bibliotek & länderesurser sökte hösten 2006 resestipendium från Svensk Biblioteksförening för att delta i konferensen.

Jag vill tacka Svensk Biblioteksförening för det resestipendium som möjliggjorde min resa till konferensen Open Repositories 2007. Jag vill också tacka det trevliga resesällskapet Carin Björklund och Jonas Gilbert från Göteborgs universitetsbibliotek.

Open Repositories 2007

Konferensen bestod av två delar; under den första halvan träffades användargrupper för DSpace, Fedora och E-prints. Personer från olika universitet höll halvtimmeslånga föredrag om projekt de genomfört eller presenterade planer för utveckling. Andra halvan av konferensen var Plenary sessions uppdelade på sex teman: Management Strategy and Policy, Preservation, User Services and Workflow, Semantic Web and Web 2.0, Interoperability och e-Science and e-Scholarships. Konferensen samlade ca 350 deltagare.

DSpace

Dspace-föredragen inleddes med att MacKenzie Smith från MIT som varit med och utvecklat DSpace berättade om användargruppens nuvarande status. För närvarande är det över 180 registrerade arkiv enligt wikin:

<http://wiki.dspace.org/index.php/DspaceInstances>. 50% av dessa är amerikanska och resten finns runt om i världen. Det finns 5 officiella användargrupper och 11 "committers"; programmerare som ansvarar för källkoden, och som utvecklar programvaran tillsammans med många av användarna. Det har släppts 5 versioner av DSpace och den senaste är 1.4.1. Under 2006 diskuterade DSpace Governance Advisory Board den framtida organisationen och behovet av en organisation som kan tillvarata användarnas intressen. Organisationen ska förvalta källkoden, söka bidrag, utveckla standarder, ge support till användarna och utarbeta tekniska strategier <http://wiki.dspace.org/index.php/DspaceGovernance>. MacKenzie Smith avslutade

med att poängtera att DSpace ska förbli stabilt och fortsätta vara ett 100% fritt och öppet system.

Därefter redogjorde John Mark Ockerbloom, Pennsylvania University för de ändringar som är på väg i DSpace arkitektur och de anpassningar som planeras för DSpace 2 som kommer inom 2 år. <http://wiki.dspace.org/index.php/ArchReview>
Scott Phillips, Texas digital library visade Manakin, som är ett ramverk för gränssnitt och det finns nu en ny speciell implementation till Dspace. Manakin är xml-baserat och Manakin 1.0 släpptes 22 januari 2007; <http://di.tamu.edu/projects/manakin>. Med Manakin kommer det att bli lättare att anpassa gränssnittet i DSpace och göra olika utseenden på olika enheter och samlingar.

Därefter följde ett flertal presentationer från olika universitet och organisationer och deras arbete med utvecklingar, några exempel:

Lisa Spiro från Rice University berättade om TIMEA (Travelers in the Middle East Archive) som innehåller stora mängder bilder och kartor och förklarade hur de löst hanteringen av dessa. <http://timea.rice.edu/index.html>

Susan Wells Parham från Georgia Tech berättade om SMARTech, Georgia Techs Dspace som de lanserade 2004. Från början fick forskarna uppgiften att registrera deras egna publikationer men eftersom de var få som prioriterade uppgiften, även bland administrativ personal och examinerade studenter, lade bibliotekspersonalen själva in uppgifterna. De ändrade syn på SMARTech från produkt till service; SMARTech innehåller en samling tjänster och är inte bara ett system. De sökte även upp allt som inte var upphovsrättsskyddat på universitetets fakulteter och registrerade det i SMARTech. Hon pratade också om hur de lagt in annat material än vetenskapligt, t ex universitetets newsletter och radioprogram vilket är mycket populärt och uppskattat av studenterna på universitetet. <http://smartech.gatech.edu/>

Jim Downing, Cambridge som även är en committer pratade om SPECTRA; ett system som bl a täcker 3 områden inom kemi. De har anpassat inmatningsformuläret så man får möjlighet att ange en embargotid vid registreringen av publikationer vilket är något som borde vara intressant för de flesta digitala arkiv. Han pratade också om vikten av att lägga ut forskningsdata för att kunna upprepa forskningsresultat. Eftersom forskningen är finansierad med statliga medel ska den också finnas tillgänglig för allmänheten. Enligt deras erfarenhet var t ex förlagen mer receptiva för Open Data (dvs möjligheten att lägga ut forskningsdata) än för Open Access eftersom det ökar värdet på deras publikationer.

Tim Donohue från Illinois Digital Environment for Access to Learning & Scholarships (IDEALS), University of Illinois hade gjort förändringar i inmatningsformuläret i DSpace. Han förklarade utförligt bakgrunden till omplaceringen av de olika stegen i inmatningsformuläret. Det är t ex inte idealiskt för en användare att godkänna ett avtal som ett sista steg i en process, först efter att man fyllt i en mängd metadata. Det vore bättre om man fick läsa avtalet direkt så man vet vad det innebär att publicera sig innan man börjat fylla i bibliografiska uppgifter. I hans lösning ligger därför Avtalet som första steg, därefter kommer Uppladdning av fil, Beskrivning, Granskning, Slutför. Anledningen till att Uppladdning av fil i kommer innan Beskrivning är att det skapar möjlighet att extrahera viss metadata ur den uppladdade filen som kan användas i steget Beskrivning.

Plenary sessions

Sally Jo Cunningham, University of Waikato, New Zealand på plenary session: User Services and Workflow berättade om en användarstudie för DSpace och E-Prints med en grupp bibliotekarier och en grupp forskare. "Librarians are not regular people" menade hon eftersom bibliotekarier i princip läser allt på en webbsida och läser även hjälptexter vilket nästan inga andra gör. Bibliotekarierna var väldigt måna om att allt skulle vara "rätt" innan de kunde gå vidare till nästa steg i inmatningsformuläret. Forskarna däremot var slarviga med metadatan och tyckte ämnesorden var för specifika. Bibliotekarierna tyckte ämnesorden var för generella. Hon påpekade på olika sätt de olika synsätt dessa två grupper har, vilket kan vara bra att förstå när man ska gå ut med information och instruktioner om digital publicering till olika grupper.

Eric Larson från University of Wisconsin- Madison Libraries på plenary session: Semantic Web and Web 2.0, beskrev BibApp som är byggt med flera olika Open Source program. BibApp fungerar som en digital mashup som hämtar in data från flera olika sidor och presenterar dem i en ny sida. Man kan till exempel skapa personaliserade sidor för forskare. Förutom kontaktinformation och samlade publikationer kan man visa funktioner som t ex "People I have been published with" vilket bygger på metadata från citeringsdatabaser. Sidor som dessa kan skapa bättre kommunikation och medvetenhet mellan institutioner. Koden till BibApp kommer att släppas på code4lib 2007 <http://www.code4lib.org/2007> .
<http://code.google.com/p/bibapp/>

Nästa konferens är Open Repositories 1-4 april 2008 i South Hampton.

<http://openrepositories.org/2008/>

Balans mellan praktiska exempel och visioner inför framtiden

På konferensen fick jag en del idéer hur vi kan arbeta vidare med Borås Akademiska Digitala Arkiv (BADA) och hur vi ska arbeta med att informera och utbilda personal på Högskolan i Borås. Det var intressant att höra hur DSpace har utvecklats och hur DSpace verkar ligga till i förhållande till de andra Open Source system som finns. Även idéer och visioner inför framtiden var intressanta. Det var överlag en bra balans mellan just praktiska exempel och idéer om vad man vill utveckla för framtiden. Man förstår att de större universitet som var representerade på konferensen har mycket resurser för att utveckla sina digitala arkiv, men det är ändå bra och roligt att vi som en liten högskola i Sverige har möjlighet att använda samma system. Samma internationella databaser och sökmotorer hämtar data från både vårt arkiv och de stora universitetens arkiv.

Borås 22 februari 2007

Sofia Arvidsson
Bibliotek & läranderesurser
Högskolan i Borås