

Reserapport från Barcelona

XML Europe 2002, 20-23/5 2002

Bengt Eriksson, Lars Brink

Allmänt om konferensen	3
Presentationer etc.....	3
Vilka sessioner beskrivs?	3
Måndag	4
Tutorial: Making knowledge management work.....	4
Allmänt	4
Innehållsligt	4
Tisdag.....	4
Keynote.....	4
XML and WORD: Downcasting and Upcasting	5
Standards Update ICE (Information and Content Exchange)	5
Onsdag, torsdag	5
There are no unstructured documents	5
e-government in the UK – e-voting Technical standards.....	5
XML as a tool for legal validity in a Security Context	5
XPIPE	6
XQuery.....	6
XML och databaser	6
Enabling Low-Cost XML-Aware Searching Capable of Complex Querying	6
Utställningen.....	7
Exegenix	7
Bra/Dåligt.....	7

Allmänt om konferensen

Presentationer etc.

Alla presentationerna har vi fått på en CD. Så småningom lägger vi in den i INFO-konferensen om möjligt, och därigenom ges också några roliga exempel på presentationsteknik.

Vilka sessioner beskrivs?

Vad beträffar de övriga sessionerna, så gick Lars (Brink) och jag delvis på samma sessioner, men när vi ansåg det motiverat så delade vi på oss och gick på det som låg närmre våra resp. specialområden.

Jag tänkte inte beskriva precis varenda session jag/vi var på, bl.a. av den anledningen att alla inte var lika bra. Tyvärr är det ju svårt att veta sådant i förväg.

Måndag

Tutorial: Making knowledge management work

Allmänt

En "tutorial" är ett annat namn på en workshop. Den kan vara mer eller mindre interaktiv, men den behandlar ett och samma ämne över en längre period, i det här fallet en hel dag. Det vanliga på dessa konferenser är att den ligger före inledningen av den egentliga konferensen, och totalt, av antalet konferensdeltagare, så kanske 20% deltar i dessa tutorials. Man betalar också extra för den.

Vad gäller interaktiviteten så kunde vi ha valt någon annan tutorial om vi hade velat ha mer interaktivitet. Problemet var bara att det då förutsattes att man skulle ha egen PC med sig, och med tanke på stöldrisken i allmänhet och i Barcelona i synnerhet, så beslöt vi oss för att försöka klara oss utan PC. Det betydde då att denna tutorial inte hade speciellt många inslag av interaktivitet – om en några – och det var litet tråkigt.

Innehållsligt

Som framgår av titeln handlade seminariet om knowledge management. "Kunskapshantering" kanske kan vara en bra svensk översättning.

Den fullständiga titeln var

"Making Knowledge Management work with XML, Xlink, Topic Maps, and AI"

Som man kan förstå av titeln handlade detta om hur man hanterar den kunskap man har inom en organisation t.ex., med användande av XML och tillhörande standarder. Detta var en ganska avancerad föreläsning.

En sak som jag tyckte var intressant här var företeelsen "Topic Maps", vilket är ett sätt att organisera sin kunskap på ett bestämt sätt. Sättet att göra detta beskrivs i en ISO-standard (ISO/IEC 13250:2000) som beskriver kunskapsstrukturer, elektroniska index och klassificeringsschemata bl.a. Ett exempel på användningen av Topic Maps finns i presentationen av själva konferensen. Presentationen ligger på en CD som kan rekommenderas för den som är intresserad av dylikt.

Att det finns oerhört många bottnar i kunskap blir man om inte annat klar över när man tittar på den omnämnda CD:n, vilken jag egentligen tycker tar det hela till en alltför hög grad av detaljering. Det är lätt hänt att detaljerna skymmer helheten. Men bortsett från det är det ändå klart att för en stor organisation som t.ex. riksdagen, så nästan krävs det ett medvetet och organiserat sätt att strukturera och hantera sin kunskap, inte minst när man vill göra det över en längre tidsperiod.

Tisdag

Keynote

På tisdag öppnade den egentliga konferensen. Som keynote-speaker var Peter Pappamikail från Europaparlamentet som höll en lång exposé över hur bra det egentligen går med XML-införande, dokumentstrukturering och för all del behovet av detta, i organisationer som våra, dvs. parlament och regeringskanslier. Ett intressant föredrag med den likaledes intressanta titeln "Of Carrots and Sticks, Chickens and Eggs".

XML and WORD: Downcasting and Upcasting

Vi var särskilt intresserade av hur man kunde gå från XML-dokument till WORD och tvärtom. Programmet Upcast äger vi sedan tidigare. Det är ett program som gör om Word-dokument till XML. En ny produkt var tänkt att visas på denna konferens. Produkten kallas Downcast, och gör motsatsen till Upcast, dvs. den gör en XML-instans till ett Word-dokument.

Tyvärr ville ödet annorlunda för denna demonstration. De personer som representerade företaget blev rånade i Barcelona, dagen innan sin demonstration och de kunde inte på den korta tid som stod till buds få fram en ersättningsdemonstration.

Standards Update ICE (Information and Content Exchange)

ICE är en XML-baserad standard som uppmärksammades som en W3C-not i slutet av 1998 och som kan användas för bl.a. för datautbyte mellan system. Rent konkret är det inte annat än ett protokoll för hur olika programvaror ska kunna förstå varandra. Men ICE kan även ses vara en arkitektur för hur man bygger lösningar för informationsförsörjning och spridning. ICE kan användas för replikering och distribution av information mellan olika system. ICE kan sägas vara en standard för PUSH/PULL av information. Ett område som ICE är tänkt att användas inom är bl.a. s.k. informationssyndikat, där flera olika informationssystem kan samordnas trots att plattformen inte är den samma. ICE 2.0 planeras släppas i september 2002 där nyheterna är stöd för WSDL, SOAP, UDDI och XSD med namnrymder (namespaces). Dr Martin, som är ordförande i gruppen som utvecklar ICE, visade även på att ICE både har stöd och kan ge utökad funktionalitet till RSS (ett open source initiativ för informationsutbyte som bl.a. används i Netscape 1999-). Fördelarna med ICE jämfört med RSS kan främst tecknas vara att ICE-överföringar kan liknas transaktioner medan RSS är mindre säkert och att ICE tillåter reläande.

Onsdag, torsdag

There are no unstructured documents

Denna föreläsning handlade om hur man utifrån ett helt vanligt dokument kunde låta ett program läsa igenom det och förstå sig på vad som var en rubrik, vad som var en punktlista, vad som var en tabell etc., genom att se hur mycket "white space" det fanns. På så sätt kunde man, utifrån ett vanligt Word-dokument som var någotsånär snygg formaterat skapa en XML-kodning för det.

e-government in the UK – e-voting Technical standards

Dessa föredrag handlade om hur långt man hade kommit i Storbritannien med att kunna erbjuda medborgarna web-baserade tjänster och man hade kommit väldigt långt, inte minst vad gäller röstning i allmänna val över Internet (allt XML-baserat). Där hade man själva infrastrukturen klar och menade att man med nödvändiga lagändringar i princip skulle kunna genomföra detta. Emellertid så fick jag aldrig riktigt klart för mig hur man skulle lösa känsliga frågor som integritet, rösthemlighet, säkerställande av att ingen röstade fler än en gång etc.

Man samarbetade med USA i den här frågan och som ett stort fälttest skulle man för USA:s utlandsbaserade styrkor låta dessa genomföra ett fullskaletest i höst i samband med amerikanska delstatskongressval.

XML as a tool for legal validity in a Security Context

Detta föredrag hölls av Cecilia Magnusson Sjöberg som är professor på institutionen för rättsinformatik på Stockholms universitet. Vi har samarbetat och samarbetar med henne i URIS-projektet. Det ämne som hon behandlade här handlade om hur man t.ex. kan skriva kontrakt, ingå avtal etc. i elektronisk form, och hålla rätt på sådant som versioner, verifiering, lagligt bindande. I alla dessa frågor spelar XML en central roll, och för Cecilias del har hon hållit på med dessa frågor även på SGML-tiden, i ett projekt som hette Corpus Legis.

XPIPE

Sean McGrath från företaget Propylon höll i en session som handlade om Xpipe. Syftet med Xpipe är att finna en metod för att utveckla skalbara och återanvändningsbara lösningar avseende transformeringar av XML-data. Istället för att göra massiva transformeringar av XML-data till sitt slutformat delas transformeringarna upp i mindre delar som kan bearbetas på flera datorer. Varje del som gör en transformation sägs vara en *Xcomponent*. Sätter hur en Xcomponent fungerar är helt öppet. Xpipe är protokollet och det som binder samman och driver transformationskedjan. Tankesättet är i sig föga nytt, men nytt är dock att varje del i kedjan kan göras till en transaktion och fördelas. Vinsterna med detta är skalbarhet, återanvändbara komponenter och feltolerans. Enligt föredragshållaren Sean McGrath har det visat sig att XSL inte är skalbart pga. att XSL är för resurs- och minneskrävande. Vidare menade han att XML är kraftlöst ur ett programmerarperspektiv. För att uppnå skalbarhet måste man ta an ett idéer från programmering. Xpipe är i dag inte en standard utan bara en not hos W3C. Xpipe finns i dag endast implementerat i en open source produkt under SourceForge (Propelx).

XQuery

Xquery heter den senaste standarden för frågespråk till XML. Tidigare hette språken Quilt och XQL. I dag är det Xpath och XQuery som gäller där Xpath är den standard som är vanligast. Xquery är liksom Xpath på gång med en uppdaterad version. Efter sett ett antal exempel på hur frågor formuleras med Xquery undrar jag om språket någonsin kommer bli spritt. Syntaxen är komplex och känns inte alls naturlig och verkar mesta vara gjort för XML-expert. Kommande version av Xquery föreslås ha stöd för att uppdatera data. I dag finns det bara stöd för att läsa data med Xquery.

XML och databaser

Oracle demonstrerade sin senaste databashanterare (version 9i release 2) som imponerade djupt vad gäller XML-stöd. XML-stödet var snyggt och naturligt integrerat i databasmotorns kärna. Oracles dialekt av SQL hade fått erforderligt tillägg för att göra XML-specifika sökfrågor. Det gick enkelt att kombinera XML-data med data i tabellform. Det gick både att skapa rena XML-tabeller och använda XML-fält i vanliga tabeller. För bästa möjliga prestanda uppmanades att man gjorde databasen medveten om datatyperna med hjälp av ett schema eller mer direkta inställningar av indexet. Databasen stödjer bl.a. både Xpath-frågor och SQLX.

Som ett *late breaking news* föredrogs status bland de rena XML-databaserna. Föredraganden gick inte närmare in på någon jämförelse med andra databashanterare men sa att produkterna kan anses mogna och bra som alternativ eller komplement i XML-lösningar. Aspekten när och hur de stora databasleverantörerna kommer hantera XML berördes inte alls vilket var synd.

Enabling Low-Cost XML-Aware Searching Capable of Complex Querying

Tre entusiastiska utvecklare berättade under om deras uppgift att för internt behov ta fram en lösning för att kunna söka bland sina XML-dokument. Projektet saknade medel för att köpa in en riktig XML-databas och utvecklarna var blev därav tvingade att titta efter en alternativ lösning. Efter en kort inventering av vad som fanns att tillgå fastnade man för att utveckla en sökmotor baserad på sökmotorn *Lucene*, som är en javabaserad sökmotor som utvecklas under delprojektet *Jakarta* tillhörande open source-projektet *Apache*. Med hjälp av Lucene lyckades man på kort tid utveckla en sökmotor som ledde till resultat som var över förväntan. Lösningen visades sig kunna erbjuda samma – om inte bättre – funktionalitet än vad många kommersiella XML-databaser erbjuder.

Lösningen kan bl.a. ge svar på frågor såsom att hitta ett visst element/attribut med ett visst innehåll eller kombination, relevansranking, sökning på hierarki och boolesk logik (bl.a. OCH och ELLER).

Arbetet finns att ladda ner från på www.isogen.com. Utvecklarna varnade dock för att det kan vara svårt att installera lösningen för en ovan person.

Utställningen

Adobe FrameMaker 7.0 visades och finns ute och har numera stöd direkt för XML och PDF.

SoftQuad, XMetal 3.0 har med denna version bl.a. stöd för XSD, Unicode, utökat scriptstöd, PDF, WebDAV och import av worddokument.

Exegenix

Företaget visade upp en intressant produkt som kunde tolka både postscript och PDF-filer för att sedan konvertera dessa till XML-format. XML-formatet sades bygga på DocBook med något tillägg. Produkten kan köra batchjobb och ligga på en separat server och är skriven i Java, vilket gör att den går att installera på ett flertal plattformar. Produkten kan mycket väl passa in i en lösning för att överföra äldre riksdagstryck till XML.

Bra/Dåligt

Det är alltid svårt att i förväg avgöra vilken kvalitet en sådan här konferens kommer att ha. Ostridigt är i alla fall att under de här dagarna så var hela Europas XML-expertis samlad i Barcelona. Det finns alltså inget mer kunskapsstätt ställe att åka till om man vill hålla sig informerad om utvecklingen på detta område. Eftersom riksdagen mer eller mindre formellt har sagt att man skall använda XML i olika aspekter av informationshantering och dokumentframställning så är man tvungen att hålla sig à jour med utvecklingen.

Beträffande själva framställningarna kan man möjligen säga, åtminstone för vissa föredrag, att kvantiteten information var större än kvaliteten, inte nödvändigtvis så att det var dåligt, men det var slående hur ofta man stötte på resonemanget att eftersom man hade så litet tid kvar så var man tvungen att köra igenom sina powerpoint-bilder fortare. En alternativ taktik kunde ju ha varit att man hade anpassat det man ville säga till den tid som man hade tillgänglig.