

IT-KOMMISSIONEN

Breddtjänster – några exempel

Rapport 51/2002

Innehåll

VISBY KUSTKORVETT	5
20 HYLLMETER PAPPER FÖRSVANN	5
KOMPLICERADE SYSTEM KRÄVER OMFATTANDE INFORMATIONSHANTERING.....	5
ALL INFORMATION I ETT XML-FORMAT	6
STRUKTURERAD OCH STANDARDISERAD INFORMATION GER FLERA FÖRDELAR:	7
SVÅRT ATT SAMORDNA OCH INTEGRERA OLIKA KÄLLOR	7
BÄTTRE INFORMATION OM LÄKEMEDEL	9
STOR VANA ATT HANTERA INFORMATION ANSVARSFULLT	9
ALL DOKUMENTATION LAGRAS CENTRALT SOM XML-FILER	10
LÖSNINGEN GER KONTROLL, ENHETLIGHET OCH AKTUALITET	11
AVTAL OCH KRAV PÅ TÄTA HÄMTNINGAR	12
PLATSANNONSER OCH CV:N.....	13
ETT EFFEKTIVARE FLÖDE AV "SÖKA JOBB-INFO"	13
OMFATTANDE INFORMATIONSHANTERING KOSTSAMT	13
FRAMTIDA PLANER.....	14
ALLA TJÄNAR PÅ PROJEKTET	16
KULTURELLA SKILLNADER OCH PUL	16
STUDERA.NU – EN AKTUELL OCH KOMPLETT KATALOG ÖVER HÖGSKOLEKURSER	18
INFORMATION FÖR LIVSAVGÖRANDE VAL	18
NY NATIONELL KURSKATALOG VARJE NATT	18
FÖRDELAR MED XML.....	19
KVALITETSGRANSKNING, LÄNKAR OCH ANDRA INTRESSENTER	20
EVENEMANGSINFORMATION	22
LIKA VILLKOR FÖR TURISTFÖRETAGEN.....	22
INGEN HAR KAPACITET ATT HANTERA ALL TURISTINFORMATION	22
SYNKRONISERING AV KALENDERDATABASER.....	23

FÖRDELAR MED STANDARDISERAT FORMAT	24
MEDVETENHET VIKTIGT FÖR RESULTATET.....	25
SVÅRT ENGAGERA SMÅ AKTÖRER	27
LÄROMEDEL OCH SKOLINFORMATION	28
EN MJUK INFRASTRUKTUR FÖR LÄROMEDEL OCH SKOLINFORMATION .	28
SKOLVÄRLDEN STARKT PÅVERKAD AV INFORMATIONSTEKNIKEN	28
BRED SATSNING PÅ GRUNDLÄGGANDE INFRASTRUKTUR OCH BASTJÄNSTER	28
XML MEDGER ÖPPENHET, INTEROPERABILITET OCH LÅNGSIKTIGHET .	31
ORDFÖRKLARINGAR	32

FÖRORD

IT-kommissionen pekar i sin rapport "Breddtjänster – ett nytt skede i IT-politiken" på att IT kan ge mycket mer än i dag.

Denna rapport visar med exempel på vad detta kan innebära. Rapporten visar också hur man praktiskt kan arbeta med information på ett strukturerat och standardiserat sätt.

Exemplen är ursprungligen framtagna av Metamatrix AB på uppdrag av IT-kommissionen. IT-kommissionen har inför denna publicering språkligt bearbetat materialet.

I samhället finns information som på olika sätt och via olika kanaler ligger till grund för nytta och nöje för enskilda, familjer, organisationer och företag. För att bli "nyttig" måste informationen bearbetas och "förpackas" som en tjänst. Om informationen förpackas på rätt sätt uppnår man en ökad nytta. Olika sådana tjänster kan skapas på nätet.

Många tjänster kräver att information från olika källor sätts samman. De olika informationskällorna måste då hanteras på ett sätt så att de är både tillgängliga och kan kombineras.

Många informationsmängder i samhället är direkt nödvändiga för att olika verksamheter skall fungera. De aktörer, ofta myndigheter, som hanterar dessa samhällskritiska informationsresurser har ett stort ansvar för att den information som de hanterar blir en integrerad del i samhällets informationsresurs.

De som äger eller har ansvar för viss information är inte alltid desamma som ser till att tjänsterna tillskapas och fungerar. För det behövs en god samverkan mellan olika aktörer och ett effektivt utbyte av information. Det kräver i sin tur att informationen hanteras och ges ett format så att den är lättillgänglig för olika

aktörer eller tjänsteproducenter. Frågan är – hur ser en sådan hantering ut?

För IT-kommissionen

Christer Marking
Kanslichef

Visby kustkorvett

20 hyllmeter papper försvann

Försvarets materielverk (FMV) och Marinen har valt ett digitalt och strukturerat informationssystem för att hantera information om kustkorvetten Visby motsvarande cirka 20 hyllmeter A4-pärmar. Detta eftersom man behöver enhetlighet, överskådlighet och beständighet i dokumentationen.

Komplicerade system kräver omfattande informationshantering

En korvett av Visbyklassen är en "plattform" för en mängd olika vapensystem. Ombord finns även ledningssystem, sensorer som lyssnar och ser både ovan och under vattenytan, en komplett ledningscentral, två olika framdrivningsmaskinerier och en helikopterhangar.

De flesta system är mer eller mindre integrerade med varandra. Utan överdrift kan man säga att en korvett av Visby-typ är en mycket komplicerad konstruktion. För att all denna materiel ska kunna förstås, användas och underhållas på rätt sätt, har en omfattande användardokumentation tagits fram.

För svenska örlogsfartyg, byggda före Visby, har denna information varit tryckt på papper och omfattade cirka 20 hyllmeter A4-pärmar. Då Marinen alltid försöker öka prestanda genom att minska vikten på fartyget, reducera brandfarligt material och dessutom förbättra rättningstjänsten av informationen så sökte FMV 1995 andra lösningar.

All information i ett XML-format

Resultatet blev att man valde att arbeta med att digitalisera informationen, vilket fick som följd att ombord på korvetten behövs ingen pappersdokumentation finnas kvar. Ett heltäckande datorbaserat informationssystem togs fram för att personalen skulle kunna nå all information från ett integrerat informationssystem. FMV valde en SGML-lösning (SGML är ISO-standard för strukturerad information) vars arkitektur baserades på en av FMV:s framtagna Grund-DTD (DokumentTypsDefinition, själva strukturen som bestämmer hur informationen struktureras) och egenutvecklade verktyg för att läsa den SGML-taggade informationen.

Mycket har hänt sedan dess och verkligheten har successivt förändrats. FMV har numera gått över till XML (eXtensible Markup Language, en familj av standarder för utveckling av strukturerad information) och har på så sätt fått tillgång till ett format som gör att information av korvetten kan läsas med vanligt förekommande webbläsare. Det gör att Marinen numera kan använda ett stort antal kommersiella produkter för att hantera och bearbeta informationen.

Figur 1 En schematisk bild av hur all dokumentation lagras centralt på ett enhetligt format och blir tillgänglig för all personal ombord.

Strukturerad och standardiserad information ger flera fördelar:

- Det är enkelt att tillföra ny information in i systemet.
- Informationen kan hållas ständigt tillgänglig för all personal.
- Det är enkelt att skapa en anpassad informationsmängd för varje fartygsindivid.
- Informationen kan läsas med en vanligt förekommande webbläsare.
- Ett stort antal kommersiella produkter finns att tillgå som kan hantera och bearbeta informationen.
- Marinen har stora krav på långsiktighet. Därför är t.ex. Word- eller HTML-filer inte praktiskt användbara, utan valet står mellan pärmar med papper och välstrukturerade standardiserade digitala format såsom SGML och XML.

Svårt att samordna och integrera olika källor

Svårigheterna att samordna och integrera informationskällor från olika typer av leverantörer blev tidigt ett problem i projektet. En av leverantörerna av system till korvetten fick därför uppgiften att utöver en dokumentation över sin egen produktion, även samla in och sammanställa material från de övriga leverantörerna. Som stöd för detta tog Marinen fram en specifikation som detaljerat beskriver hur den inlevererade informationen ska vara strukturerad. Här låg definitivt en av de största utmaningarna i projektet, nämligen att få leverantörer spridda över norra halvklotet att tolka och följa specifikationen på ett likartat sätt och därmed bortse från tidigare framtagna lokala lösningar.

Den tekniska utvecklingen går oerhört snabbt inom teknikinformationsområdet, så det finns hela tiden en önskan om att alltid "hänga på" det senaste och surfa på utvecklingsvågens framkant. Faran med att göra det är att man aldrig blir riktigt färdig. Därför valde man länge att hålla fast vid den gamla GRUND-DTDn. Nyligen har man dock övergivit denna till förmån för AECMA, en internationellt framtagen standard (www.aecma.com och www.tpsmq.org).

Figur 2 Dokumentationen till en Visby kustkorvett utskriven på A4 och satt i pärmar.

Visbyprojektet har bedrivit en pionjärverksamhet med alla de bekymmer det för med sig. Med facit i hand anser Marinen resultatet vara väl värt de resurser som använts för att förse Visbysystemet med ett komplett, integrerat och modernt utformat informationssystem, där mängden information dessutom är möjlig att hålla aktuell. Genom att lägga informationen på en server ombord blir den lätt tillgänglig för alla berörda. En bonuseffekt har kunnat konstateras; nämligen att man får möjlighet att skapa anpassad informationsmängd för varje fartygsindivid.

FASS.SE

Bättre information om läkemedel

Läkemedelsbeskrivningarna som bl.a. ligger till grund för den populära läkemedelskatalogen Fass hanteras numera helt som XML-dokument som lagras centralt. Fass.se är därför alltid aktuell, godkännandeprocessen hos Läkemedelsverket är snabbare, organisationerna är redo för samverkan internationellt tack vare en standardiserad dokumentstruktur och dessutom sker presentation av informationen smidigt genom många olika kanaler.



Stor vana att hantera information ansvarsfullt

Läkemedelsindustriföreningen, (LIF) www.lif.se, har sedan 1966 publicerat en katalog över godkända läkemedel i Sverige. Katalogen kallas Fass och har stor spridning bland både läkare och patienter.

Få känner till vilket företag som tillverkar ett givet läkemedel. Därför är Fass en bra säljkanal för läkemedelsföretagen och den är på så sätt ett värdefullt varumärke för hela branschen. I Fass presenteras alla produkter på ett seriöst och likvärdigt sätt. Fass finns i tre versioner - en upplaga riktad till läkare, en annan till patienterna. En tredje innehåller veterinär-läkemedel.

Fass-texterna bygger på två informationsmängder: en produktresumé som sammanfattar läkemedlets egenskaper för läkare m.fl. och en så kallad bipacksedel som sammanfattar läkemedlets egenskaper för patienter. Bipacksedeln är det informationsblad som finns i förpackningen.

När läkemedlet godkänns av Läkemedelsverket baseras beslutet bland annat på innehållet i dessa dokument, som alltså båda finns på webbplatsen fass.se. Felaktig information om t.ex. biverkningar eller dosering kan få allvarliga konsekvenser. Därför ställs höga krav på informationskvaliteten.

Att finna information om läkemedel på Internet är inte alltid så lätt. En sökning på "Viagra" ger till exempel nästan uteslutande pornografiska hemsidor som resultat, och en sökning på "Ipren" ger en hel del musiksnuttar om Iprenmannen.

I samband med att Läkemedelsverkets godkännande av ett nytt läkemedel har läkemedelsföretagen tidigare använt t.ex. Word-dokument. Dessa har skickats till Läkemedelsverket, som vid godkännande av en produkt skrivit ut dokumentet på vattenstämplade papper vilka skickats till läkemedelsföretaget. Eftersom godkännandeprocessen ibland tar tid, har det hänt att den dokumentation, som blivit godkänd, inte längre är identisk med den information som finns hos läkemedelsproducenten.

All dokumentation lagras centralt som XML-filer

LIF har tillsammans med tre läkemedelsföretag och Läkemedelsverket i ett pilotprojekt hanterat all läkemedelsinformation med hjälp av XML. Originalinformationen lagras hela tiden hos LIF och redigeras med hjälp av specialskrivna redigeringsprogram. Ca 350 handläggare från de 130 läkemedelsföretagen har på LIF:s bekostnad fått utbildning i redigering av informationsstrukturen. Projektet har fallit ut till belåtenhet för samtliga inblandade parter och branschen har därför bestämt sig för att börja jobba på detta sätt.

Den struktur (DTD) som används är en tidig version av en specifikation (<http://esubmission.eudra.org/pim>) som håller på att

standardiseras inom den europeiska läkemedelsorganisationen EMEA (www.emea.eu.int).

Vid övergången till den nya strukturen konverterades alla gamla dokument med hjälp av ett konverteringsprogram, men man var, inte helt oväntat, tvungen att manuellt gå igenom och korrigera de resulterande XML-dokumenterna.

Figur 3 Schematisk bild av hur läkemedelsinformationen flödar från läkemedelsföretagen till LIF, kontrolleras av Läkemedelsverket och slutligen hamnar i Fass, Patientfass, Fass.se, på läkemedelsföretagens webbsidor, i internationella databaser mm.

Lösningen ger kontroll, enhetlighet och aktualitet

Originalfilerna hanteras i ett gemensamt system och det är alltid helt tydligt vilken version av en text som gäller. Läkemedelsverkets godkännanden blir dessutom omedelbart synliga i systemet.

Eftersom all information lagras på ett gemensamt format kan man presentera informationen från de många olika informationsägarna på ett enhetligt sätt som understödjer jämförelser och navigation. Dessutom kan informationen presenteras i flera olika vyer. Till exempel presenteras informationen i trycksaken Fass, på webbplatsen Fass.se, och på vissa av läkemedelsföretagens egna webbsidor eller som en informationskomponent i sjukhusens vårdssystem som stöd vid receptskrivning mm. Exakt samma information men på många olika format.

Trycksaken Fass ges ut en gång per år. Rättelser och ändringar ges ut i form av tilläggsblad en gång per kvartal.

Webbpresentationen ska alltid innehålla den senast aktuella informationen.

Normalt tar det ca 10 år att ta fram ett läkemedel. Under en sådan tidsperiod hinner ordbehandlingsprogrammen genomgå många programgenerationer. Arkiveringsbehoven var ett viktigt skäl att lämna Word-formatet. Vid konvertering av gamla word-dokument till nya versioner av Word tappar man ofta viss information.

Avtal och krav på täta hämtningar

På grund av de höga kraven på informationskvalitet lämnar man inte ut originaldokument ur Fass till vem som helst. De är alltför lätta att presentera på ett till synes professionellt och därmed pålitligt sätt. LIF upprättar istället avtal med den som vill fungera som distributör av Fass-innehåll.

För att garantera aktualiteten i den information som når den slutliga mottagaren ställer LIF bland annat krav på tillräckligt täta hämtningar av senaste versionen av allt innehåll,. Man kräver även att få se hur informationen slutligen ska presenteras.

Prenumeranter inom allmännyttan betalar en symbolisk administrationsavgift till LIF, medan kommersiella aktörer får betala ett marknadsmässigt pris för Fass-informationen.

Läkemedelsbranschen har krav på öppen priskonkurrens och därför har det inte varit några problem att få leverantörerna att tydligt ange priser i XML-filerna.

PLATSANNONSER OCH CV:N

Ett effektivare flöde av ”söka jobb-info”

Arbetsmarknadsstyrelsen, AMS, håller på att tillsammans med andra aktörer på arbetsmarknaden ta fram en modell för information om lediga arbeten som bygger på ett befintligt internationellt standardramverk. Ramverket används för informationsutbyte av platsannonser och CV:n. När arbetsgivare, bemannings- och rekryteringsföretag och AMS anpassar sina interna rekryteringssystem till ett standardiserat och strukturerat format kan de utbyta information om platsannonser och CV utan manuell hantering.



Omfattande informationshantering kostsamt

AMS är central myndighet för arbetsförmedlingen, som totalt i landet tar emot ca 50 000 platsannonser per månad, och har ca en halv miljon platssökande i sina databaser. För aktuell statistik och mer information se www.ams.se.

Hantering av platsannonser och CV har hittills skett manuellt. De uppgifter som lämnas in är på papper, via fax, per telefon, eller som e-postmeddelanden eller tillsända ordbehandlingsfiler. Denna hantering är kostsam trots att den endast i begränsad utsträckning tillför ett värde. Kontroll av informationens rimlighet och fullständighet skulle kunna ske direkt hos uppgiftslämnaren om denne hade bra mallar.

Figur 4 Från manuell inskrivning till automatisk överföring mellan databaser.

AMS leder ett pilotprojekt med ett tiotal representanter, bl.a. bemanningsföretaget Manpower (www.manpower.se), personalrekryterare på Göteborgs stad (www.goteborg.se/prod/G-info/Ledigajobb.nsf) och rekryteringssajten Offentliga jobb (www.offentligajobb.se/). Målgruppen för projektet är övriga aktörer på arbetsmarknaden och syftet att pröva ett elektroniskt flöde av information.

Projektet använder den internationella branschstandarden HR-XML (Namnet kommer av den engelska branschbenämningen, human resources, se www.hr-xml.org). HR-XML-konsortiet har publicerat specifikationer för hur de flesta typer av arbetsmarknadsinformation kan representeras i XML. I det svenska pilotprojektet har man fokuserat på rekrytering men också tagit fram en modell för hur de informationssystem som används i branschen ska kunna utväxla och hantera sådan information.

Figur 5 Enklare med XML och strukturerad information..

En mycket viktig faktor när man börjar kommunicera automatiskt mellan organisationer är att samtliga involverade menar samma sak med de begrepp som används. AMS har sedan länge ett väl fungerande register över viktiga termer och begrepp, t.ex. standardtermer för yrken, kompetenser, utbildningar, länder och språk. Detta register, "AMS yrkesmallar", kommer väl till pass när mer och mer information hanteras strukturerat och standardiserat. Eftersom AMS publicerar registret elektroniskt och håller det ständigt tillgängligt kan alla i branschen använda samma begrepp i sina dokument.

Framtida planer

Det nuvarande pilotprojektet är inriktat på att sprida platsannonser. I ett senare skede planeras att även distribuera CV:n.

Syftet är att göra ett formulär på en webbsida där användaren får mata in sin CV-information och välja till vilka mottagare informationen ska distribueras. Informationen lagras sedan på en server och från den lagrade informationen genereras ett utskick till samtliga mottagare. Eventuellt skall även skapas möjligheter till återutsändning och tillbakadragande av informationen.

Eftersom ett CV är ett långlivat dokument som ofta behöver aktualiseras, kan det i detta sammanhang vara intressant att tillhandahålla maskingränsnitt mot centrala CV-databaser istället för att distribuera kopior av dokumenten som ju riskerar att bli inaktuella. Sådana maskingränsnitt kan med fördel bygga på de standarder och metoder som utvecklas och blivit kända under namnet Web Services.

AMS har även inlett samarbeten med motsvarande myndigheter både i de nordiska länderna och inom EU med syfte att gemensamt fastställa specifikationer som möjliggör ett effektivt flöde av arbetsmarknadsinformation mellan länderna. Man har även samarbeten med Skolverket m.fl. inom utbildningssektorn i syfte att kunna knyta samman utbildningsinformation med arbetsmarknadsinformation.

Figur 6 En schematisk bild av hur CV- och platsannons-information flödar från platssökande resp. företag via olika platsförmedlingstjänster och till mottagaren.

Alla tjänar på projektet

Projektet som ännu är på pilotstadium, visar redan på många potentiella vinster både för samhället i stort och för de aktivt inblandade parterna. Några av fördelarna är:

- Standardiseringen och användningen av ett gemensamt format gör att man undviker onödigt dubbelarbete i form av inskrivning i databaser.
- De XML-scheman som styr hur informationen är strukturerad minskar risken för att viktiga upplysningar glöms bort eller får orimligt innehåll av misstag.
- Det automatiserade flödet ökar förutsättningarna för alla aktörer att få en mer heltäckande bild av den aktuella situationen på arbetsmarknaden. Arbetssökande riskerar inte att missa information om ett passande jobb och arbetsgivare får större möjligheter att hitta rätt personal.
- Dokumentformatet är en öppen specifikation och ger olika användare goda möjligheter att ta del av och dra nytta av informationsflödet.
- En internationell standard förenklar både för företag och för arbetstagare att förmedla sin information utomlands. De programvaror som redan används i branschen kommer successivt få ett inbyggt stöd för detta format. Redan idag har flera programvaror sådant stöd. Några exempel är mohoResume Extractor (www.mohomine.com/products/resumeextractionfeatures.asp) och TalentFolio (www.alao.com/TFX_features.asp).

Kulturella skillnader och PUL

Vid anpassning av HR-XML från nordamerikanska till svenska förhållanden har man stött på ett visst kulturellt betingat problem: I USA har man lagstadgad positiv särbehandling av arbetssökande med avseende på etnisk tillhörighet, medan vi i Sverige har en lag som förbjuder arbetsmarknadens parter från att överhuvudtaget registrera information om ras, religion eller etnisk bakgrund.

Sådana problem kan till exempel lösas med hjälp av nationella rekommendationer och riktlinjer för hur en teknik ska användas. XML är, som namnet antyder (extensible=utbyggbart, uttänjbart), gjort för att vara anpassningsbart, och lämpar sig därför bra för att på detta sätt göra lokala specialvarianter av internationella standarder. (X:et står för eXtensible alltså utbyggbart: Extensible Markup Language, utbyggbart märkspråk).

En annan fråga som kommer att bli aktuell är hur Personuppgiftslagen (PUL, www.lagrummet.se och andra integritetsfrågor skall tillämpas. Ett effektivt och automatiserat informationsflöde om till exempel människors kompetenser och egenskaper innebär, samtidigt som det har stora nyttoaspekter, vissa risker. AMS bedömning är dock att det finns möjlighet att utforma säkra system, till exempel med hjälp av olika krypteringstekniker.

STUDERA.NU – EN AKTUELL OCH KOMPLETT KATALOG ÖVER HÖGSKOLEKURSER

Högskoleverkets webbplats www.studera.nu är en ingång till information både om högskolestudier och livet som student i Sverige. Lösningen utnyttjar strukturerad och standardiserad information och metadata för att automatisera informationsöverföringen från enskilda högskolor och universitet till studera.nu. Därmed kan dagsaktuell information om kursutbud från landets universitet och högskolor presenteras på ett sätt som underlättar jämförelser och val, samlat på en webbplats.

Figur 7 Webbplatsen www.studera.nu.

Information för livsavgörande val

Ett par gånger per år kan man ansöka till universitet och högskolor i Sverige. Tidigare har den som vill studera varit hänvisad till de broschyrer som respektive högskola publicerar, eventuella personliga kontakter eller i bästa fall sakliga jämförelser av olika utbildningsalternativ gjorda av någon ambitiös dagstidning.

Det finns dock ca 20.000 universitets- och högskoleutbildningar i Sverige. Att någon student skulle orka skaffa alla broschyrer och kartlägga utbudet på egen hand tycks orimligt. Inte heller torde någon tidning ha resurser att presentera samtliga kursvägar i en rättvis och enkelt överskådlig jämförelse. Därtill kommer att information om kurserna förändras varför den mest aktuella informationen i praktiken endast finns hos respektive lärosäte.

Ny nationell kurskatalog varje natt

Högskoleverket har tagit fram en sökmotor som varje natt hämtar aktuell kursinformation från universitetens och högskolornas webbplatser. Med hjälp av en enkel struktur och några metadatas (information om information) kan den hämtade informationen automatiskt läggas in i en databas som direkt blir sökbar för allmänheten på webbplatsen studera.nu.

Det enda som krävs av de enskilda universiteterna och högskolorna är att de ser till att ständigt ha aktuell information om kurserna på sina egna webbplatser och att de använder sig av den struktur som Högskoleverket tagit fram.

Figur 8 En schematisk bild av hur kursinformation flödar från Högskoleverkets webbplats till databaser (Studera.nu m.fl.) som ger studenten tillgång till det samlade kursutbudet hos alla högskolor och universitet.

I och med att högskolornas och universitetens webbplatser konkurrerar om studenter är de alla angelägna om att synas i studera.nu, och eftersom lärosäten ändå måste hålla sin egen webbplats uppdaterad innebär inte informationslämnandet till Högskoleverket någon större belastning.

Fördelar med XML

I och med att sökmotorn arbetar med metadatas hos webbsidor (som är kodade med HTML) är det inte nödvändigt för högskolorna att publicera informationen som rena XML-dokument. Hanteringen ställer ändå krav på metadata-standardisering och struktur. Under insamlingen konverteras meta-taggar från HTML-dokument till XML innan informationen läggs i databasen. Lösningen har följande fördelar:

- Information från högskolorna kan maskinellt sammanställas till en unik databas, trots att uppgiftslämnarna internt använder en stor mängd olika datorsystem och många olika programvaror.
- Hög aktualitet och därmed hög kvalitet uppnås tack vare automatiseringen.
- Sökande har nu förutsättning att hitta ALL information!
- Standardiserad metadata gör att kurserna kan jämföras på ett sakligt och överskådligt sätt.
- Standardiserad metadata bidrar också till att informationen kan komma till nytta även i andra sammanhang än det huvudsakliga. Till exempel i den metadatakunniga sökmotorn Svesök (www.svesok.kb.se) som tagits fram av Kungliga Biblioteket.

Kvalitetsgranskning, länkar och andra intressenter

Högskolor och universitet har kunnat samarbeta och komma överens om terminologier och strukturer och gemensamt bidragit till en mycket nyttig central nationell medborgarservice. Men det finns också en del lärdomar som kommit ut ur arbetet.

Det visade sig att en ganska stor andel av den inkommande kursinformationen var felaktigt uppmärkt. Detta fick naturligtvis till följd att de felmärkta kurserna saknades i studera.nu. Man kan tro att universitet och högskolor så starkt önskar att deras information ska finnas med att de kontrollerar om kurserna verkligen dykt upp, men när en högskola eller ett universitet kan ha många hundra kurser, och bara ett fåtal av dem fallit bort så är det inte säkert att handläggarna märker detta. Därför krävs att systemet ger återkoppling till universiteten och högskolorna, så att de kan se hur

många och vilka kurser som kommit in i systemet, och orsak till att de eventuellt inte kommit in.

Universiteten och högskolorna vill naturligtvis inte bara synas i den gemensamma katalogen, utan även ha möjlighet att genom attraktiva presentationer på sina egna webbplatser framhäva sina egna speciella styrkor. Därför fungerar studera.nu så att sökningar och tabell-listningar sker i den gemensamma databasen, men när den sökande vill få en närmare presentation av en specifik kurs hänvisas denne till kursbeskrivningen på den högskolas webbplats som håller kursen.

En annan iakttagelse är att det varit möjligt att efterhand förfinas den metadatamodell man använt. Man ska alltså inte vara rädd för att starta i liten skala, men hela tiden sträva efter att de system och modeller man bygger upp ska gå att utöka och förändra.

Det visade sig tidigt att fler än Högskoleverket är intresserade av den strukturerade information som genom studera.nu-projektet gjorts tillgänglig på Internet. Den nystartade Myndigheten för Sveriges Nätuniversitet (www.netuniversity.se) samlar in samma information, men gör ett mycket snävare urval bland kurserna och presenterar dem på ett helt annat sätt. Likaså finns planer på att exponera delar av kursinformationen i Skolverkets nav för utbildningsinformation, hos AMS m.fl.

EVENEMANGSINFORMATION

Lika villkor för turistföretagen

Sveriges Rese- och Turistråd med flera aktörer i turistnäringen har med hjälp av ett gemensamt schema för beskrivning av evenemang skapat ett automatiserat flöde av information om aktiviteter som konserter, teaterföreställningar och idrottsevenemang. Den samlade mängden information i turistportaler såsom www.gotoscandinavia.com har därmed kunnat ökas markant, utan manuellt arbete.

Ingen har kapacitet att hantera all turistinformation

Sverige har ett närmast oräkneligt antal evenemangskalendrar i tidningar, på webbplatser, på anslagstavlor osv. Tyvärr finns det väldigt få kalendrar som återspeglar hela det ofta mycket rika utbudet av evenemang som faktiskt finns. Orsaken är att den som arrangerar ett evenemang, låt säga en konsert, aldrig har möjlighet att vända sig till alla relevanta kalendrier med sin information.

Turistnäringen är naturligtvis betjänt av en rik och allmänt tillgänglig information om det breda utbudet. Förutsättningen för att en turist ska välja en specifik region (eller ett land) framför en annan är ju inte bara att regionen erbjuder någonting som tilltalar turisten, utan också att turisten får veta detta. Och en potentiell turist har ingen möjlighet att läsa alla de många kalendarierna.

Traditionellt har problemet lösts genom att redaktörer och turistbyråer gör urval av turistinformation och presenterar det man bedömer som intressant för de olika marknaderna. Det finns dock flera brister med denna metod. En är själva kostnaden för detta

redaktionella arbete. En annan är regionens uteblivna intäkter som blir resultatet av att redaktören tvingas göra ett urval som bara motsvarar intresset hos (förhoppningsvis) majoriteten av de potentiella besökarna. Minoritetsintressen blir av nödvändighet försummade, och motsvarande intäkter alltså förlorade.

Figur 9 Informationsspridning idag.

Synkronisering av kalenderdatabaser

Sveriges Rese- och Turistråd (www.swetourism.se/) har sedan 1997 varit engagerade i projektet SKiCal (www.skical.org). SKiCal-gruppen, som tidigt involverade representanter för Riksidrottsförbundet, handikapprörelsen, Riksteatern, Kulturrådet m.fl. tog 1999 fram ett schema för maskinläsbar händelserelaterad information.

Schemat beskriver i detalj hur information såsom starttid, pris, adress, arrangör mm hanteras. Organisationer som har databaser med evenemangsinformation kan använda en programmodul som exponerar innehållet i databasen på SKiCal-format, och/eller en annan programmodul som klarar av att hämta in motsvarande information från andra organisationers databaser. På så sätt kan alla kalendrier som väljer att utväxla information med varandra få förutsättning att bli långt mer innehållsrika utan behov av ytterligare manuellt arbete.

Figur 10 Informationsspridning med XML.

Tidningen Aftonbladet var först ut 1999 med att prova på SkiCal (www.aftonbladet.se/puls/9906/16/ski.html), därefter följde

Stockholms Stad (www.stockholmtown.com), tidningen Metro (www.metro.se), Kulturnät Sveriges "Dettahänder.nu" (www.dettahander.nu), Eskilstunakuriren (www.ekuriren.se) m.fl. År 2001 öppnades, med stöd av Turistrådet, webbplatsen Sverigeresan (www.sverigeresan.se eller mot utlandet www.gotosvandinavia.com), som hämtar sin evenemangsinformation i SKiCal-format¹ från turistbyråerna ute i landet.

Därmed är alla rådata om evenemangen exakt samma på lokal nivå (där man finns nära informationens ursprung) som på den centrala nationella webbplatsen. Men det innebär inte att informationen presenteras på samma sätt, eller att samma urval av detaljinformation visas på de olika webbplatserna. Sådant anpassas av respektive publicist för att motsvara just dennes aktuella målgrupp och syfte.

Figur 11 En schematisk bild av hur evenemangsinformation flödar från arrangörer till tidningar, turistbyråer och andra portaler (som även utväxlar information med varandra), och slutligen hamnar hos den intresserade allmänheten.

Fördelar med standardiserat format

¹ SKiCal-schemat bygger inte på XML utan på en annan Internet-standard som heter iCalendar, detta för att uppnå kompatibilitet med ett äldre format som många kalenderprogramvaror kan hantera. Evenemangsinformation på SKiCal-format kan, och har även hanterats på XML-format.

- Möjligt för evenemangsarrangörer att synas i fler kalendrar och kataloger
- Möjligt för kalendrar och kataloger att bli näst intill fullständiga
- Mindre risk glömma viktiga data
- Mindre risk för orimlig information
- Mindre risk för felkopiering

Medvetenhet viktigt för resultatet

Erfarenheten är dock att flertalet evenemangsdata-baser konstruerats uteslutande för att ligga till grund för en specifik webbplats eller trycksak - inte för att förutsättningslöst lagra information om evenemangen. Detta gör att vissa av databaserna är dåligt förberedda för att presentera sin information på nya format, eller för att automatiskt utbyta information med andra databaser.

Exempelvis är det ganska vanligt att klockslag och datum representeras som fritext i databaserna. (Sannolikt för att ge redaktören frihet att till exempel lägga in en liten kommentar i anslutning till klockslaget.)

Detta ger dock upphov till problem: Om man vill presentera databasens innehåll i en engelskspråkig vy så bör datum formges på det sätt som amerikanska turister är vana vid (t.ex. 08/21/02, och inte 2002-08-21 som vi i skriver). Databaser som modellerats väl kan utan problem anpassa presentationen till ett nytt format, men när fritextfält används och ovanstående datum kanske skrivits in så här: "21 aug (skolstartdagen)" så blir det hopplöst för datorer att göra en automatisk anpassning.

XML:s intåg på arenan ger alltså nya möjligheter, men de kommer först till sin rätt när informationshantering sker med en

medvetenhet att informationen kan komma att användas för många olika syften. Även personer som jobbar med databasteknik har alltså hittills haft en tendens att fokusera på den visuella presentationen i ett givet medium. En viktig uppgift för teknikansvariga är alltså att långsiktigt skapa en medvetenhet om att information i databaser kan komma att dyka upp inte bara i det sammanhang som databasen ursprungligen skapats för.

Precis som i alla andra branscher behöver turistnäringen komma överens om terminologier för att med automatik kunna utbyta information. Turistrådet har sedan länge arbetat med dessa frågor och branschen har tagit del av denna insikt och de förslag till termlistor som producerats. Därmed inte sagt att alla redan använder samma termer, men det finns en vilja att "göra rätt" och en förståelse för behovet, trots att det ju alltid smärtar att inte själv kunna definiera begrepp och kategorisera produkter efter eget tycke.

Vid synkronisering av evenemangsinformation har ett antal olika policy-beslut fattats på olika håll. Orsakerna till varje val kan vara tekniska, politiska, ekonomiska eller någonting annat. Sannolikt återkommer fenomenet i andra branscher. Följande är ett urval av behov och användningsmönster:

- En stad vill ha automatiskt insamlad information i en databas, tillgängligt för vem som helst att kopiera från, och samtidigt ha en databas med redaktionellt förädlad material som man kan ta betalt för.
- En annan stad tar betalt för all information som exporteras.
- Ytterligare en annan stad exporterar gärna all sin information gratis, bara någon annan betalar kostnaden för att bygga exportmodulen.

- En portal publicerar vilken evenemangsinformation som helst på sin webbplats, men gör en kvalitetskontroll av det urval som ska presenteras i en tryckt tidning.
- En portal samlar in så mycket information som möjligt med automatik, men granskar allt inkommet material manuellt innan det syns på webbplats och i trycksaker.
- En portal samlar bara in information från betrodda leverantörer, och har ingen manuell granskning av det som hämtats in.
- En portal samlar in från andra, men exporterar ingenting.
- En portal exporterar allt, men samlar inte in från andra.

Svårt engagera små aktörer

När projektet SKiCal startade fanns en förväntan att de som främst skulle komma att ivra för denna teknik var de små arrangörerna. Teoretiskt ger ju tekniken lika god exponering för deras information som de med stora marknadsföringsbudgetar, trots en minimal egen arbetsinsats. Det har dock visat sig att det främst är på nationell nivå som kraften finns att driva fram en fungerande infrastruktur, och innan en sådan är på plats är det svårt att överhuvudtaget väcka intresse hos de små aktörerna. Men när infrastrukturen börjar bli etablerad torde fler och fler av de små börja upptäcka värdet i detta sätt att sprida sin information.

LÄROMEDEL OCH SKOLINFORMATION

En mjuk infrastruktur för läromedel och skolinformation

Skolverket (www.skolverket.se) driver ett utvecklingsprojekt som syftar till att bygga upp en mjuk infrastruktur för skolan. Projektet berör många typer av information (t.ex. kursinformation, läromedelsmoduler, elevinformation och lagring av arbetsdokument) och den gemensamma nämnaren är att Skolverket strävar efter att för denna omfattande informationshantering använda öppna standarder och metadataterminologier.



Skolvärlden starkt påverkad av informationstekniken

Sedan 1994 har landets skolor gått från i princip ingen extern datakommunikation till att i stort sett alla skolor har en Internetanslutning. För en mycket stor del av eleverna är begrepp såsom Lunarstorm (www.lunarstorm.se), ICQ (www.icq.com) och Hotmail (www.hotmail.com) minst lika vardagliga och viktiga fenomen som fotbollsklubben (www.degerforsif.se) och mobiltelefonen. Detta påverkar naturligtvis skolan.

Skolverket arbetar för att stödja skolorna i den pågående förändringsprocessen. Precis som i andra branscher finns ett behov av att ta fram gemensamma terminologier och scheman för att möjliggöra effektiv information.

Bred satsning på grundläggande infrastruktur och bastjänster

Ett av Skolverkets regeringsuppdrag är särskilt relevant i detta sammanhang. "Mjuk infrastruktur på utbildningsområdet" kallas projektet, består av tre nivåer:

1. Ett långsiktigt grundläggande arbete med att utforma olika standarder och rekommendationer för att såväl innehåll som tekniska lösningar ska kunna integreras med varandra
2. På nationell nivå skapa tekniska förutsättningar för olika tjänster på nätet.
3. Ta fram ett antal prototyper av nyttotjänster för skolan.

Inom ramen för den första nivån har Skolverket framförallt bidragit till att ta fram strukturer, d.v.s. terminologier och scheman för information:

- kompetensutveckling av pedagogisk personal
- läromedel och lärkomponenter
- kurs- och annan utbildningsinformation
- pedagogiska texter

Detta arbete har skett inom ramen för en standardiseringsgrupp inom SIS (www.sis.se), "TK450 Lärande", och i en internationell ISO-standardiseringsgrupp benämnd JTC1/SC36 (<http://jtc1sc36.org>). Skolverket har också tagit fram publikationer som vägleder och undervisar om betydelsen av metadata och infostrukturer.

På den andra nivån, grundläggande informationstjänster, som i stor utsträckning bygger på resultat från den första nivån, har Skolverket:

- Tagit initiativ till en distribuerad nationell e-postkatalog.
- Genomfört en förstudie om ett "nav" för information om hela utbudet av fortbildningar för lärare. Studien resulterade i en

rekommendation att bygga navet som en utökning av Högskoleverkets studera.nu, och gav även en konkret lista med förslag på utökningar av metadatamodell och vokabulär för studera.nu.

- Byggt en "verktygslåda" bestående av ett antal programkomponenter såsom kan användas för samverkan inom distribuerade arbetsgrupper. (Diskussionsforum, webbpubliceringsverktyg, kalendarium mm.)
- I samarbete med Centrum för flexibelt lärande (www.cfl.se) tagit fram ett arkiv för hantering av rikligt metadatauppmärkta lärkomponenter, som kallas SCAM (Standardized Content Archive Management, <http://sourceforge.net/projects/scam>). Skolverket har även provat på att integrera så kallad peer-to-peer-teknik, d.v.s. samma teknik som användes i det omtalade systemet Kazaa (www.kazaa.com), för att bygga ett "nav utan centraldator". Detta delprojekt kallas Edutella (<http://edutella.jxta.org>).
- I samarbete med Utbildningsradion (www.ur.se) har tagits initiativ till att UR:s material skall göras tillgängligt i ett Mediearkiv. Arbetet bygger tekniskt på W3C såsom XML och RDF som skolorna har tillgång till.

Den tredje nivån är ännu på planeringsstadiet, men här tänker man sig att till exempel "personliga kompetensportföljer" och andra högnivå-tjänster ska tas fram. Ett samarbetsprojekt med läromedelsförlagen för att upprätta en mäklartjänst för lärkomponenter ligger till exempel i startgroparna.

Projektet är således både djupgående och mångfacetterat och det är uppenbart att Skolverket anstränger sig för att skolväsendet ska kunna dra nytta av den tekniska utvecklingen.

Figur 12 En schematisk bild av hur utbildningsinformation flödar från läromedelsproducenter (lär-objekt, läromedelsinfo), kursarrangörer (kursinformation) m.fl. via ett planerat nav ("Lärtorget") och vidare in i olika informationssystem hos skolorna.

XML medger öppenhet, interoperabilitet och långsiktighet

Syftet med att bygga upp en mjuk infrastruktur för skolan är att utveckla skolans IT-användning så att lärare och elever kan fungera både som informationskonsumenter och som informationsproducenter i en rik informationsmiljö.

Återkommande nyckelord i all dokumentation från projektet är "tillgänglighet", "öppna standarder", "metadata" och "återanvändbarhet". Den som känner till XML förstår direkt kopplingen.

ORDFÖRKLARINGAR

DTD

DokumentTypsDefinition - en beskrivning av strukturen i de XML- eller SGML-dokument man hanterar. Som alternativ till DTD kan man använda t.ex. W3C XML Schema

Dublin Core (DC)

En uppsättning av centrala 15 metadataelement, som uppnått internationell standardstatus. Exempel på DC-element är författare, datum, ämne och rättighetsinnehavare.

Metadata

Information om information. Vanliga metadata om ett dokument är t.ex. skapat-datum, författare, ämnesord. Se även Dublin core.

RDF

Resource Description Framework - ett generellt XML-format för kunskapsrepresentation.

Schema

Ett schema är ett dokument som definierar strukturen hos en viss typ av information. Till exempel är hr-xml:s JobPositionPosting ett schema som definierar hur en platsannons ska vara strukturerad. Scheman kan skrivas på några olika format. Vanliga schema-scheman inom xml-världen är W3C XML Schema och DTD.

SGML

Standard Generalized Markup Language - en föregångare till XML

XML
eXtensible Markup Language

XSL
Språk för att beskriva transformationer av XML-data, kan användas för att utifrån ett givet XML-content göra en presentation på ett annat XML-format eller på t.ex. HTML-format.

W3C XML Schema
Dokument som definierar en godkänd struktur hos en given typ av xml-dokument. Se även DTD.