



2004:23

Vägledning: XML-teknik och metadata

Nummer och utgåva: SV1:4
Fastställd 2004-09-07



Förord

Inom informationshanteringen har XML och metadata blivit signalord som ligger i tiden och symboliserar modern systemutveckling.

Användning av XML och metadata är viktiga tekniska förutsättningar som behöver sättas in i sitt sammanhang för att effektivisera informationsförsörjningen inom förvaltningen. Statskontoret har därför tagit fram denna vägledning. Här föreligger nu den fjärde utgåvan, den första publicerades 2001.

Den ingår i Statskontorets serie Vägledningar för IT-området. Vägledningarna publiceras elektroniskt på www.statskontoret.se, bland annat för att enkelt kunna utnyttja de talrika länkarna till fördjupad information.

Syftet är att bidra till en effektiv och samordnad utveckling och användning av IT, i första hand inom den offentliga sektorn, genom att erbjuda aktuella råd och rekommendationer inom utvalda områden. Målgruppen är i första hand IT-strateger, IT-ansvariga, projektledare, konsulter och IT-leverantörer.

Denna fjärde version av vägledningen har en omarbetad innehållsstruktur och den har utökats med mer information om aktuella användningsområden och pågående initiativ. Sedan version tre kom ut (december 2003) har också flera viktiga XML-specifikationer blivit fastställda som officiella W3C Recommendations, uppdaterats eller kommit i nya versioner.

Generaldirektör Knut Rexed har beslutat i detta ärende. Avdelningsdirektör Karl Wessbrandt, föredragande, var närvarande vid den slutliga handläggningen.

Enligt Statskontorets beslut

Karl Wessbrandt

Innehållsförteckning

1. VÄGLEDNING XML-TEKNIK OCH METADATA	7
1.1. SYFTE OCH MÅLSÄTTNING	7
1.2. MÅLGRUPPER	8
1.3. AVGRÄNSNINGAR	8
1.4. INTRODUKTION	9
2. ANVÄNDNINGSSOMRÅDEN.....	11
2.1. INFORMATIONSENTBYTTE.....	11
2.2. ENKLA TRANSFORMATIONER GER STÖRRE INFORMATIONSVÄRDE.....	12
2.3. INFORMATIONSTRUKTURERING	13
2.3.1. Metadata	13
2.3.2. Datamodeller	14
2.3.3. Vokabulärer	15
2.3.4. Digital publicering.....	15
2.3.5. Arkivering	16
2.4. SÖKNING	17
2.5. KONTORSDOKUMENT.....	17
2.6. DATABASLÖSNINGAR.....	18
3. STANDARDISERINGSARBETE.....	19
3.1. FAKTORER SOM PÅVERKAR STANDARDERS ACCEPTANS	19
3.2. STANDARDISERINGSORGANISATIONER.....	20
3.2.1. World Wide Web Consortium (W3C).....	20
3.2.2. Dublin Core Metadata Initiative (DCMI).....	21
3.2.3. Web Services Interoperability Organization (WS-I).....	21
3.2.4. Organization for the Advancement of Structured Information Standards (OASIS).....	22
3.2.5. International Organization for Standardization (ISO).....	22
3.3. VIKTIGA XML-STANDARDER	23
3.3.1. Extensible Markup Language (XML) 1.0 (Third Edition).....	23
3.3.2. Extensible Markup Language (XML) 1.1.....	24
3.3.3. Namespaces in XML Version 1.0.....	24
3.3.4. Namespaces in XML Version 1.1	25
3.3.5. XML Base.....	25
3.3.6. XML Linking Language (XLink) Version 1.0.....	25
3.3.7. XML Path Language (XPath) Version 1.0.....	26
3.3.8. XML Pointer Language (XPointer) Version 1.0	26
3.3.9. XML Schema Version 1.0.....	27
3.3.10. Extensible Stylesheet Language (XSL) Version 1.0	27
3.3.11. XSL Transformations (XSLT) Version 1.0	28
3.3.12. XHTML Version 1.0 (2 nd edition).....	29
3.4. VIKTIGA WEB SERVICE STANDARDER	30
3.4.1. Introduktion.....	30
3.4.2. SOAP Version 1.2	31
3.4.3. Web Services Description Language (WSDL) Version 2.0.....	32
3.4.4. Universal Description, Discovery, and Integration (UDDI)	32
3.4.5. WS-I Basic profile Version 1.0.....	32

3.5.	VIKTIGA STANDARDER FÖR METADATA.....	33
3.5.1.	<i>Dublin Core Metadata Element Set (DCMES) Version 1.1</i>	34
4.	PÅGÅENDE INITIATIV.....	35
4.1.	STANDARDMEDDELANDEN	35
4.2.	EU:S IDA OCH NATIONELLA EUROPEISKA INITIATIV.....	36
4.2.1.	<i>Danmark</i>	36
4.2.2.	<i>Storbritannien</i>	37
4.2.3.	<i>Tyskland</i>	37
4.3.	BRANSCHINITIATIV	38
4.3.1.	<i>ebXML</i>	38
4.3.2.	<i>XBRL</i>	39
4.3.3.	<i>GML</i>	39
4.3.4.	<i>HR-XML</i>	40
4.3.5.	<i>SCORM</i>	40
5.	ORDLISTA.....	43
5.1.	AKRONYMER	43
5.2.	ENGELSKA TERMER	44
5.3.	REFERENSER.....	45
5.4.	KONTAKTPERSON.....	46

1. Vägledning XML-teknik och metadata

1.1. Syfte och målsättning

XML och metadata ligger i tiden och dessa begrepp har blivit något av symboler för modern informationshantering. Det pågår ett stort antal utvecklingsprojekt som bygger på användning av XML och metadata, såväl inom offentlig som privat sektor.

Men tekniken i sig själv är inte nog och ger ingen garanti för framgång. Det handlar om att integrera gammalt och nytt på ett optimalt sätt. Mycket av utvecklingen styrs av IT-branschens utveckling av standardiserade programvaror och tillgången till öppna programgränssnitt. I takt med utbyggnaden av 24-timmarsmyndigheten ökar behovet av att också använda enhetliga metoder och standarder för informationsutbyte. En effektiv och samverkande informationshantering är en grundförutsättning för att öka tillgänglighet och öppenhet hos myndigheterna.

Denna vägledning behandlar XML-teknik och metadata, och hur dessa teknikområden i samverkan kan bidra till att bygga en effektiv informationsstruktur.

Strukturerad informationshantering är viktig som möjliggörare för en effektiv informationsförsörjning och XML-teknik och metadata har många användningsområden inom modern systemutveckling. Samtidigt måste verksamhetsprocesserna utvecklas och tillämpningen av den nya tekniken behöver samordnas och samverka med befintliga lösningar för att nå önskat resultat.

Målsättningen med denna vägledning är att stärka beställarkompetensen inom offentlig sektor vid planeringen av informationsförsörjningslösningar. Den skall också förbättra förutsättningarna för beställarna att kommunicera med IT-leverantörer och XML-specialister.

Denna utgåva av vägledningen är den fjärde versionen, redigerad sommaren 2004. Den försöker ge en aktuell bild av utvecklingsläget inom utvalda områden, men i takt med den snabba teknik- och marknadsutvecklingen på området behöver innehållet egentligen ständigt uppdateras och justeras. Statskontoret rekommenderar därför alla berörda att själva kontinuerligt följa vad som sker, t.ex. på de webbplatser som vägledningen hänvisar till.

1.2. Målgrupper

Vägledningens målgrupper och definitioner av dessa är:

IT- och informationsstrateger - Personer som har ett långsiktigt och övergripande perspektiv på sin myndighets informationshanteringslösningar.

Projektledare - Mer eller mindre teknikorienterade personer som ansvarar för projekt som skall lösa specifika informationshanteringsuppgifter.

Systemarkitekter - Personer med fokus på övergripande systemdesign och metoder.

Systemkonstruktörer - Personer som arbetar med att lösa specifika systemutvecklingsuppgifter med hjälp av olika metoder och utvecklingsverktyg.

Läsaren förutsätts ha grundläggande kunskaper om informationshantering, XML och metadata.

1.3. Avgränsningar

Denna vägledning tar upp ett urval av generella, brett tillämpbara standarder och initiativ som bedöms vara relevanta och gör inga anspråk på fullständighet.

Vidare bör noteras att för effektiv informationsstrukturering och metadatamärkning behövs s.k. kontrollerade vokabulärer. Att skapa sådana kräver ofta långsiktiga terminologiutvecklingsprojekt. Att bevaka vilka vokabulärer som finns och kan användas inom specifika sektorer ligger dock utanför denna vägledning.

Anmärkning

I denna vägledning används termen "standard" för det som brukar kallas "de facto standard", "öppen industristandard" eller "rekommendation". De "standarder" som nämns är alltså ofta inte formellt fastställda av ett officiellt standardiseringsorgan, utan de är rekommendationer som produceras och förvaltas av industrigrupperingar med bred förankring i såväl användar- som leverantörsled, t.ex. W3C.

I praktiken används de på motsvarande sätt som formella standarder.

1.4. Introduktion

Strukturerad informationshantering ligger i tiden. Det ständigt växande informationsflödet gör det uppenbart att man behöver nya hjälpmedel för att kunna navigera rätt och hitta fram till den information som man behöver i en viss situation. En viktig målsättning med allt informationsstruktureringsarbete är att framtidssäkra den investering som arbetet innebär. Med andra ord, att sträva efter att hitta lösningar vilka minimerar den mängd underhållsarbete som kan krävas i framtiden, då omvärlden, verksamheten eller processerna ändras. Framtida kostnader beror på hur analys, strukturering och modellering genomförs, på vilka tillämpningar som väljs ut och på hur lösningarna integreras med befintliga system.

XML är en grupp standarder som har starkt stöd av företag och organisationer, i Sverige och internationellt. XML är inte en enda standard eller en färdig lösning utan en verktygslåda som innehåller en uppsättning specifikationer och verktyg, som kontinuerligt utvecklas. Denna vägledning presenterar olika XML-standarder samt XML-baserade initiativ som kan användas som utgångspunkter.

I övrigt rekommenderas den princip (fritt översatt) som återfinns på XML-grundarnas webbplats (World Wide Web Consortium, www.w3.org):

"XML är inte alltid den bästa lösningen, men den är alltid värd att överväga".

För att själv utveckla XML-baserade tillämpningar krävs planering och förberedelser. Informationsstrukturer och terminologier behöver fastställas och ändamålsenliga programvaror anskaffas eller tas fram.

Att "satsa på XML" behöver därför motiveras och förtydligas.

Det är värt att peka på möjligheten att i praktiken använda sig av en delmängd av en standard. Generella och uttrycksfulla standarder är i vissa fall omfattande, just för att de försöker vara kraftfulla på ett generellt sätt. Konkreta behov i en organisation kan ofta vara mer begränsade. Det finns då möjligheten att man anammar en standard men bara använder en del av det som standarden tillåter. Detta kan vara ett sätt att få en större enhetlighet i standardanvändningen inom en organisation, men kräver naturligtvis att man klart identifierar den delmängd av standarden som man tillåter användas.

XML som idé har etablerat sig snabbt och stöds av åtskilliga offentliga projekt och organisationer, både i Sverige och internationellt, liksom av de flesta tongivande aktörer inom IT-industrin. Även inom det privata

näringslivet utanför IT-branschen pågår många intressanta XML-initiativ, till exempel i sektorer som handel, finans och resor.

Med de många fördelar som XML-konceptet utlovar så är sannolikheten stor för att XML-baserade informationsstrukturer kommer att bli dominerande. XML skall inte ses som en ersättare för HTML på webben utan snarare som fundamentet för en helt ny teknikgeneration som successivt kan komma att ersätta exempelvis HTML, SQL, Edifact och många andra dokument- och filformat, kommunikationsprotokoll, programmeringsgränssnitt osv. På vissa områden går det fort men tidsperspektivet för att nå fullt genomslag är långsiktigt. Från början var XML-utvecklingen inriktad på digital publicering, men efter hand har även XML-baserat informationsutbyte mellan heterogena system blivit ett viktigt tillämpningsområde. Det handlar då inte om att byta ut de gamla proprietära lösningarna utan om att förlänga livslängden genom att anpassa och integrera dem med nya, öppna och standardiserade system- och kommunikationsgränssnitt.

Frågan om att satsa på XML blir därför till slut en fråga som den globala IT-branschen avgör genom sina val av tekniska infrastrukturplattformar. Då nästan alla ledande leverantörer för närvarande visar stort intresse för XML, Web Services m.m., så är det en god strategi att även i användarledet ligga långt framme. Genom att bygga upp egna kunskaper och skaffa egna erfarenheter så skapas förutsättningar att kunna bedöma och påverka marknadsutvecklingen för att optimalt utnyttja de tekniska lösningar som marknader erbjuder. Man bör alltid noga överväga vilka alternativ som finns i form av standard- eller open source-lösningar innan man tar beslut om att inleda egen systemutveckling. Vid sidan av kostnads- och underhållsaspekterna så riskerar man att inte fullt ut kunna dra nytta av XML:s samverkanspotential om man skapar egna, ”proprietära” XML-lösningar.

Det som tar längst tid och kräver mest resurser är ändå inte den tekniska systemutvecklingen. Det är den verksamhetsinriktade informationsanalysen och det därpå följande arbetet med att strukturera, standardisera och anpassa informationshanteringen inom och mellan berörda organisationer.

2. Användningsområden

I detta kapitel listas några exempel på praktiska tillämpningsområden där det idag kan vara lämpligt att överväga en XML-baserad lösning. Kapitlet berör också kort några områden där de större systemleverantörerna valt att inkludera XML-stöd i sina lösningar.

2.1. Informationsutbyte

Begreppet information syftar här till någon verksamhetsknuten datamängd. Dokument, blankett och registerpost är några exempel på information som kan behöva utbytas, t.ex. vid ärendehantering. Ett informationsutbyte kan vara enkelt eller ingå i en mer komplicerad ärendehantering. För att informationsutbytet skall fungera krävs det att avsändare och mottagare har en gemensam definition och klassificering av vad som ska utbytas. Historiskt sett har samverkan mellan olika verksamheter och system oftast förutsatt specifika integrationsinsatser, beroende på utgångsläget med heterogena system och slutna datadefinitioner.

Genom att använda XML för att specificera informationen kan man skapa en gemensam, oberoende beskrivning av data. Dessutom är det möjligt att nyttja den verktygslåda som ingår i XML för att t.ex. automatiskt verifiera att informationen är korrekt och fullständig. Idag är XML en etablerad standard och standardsystem erbjuder ofta XML som ett möjligt gränssnitt för in- och utdata. Detta innebär att man har öppnat en tidigare slutna datadefinition med en XML-baserad anpassning. XML är en bra bas för informationsutbyte mellan verksamheter med öppna gränssnitt för sammankoppling med andra stödsystem samt det stöd som verktygslådan innehåller för vidare informationsbearbetning.

Den stora vinsten med att använda XML vid informationsutbyte är dock om många använder samma datadefinition. Då har man skapat ett domän-gemensamt språk för informationsutbyte som kan användas av samtliga aktörer inom denna domän. Exempel på initiativ för gemensamma definitioner är bland annat HR-XML, ebXML, XBRL och GML.

Statskontorets initiativ till utveckling av XML-baserade standardmeddelanden för frekvent informationsutbyte mellan myndigheter är ett annat aktuellt exempel.

XML är en bra teknik för att strukturera information vid informationsutbyte. Hur och var själva informationsutbytet skall realiseras beror på andra faktorer. Det är viktigt att hålla fokus på verksamhetslogiken och infor-

mationsstrukturen; teknikvalet är en möjliggörare men löser inte hela informationsintegrationen.

2.2. Enkla transformationer ger större informationsvärde

Transformationer kan vara mycket värdefulla. Det vet elbolagen som måste transformera elektriciteten för att kunna transportera den över landet, och sedan transformera den igen för att kunna leverera den till oss kunder. Också i den digitala världen finns stora behov av att transformera. Det kan handla om att sändare och mottagare har olika programvaror - informationen behöver transformeras så att programmen kan kommunicera.

Det kan också handla om att olika mottagare behöver olika perspektiv på informationen. En chef vinner tid på att få sammanfattningar på alla ärenden, åtkomliga via en överskådlig meny, medan en specialist kanske vinner tid på att få ett snävt urval av ärenden presenterade i tidsordning, med alla detaljer synliga samtidigt på skärmen.

Det finns många exempel på traditionellt skräddarsydda program som transformerar information så att precis rätt aspekt blir tillgänglig för rätt person. Men sådana system är ofta dyra och ganska hårt knutna till det företag som utvecklat programmet.

I XML-världen erbjuds en mer generell lösning, nämligen ett språk - XSLT - för att beskriva transformationer av XML-information.

En stor fördel med XSLT är att man inte behöver köpa nya programvaror för varje nytt sätt man vill kunna presentera sin information. Man behöver bara en XSLT-processor, och dessutom, eftersom XSLT är en standard, går det bra med vilken XSLT-processor som helst (det finns bl.a. flera gratisvarianter). En XSLT-processor använder stilmallar skrivna i XSLT för att bearbeta/transformera XML-information.

Antag att en turistbyrå har en mängd information om hotellrum. För olika målgrupper vill man mata ut informationen på olika sätt:

- * På webben till webbsidor
- * På webben som PDF-fil lämplig för utskrift
- * Till litet html-avsnitt med aktuella erbjudanden för att kunna fungera som annonsering på andra webbplatser
- * Till ett visst XML-format som hotellbokningscentralens program kan läsa för att informationen ska kunna exporteras dit

- * Till ett annat format för att ett litet urval aktuell information ska komma med i några olika resebroschyrer i USA
- * Till tab-separerad textfil för att kunna rapportera aktuellt rumsbestånd till den nationella turistmyndigheten osv.

Med en traditionell databaslösning kan man åstadkomma allt detta. Men när man vill kunna mata ut till ett ytterligare XML-format för att ett visst urval ska komma med i en annons med standard-layout som skickas till tidningarna varje vecka - då behöver databaslösningen ändras. Om det är en skräddarsydd databaslösning så kan databaskonsulten bygga om den, men om databaslösningen är en produkt som kommer i ny version en gång per år så blir det svårare och dyrare.

Om all information om hotellen fanns tillgänglig på XML-format så skulle man kunna använda XSLT. Man beskriver varje önskat format i en ny stilmall i XSLT och så kan godtycklig XSLT-processor sköta transformeringen.

Om man sedan vill lägga till nya presentationer, kanske i blindskrift eller i sammandrag sorterat efter pris, så är det bara att skriva ytterligare XSLT-filer.

Huvudpoängen med XSLT är alltså att det blir ekonomiskt överkomligt att kunna skapa många olika presentationsformat, inte bara avseende grafisk form utan även gällande urval, sortering och dataformat.

2.3. Informationsstrukturering

2.3.1. Metadata

Metadata, dvs. data om data, har länge använts i arkiv, bibliotek och i datasammanhang. Med växande informationsmängder ökar intresset för metadata, framför allt för att de underlättar informationssökning på webben, men metadata har också betydelse för att hålla reda på statusinformation, vilket t.ex. behövs vid ärendehantering.

Metadata är i sig teknikneutrala och kan hanteras såväl med hjälp av kortlådor eller databaser som i HTML eller XML. Man behöver därför skilja på innehåll och form. Däremot finns en naturlig koppling till XML, eftersom XML-formatet i princip bygger på strukturer av metadataelement.

Den mest spridda metadatamodellen, Dublin Core, är numera en internatio-

nell standard. Dublin Core har 15 grundelement i basversionen, den så kallade Simple Dublin Core, plus möjlighet till specialisering av dessa till det som kallas Qualified Dublin Core. Dublin Core (DC) används när begrepp som titel, författare, datum, ämne, typ, format m.m. är relevanta. DC-element som inte behövs kan utelämnas.

Kungliga bibliotekets enhet för bibliografisk utveckling och samordning (BUS) har sidor om metadata, se <http://www.kb.se/Bus/Metadata/>. Där finns praktiska råd, t.ex. länkar till program som t.ex. skapar HTML-kod för Dublin Core metadata.

Dublin Core täcker dock inte alla behov. Vidareutveckling och standardisering pågår på olika håll och inom vissa sektorer etableras andra, mer branschanpassade metadatamodeller. För statusorienterade metadata, t.ex. inom dokumenthantering, ärendehantering och arbetsflöden, ställs åter andra krav. Här finns olika initiativ med idéer till lösningar. Inom olika nationella eGovernment-initiativ, t.ex. i Storbritannien, Danmark och Finland, har Dublin Core utvidgats med ytterligare element och inom CEN finns en workshop som arbetar med målet att samordna dessa olika tillägg till en gemensam Dublin Core Application Profile, som skulle täcka de flesta identifierade behoven för eGovernment inom EU. Dessa pågående arbeten behandlas dock ej närmare här.

2.3.2. Datamodeller

En kategori av metadata är sådana data som dokumenterar datamodeller (informationsmodeller, informationsstrukturer). Den här typen av metadata anger krav som konkreta informationsobjekt måste uppfylla och beskriver hur olika informationsobjekt kan länkas samman.

När det gäller frågor om språk för representation av verksamhetsinformation är grundrekommendationen att XML-baserade standarder skall användas. Speciellt för informationsutbytesgränssnitt mellan system och organisationer bör XML-lösningar övervägas. Den centrala standarden för detta är W3C:s XML Schema. Då XML Schema-standarderna är mycket omfattande finns det även initiativ som siktar på att ta fram mindre omfattande schemadefinitioner för XML. Som exempel kan OASIS RelaxNG nämnas, vilket anses vara ett mer överskådligt format för schemadefinitioner.

Att införa metadata och datamodeller av god kvalitet är resurs- och kompetenskrävande. Först behövs noggranna analyser av det aktuella materialet. Sedan krävs att terminologier fastställs, sprids och används, att rollfördelning, arbetsrutiner och programvaror för inmatning av metadata anpassas och att regelverket därefter följs av alla berörda med god disciplin. Det krävs utbildning, träning och motivation, både för att lägga in och för att framgångsrikt kunna söka med hjälp av metadata.

På sikt kan informationskvaliteten kraftigt förbättras genom att införa mer standardiserade terminologier. Att fastställa teknik, format och strukturer är viktigt, men svårast och mest avgörande för att bygga upp en framgångsrik informationshantering är att fylla strukturerna med rätt innehåll!

2.3.3. Vokabulärer

För att uppnå hög kvalitet och effektivitet vid användning av metadata så bör man helst ha s.k. kontrollerade vokabulärer eller fastställda terminologier, kodlistor m.m. till sitt förfogande.

Att man exempelvis har kommit överens om att använda Dublin Core som metadataformat innebär inte i sig att man vet hur man ska tolka innehållet i ett specifikt element. För att effektivisera dessa typer av överenskommelser har man därför inom Dublin Core förutom grundelementen även fastslagit ett antal vokabulärer (encoding schemes) för innehållet i de olika elementen. Vissa av dessa är formella standarder som exempelvis ISO:s språkkoder och datumformat. Andra vokabulärer är publikt tillgängliga kontrollerade vokabulärer som exempelvis TGN (The J. Paul Getty Trust:s Thesaurus of Geographic Names) för geografiska platser eller IANA:s (The Internet Assigned Numbers Authority) lista över Internet Media Types (MIME).

Arbetet med att skapa och fastställa vokabulärer bedrivs ofta inom ramen för långsiktiga utvecklingsprojekt, kräver tillgång till specialistkompetens och bör därför gärna hanteras sektorsvis. Terminologifrågorna är viktiga för att åstadkomma en väl strukturerad informationshantering, men dessa behandlas inte närmare i denna vägledning. Statskontoret har i andra sammanhang utrett och föreslagit hur myndigheternas informationsutbyte kan underlättas med hjälp av XML-baserade standardmeddelanden och hur ett nationellt samverkansregister skulle kunna upprättas, där t.ex. återanvändbara strukturer, scheman, modeller, metoder och terminologier då kunde publiceras, kommenteras och vidareutvecklas. Något beslut om realisering har dock ännu inte tagits.

2.3.4. Digital publicering

Digital publicering var det ursprungliga huvudsyftet med XML och detta är ett centralt användningsområde. Här ligger fokus på den flexibilitet som XML åstadkommer, bland annat genom att XML gör det möjligt att separera innehåll, struktur och format. Inom detta område drar man också nytta av XML för att definiera gemensamma metadata och datamodeller.

Ett internationellt initiativ inom området är Open Archives Initiative (<http://www.openarchives.org/>) som bland annat används av det Digitala Vetenskapliga Arkivet (<http://www.diva-portal.se/>) som publicerar avhandlingar och andra publikationer i fulltext från ett antal nordiska uni-

versitet och högskolor. Open Archives Initiative har definierat ett standardprotokoll för arkiv som ska möjliggöra insamling av metadata från samtliga arkiv som stödjer protokollet. Man har fastställt Dublin Core som grundläggande metadatamodell men tillhandahåller även stöd för användande av andra, mer specialiserade metadatamodeller, t.ex. MODS som i sin tur bygger på MARC21.

Det gemensamma protokollet gör det möjligt för exempelvis söktjänster att kunna samla in information från alla arkiv som stödjer OAI. Den gemensamma metadatamodellen gör det teoretiskt möjligt för en söktjänst att erbjuda en detaljerad matchning på Dublin Cores olika metadataelement tvärs över data från olika arkiv. Dock erbjuder Dublin Cores grunduppställning av element en stor frihetsgrad så till vida att alla element kan förekomma noll till många gånger och ofta innehålla vilka strängar som helst. Detta gör att även metadata som formellt följer Dublin Cores metadatamodell kan ha väldigt skiftande kvalitet. Det är därför viktigt att göra överenskommelser om vokabulärer och obligatorier som kan höja informationskvalitén. Man kan exempelvis använda kvalificerad Dublin Core eller någon annan branschspecifik datamodell som har en mer specifik datatypning.

Ett svenskt initiativ som kan relateras till området är EMIL (Education Information Markup Language). EMIL tas fram som en gemensam metadatamodell för utbildningsinformation och syftar bland annat till att förenkla sammanställningen av publicerad information om utbildningar från flera olika myndigheter. För mer information om projektet se : <http://mjukis.skolutveckling.se/mjukis/start.jsp?vp=ZA21>

2.3.5. Arkivering

XML blev tidigt intressant för digital långtidslagring av dokument eftersom användande av XML som lagringsformat minskar risken för att en fil med information skall bli oläsbar när dess format inte längre stöds av någon tillgänglig programvara. En fil i XML-format är en vanlig textfil och kan läsas som en sådan. Information i XML-format kan också transformeras exempelvis till nya formatversioner för vilka det finns fungerande programvara. Det finns också ett flertal standarder och förslag till XML-baserade arkivstrukturer. Statskontoret har publicerat en förstudie som behandlar området "Framtidens elektroniska arkiv" (2003/67).

Det formella myndighetsansvaret för hur XML ska användas vid digital arkivering ligger hos Riksarkivet i dess roll som arkivmyndighet.

2.4. Sökning

Metadata och väldefinierade datamodeller är en nyckel till konstruktionen av mer avancerade söktjänster som inte bara erbjuder fritextsökning. Detta omfattar möjligheter att:

- söka i specifika metadataelement
- kombinera sökningar i flera metadataelement
- för specifika metadataelement ge fördefinierade sökkriterier baserade på tillåtna vokabulärer
- implementera anpassade sökmekanismer utgående från en metadataelementdefinition, exempelvis sökning på tidsperioder om metadataelementet är definierat som ett datum eller kartbaserad sökning om metadataelementet innehåller geografisk data.

I och med XML:s täta koppling till metadata så finns en förväntning att en XMLifiering av en datamängd genast gör den bättre sökbar. I teorin stämmer detta då en mer detaljerad uppmärkning av data ger mer möjligheter till avancerade sökfunktioner. I praktiken uppstår dock lätt problem då många existerande sökstöd, som exempelvis sökmotorer för webbplatser, inte stödjer XML-formatet. Man måste då använda anpassade sökverktyg för att få tillgång till de mer effektiva sökmöjligheterna. Dessa anpassade sökverktyg stödjer en specifik datamodell. Detta medför att när man definierat en gemensam datamodell måste man även implementera ett anpassat sökstöd.

Det pågår olika initiativ som syftar till att göra strukturerad information till grunden för den "nya" webben, och därmed möjliggöra mer avancerade sökmöjligheter. RDF (Resource Description Framework) är ett arbetsområde inom W3C centrerat kring standarder vilka definierar en grund för hur metadata bör struktureras och användas. "The Semantic Web" (läs mera på <http://www.w3.org/2001/sw/>), som är ett annat arbetsområde inom W3C, är en framtidsvision för intelligent informationshantering, bl.a. informationssökning. En hel del utveckling återstår dock innan detta kan bli praktiskt verklighet.

2.5. Kontorsdokument

XML har börjat bli aktuellt som filformat för ordbehandlare och kalkylprogram. Detta öppnar möjligheter till förenklat utbyte av information mellan användare utan att hänsyn behöver tas till vilket ordbehandlingsprogram som har använts vid framtagningen av informationen.

OASIS driver ett Open Office initiativ, baserat på SUNs StarOffice, där man tar fram ett kontorsapplikationspaket som bygger på öppen källkod och som

använder ett XML-baserat lagringsformat som standard.

I Microsofts Office 2003 finns stöd för XML-format och Microsofts XML-scheman för kontorsdokument finns publicerade och tillgängliga under vad Microsoft benämner ett royalty-fritt licensierings program.

I nuläget är dock inte XML-formaten från de olika programpaketerna kompatibla och det är oklart om och när de kommer att bli utbytbara, men ett förbättrat och utvidgat programvarustöd för olika XML-format bör på sikt komma att minska beroendet av att använda specifika kontorsapplikationer.

EU-kommissionens IDA-kommitté (Interchange of Data between Administrations) publicerade i maj 2004 en rekommendation till medlemsländerna och till IT-industrin om fortsatt utveckling och ökad användning av öppna, XML-baserade dokumentformat. Se http://europa.eu.int/ISPO/ida/jsps/dsp_showDocument.jsp?printerVersion=1&documentID=2592

2.6. Databaslösningar

När XML börjar bli alltmer använt uppstår frågor kring hur information i XML-format ska lagras av verksamhetens informationssystem. Ett XML Schema är en definition av en datamodell på samma sätt som ett databas-schema är det.

När en verksamhets informationssystem innehåller system som baseras på både XML-schemats informationsmodell och på databasens datamodell så uppstår behov av koordinering och integration. Denna integration måste främst göras verksamhetsmässigt genom att man bestämmer vilken datamodell som är styrande för vilken information och hur systemen kan anpassas efter detta.

Tekniskt finns i dag stöd för hantering av XML vid databaslagring. De flesta större databashanterare erbjuder utökningar som gör att man kan överföra data mellan XML-dokument och databaser.

Det finns också rena XML-databaser där DOM (Document Object Model) används som gränssnitt för databasåtkomsten, vilket möjliggör en lösning utan traditionella relationsdatabaser. Vilken lösning som väljs kommer oftast att vara beroende av hur existerande system kan och bör påverkas.

3. Standardiseringsarbete

Anmärkning:

I denna vägledning används termen "standard" för det som brukar kallas "de facto standard", "öppen industristandard" eller "rekommendation". De "standarder" som nämns är alltså ofta inte formellt fastställda av ett officiellt standardiseringsorgan, utan de är rekommendationer som produceras och förvaltas av industrigrupperingar med bred förankring i såväl användar- som leverantörsled, t.ex. W3C. I praktiken används de på motsvarande sätt som formella standarder.

För XML-området är grundläggande standarder fastställda av World Wide Web Consortium (W3C). Ytterligare standarder finns fastställda av organisationer som Dublin Core Metadata Initiative (DCMI) och International Organization for Standardization (ISO). Se vidare 4.3 för mer information om standardiseringsorganisationer.

Denna vägledning presenterar ett urval standarder som bedöms vara relevanta. Att en fastställd standard inte tas upp i vägledningen innebär bara att Statskontoret av olika skäl inte har tagit ställning, det skall inte tolkas som att Statskontoret avråder från att tillämpa den. Vissa, ännu ej fastställda förslag till kommande "W3C Recommendations", har redan nått stor spridning och användning. Det kan finnas goda skäl att redan nu använda dem, men man bör då vara medveten om att risken finns att förslagen kan komma att förändras.

3.1. Faktorer som påverkar standarders acceptans

Ett antal faktorer påverkar om, hur snabbt och på vilket sätt en standard får stöd i kommersiell programvara.

Tidsfördröjningen mellan en standards publicering och kommersiell produkttillgänglighet varierar från område till område. Att leverantörer aktivt deltar i standardiseringsarbetet skapar förutsättningar för att produkter snabbt kommer till marknaden. Produkter för större och mer kritiska marknadssegment utvecklas naturligtvis snabbare.

Produktstabilitet stöds av (1) stabilitet i standarder, och (2) sätt att avgränsa standarder (dels allmänt, men även gränsdragning gentemot andra standarder). Ett omsorgsfullt och öppet standardutvecklingsarbete och ett arbete där grupper av standarder identifieras och specificeras som bitar i ett

större pussel borgar för mer stabila, öppna standarder och därmed stabila produkter på marknaden.

Betydelsen av öppen programvara (svensk översättning av den etablerade beteckningen Open Source Software) skall inte underskattas. Tidigt implementeringsarbete i open source-organisationer stöder i praktiken standardiseringsarbetet genom identifiering av möjliga inkonsekvenser och gap i föreslagna standarder. De identifierar även möjliga tekniska implementeringsansatser som förbereder för efterfrågan genom att tillhandahålla tidiga implementationer. Ett exempel på detta är Apache-gruppen (<http://www.apache.org>), där många XML-orienterade programprototyper har tagits fram (såsom webb-servern Apache och processorer för SOAP, XSLT och SVG).

3.2. Standardiseringsorganisationer

3.2.1. World Wide Web Consortium (W3C)

Organisationen W3C bildades 1994 för att leda utvecklingen av gemensamma webbstandarder och därmed säkra framtida interoperabilitet över webben. W3C har idag cirka 350 medlemsorganisationer från hela världen, såväl myndigheter och användar- och forskningsorganisationer som produktproducerande företag.

Arbetet i W3C drivs i arbetsgrupper, huvudsakligen bemannade med personal från W3C:s medlemsorganisationer. Processen är konsensusbaserad, dvs. resultat som produceras stöds av alla inblandade organisationer.

Följande statusbeteckningar används på vägen fram till en fastställd "Recommendation"¹:

- Working draft (WD)
Arbetsversion av eventuell framtida standard.
- Candidate recommendation (CR)
Specifikation som aspirerar på att bli en standard och väntar på kommentarer från andra arbetsgrupper inom W3C.
- Proposed recommendation (PR)
Förslag till standard som nått en viss förankring och som nu är uppe för granskning.

¹ <http://www.w3.org/2003/06/Process-20030618/tr.html>

- Recommendation (REC)
Av W3C antagen standard, vilket innebär att W3C:s medlemsorganisationer granskat och godkänt förslaget.

Det finns förslag på att införa ytterligare en statusklassificering för dokument, Rescinded Recommendation (RR). Klassificeringen skall användas på standarder som W3C inte längre vill förordna, t.ex. därför att teknikutvecklingen tagit en annan riktning.

XML är namnet på den grundläggande XML-specifikation som publicerades som "Recommendation" av W3C år 1998, men XML används också som ett samlingsbegrepp för hela XML-familjen som omfattar ett stort antal relaterade standarder och initiativ.

3.2.2. Dublin Core Metadata Initiative (DCMI)

DCMI är en internationell, ämnesöverskridande gruppering som syftar till att utveckla mekanismer för att beskriva resurser i elektronisk miljö. Arbetet med Dublin Core påbörjades 1995 vid en workshop i Dublin (Ohio) och fokuserade initialt på metadata som stöd för sökningar i biblioteksvärlden, dvs. sökningar i det man traditionellt kallar publikationer. I takt med att nya typer av objekt blir tillgängliga på webben, har DCMI:s resultat kommit till användning även utanför biblioteksvärlden för att uttrycka metadata om andra typer av informationsobjekt.

DCMI har för närvarande 18 aktiva arbets- och intressegrupper. Det finns runt 80 olika implementeringsprojekt av DCMI registrerade på DCMI-projektets hemsida, hemmahörande i alla olika världsdelar och inom många olika ämnesområden.

DCMI använder statusbeteckningarna Note, Working Draft, Proposed Recommendation och Recommendation på liknande sätt som W3C (se ovan).

3.2.3. Web Services Interoperability Organization (WS-I)

WS-I, grundat 2001, är en annan internationell sammanslutning av företag, med målsättning att arbeta för interoperabilitet mellan webbtjänster ("web services") baserade på accepterade standarder.

Det huvudsakliga resultatet av WS-I:s arbete är specificerade profiler, där en webbtjänst skall kunna deklarerat att den understödjer en viss profil. En sådan profil definieras i termer av ett antal specifika webbstandarder (t.ex. SOAP och WSDL), en nivå för varje standard (t.ex. SOAP 1.1) och hur vissa specifika element i denna nivå skall användas. Man kan uttrycka det

som att en WS-I profil är en standard definierad som en konfiguration av explicita delmängder av existerande standarder.

Arbetet bedrivs i arbetsgrupper, vilka skapas för att ta fram konkreta resultat, t.ex. testprogramvara, testapplikationer eller specifika standarder. Man bör notera att WS-I:s insatser är ett komplement till andra organisationers arbete inom området webbstandarder, och WS-I samverkar med dessa andra organisationer. Denna samverkan underlättas av att WS-I:s medlemmar till största delen också är W3C-medlemmar.

Följande statusbeteckningar används av WS-I:

- Draft specification
Arbetsversion av en eventuell framtida standard.
- Final specification
Fastställd standard, granskad och godkänd av WS-I:s medlemmar.

3.2.4. Organization for the Advancement of Structured Information Standards (OASIS)

OASIS är ytterligare ett globalt konsortium, bildat 1993, som stöder utveckling, praktisk tillämpning och spridning av standarder. OASIS har över 3000 medlemmar som representerar 600 organisationer och individuella medlemmar i 100 länder och arbetet är organiserat i ett stort antal tekniska kommittéer. Exempel på uppmärksammade projekt som drivs inom ramen för OASIS är ebXML, LegalXML och UDDI.

Två relativt nystartade tekniska kommittéer är av särskilt intresse eftersom de arbetar med Open Office XML Format respektive med standardiseringsfrågor kring e-Government.

3.2.5. International Organization for Standardization (ISO)

ISO är världens största standardiseringsorganisation. Sekretariatet är förlagt till Genève men ISO har medlemmar från 147 länder. Sedan 1947 har ISO publicerat fler än 13 700 internationella standarder inom de flesta område, dess främsta uppgift är dock att ta fram tekniska standarder.

Standardiseringsaktiviteter förekommer även inom många andra standardiseringsorgan: Internet Engineering Task Force (IETF), European Committee for Standardization (CEN) och amerikanska National Information Standards Organization (NISO) är några exempel.

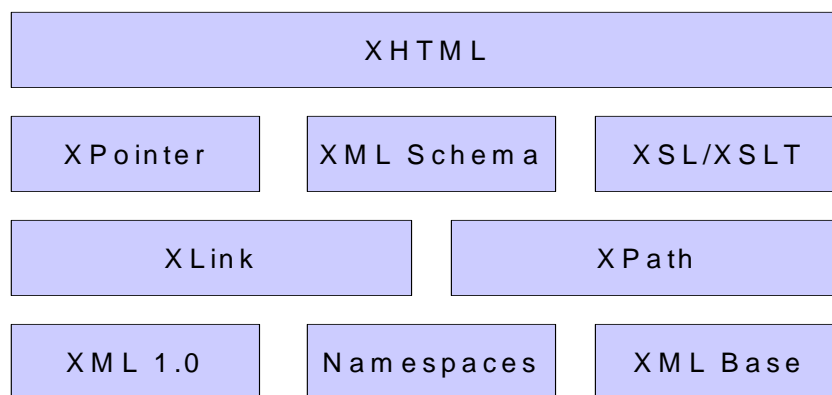
3.3. Viktiga XML-standarder

XML är inte en färdig lösning utan ett koncept med en verktygslåda, vilket framgår av namnet: XML = eXtensible Markup Language:

Alltså ett *språk*, avsett för att *märka upp* information, konstruerat för att *utvidgas*.

Den grundläggande XML-standarden fastställdes så sent som 1998 och även om målen och visionerna är formulerade och en uppsättning generella standarder är fastlagda så återstår det en hel del arbete för att skapa kompletta, branschspecifika ramverk. Först skall W3C samordna och besluta om alla pågående förslag till standarder. Sedan skall marknadens aktörer dels definiera och enas om olika branschapplikationer, dels utveckla nödvändiga strukturer och programvaror som bygger på de överenskomna standarderna.

Statskontoret har i detta dokument identifierat några viktiga XML-baserade standarder, vilka beskrivs kortfattat nedan. Eftersom dessa standarder bygger på varandra kan en figur illustrera hur de är uppbyggda på ett komponentorienterat sätt.



Figur 1. I denna figur gäller att standarder i ett visst lager utnyttjar standarder i underliggande lager.

3.3.1. Extensible Markup Language (XML) 1.0 (Third Edition)

W3C Recommendation (februari 1998, 2nd edition: oktober 2000, 3rd edition februari 2004)

Grunden i hela XML-ramverket för strukturerad informationshantering är denna Recommendation. Den specificerar bland annat hur XML-dokument får se ut och skall byggas upp, dvs. syntaxen. Mot dessa krav kontrolleras att ett XML-dokument följer den gällande XML-syntaxen. XML 1.0 specificerar XML-dokuments logiska struktur (element/attribut-struktur), deras fysiska struktur (lexikala/syntaktiska form), och begreppet validitet (gentemot en dokumenttyp).

Den tredje upplagan (februari 2004) ska inte ses som en ny version av XML utan den innehåller de rättningar som inkommit till den andra upplagan (2nd edition, oktober 2000). Rättningarna berör främst språkliga fel i specifikationstexten och ska inte ses som en ny teknisk standard. För en beskrivning av rättningarna se <http://www.w3.org/XML/xml-V10-2e-errata>.

Läs mer på: <http://www.w3.org/XML/>
Teknisk specifikation: <http://www.w3.org/TR/REC-xml>

3.3.2. Extensible Markup Language (XML) 1.1

W3C Recommendation (februari 2004)

I februari 2004 kom en uppdatering av XML, version 1.1, även den med status Recommendation. Uppdateringen orsakas i huvudsak av den utveckling som Unicode-standarden genomgått sedan tidpunkten då XML 1.0 blev Recommendation. Utvecklingen får effekt på vad som är legala namn i XML-dokument eftersom antalet definierade Unicode-tecken har förändrats. XML 1.1 uppdaterar XML så att man inte längre är beroende av den specifika Unicode-versionen.

Om man inte behöver de nya delarna i XML 1.1 går det fortfarande bra att använda XML 1.0, då XML Parsers förstår båda versionerna.

Teknisk specifikation: <http://www.w3.org/TR/xml11/>

3.3.3. Namespaces in XML Version 1.0

W3C Recommendation (januari 1999)

Det ökande behovet av att sammanställa information från oberoende källor ställer krav på att olika uppmärkningsspråk (som används i olika dokument) inte sammanblandas med varandra. Det måste vara möjligt att entydigt, men samtidigt effektivt och hanterbart, kunna särskilja en uppsättning XML-element/attribut från en annan uppsättning element/attribut. I XML-familjen åstadkommes detta genom användning av namnrymder ("namespaces").

"Namespaces in XML" är en W3C Recommendation som definierar hur namn på element och attribut får uttryckas, och hur två olika "namnförekomster" kan stå för samma namn. Användningen av Namespaces gör det möjligt att hitta det element eller attribut som specifikt används i ett XML dokument, även om det finns element med samma elementnamn, t.ex. "ID". Namespaces åstadkommer detta genom att man för varje element anger vilken definition av elementet som gäller, dvs. vilken namnrymd det tillhör. Ett prefix sätts in varje gång ett begrepp från en viss applikation används och prefixet fungerar då som referens till rätt namnrymd.

Läs mer på: <http://www.w3.org/TR/REC-xml-names/>

3.3.4. Namespaces in XML Version 1.1

W3C Recommendation (februari 2003)

I februari 2003 blev Namespaces version 1.1 en Recommendation. Skillnaden gentemot Namespaces 1.0 är, utöver rättelser av några fel, tillägg av en enda funktion: möjlighet att avdeklarerera ("undeclare") ett prefix. Namespaces 1.1 avses gälla från och med XML 1.1.

Läs mer om Namespaces 1.1 på: <http://www.w3.org/TR/2004/REC-xml-names11-20040204/>

3.3.5. XML Base

W3C Recommendation (juni 2001)

XML Base tillhandahåller uttrycksmedel för att ange en referenspunkt för relativa länkar i ett XML-dokument på ett sätt som motsvarar Base-elementets roll i ett HTML-dokument. Dess huvudsakliga användningsområde är att ge stöd åt XLink, men genom att definieras på ett modulärt sätt kan XML Base även användas i andra XML-sammanhang där relativa länkar används.

Denna W3C Recommendation, en av de textmässigt sett minsta i XML-familjen, definierar ett enda attribut, xml:base, samt specificerar hur relativa länkar skall omtolkas för att bli absoluta länkar.

Teknisk specifikation: <http://www.w3.org/TR/xmlbase/>

3.3.6. XML Linking Language (XLink) Version 1.0

W3C Recommendation (juni 2001)

Ett av de områden där aktuella och framtida behov går utöver vad som traditionellt erbjuds av HTML är behovet av mer avancerade länkar mellan

olika dokument (resurser). HTML erbjuder några språkkonstruktioner, varmed enkla riktade länkar kan specificeras. XLink erbjuder möjligheter att uttrycka sådant som att en länk kan koppla till mer än en resurs, vilken som visas och/eller väljs är sedan upp till webbläsaren. XLink gör det också möjligt att knyta metadata till en länk och att specificera länkar på platser som är skilda från de resurser de sammanlänkar.

För att förenkla hanteringen av HTML-liknande länkar (enkla länkar täcker en stor del av dagens behov av länkar), definierar XLink två kategorier av länkar: "extended links" och "simple links". Extended links erbjuder den fulla XLink-funktionen men kräver något mer omfattande specifikationer av länkar. Simple links har inte så rika uttrycksmöjligheter men möjliggör att länkar specificeras på ett sätt som påminner mycket om vanliga HTML-länkar.

Teknisk specifikation: <http://www.w3.org/TR/xlink/>

3.3.7. XML Path Language (XPath) Version 1.0

W3C Recommendation (november 1999)

XPath används för att referera till data inom ett XML-dokument. Dess primära användningsområde är att utgöra en gemensam grund för att specificera delar av XML-dokument, vilket är funktioner som behövs i t.ex. XSL Transformations (XSLT) och XPointer.

Data som kan refereras med hjälp av XPath är t.ex. element, attribut, text och andra komponenter. Notationen liknar sökvägar ("paths") i en katalogstruktur. XPath bygger förutom XML 1.0 även på XML Namespaces.

XPath 2.0 är publicerad i november 2003 som "Working Draft". Målsättningen med denna efterföljare till XPath 1.0 är huvudsakligen att förenkla XPaths användning, samtidigt som bakåtkompatibilitet bevaras. Av detta skäl kan användning av XPath 1.0 anses vara framtidssäkrad.

Teknisk specifikation: <http://www.w3.org/TR/xpath>

3.3.8. XML Pointer Language (XPointer) Version 1.0

W3C Recommendation (mars 2003)

XPointer, XML Pointer Language, är det språk som skall användas då man behöver uttrycka en länk till en bestämd del av ett XML-dokument. XPointer bygger vidare på XPath, och används för att identifiera delar av dokument. Identifieringen baseras på olika egenskaper i dokumentet, t.ex. element-typer, attributvärden, en position relativt ett annat element m.m.

Teknisk specifikation: <http://www.w3.org/TR/xptr-element/>,
<http://www.w3.org/TR/xptr-framework/> samt <http://www.w3.org/TR/xptr-xmlns/>

3.3.9. XML Schema Version 1.0

W3C Recommendation (maj 2001)

XML Schema är ett modernare och mer fullständigt sätt att definiera strukturen för XML-dokument och kan användas istället för en DTD. XML Schema erbjuder mer stöd för modularitet och datatypning än vad man kan specificera i en DTD, dock blir schemat i regel längre och mindre överskådligt än motsvarande DTD. Ett XML Schema ser ut som ett XML-dokument, vilket beror på att det i sig är en XML-applikation, till skillnad från en DTD som uttrycks i ett specialiserat språk. Förutom XML 1.0 bygger XML Schema även på XML Namespaces. Många XML-applikationer använder fortfarande DTD:er och många programvaror saknar fullt stöd för XML Schema, men användningen av XML Schema ökar, framför allt tack vare dess vidgade möjligheter att uttrycka krav på datatypning.

Läs mer på: <http://www.w3.org/XML/Schema>
<http://www.w3.org/TR/xmlschema-0/> ("XML Schema Part 0: Primer")

Teknisk specifikation:
<http://www.w3.org/TR/xmlschema-1/> ("XML Schema Part 1: Structures")
<http://www.w3.org/TR/xmlschema-2/> ("XML Schema Part 2: Datatypes")

3.3.10. Extensible Stylesheet Language (XSL) Version 1.0

W3C Recommendation (oktober 2001)

XSL är en sammansättning av tre olika XML komponenter som tillsammans används för att beskriva hur XML-dokument ska se ut när de ska presenteras, t.ex. i en webbläsare. En XSL-formatmall, ett "stylesheet", beskriver hur ett XML-dokument av viss typ skall formateras. I detta avseende har XSL samma målsättning som, och är i många avseenden kompatibelt med, CSS2 (Cascading Style Sheets level 2), men XSL använder XML-syntax, vilket gör det mer uttrycksfullt än CSS2.

Vid användning av XSL för formatering så utnyttjas tre olika komponenter:

1. XSLT
Ett språk för att transformera XML-dokument
2. Xpath
Ett språk för att referera till delar av dokument
3. XSL Formatting Objects
En notation för att specificera typografisk presentation

Dessa tre komponenter stöder två huvudsteg i presentationen:

- transformering (använder XSLT och XPath) och
- formatering (använder XSL Formatting Objects).

XSLT och XPath togs fram som komponenter för att stödja presentation av XML-dokument, men genom att på lämpligt sätt avgränsa och generalisera dessa två komponenter, så har de även kommit till användning vid allmän transformering och sökning av XML-dokument. XSL Formatting Objects är en del i XSL Version 1.0.

Cascading Stylesheet level 2, CSS2, framtaget för användning tillsammans med HTML, är föregångare till XSL och kan också användas för formatering av XML-dokument.

Teknisk specifikation XSL: <http://www.w3.org/TR/xsl/>

Läs mer om CSS på: <http://www.w3.org/Style/CSS/>

3.3.11. XSL Transformations (XSLT) Version 1.0

W3C Recommendation (november 1999)

Ursprungligen en del av XSL (se ovan), men är sedan 1999 en egen XML-specifikation. XSLT används för att beskriva hur XML-dokument kan översättas till andra XML-dokument. Här avses översättning av själva uppmärkningen och inte av textinnehållet i dokumentet. En XSLT-översättning kallas för en transformation och översätter uppmärkningen från en XML-applikation till en annan.

Teknisk specifikation: <http://www.w3.org/TR/xslt>

XSLT 2.0 fick i november 2003 status "Working Draft, last call". Målsättningen med denna efterföljare till XSLT 1.0 är huvudsakligen att förenkla XSLT-användning, samtidigt som bakåtkompatibilitet bevaras.

Teknisk specifikation: <http://www.w3.org/TR/xslt20/>

3.3.12. XHTML Version 1.0 (2nd edition)

W3C Recommendation (januari 2000, reviderad augusti 2002)

Extensible HyperText Markup Language (XHTML) återskapar och utvidgar HTML, uttryckt i XML. XHTML 1.0 är den första XHTML Recommendation och den som främst används idag. Den omformulerar HTML 4.01 i XML.

Det finns tre varianter definierade i XHTML 1.0: Strict, Transitional och Frameset.

1. XHTML 1.0 Strict möjliggör en ren struktur-uppmärkning, fri från presentations-uppmärkning, så att presentationsegenskaper helt kan uttryckas i CSS.
2. XHTML 1.0 Transitional innehåller utvidgningar för att kunna uttrycka vissa presentationsegenskaper, t.ex. bakgrundsfärg. Avsikten med varianten "Transitional" är att under en övergångstid kunna stödja användning av äldre webbläsare som inte kan hantera CSS, och som därför behöver extra stöd för XHTML.
3. XHTML 1.0 Frameset möjliggör användandet av ramar vid presentation.

Andra XHTML- projekt

XHTML 1.1 eller som det också heter XHTML Modularization (W3C Recommendation, april 2001), specificerar en modularisering av XHTML-språket vilket gör det möjligt att olika delmängder av XHTML kan definieras och kombineras på ett systematiskt sätt. Därigenom är det möjligt att specificera vilka moduler som en plattform stöder. Se <http://www.w3.org/TR/xhtml-modularization/conformance.html> för mer information om XHTML Modularization.

XHTML Basic (W3C Recommendation, december 2000), avsett som uppmärkningspråk för mobiltelefoner, PDA:er, "set top boxes" (digitalTV-boxar) och liknande, är ett exempel på ett XHTML-språk specificerat med hjälp av XHTML Modularization.

XHTML 2.0 är under utveckling och baserat på XHTML 1.0 men specificerat med hjälp av XHTML Modularization.

XHTML och webbläsare

XHTML bör användas för webbsidor och stöds av de flesta moderna webbläsare.

Senare versioner av Internet Explorer, Netscape och Mozilla har stöd för XHTML, se www.microsoft.se, www.netscape.com och www.mozilla.org för mer detaljerade uppgifter. Opera 7 stödjer XHTML Basic men har ett antal undantag i sitt stöd för XHTML 1.0 samt XHTML 1.1. Även tidigare versionen av Opera har ett visst stöd för XHTML, se www.opera.com för mer detaljerade uppgifter.

Teknisk specifikation XHTML 1.0: <http://www.w3.org/TR/xhtml1/>

Teknisk specifikation XHTML 1.1 (Modularization):
<http://www.w3.org/TR/xhtml11/>

Teknisk specifikation XHTML Basic:
<http://www.w3.org/TR/xhtml-basic/>

3.4. Viktiga Web Service standarder

3.4.1. Introduktion

Web services (webb-tjänster) har under senare år varit ett av de hetaste IT-områdena och bygger på XML. Alla betydelsefulla teknologileverantörer har satsat på att leverera infrastruktur som skall göra det enkelt att såväl bygga upp som förvalta en organisations utbud av webb-tjänster. Många användarorganisationer ser också webb-tjänster som en infrastruktur för integration mellan egna tillämpningsprogram på ett intranät, vilket kan bidra till att minska beroendet av leverantörsspecifika integrationsplattformar.

Utvecklingstakten har bromsats upp något, bland annat beroende på att standarderna inte varit färdigutvecklade, på en osäkerhet vid valet mellan olika initiativ från konsortier och leverantörer samt på frågetecken kring säkerheten. I takt med att standarder och produkter mognar fortsätter dock utvecklingen och spridningen av Web services till nya områden.

Web services är inte ett helt entydigt definierat begrepp men de brukar allmänt anses bygga på XML, SOAP, WSDL och UDDI eller motsvarande.

Med en webbtjänst menas en automatisk tjänst som typiskt levereras över HTTP. För att kunna utnyttja en tjänst måste man veta vad tjänsten kan utföra och hur man kan anropa den. WSDL, "Web Service Description Language", är ett språk i vilket man kan uttrycka sådana beskrivningar. Med andra ord, WSDL stöder upptäckt ("discovery") och beskrivning

("description") av tillgängliga webbtjänster. Sådana tjänstebeskrivningar publiceras av tjänsteleverantören, i någon publikt sökbar databas, så att klienter kan hitta lämpliga tjänster (t.ex. via UDDI, "Universal Description, Discovery, and Integration"). När väl klienten har fått WSDL-beskrivningen av tjänsten så kan klienten anropa denna genom att sända ett meddelande till tjänsten (t.ex. till en URL som angivits i tjänstens WSDL-beskrivning). Såväl det sända meddelandet som det svar klienten får tillbaka uttrycks i SOAP, som är ett språk varmed man kan överföra strukturerad information mellan tillämpningsprogram.

Flexibiliteten i de standarder som webb-tjänster bygger på gör det möjligt att definiera såväl mycket komplexa som mycket enkla tjänster. För att underlätta ett snabbt införande av dessa standarder har några enklare kombinationer av grundstandarderna fastlagts i form av "webb-tjänst-profiler". Genom att en klient direkt kan identifiera vilken profil tjänsten utnyttjar så kan klienten enkelt bygga de meddelanden som sändes till tjänsten och enkelt avkoda det svar som erhålles i retur. Sådana profiler definieras t.ex. av en till W3C komplementär gruppering, WS-I (se 4.3.3).

SOAP 1.2 har nu blivit en W3C Recommendation och UDDI v2 kan betraktas som en öppen standard. Den återstående centrala komponenten för web services som ännu inte har uppnått en fastställd status är beskrivningsspråket WSDL.

3.4.2. SOAP Version 1.2

W3C Recommendation (juni 2003)

SOAP är en central komponent i W3C:s utvecklingsaktiviteter inom Web services, se även följande avsnitt samt <http://www.w3.org/2002/ws/Activity>

SOAP är ett språk för att kommunicera funktionsanrop och strukturerad information i XML-format mellan tillämpningsprogram i en distribuerad och decentraliserad miljö.

SOAP 1.2 består av fyra delar. Del 0: Primer, Del ett: Messaging Framework, Del två: Adjuncts, och slutligen ett dokument som beskriver Specification Assertions and Test Collection.

Tekniska specifikationer:

<http://www.w3.org/TR/soap12-part0/>,
<http://www.w3.org/TR/soap12-part1/>,
<http://www.w3.org/TR/soap12-part2/> samt <http://www.w3.org/TR/soap12-testcollection/>.

3.4.3. Web Services Description Language (WSDL) Version 2.0

W3C Working Draft (november 2003)

WSDL är ett XML-baserat språk varmed man kan beskriva webb-tjänster, "web services". Beskrivningsmodellen baseras på meddelandeformat och protokoll, operationer och tjänster ("services"). WSDL definierar också bindningar till ett antal protokoll, såsom SOAP, HTTP och MIME.

WSDL 1.1 är en tidigare arbetsversion av WSDL-språket, ett steg på vägen mot WSDL 2.0 (föret kallad 1.2), med målet att bli fastlagd W3C-standard. Det nya i version 2.0 (gentemot version 1.1) är att den är bättre integrerad i grundläggande specifikationer (såsom XML Schema), "service descriptions" kan uttryckas enklare, vissa potentiella interoperabilitetsproblem har eliminerats, och att bindningarna mot HTTP har förbättrats. Detta betyder att version 2.0 bör betraktas som mer än en ren utvidgning av version 1.1.

Läs mer på <http://www.w3.org/2002/ws/desc/>

WSDL 2.0 är för närvarande en Working Draft.

Teknisk specifikation: <http://www.w3.org/TR/wsdl20/>

3.4.4. Universal Description, Discovery, and Integration (UDDI)

OASIS Open Standard

Universal Description, Discovery and Integration är en specifikation för distribuerade, webbaserade informationsregister över tillgängliga webb-tjänster. UDDI finns också som publikt tillgänglig implementering av specifikationen, som gör det möjligt att registrera information om de webbtjänster man erbjuder så att andra kan finna dem.

UDDI-specifikationen tas fram inom ramen för organisationen OASIS öppna standardisering. För närvarande gäller UDDI version 2 som standard, men UDDI version 3 publicerades i augusti 2003.

Läs mer på: <http://www.uddi.org/>

Tekniska specifikationer:

<http://www.oasis-open.org/committees/uddi-spec/doc/tcspecs.htm#uddiv2>

3.4.5. WS-I Basic profile Version 1.0

WS-I Final Specification (april 2004)

Denna specifikation tas fram av WS-I (se avsnitt 4.3.3), till skillnad från de specifikationer som nämnts ovan, vilka tas fram av W3C. Basic Profile är den första webbservice profilen som presenterats av WS-I, men kommer att följas av flera mer avancerade profiler. Genom att webbservices skapas efter en vald WS-I-profil hoppas man undvika de interoperabilitetsproblem som kan vara resultatet av att olika standarder används eller att de tolkas olika.

Basic Profile är en konfiguration definierad på standarderna XML Schema 1.0, SOAP 1.1, WSDL 1.1 och UDDI 2.0.

WS-I planerar att presentera fler profiler allteftersom det blir tydligare vilka behov som finns. Förmodligen kommer nya profiler att inkludera nya versioner av XML Recommendations, när dessa har fått en större spridning och acceptans.

Det föreligger således ett visst timing-problem, då W3C redan har publicerat uppdateringar av de specifikationer som WS-I har haft som grund för sitt arbete.

Eftersom det finns ett antal viktiga skillnader mellan SOAP 1.1 och SOAP 1.2 är det troligt att SOAP 1.2 kommer att inkluderas i en eller fler framtida profiler.

WSDL-komponenten i Basic Profile 1 är version WSDL 1.1 som får betraktas som en provisorisk version på vägen mot en fastlagd WSDL-rekommendation. Som beskrivs under avsnittet om "WSDL", ovan, så är WSDL 2.0 inte en ren utvidgning av WSDL 1.1, vilket kan medföra att rekommendationerna i Basic Profile inte blir direkt tillämpliga på WSDL 2.0, som troligen kommer att ligga närmare en kommande W3C-rekommendation. Det är rimligt att anta att det kommer att produceras en ny WS-I profil om hur WSDL 2.0 skall användas, men det är i dagens läge inte känt om och när detta kommer att ske.

Teknisk specifikation: <http://www.ws-i.org/Profiles/BasicProfile-1.0-2004-04-16.html>

3.5. Viktiga standarder för Metadata

Inom metadataområdet pågår en intensiv utveckling. Leverantörsspecifika lösningar positionerar sig på marknaden och det förekommer många sektors- eller branschspecifika grupperingar som tar olika samordnings-initiativ. Metadatastandardisering ingår också ofta som en integrerad del i större projekt för informationsstrukturering och terminologisamordning och kan därför vara svår att särskilja. I detta sammanhang har vi därför för närvarande bara identifierat en generell metadatastandard, som vi bedömer

som en de facto standard och även som en utgångspunkt för många andra metadata-initiativ, nämligen Dublin Core:

3.5.1. Dublin Core Metadata Element Set (DCMES) Version 1.1

DCMI Recommendation (2003)

DCMES har också fastställts som internationell standard, ISO 15836-2003.

Denna standard är framtagen av DCMI, till skillnad från flertalet standarder som nämnts ovan, vilka tagits fram av W3C.

Dublin Core Metadata Initiative är ett internationellt, ämnesöverskridande arbete som syftar till att utveckla mekanismer för att beskriva resurser i elektronisk miljö. Ett antal metadata-element har specificerats för detta ändamål.

För att underlätta återsökning av dokument rekommenderas i första hand användning av relevanta delar av de 15 elementen i DCMES.

Läs mer på: <http://dublincore.org/>

Teknisk specifikation: <http://dublincore.org/documents/2003/02/04/dces>

4. Pågående initiativ

Det finns ett stort antal andra standarder och initiativ som bygger på XML och metadata, t.ex. applikationer för att stödja olika sektorer speciella behov. Att kontinuerligt bevaka och värdera dessa branschlösningar ligger utanför denna väglednings ambitioner, men många av dessa är väl etablerade inom respektive område.

Varje organisation bör sätta sig in i möjligheterna med XML, men att på egen hand börja utveckla ”proprietära” XML- och metadata lösningar är sällan det mest effektiva. Kostnaderna kan bli höga och om man sedan är ensam om sin XML-tillämpning så har man kanske missat potentiella samordningsvinster. Statskontoret rekommenderar därför myndigheterna att i sina projekt aktivt undersöka om gemensamma branschinitiativ finns inom respektive myndighets verksamhetssektor. I sådana fall bör myndigheten pröva möjligheten och lämpligheten av att ansluta sig till dessa initiativ. Exempel på där så har skett är t.ex. GML (geografisk information) och HR-XML (arbetsmarknad). Sökingångar till sådana samarbetsmöjligheter kan t.ex. finnas via några av de webbplatser som anges under referenser (se nedan).

Nedan följer information om några initiativ som är generella och etablerade och bör kunna vara av intresse för ett flertal myndigheter. Dessutom beskriver de initiativ där svenska myndigheter har engagerat sig i det internationella arbetet. Myndigheter som har påbörjat ett arbete med informationsstrukturer kan även titta på initiativ inom andra områden, då dessa ibland kan ge tillgång till lösningar och arbetssätt som kan fungera även inom det egna området. De olika e-myndighetsinitiativen runt om i världen är ofta ambitiösa och öppna när det gäller att dokumentera möjliga arbetsmetoder inom området informationsstrukturering.

XML-akademin är ett informellt nätverk för XML inom offentlig sektor, med syfte att utbyta erfarenheter samt att underlätta och påskynda användningen av XML inom förvaltningen. XML-akademin ger information, tips och idéer samt rekommendationer inom området och vänder sig framför allt till IT- och informationsstrateger inom offentlig sektor.

<http://www.xmlakademin.nu>

4.1. Standardmeddelanden

Tillsammans med representanter från ett urval av registeransvariga myndigheter och IT-leverantörer arbetar Statskontoret med att utforma ett koncept för att samordna och förenkla informationsutbytet av definierade informationsmängder mellan organisationer. Förslaget bygger på att

utveckla och publicera återanvändbara, komponentbaserade informationsstrukturer, som definieras med hjälp av XML Schema och Namespaces. Målet är också att harmonisera dessa rekommendationer med motsvarande i olika internationella initiativ, se nedan. Resultatet av arbetet med standardmeddelanden skall redovisas i en separat rapport. Läs mer på <http://www.24-timmarsmyndigheten.se/DynPage.aspx?id=442>

4.2. EU:s IDA och nationella europeiska initiativ

IDA-programmet (Interchange of Data between Administrations) är ett strategiskt initiativ som drivs av EU-kommissionen i syfte att effektivisera informationsutbytet inom EU och mellan medlemsländernas förvaltningar. I ett inledningsskede fokuserade man på att initiera tekniska lösningar men idag kan IDA ses som drivkraften bakom EU:s arbete inom e-Government. IDA:s verksamhet baseras på två beslut i rådet:

- "Guidelines" – Att ta fram riktlinjer för hur nätverk i olika sektorer ska byggas upp samt vilka regler de ska följa.
- "Interoperability" – Att ta upp vilka generella tjänster som ska finnas samt allmänna verktyg, tekniker och riktlinjer för interoperabilitet.

I april 2004 beslutade rådet också om det nya IDABC-programmet, som omfattar ”leverans av PAN-europeiska e-Government-tjänster för Public Administrations, Businesses and Citizens”.

Läs mer om IDA:s arbete på:

<http://europa.eu.int/ISPO/ida/jsps/index.jsp?fuseAction=home>

Det finns även ett flertal exempel på nationella initiativ i Europa som syftar till att främja användningen av XML inom offentlig förvaltning och ett par av dem presenteras kortfattat nedan.

4.2.1. Danmark

I Danmark bedrivs ett projekt, OIOXML, som syftar till att förbättra utväxling av data både inom den offentliga sektorn och mellan den offentliga och privata sektorn. Andra visioner är att förbättra databehandling, ge lättare åtkomst och återanvändning av insamlad data samt underlätta implementering av e-tjänster. Medlet för att kunna uppnå dessa visioner är att låta XML utgöra ramarna för datautväxling i offentlig förvaltning.

XML-initiativet består av två delprojekt: standardisering av offentliga data och etablering av Infostrukturbasen. Standardiseringsprocessen har till mål att fastlägga XML-baserade standarder för utväxling av data mellan

offentliga myndigheter och mellan offentliga och privata institutioner. Etableringen av Infostrukturbasen har som mål att skapa en databas som består av information om innehållet i offentliga databaser samt hur man får tillgång till dessa data.

De båda delprojekten startade under 2001 och tanken är att de två processerna fortlöpande ska samordnas.

Mer information om det danska initiativet finns på:
OIOXML, www.oio.dk/XML

4.2.2. Storbritannien

Govtalk är ett brittiskt projekt som likt det danska fokuserar på användningen av XML i offentlig förvaltning.

Inom ramarna för projektet har det utvecklats ett antal tekniska standarder som ska utgöra grunden inom e-Government-området. Dessa standarder möjliggör ett enklare informationsflöde inom den offentliga förvaltningen och ger medborgare och företag bättre åtkomst till statliga tjänster.

Standarder som utvecklas berör bl.a. XML-scheman, metadatastandarder (e-Government Metadata Standard) och ramverk för e-tjänster. I projektet ingår även ett ramverk för interoperabilitet, "e-Government Interoperability Framework", e-GIF. Den huvudsakliga tanken med e-GIF är att godkänna och anta Internet- och world wide web-standarder för utveckling av offentliga system. Syftet är att minska kostnaderna och riskerna med att hantera IT-system samtidigt som man inom den offentliga sektorn vill hålla jämna steg med den globala IT-utvecklingen.

Läs mer om Govtalk på: <http://www.govtalk.gov.uk/>

Läs mer om e-GIF på:

<http://www.govtalk.gov.uk/schemasstandards/egif.asp>

4.2.3. Tyskland

Även i Tyskland pågår ett federalt statligt initiativ under namnet SAGA för att främja XML-användningen. Ett dokument har tagits fram, "Standards and Architectures for e-Government Applications 2.0", som nu finns att tillgå på engelska. Vägledningen presenterar relevanta standarder, processer och metoder inom området.

Läs mer om SAGA på: www.kbst.bund.de/SAGA

4.3. Branschinitiativ

4.3.1. ebXML

EbXML (Electronic Business using eXtensible Markup Language) är ett ramverk för affärskommunikation, s.k. business-to-business (B2B). Bakom ebXML står Förenta Nationerna (FN) genom organet United Nations Centre for Trade Facilitation (UN/CEFACT) och Organization for the Advancement of Structured Information Standards (OASIS). Flera ebXML-specifikationer har nyligen också publicerats som tekniska rapporter av ISO i serien ISO 15000.

Projektet startade 1999 och inkluderade så småningom XML-standarder för bland annat affärsprocesser, samarbetsprotokoll, kärndataskomponenter (Core Components) och utbyte av affärsmeddelanden.

ebXML innehåller modeller för hela processen i elektronisk handel, från att hitta en affärspartner till att utföra löpande affärer. ebXML baseras på FN:s välbeprövade standard Electronic Data Interchange (EDI) och som meddelandeformat används Simple Object Access Protocol (SOAP). ebXML-konceptet är inte begränsat till elektronisk handel utan kan också tillämpas för t.ex. offentlig ärendehantering, vilket bl.a. utreds inom EU:s IDA-program.

I Sverige har GEA (Gemenskapen för Elektroniska Affärer, som dock har avvecklats och i juni 2004 ombildats som NEA – Nätverket för Elektroniska Affärer) tagit initiativ till en svensk intressegrupp för ebXML-frågor. Tanken är att ebXML-ramverket och dess användning ska diskuteras och olika aktiviteter genomföras. Områden som intressegruppen kan ägna sig åt är ställningstagande till pågående internationell utveckling, nationella frågor i anslutning till ebXML, agera som referensgrupp för deltagare i internationellt arbete. Vidare skall intressegruppen vara ett forum för frågor kring funktioner och utveckling samt för planering och genomförande av utbildningar och seminarier.

Svenska Kommunförbundet, Landstingsförbundet, Statskontoret och Kammarkollegiet bedriver gemensamt sedan flera år tillbaka ett arbete med att främja förutsättningarna för elektronisk handel för att skapa kostnads-effektivare hantering av inköps- och fakturaprocesserna. Arbetet inriktas på öppna standarder för elektronisk affärskommunikation under begreppet SFTI (Singel Face To Industry) och i juni 2004 fastställdes en enkel elektronisk faktura.

Läs mer om de svenska initiativen på: <http://www.ebxml.nu/se/>, http://www.gea.nu/ebxml/ebxml_intressegrupp.html och <http://www.nea.nu>

Läs mer om SFTI: <http://www.eh.svekom.se>

Läs mer om ebXML: <http://www.ebxml.org/>

Hos XML Cover Pages: <http://www.oasis-open.org/cover/ebXML.html>

Läs mer om UN/CEFACT <http://www.unece.org/cefact/>

Läs mer om OASIS <http://www.oasis-open.org/>

4.3.2. XBRL

XBRL (Extensible Business Reporting Language) är en global informationsstandard avsedd för bearbetning och utbyte av information av ekonomisk redovisningskaraktär. Standarden möjliggör med hjälp av Internet effektiv elektronisk kommunikation av finansiell information oavsett vilken teknisk plattform aktörerna arbetar med.

Arbetet med XBRL startades under 1998 och det är framförallt AICPA (American Institute of Certified Public Accountants) som står bakom. I dagsläget sker utvecklingen av ett konsortium bestående av cirka 250 företag (bland annat de fyra största revisionsbyråerna), myndigheter och organisationer.

I Sverige har redovisnings- och revisionsbranschen visat intresse för att driva utvecklingen och FAR (Föreningen för auktoriserade revisorer), SRS (Svenska Revisorssamfundet) och SRF (Sveriges Redovisningskonsulters förbund) lanserade i mars 2004 den ideella föreningen XBRL Sweden. Inom den offentliga sektorn har bl.a. Bolagsverket ställt sig positivt till utvecklingen och användandet av XBRL.

Läs mer om XBRL på: www.xbrl.org och www.xbrl.se (det svenska initiativet)

Läs mer om AICPA på: www.aicpa.org

4.3.3. GML

GML (Geography Markup Language) är den mest spridda öppna specifikationen för modellering, transport och lagring av geografisk information. Det är en XML-vokabulär, skriven i XML. Projektet drivs av The Open GIS Consortium (OGC), som består av 258 företag och organisationer från offentlig sektor och universitet.

De huvudsakliga begreppen som används i GML har sitt ursprung i OpenGIS Abstract Specification och ISO 19100 serien.

Läs mer på <http://www.opengis.org/specs/>

4.3.4. HR-XML

HR-XML (Human Resources) är en självständig, ideell organisation engagerad i utveckling och främjande av en standardserie av XML-specifikationer som kan möjliggöra e-business och automatisering av datautbyte inom personalbranschen. Bland annat finns standarder för platsannonser och meritförteckningar i XML. Ansvariga för utvecklingen av detta ramverk är det amerikanska konsortiet HR-XML Consortium. Konsortiet har ett hundratal organisationer som medlemmar, från 22 olika länder. Medlemmarna kan själva föreslå att konsortiet åtar sig en standard-uppgift, förutsatt att minst tre organisationssponsorer finns, samt att vissa andra krav är uppfyllda.

Arbetet med HR-XML syftar till att möjliggöra e-business och automatisering av personalrelaterad informationstransaktion. För att utbytet av dessa data ska kunna ske krävs en överenskommelse mellan de inblandade parterna om hur transaktionen ska utföras. Syftet med konsortiets arbete är att utveckla öppna standarder baserade på XML som underlättar förhandlingen och överenskommelsen mellan de inblandade parterna.

HR-XML konsortiet har framställt ett bibliotek med cirka 75 XML-scheman, som finns att ladda hem på deras webbplats. Dessa scheman definierar dataelement för specifika HR-transaktioner, likaså möjligheter och begränsningar rörande elementens användning.

Ett svenskt initiativ har tagits från Arbetsmarknadsstyrelsen (AMS), som med den amerikanska modellen som grund utarbetar ett format anpassat för den svenska arbetsmarknaden.

Läs mer om HR-XML <http://www.hr-xml.org/>

Läs mer om AMS <http://www.ams.se/>

4.3.5. SCORM

ADL (Advanced Distributed Learning initiative) bildades 1997 av det amerikanska försvarsdepartementet för att skapa en gemensam referensmodell för e-Learning. Den modell (SCORM –Shareable Content Object Reference Model) som tagits fram bygger på anpassningar av existerande specifikationer som AICC, IMS och IEEE. Utvecklingen sker i nära samarbete med dessa organisationer.

SCORM är ett XML-baserat ramverk som används för att definiera och ge tillgång till information om lärandeobjekt, så att de kan bli delade mellan olika learning management systems.

Drivkraften för arbetet med SCORM är att öka utvecklingen av kostnadseffektivt webbaserat lärande. Med detta syftar man bl.a. på att skapa en marknad för högkvalitativa lärmoduler som är lätt sökbara och återanvändbara. Den egna organisationen skall inte behöva nyutveckla för alla utbildningsbehov.

Det övergripande målet med arbetet har varit att sänka kostnaderna för att ta fram effektivt kursmaterial för distribution över Internet. ADL uppskattar att kostnaderna kan sänkas med 50-80 % genom att skapa innehåll som är hållbart (inte behöver uppdateras p.g.a. nya versioner av mjukvara), interoperabelt (kan fungera över en stor bredd av hårdvara, operativsystem, webbläsare), tillgängligt och återanvändbart.

SCORM utgörs av en samling specifikationer och standarder som kan ses som separata ”böcker” samlade till ett gemensamt växande bibliotek. I stort sett alla specifikationer och riktlinjer är tagna från andra organisationer.

I Sverige engagerar sig myndigheter som Myndigheten för Skolutveckling och Centrum för flexibelt lärande i olika projekt kring SCORM bland annat i EU-projekt som CELEBRATE.

För vidare läsning om SCORM se ADL:s webbplats www.adlnet.org

5. Ordlista

5.1. Akronymer

Följande lista beskriver några tekniska akronymer som refererats i detta dokument.

- **CSS:** Cascading Style Sheets - beskrivning av hur dokument skall presenteras på datorskärm, hur de skall skrivas ut, skall översättas till tal etc. De mest kända standarderna är *CSS1* (W3C Recommendation, december 1996) och *CSS2* (W3C Recommendation, maj 1998). Ett subset av *CSS2*, *CSS Mobile Profile*, avsett för apparater som mobiltelefoner, har nått status W3C Candidate Recommendation i oktober 2001. CSS kan användas för såväl HTML-dokument som för XML-dokument. Mer information: <http://www.w3.org/Style/> . CSS har en speciell syntax, inte XML-baserad. Ett XML-baserat alternativ är XSL (se XSL ovan).
- **DOM:** Document Object Model. DOM är framtaget av W3C för att tillhandahålla ett standardiserat, plattform- och programberoende gränssnitt, att användas mot XML-dokument. Med hjälp av detta programmeringsgränssnitt kan man använda olika script för att dynamiskt manipulera HTML och XML-dokument. Mer information: <http://www.w3.org/DOM/>.
- **DTD:** Document Type Definition - beskrivning av hur en viss typ av dokument får representeras, i termer av element, attribut och attributvärden. Den typ av DTD som XML understöder skiljer sig i vissa avseenden från DTD:er i SGML. I samband med att XML Schema standardiserats (se ovan) har behovet av DTD:er i XML minskat.
- **RDF:** Resource Description Framework - ett ramverk för att representera metadata. Det erbjuder interoperabilitet mellan applikationsprogram som utbyter information i XML-format. RDF kan användas i många olika sammanhang, såsom resurs-upptäckt ("resource discovery") vid intelligent sökning, för att beskriva innehåll på enskilda informationsobjekt och relationer mellan objekt samt för att beskriva egenskaper hos samlingar av informationsobjekt. Mer information: <http://www.w3.org/RDF/> .
- **Semantic Web:** arbetsområde i W3C. Är en framtidsvision för intelligent informationshantering, bland annat informationssökning. Mer information: <http://www.w3.org/2001/sw/>

- SOAP (Simple Object Access Protocol (SOAP) bygger på XML. Denna teknik möjliggör anrop mellan olika program på ett löst sammansatt standardsätt, vilket i sin tur gör det möjligt att bygga program som distribueras över Internet. Både XML och SOAP har officiellt stöd av i stort sett alla större programvaruföretag i branschen, inklusive Microsoft, IBM, HP, Sun och Oracle.
- UDDI (Universal Description, Discovery and Integration): en webbaserad distribuerad katalog som kan användas för att lista företag på Internet och gör att de kan hitta varandra, ungefär som de gula och vita sidorna i en traditionell telefonkatalog.
- WSDL (Web Service Description Language): ett språk som stöder upptäckt ("discovery") och beskrivning ("description") av tillgängliga webbtjänster. Sådana tjänstebeskrivningar publiceras av tjänstleverantören, i någon publikt sökbar databas, så att klienter kan hitta lämpliga tjänster (t.ex. via UDDI, "Universal Description, Discovery, and Integration"). När väl klienten har fått WSDL-beskrivningen av tjänsten så kan klienten anropa denna genom att sända ett meddelande till tjänsten (t.ex. till en URL som angivits i tjänstens WSDL-beskrivning). Såväl det sända meddelandet som det svar klienten får tillbaka uttrycks i SOAP, som är ett språk varmed man kan överföra strukturerad information mellan tillämpningsprogram.

5.2. Engelska termer

Följande lista beskriver hur vissa engelska termer har översatts till svenska. Avsikten med denna lista är att minska riskerna för missuppfattningar, beroende på att svensk teknisk vokabulär inte har stabiliserats.

- apparatoberoende = device independence. Förekommer i sammanhang som "apparatoberoende uppmärkning", vilket betyder att information med denna uppmärkning kan presenteras på många olika typer av apparater, som webb-läsare, mobiltelefon, PDA, etc.
- attribut [hos ett element]= attribute [of an element]. Egenskapsvärde hos ett element. T.ex. "postnummer-attribut" i ett "adress-element".
- datatypning = data typing. Specifikation av vilka typer av data som t.ex. ett attribut får ha. T.ex. att attributet svenskt postnummer måste vara ett fem-siffrigt heltal.
- element [i ett dokument] = element [in a document]. Ett objekt som i sig är ett dokument eller är en sammanhängande del av ett dokument.

T.ex. adress-element, beskrivande adress-information med attribut som postnummer, m.m.

- formatmall = style sheet. Se XSL, XSLT och CSS
- presentationsuppmärkning = presentation markup. Den uppmärkning som ett dokument har då det skall presenteras på någon apparat, såsom en webb-läsare. T.ex. att postnummer skall visas i fetstil.
- strukturuppmärkning = structural markup. Den uppmärkning som beskriver den innehållsliga strukturen av information. T.ex. att en faktura har ett fakturanummer, adressinformation, datum, m.m.
- uppmärkning = mark up. Förekommer i sammanhang som "uppmärkningsspråk" = "markup languages" och "strukturuppmärkning".
- webbtjänster = web services. Webbtjänster ger information och tjänster till andra program. Programmen får åtkomst till XML-webbtjänster via standardwebbprotokoll och dataformat som t ex HTTP, XML och SOAP.
- XML-applikation = XML application. Förekommer i sammanhang som "MathML är en XML-applikation", vilket betyder att MathML är ett XML-baserat språk. Skall ej sammanblandas med "applikationsprogram".

5.3. Referenser

Följande webbplatser kan rekommenderas för att söka aktuell och fördjupad information:

<http://www.w3c.org/> - W3C:s officiella webbplats.

<http://www.w3.org/TR/#Recommendations> - Aktuell statuslista för W3C:s Recommendations.

<http://w3c.sics.se> - Svenska W3C-kontorets webbplats. Här återfinns bl.a. svenska översättningar av officiella W3C-dokument.

<http://www.dublincore.org/> - Dublin Core Metadata Initiative's officiella webbplats.

<http://www.ws-i.org/> - Web Services Interoperability officiella webbplats.

<http://www.oasis-open.org/> - OASIS officiella webbplats.

<http://www.xml.org/> - Register sponsrat av OASIS, internationellt konsortium som stöder användning av olika standarder. Innehåller både teknikorierad information och branschorienterad information.

<http://xml.coverpages.org/> - Register sponsrat av OASIS, med information om såväl grundläggande XML-standarder som branschorienterade XML-tillämpningar. Här finns en relativt omfattande samling länkar.

http://dmoz.org/Computers/Data_Formats/Markup_Languages/XML/Applications/ - dmoz Open Directory Project, listar ett urval XML-tillämpningar.

<http://www.xmlsoftware.com/> - Översikt över XML-programvaror, privat initiativ.

<http://www.kb.se/> - Kungliga Biblioteket, har Dublin Core översatt till svenska och ger bra praktiska hänvisningar.

<http://www.xmlakademin.nu/> - basinformation om XML och ett informellt nätverk för informationsstrateger inom svensk offentlig sektor.

<http://www.statskontoret.se/> - Statskontoret rapportserie: Vad är XML-familjen? (2000:30-36).

<http://it-upphandling.statskontoret.se/Uhw/> - se Statskontorets ramavtal för informationsförsörjning 2002 som bland annat erbjuder sökmotorer och XML-produkter och tjänster.

<http://www.24-timmarsmyndigheten.se> - översiktlig information om olika delprojekt och aktiviteter kring utvecklingen av 24-timmarsmyndigheten. Här publiceras också nyheter från e-nämnden (Nämnden för elektronisk förvaltning).

5.4. Kontaktperson

Karl Wessbrandt, Statskontoret, tel. 08-454 47 12
e-post: karl.wessbrandt@statskontoret.se