

XML-documentschema's en e-learning

Samenvatting

Een documentschema legt de structuur van een document vast. Veel documentschema's zijn al beschikbaar. Ze worden ontwikkeld in consortia of als R&D-project. Een goede veranderprocedure is daarbij gewenst. EML, de Education Modelling Language, is een voorbeeld van een documentschema. Dit is als R&D-project ontwikkeld en wordt gebruikt binnen verschillende instellingen voor hoger onderwijs.

XML kan gebruikt worden om ieder type document op een eigen, uniforme wijze te beschrijven, passend bij de aard van het document. Het document kan daartoe geassocieerd worden met een zogenoemd 'documentschema' waarin is vastgelegd welke elementen, attributen en notaties geldig zijn in het document en in welke volgorde ze moeten of mogen voorkomen.

Rob Koper

Met schema's kan de structuur, inhoud en semantiek van XML-documenten worden gedefinieerd (Bray e.a., 1998). Schema's kunnen een vaste structuur bij een document afdwingen, zoals bij formulieren het geval is, maar kunnen ook veel flexibiliteit toelaten in de opbouw van documenten, zoals bijvoorbeeld nodig is bij romans.

De meest gebruikte – en nog steeds de enige standaard – schematechniek voor XML is momenteel de DTD (Document Type Definition). Er wordt hard gewerkt aan nieuwe schematechnieken die een aantal tekortkomingen van de DTD's opvangen en die geen speciaal taaltje vereisen, maar die gewoon in XML zijn gedefinieerd. Een kort voorbeeld. Neem het volgende in XML gestructureerde FAQ-document:

```
<FAQ>
  <Titel>FAQ over XML</Titel>
  <Vraag>Wat is een DTD?</Vraag>
  <Antwoord>Een documentschema gespecificeerd in een speciale taal</Antwoord>
  <Vraag>Zijn er nog andere documentschema's?</Vraag>
  <Antwoord>Ja, de W3C werkt aan een nieuwe schemataal genaamd XML-schema</Antwoord>
</FAQ>
```

Het documentschema dat in DTD-vorm met dit XML-document geassocieerd zou kunnen zijn, is het volgende:

```
<!ELEMENT FAQ (Titel, Beschrijving?, (Vraag,Antwoord)+) >
<!ELEMENT Titel (#PCDATA) >
<!ELEMENT Beschrijving (#PCDATA) >
<!ELEMENT Vraag (#PCDATA) >
<!ELEMENT Antwoord (#PCDATA) >
```

Een vraagteken betekent dat het element optioneel is, een plusje betekent dat het element of de groep een of meer keer mag voorkomen. #PCDATA betekent dat een regel tekst verwacht wordt (in Unicode).

In gewoon Nederlands drukt deze DTD uit dat een FAQ bestaat uit precies een titel, mogelijk gevolgd door precies een beschrijving en dan gevolgd door een of meer paren van vraag en antwoord.

De kracht van documentschema's komt voort uit het feit dat het mogelijk is om XML-documenten te valideren op zo'n schema. In het voorbeeld voldoet het document aan het schema. Als de titel zou ontbreken, voldoet het niet. Deze validatie kan geautomatiseerd plaatsvinden. Op deze manier kan een goede structurering van documenten gewaarborgd worden: een brief wordt een brief, een gedicht een gedicht. Dit is uitermate nuttig voor verschillende doeleinden, zoals voor automatische bewerking, standaardisering en conservatie. Door documentschema's te gebruiken kunnen auteurs geholpen worden om documenten op een consistente manier te schrijven of samen te stellen; kunnen documentverzamelingen van eenzelfde type worden opgebouwd, systematisch worden doorzocht en bewerkt en kan uitwisseling van gelijke documenttypen tussen applicaties worden bevorderd. Iemand die een of meer van deze wensen of problemen heeft, doet er goed aan XML te gebruiken in combinatie met een geschikt documentschema.

Enkele jaren geleden kregen ook wij de behoefte om te werken met XML om e-learning (leren via internet) op een effectieve en efficiënte manier te realiseren. Na een grondig vooronderzoek is uiteindelijk besloten zelf een documentschema te ontwikkelen. Dat schema staat nu bekend als EML (Educational Modelling Language, zie: eml.ou.nl). Hierna wordt stilgestaan bij de problematiek van het ontwikkelen en gebruiken van documentschema's, waarbij EML als voorbeeld dient.

Ontwikkelen van schema's

Ieder die besluit om XML te gebruiken staat voor de vraag welk schema daarbij moet worden gebruikt. Het is weliswaar mogelijk om XML te gebruiken zonder schema, maar dan vervallen veel van de voordelen. Als men kiest voor het gebruik van een schema, dan is de eerste stap om te kijken of er al een geschikt schema voorhanden is. Veel organisaties die schema's ontwikkelen stellen deze publiek beschikbaar via het internet. Een goede start is het zoeken bij Oasis (www.oasis-open.org/) of sites als [schema.net](http://www.schema.net/) (www.schema.net/). Meer algemene schema's zijn te vinden bij W3C (w3c.org). In vele vakgebieden zijn er specifieke organisaties die zich bezighouden met het ontwikkelen en testen van schema's (voor e-learning is dat bijvoorbeeld IMS: imsproject.org, IEEE LTSC LOM: ltsc.ieee.org/, CEN/ISSS: www.cenorm.be/iss/Workshop/lt/), ADL: www.adlnet.org/). Zijn er geen geschikte schema's voorhanden, dan zal men deze moeten (laten) ontwikkelen. Daarbij kan men twee wegen bewandelen: een consortiumbenadering of een R&D-benadering. Bij de consortiumbenadering probeert men zo veel mogelijk partijen bijeen te brengen die behoefte hebben aan of belangen hebben in de betreffende documentschema's. Gezamenlijk ontwikkelt men dan volgens bepaalde procedures de benodigde schema's. Het IMS-consortium volgt bijvoorbeeld zo'n werkwijze. Er zijn vele consortia in allerhande vakgebieden actief. Het voordeel van deze benadering is dat de schema's die worden opgeleverd een breed draagvlak hebben. Dat is nodig voor eventuele latere (ISO) standaardisering. Het nadeel is dat dit proces lang duurt en vaak schema's oplevert die nogal wat compromissen bevatten die strikt genomen niet nodig zijn of anders zouden worden gemodelleerd als er geen heilige huisjes werden beschermd.

De R&D-benadering kent deze nadelen niet. In deze benadering, die de Open Universiteit gekozen heeft, wordt een schema zorgvuldig ontworpen en getoetst, uitgaande van de eisen die voortkomen uit het op te lossen probleem en van de kennis die door onderzoek en de praktijk is opgebouwd. Het nadeel van deze benadering is dat het draagvlak niet vooraf gegeven is, maar later zal moeten worden verworven. Deze benadering is dan ook alleen geschikt voor problemen die niet primair gericht zijn

op de uitwisseling van documenten, maar die meer zijn gericht op de betekenisvolle structurering van documenten om intrinsieke problemen in het veld op te lossen. Zo wilden wij problemen oplossen die te maken hadden met het verbeteren van leerprocessen. De vraag is of het mogelijk is om leren flexibeler, effectiever en persoonlijker te maken. Als het ontwikkelde schema later navolgers krijgt, ontstaat daarbij de mogelijkheid om documenten uit te wisselen tussen de gebruikers.

Voor het ontwikkelen van schema's kan een reguliere ontwikkelmethodiek worden gevolgd, bestaande uit: analyse, ontwerp, ontwikkeling, implementatie en evaluatie. In onze ervaring is het bij schemaontwikkeling specifiek van belang om steeds naar het juiste abstractieniveau van de modellering te zoeken. Als het schema te concreet is, dan wordt het veel te uitgebreid om alle voorkomende gevallen te kunnen modelleren. Ieder nieuw geval leidt ook weer tot nieuwe aanpassingen aan het schema. Als het schema te abstract wordt, dan zegt het op den duur niets meer over de betekenis van het document hetgeen het schema overbodig zou maken. In ons geval zijn we begonnen met het inventariseren van alle mogelijke documenttypen in e-learning. Voorbeelden zijn: studietaken, probleembeschrijvingen, toetsen, studieteksten en cases. In de eerste modelleringslag is eerst voor alles een specifiek schema gemaakt. Daarna is abstractie uitgevoerd op al deze modellen, waardoor het aantal elementen en structuren zeer sterk werd gereduceerd en de toepasbaarheid voor verschillende situaties toenam.

Naast het zoeken naar het juiste abstractieniveau is het van belang steeds prototypisch te bekijken of automatische verwerking mogelijk en handig is. Niet alle schema's zijn namelijk even gemakkelijk te verwerken. Uiteraard zijn ook tests met gebruikers aan de invoer- en aan de uitvoerkant noodzakelijk. Bij het ontwikkelen van schema's is het nuttig om waar mogelijk aan te sluiten bij bestaande schema's. Zo bestaan er schema's voor tabellen, formules en metadata. Deze kunnen als modules worden gebruikt in het nieuw te ontwikkelen schema.

Nog een waarschuwing: het ontwikkelen van een documentschema is geen sinecure! Het proces duurt al snel enkele jaren. Doet men het niet zorgvuldig, dan ontstaat er gedurende het gebruik veel behoefte aan aanpassing waardoor de compatibiliteit van de reeds beschikbare XML-documenten in het geding is. Daarmee vervalt een belangrijk voordeel van het gebruik van XML. Ga dus pas tot grootschalig gebruik van een schema over nadat dit uitvoerig is getest en zo stabiel mogelijk is. Let er daarbij op dat vele specificatiecommissies weliswaar schema's vaststellen, maar deze niet testen. Dat laten ze vaak aan de industrie over. Richt ook een goede veranderingsprocedure in, waarbij gegarandeerd wordt dat alleen in het uiterste geval wijzigingen worden doorgevoerd die de compatibiliteit van bestaande materialen in gevaar brengt. Dit betekent dus niet dat schema's niet kunnen of moeten wijzigen: vele wijzigingen kunnen worden doorgevoerd die de compatibiliteit niet bedreigen, mits de basis goed ontwikkeld is.

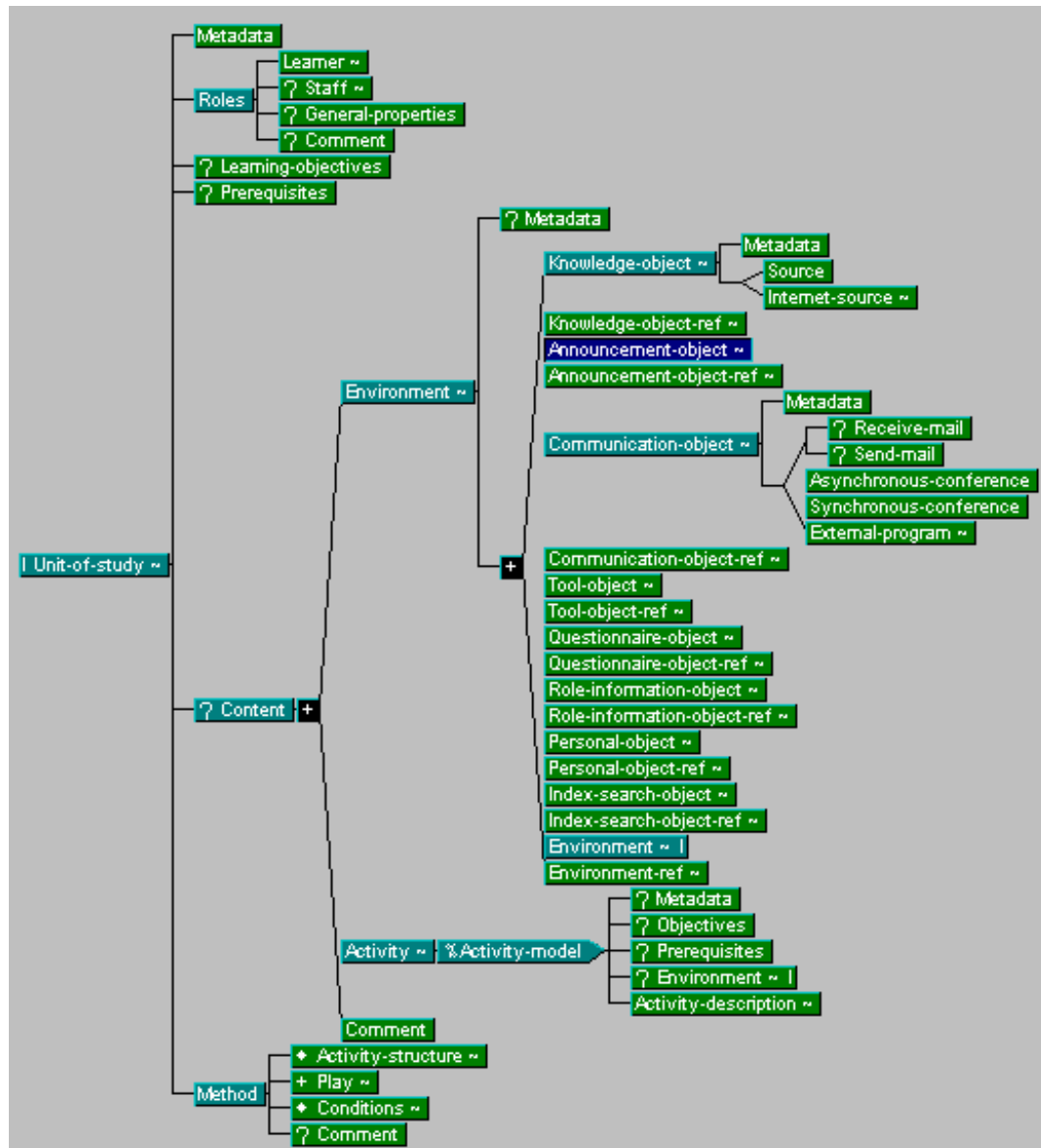
EML, een schema voor e-learning

Wat hebben wij eigenlijk ontwikkeld, ofwel wat is EML? Het doel bij het maken van EML was het creëren van een serie documentschema's waarmee alle in onderwijs en opleiding voorkomende documenttypen kunnen worden beschreven en kunnen worden samengebracht tot een samenhangend geheel. Dat wordt een 'studie-eenheid' genoemd. Voorbeelden van studie-eenheden zijn: cursus, module, curriculum, les, blok, workshop, practicum. Voorbeelden van componenten met een eigen schema die worden geïntegreerd in een studie-eenheid zijn: leeractiviteiten, kennisobjecten, communicatieobjecten, instrumenten, toetsobjecten en zoekobjecten.

De term documenttype wordt hier breed opgevat, het bevat alle (referenties naar) multimediale entiteiten en communicatiefaciliteiten. Eigenlijk wordt alles als document opgevat dat *vooraf* gepland en gemaakt wordt, inclusief de procesbeschrijvingen die in het onderwijs plaatsvinden. Echter, alles wat tijdens de uitvoering van het onderwijs in real time ontstaat wordt niet meegemodelleerd.

Uitgangspunt van EML was het realiseren van onderwijsvernieuwing waarbij ICT een belangrijke ondersteunende rol speelt. Trefwoorden waren: personalisatie, moderne onderwijsvormen (o.a. competentiegericht onderwijs, probleemgestuurd onderwijs), uitlevering in verschillende settings (afstandsonderwijs, contactonderwijs, integratie werken/leren), uitleveren via verschillende media (o.a. internet, papier), ondersteunen van nieuwe vormen van assessment (portfolio assessment). Zie voor meer informatie hierover Koper (2000, 2001).

EML bestaat uiteindelijk uit zo'n 380 elementen. In figuur 1 zijn de hoogste niveaustucturen afgebeeld.



[Figuur 1] De structuur van EML op hoofdniveau

Het schema is modulaair van opzet. Een aantal modules kan naar wens worden omgewisseld door andere modules, omdat er verschillende alternatieve schema's voor zijn. Voor tabellen kan bijvoorbeeld het CALS- of het HTML-tabellenschema worden gebruikt. Het aantal verplichte elementen is vrij beperkt. Een minimum studie-eenheid in EML heeft de structuur zoals in figuur 2.

```

<Unit-of-study>
  <Metadata>
    <Title>Cursus Basgitaar</Title>
  </Metadata>
  <Roles>
    <Learner Id="student"/>
  </Roles>
  <Content>
    <Activity Id="studeren">
      <Environment>
        <Knowledge-object Id="boek">
          <Metadata>
            <Title>Boek Basgitaar</Title>
          </Metadata>
          <Internet-source Url= "http://boek.nl/bas.html"/>
        </Environment>
        <Activity-description>
          <What>Bestudeer <EML-ref Id-ref="boek">het boek</EML-ref></What>
          <Completed><User-choice/></Completed>
        </Activity-description>
      </Activity>
    </Content>
    <Method>
      <Play>
        <Role-ref Id-ref="student"/><Activity-ref Id-ref="studeren"/>
      </Play>
    </Method>
  </Unit-of-study>

```

[Figuur 2.] De minimumstructuur voor een studie-eenheid in EML

In het voorbeeld is de inhoud van het kennisobject (knowledge object) op internet te vinden. Het kan ook geheel in EML gespecificeerd worden. Het is ook mogelijk om de samenstellende delen los te gebruiken zonder dat ze geïntegreerd worden tot een studie-eenheid. Dat is vooral van belang voor kennisobjecten (boeken, artikelen, interactieve applicaties, e.d.), maar ook voor studietaken. Deze kunnen afzonderlijk gemaakt en beheerd worden en in verschillende studie-eenheden worden gebruikt.

U kunt de gehele DTD downloaden via eml.ou.nl.

Gebruik van de schema's

Uiteraard zijn voor het maken, beheren en gebruiken van XML-documenten verschillende tools nodig. Daarbij bestaat principieel steeds een keuze tussen het gebruik van generieke XML-tools (zie bijvoorbeeld www.xmlsoftware.com of xml.com), zoals XML-editors, contentmanagementsystemen, conversie- en publicatiesystemen of het programmeren van specifieke oplossingen. Beide hebben zo hun voor en tegen. Wij hebben gekozen voor de integratie van een verzameling generieke tools om EML-documenten te kunnen bewerken en publiceren (zie edubox.nl). Het nadeel is wel dat vooral de generieke invoermiddelen nog niet erg gebruiksvriendelijk zijn. Daarnaast ontbreekt nog een aantal zaken (met name aan de publicatiezijde), waardoor nog met niet-standaard oplossingen moet worden gewerkt.

We hebben de afgelopen tijd al behoorlijk wat ervaring kunnen opdoen in het gebruik van EML en de tools om EML te verwerken (Edubox). Vele honderden studie-uren zijn nu gemaakt en gebruikt. De ervaringen zijn dermate positief dat de Open Universiteit besloten heeft XML (met het EML-schema) in te voeren als de standaard manier van werken. De eerste cursussen zijn al ontwikkeld. Ook de Unisa, de Open Universiteit van Zuid-Afrika, kiest deze route. De Hoge Hotelschool in Maastricht heeft al een flink aantal modules in EML ontwikkeld en gebruikt. Daarnaast is XML gekozen als de standaard voor het ontwikkelen van cursussen in het nieuwe Consortium Digitale Universiteit (een consortium van vier universiteiten en vijf hbo-instellingen).

Onderdeel van het huidige werk rond EML is de actieve participatie in diverse internationale groepen die zich met standaarden rond e-learning bezighouden, zoals Prometheus (www.prometeus.org), IMS en de CEN/ISSS.

Toch zijn de ontwikkelingen rond XML en EML nog nieuw en zal er nog veel in de markt moeten gebeuren om het geheel routinematig, gebruiksvriendelijk en soepel te laten werken. De eerste schreden in de goede richting zijn echter gezet.

Literatuur

Bray, T., Paoli, J., & Sperberg-McQueen, C. M., *Extensible Markup Language (XML) 1.0*, World Wide Web Consortium (W3C), 1998 (www.w3.org/TR/REC-xml.html).

Koper, E.J.R., Van verandering naar vernieuwing: onderwijstechnologische grondslagen van elektronische leeromgevingen, Open Universiteit Nederland, 2000 (verkrijgbaar op eml.ou.nl/introduction/articles.htm).

Koper, E.J.R., Elektronische leeromgevingen: gewaarwordingen en perspectief, *Opleiding en Ontwikkeling* 3, 2001, p. 7-12.

Rob Koper

is Hoogleraar Onderwijstechnologie bij het Onderwijstechnologisch Expertise Centrum (OTEC) van de Open Universiteit Nederland.