

ARTIKEL

auteur Reinier Balt en Guido van der Harst
tijdschrift Automatisering Gids

7-12-2002

Keuze voor XML is een non-issue

Is XML de magische oplossing voor het realiseren van elektronische gegevensuitwisseling? Hier bestaan nogal wat misverstanden over, concluderen Reinier Balt en Guido van der Harst. Zij bespreken er vier. XML is slechts één van de bouwstenen.

De uitwisseling van gegevens tussen toepassingen binnen organisaties is langzamerhand gemeengoed geworden. Verkoopgegevens gaan van backoffice-systemen naar het CRM-pakket, financiële gegevens van de voorraadadministratie naar financiële administratie et cetera. Met de komst van internetstandaarden en de druk op efficiency zien we dat ook de elektronische uitwisseling van gegevens met systemen van buiten de organisatie aan populariteit wint. Steeds meer organisaties krijgen te maken met afnemers en leveranciers die elektronisch gegevens willen uitwisselen.

Als we de stroom aan publicaties in de vakpers mogen geloven dan lijkt XML de magische oplossing te zijn voor het realiseren van elektronische gegevensuitwisseling (zowel tussen interne toepassingen als tussen interne en externe toepassingen). Maar is dat wel zo? Is de interoperabiliteit van de organisatie bij het gebruik van XML in kunnen en kruiken? Op dit punt bestaan nogal wat misverstanden. XML is slechts één van de bouwstenen voor elektronische gegevensuitwisseling. De keuze voor XML is in feite een non-issue.

Misverstand 1

XML is een programmeertaal met ongekeerde mogelijkheden

Een programmeertaal is een taal waarmee we de computer logica kunnen laten uitvoeren in de vorm van een serie van instructies. XML is een taal om het formaat van gegevensberichten te definiëren. Het formaat bepaalt uit welke elementen een bericht kan zijn opgebouwd, bijvoorbeeld de elementen NAW-gegevens, bestellingen, het afleveradres, betalingsgegevens, et cetera. Ook bepaalt het formaat hoe deze gegevens zijn ingedeeld, bijvoorbeeld dat een element 'aantal' in een bestelregel een geheel getal is of dat het eindbedrag van een factuur een gebroken getal is met twee cijfers na de komma.

Kortom: XML is helemaal geen programmeertaal. Wel zien we een groeiende ondersteuning in verschillende programmeertalen om XML-berichten te kunnen interpreteren en manipuleren. Zo levert Microsoft tegenwoordig standaard een bibliotheek met specifieke XML-functies bij het Windowsplatform (de msxml3-DLL in de system-directory).

Misverstand 2

XML is het middel voor elektronische gegevensuitwisseling

Doordat XML breed geaccepteerd is als standaard voor het formaat van elektronische berichten, maakt zij elektronische gegevensuitwisseling een stuk eenvoudiger. Dat is zeker waar. Echter, XML lost het grote probleem rondom gegevensuitwisseling niet op. Om iets zinvol met de uitgewisselde gegevens te kunnen doen, zijn er immers ook afspraken nodig over de betekenis van de gegevens. De betekenis, de semantiek, bepaalt hoe de toepassingen de gegevens in een bericht moeten interpreteren. Bijvoorbeeld dat de NAW-gegevens tevens het factuuradres vormen. Of dat het eindbedrag van een factuur in euro's is en niet in dollars. Of dat code '6' betekent dat de klant niet kredietwaardig is.

XML geeft enkel aan hoe het formaat van een bericht eruit moet zien. XML doet geen uitspraak over de betekenis van de elementen in een bericht. Ergo: een keuze voor XML is onvoldoende basis om elektronisch gegevens uit te wisselen. Organisaties moeten ook nog invulling geven aan de semantische afspraken. Net als in het pre-XML tijdperk kunnen organisaties dit slechts op één manier oplossen: door het maken van onderlinge afspraken. XML verandert daar weinig aan. Dit kunnen afspraken zijn tussen verschillende organisaties maar ook afspraken binnen de eigen organisatie. In beide gevallen kunnen organisaties twee mogelijke routes bewandelen:

- zelf definiëren van een semantiek;
- gebruiken van een bestaande semantiek.

Het zelf definiëren van de semantiek lijkt een vruchtbare route. De organisatie is zelf 'in charge' van wat er gebeurt, is niet afhankelijk van derden en kan de semantiek perfect laten aansluiten op haar specifieke situatie. De keerzijde is dat het definiëren van een eigen semantiek een niet te onderschatten inspanning vergt, nog los van het onderhoud dat in de gebruiksfase volgt. Bovendien profiteert de organisatie niet van de kennis die beschikbaar is in bestaande semantieken.

Deze nadelen vervallen wanneer een organisatie kiest voor het gebruik van een bestaande semantiek. Toch klinkt ook dit makkelijker dan het is. Er zijn immers honderden semantische raamwerken op basis van XML beschikbaar (zie kader). Welke sluit nu het beste aan? Hiervoor is een goed overzicht in de beschikbare semantieken vereist. Op basis van criteria als dekkingsgraad, gebruik door organisaties, ondersteuning door leveranciers, beschikbaarheid van kennis en de continuïteit van de beherende partij, kan een organisatie een weloverwogen keuze maken.

Misverstand 3

XML is de oplossing voor EAI

Enterprise Application Integration (EAI) is het integreren van verschillende toepassingen binnen een organisatie. Vaak wordt gesteld dat met de komst van XML alle integratievraagstukken zijn opgelost. Zeker geldt, dat XML door haar brede acceptatie de integratievraagstukken vereenvoudigt. Maar ook hier hebben de voordelen alleen betrekking op het formaat van de berichten. Een openstaand vraagstuk blijft: volgens welk *protocol* wisselen toepassingen nu berichten uit?

Op deze vraag moeten organisaties los van de keuze voor XML een antwoord vinden om EAI succesvol te maken. Een belangrijke ontwikkeling op het vlak van uitwisselingsprotocollen is de opkomst van de standaarden die bekend zijn onder de noemer web services. Deze ontwikkeling is zo bijzonder omdat het nu voor het eerst in de geschiedenis van de automatisering mogelijk is om toepassingen met elkaar te laten samenwerken die geschreven zijn in verschillende

programmeertalen en die draaien op verschillende systeemplatformen. Dit alles op basis van open standaarden waar veel partijen zich aan conformeren.

Een andere openstaande vraag op het vlak van EAI is hoe een organisatie haar bestaande (legacy) systemen geschikt kan maken voor integratie. Deze systemen zijn vaak minder open. XML helpt daar niet bij. Wat wel helpt, zijn 'integration brokers' (zie kader). Een integration broker is een afzonderlijke toepassing die gangbare standaarden als XML en web services kan emuleren. Dit is het nabootsen van de gangbare standaarden naar de buitenwereld. De organisatie legt als het ware een schil om de bestaande toepassingen heen. Het voordeel van emulatie is dat het aansluiten op de gangbare standaarden slechts een beperkte invloed heeft op de bestaande systemen. Eveneens heeft de buitenwereld door de schil geen hinder van wijzigingen in de bestaande systemen.

Misverstand 4

XML-databases zullen de plaats innemen van relationele databases

De structuur van een XML-bericht is een hiërarchische boomstructuur. Dit is natuurlijk een winst ten opzicht van het Edifact-formaat omdat dit formaat slechts een sequentie van gegevens ondersteunt. Maar goedbeschouwd is het wel een achteruitgang ten opzichte van relationele databases. Niet voor niets hebben we afscheid genomen van de hiërarchische databases uit de jaren zestig. De objecten in de werkelijke wereld laten zich niet altijd hiërarchisch indelen, maar zijn veelal via allerlei kruisverbanden aan elkaar gerelateerd. Relationele databases hebben zich bewezen als de meest praktische technologie om gestructureerde gegevens vast te leggen. Het vastleggen van gegevens in XML-formaat biedt hier geen toegevoegde waarde.

Wel is het zo dat vanwege de brede ondersteuning van XML het bijvoorbeeld handig is om het resultaat van een query op een relationele database in XML-formaat te kunnen ontvangen. Alle bekende relationele databases bieden hier inmiddels ondersteuning voor.

Technici van bedrijven als Microsoft en Sun hebben de XML-standaard ontwikkeld onder de vlag van het World Wide Web Consortium (W3C). Deze organisatie heeft zichzelf verantwoordelijk gemaakt voor het volledig ontwikkelen van het world-wide web. De reden voor het opstellen van de XML-standaard is dat er geen goede taal beschikbaar was om informatie gestructureerd uit te wisselen. Zo richt de HTML-standaard zich teveel op de opmaak van webpagina's en is een standaard als Edifact te beperkt en te weinig flexibel. Tot slot is de aan HTML ten grondslag liggende SGML-standaard veel te complex. Tot op de dag van vandaag zijn er geen implementaties die de volledige SGML-standaard ondersteunen.

Op 10 februari 1998 heeft W3C XML tot aanbevolen standaard verheven. Vanaf dat moment is het snel gegaan. Nagenoeg alle softwareleveranciers die oplossingen bieden op het vlak van gegevensuitwisseling, bieden ondersteuning voor XML. Ook organisaties passen XML steeds vaker toe. We kunnen rustig stellen dat zij is uitgegroeid tot dé standaard voor het definiëren van het formaat van elektronische berichten. Een organisatie moet wel een heel goede reden hebben om niet voor deze standaard te kiezen als het gaat om het formaat waarin de gegevens worden uitgewisseld.

Reinier Balt (balt@vka.nl) en Guido van der Harst (harst@vka.nl) zijn respectievelijk als consultant en senior consultant werkzaam bij het onafhankelijke ICT-adviesbureau Verdonck, Klooster & Associates te Zoetermeer.

[Kader 1]

Semantische raamwerken

Er is in het verleden reeds hard gewerkt aan het maken van standaarden voor de semantiek. Denk bijvoorbeeld aan de Edifact-standaard van UN/Cefact of de Corba Domain Services. Ook op het vlak van XML zijn er verschillende initiatieven om semantische raamwerken op te zetten. Deze raamwerken zijn vaak gericht op een specifiek domein.

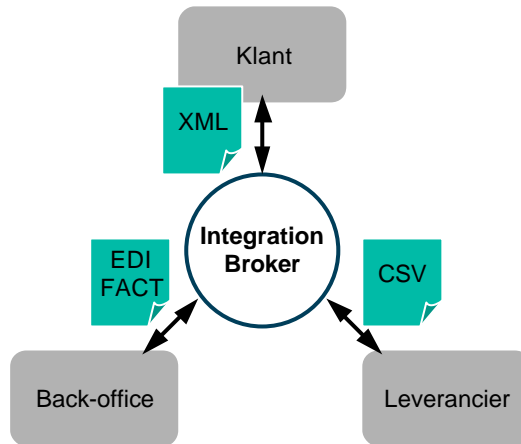
Oasis beheert een groot aantal van deze standaarden waaronder klantinformatie, documenten en rapporten en rechtspraak. Oasis werkt samen met UN/Cefact aan de ambitieuze ebXML-standaard voor elektronisch zakendoen. Het ambitieniveau van deze standaard is zo hoog dat de vraag is of deze niet door leveranciersstandaarden wordt ingehaald. Andere semantische raamwerken zijn eSpeak van HP, BizTalk van Microsoft, RosettaNet voor de technische industrie en HL7 voor de gezondheidszorg.

Een interessante ontwikkeling in deze is het 'semantic web' van W3C. Dit moet de opvolger worden van het world-wide web zoals we dat nu kennen. Met de semantic web standaarden kunnen aan bronnen op Internet (bijvoorbeeld webpagina's) annotaties worden gekoppeld. Deze annotaties doen een uitspraak over de inhoudelijke betekenis van de bron. Naast webpagina's kan je ook aan andere bronnen denken, zoals mp3-bestanden, databases, software, webservices, films, et cetera. Het fraaie van de semantic webstandaarden is dat betekenissen naar elkaar kunnen verwijzen in de zin van 'deze term betekent op deze webpagina hetzelfde als die term op die pagina'. Vandaar ook de naam 'semantic web'. De standaarden zijn momenteel in ontwikkeling. Als het semantic web van de grond komt, kunnen bronnen aan elkaar gerelateerd worden op een wijze die vooraf niet was voorzien. De mogelijkheden zijn onbeperkt.

[Kader 2]

Integration broker

Een integration broker is een end-to-end platform dat het mogelijk maakt bedrijfsprocessen volledig te automatiseren, zelfs over de grenzen van het bedrijf heen. Integration brokers maken hiervoor gebruik van traditionele middlewaretoeepassingen zoals 'request brokers', 'messaging middleware', 'workflow' en transactiemonitoren.



Figuur De integration broker

De belangrijke kracht van een integration broker is de mogelijkheid berichten te vertalen van het ene formaat naar het andere. Zo kan het platform een binnenkomende bericht met een bestelling in XML-formaat, volledig transparant omzetten naar Edifact-formaat en routeren naar de backoffice-applicatie. Deze applicatie stuurt misschien een Edifact-bericht naar een leverancier. De integration broker weet dat de leverancier alleen bestellingen accepteert via het CSV-formaat. Zonder aanpassingen aan de back-office-applicatie kan het platform het Edifact-bericht omzetten in een CSV-bestand en dit versturen aan de leverancier. Voor de buitenwereld lijkt het dus alsof gecommuniceerd wordt met behulp van XML, terwijl in de back office gewoon gewerkt wordt met het traditionele Edifact.

Deze omzetting van het ene formaat naar het andere gaat goed zolang de semantiek van het bericht dat binnenkomt overeenkomt met de semantiek die de back office verwacht. Voor het succesvol toepassen van integration brokers is het maken van semantische afspraken weer essentieel.