

XML – Ursprung und Funktionsweise

Seit der Einführung des Electronic Business gelten neue Regeln in der Informationstechnologie. Internet-basierte Anwendungen mit einer unbegrenzten Anzahl von Benutzern benötigen skalierbare Systeme, die Spitzenleistung und hohe Transaktionsraten bieten. Die dabei verarbeiteten Informationen sind heterogener Natur und können sich irgendwo auf der Welt befinden. Ermöglicht wurde dieses neue Paradigma der Informationsverarbeitung durch offene Kommunikationsstandards, wie HTTP, TCP/IP, HTML und XML.

Man geht heute davon aus, daß sich die momentane Zahl von 150 Millionen per Internet vernetzten Computern in den nächsten fünf bis acht Jahren (vielleicht auch schon früher) auf über eine Milliarde weltweit erhöhen wird. Mit dieser rasanten Zunahme an Rechnern, die für Electronic Business ausgerüstet sind, wird die Zahl der Transaktionen, die über das Internet abgewickelt werden, exponentiell zunehmen.

Ein anderer Aspekt, der zu einer Zunahme des Netzverkehrs führen wird, ist die große Zahl heterogener Datenbanksysteme, auf die im Zuge immer offener Transaktionen zugegriffen werden kann. Mit großer Wahrscheinlichkeit werden die heute installierten Backend-Systeme Datenbankabfragen und Transaktionen nicht schnell genug bearbeiten können. Was wiederum an der unendlich großen Zahl neuer Datentypen (BLOBs wie z.B. Video oder Audio) und unterschiedlicher Datenformate liegt, die im Zuge einer komplexen Transaktion zu verarbeiten sind. Die Konsequenz: Unannehmbare Wartezeiten - im schlimmsten Fall auch Transaktionsabbrüche, wodurch sich das Electronic Business selbst zum Erliegen brächte.

Nach Angaben renommierter Marktforschungsinstitute scheitern derzeit drei von vier Electronic Business Projekten. Häufig liegen die Ursachen dafür bei der Integration der Systemplattformen innerhalb der Lieferketten. Bei genauerem Hinsehen zeigt sich nämlich: Viele sogenannte Supply Chain Ketten davon leben, dass die einzelnen Glieder dieser Ketten durch manuellen Übertrag der Aufträge zusammengehalten werden.

Die Zukunft des Electronic Business heißt XML. Im Gegensatz zu der weitverbreiteten Meinung liegt das große Potenzial dieser lingua franca des Internets aber nicht in der Informationsrecherche auf dem World Wide Web, sondern beim Informations- und Datenaustausch über Unternehmensgrenzen hinweg. Wenn es zukünftig um die Integration der Geschäftsprozesse unterschiedlicher Unternehmen geht, ist XML erste Wahl. Das geht hin bis zur direkten Integration von Unternehmenssoftware, denn alle wichtigen Hersteller von ERP-Systemen unterstützen bereits heute dieses universelle Datenformat.

Relationale Datenbanksysteme können nicht alle Anforderungen des Electronic Business erfüllen, da sie Daten und ihre Bedeutung getrennt verwalten. Nachträglich hinzugefügte XML-Schnittstellen sind nur Notlösungen, die Business-Anforderungen in punkto Datenintegrität und Performance nicht gerecht werden. Auch das gängige Web-Format HTML bietet keine Lösung, da es nur die richtige Darstellung, nicht aber die korrekte Verarbeitung von Daten sicherstellt.

Der Durchbruch von XML als "lingua galactica", als Standardsprache für Geschäftstransaktionen über das Internet, zeichnet sich als Lösung für dieses Problem ab: Wenn Daten in nativem XML-Format gespeichert werden, sind zeitraubende Konvertierungen von einem Datentyp in einen anderen nicht mehr notwendig. XML wird Electronic-Business-Transaktionen zu der notwendigen Geschwindigkeit verhelfen, und unnötiger Netzverkehr wird dadurch spürbar abnehmen. Der schnelle Zugriff auf beliebige Datentypen wird vor diesem Hintergrund zu einem Faktor, der über den Erfolg oder das Scheitern der Electronic-Business-Aktivitäten eines Unternehmens entscheiden kann. Alle Unternehmen egal, ob groß oder klein, stehen im Internet im Wettbewerb. Mitbewerber sind heute nur noch einen Mouseklick entfernt, wodurch sich viele monopolistisch geprägten Märkte in hart umkämpfte virtuelle Märkte verwandeln werden. Und nur wer die Erwartungen seiner Kunden erfüllt, wird überleben. Der frühzeitige und fachkundige Einrichtung von Geschäftsabläufen, die auf Electronic-Business-Transaktionen abgestimmt sind, können hier zu einem entscheidenden Wettbewerbsfaktor werden. Je höher der Automatisierungsgrad, desto geringer die Transaktions- und Lagerhaltungskosten. Und schnelles automatisiertes Reagieren auf Kundenprobleme fördert die Kundenzufriedenheit. Unternehmen, die es schaffen, diese Prozesse erfolgreich aufzubauen, stehen hervorragende Wachstumschancen bevor.

XML ist ein vom World Wide Web Consortium (W3C) verabschiedeter Industriestandard, der von allen führenden Softwareanbietern unterstützt wird. Die Einführung von XML markiert den Beginn einer neuen Ära in der Geschichte des Internet: Die Transformation des Web von einem Informationsnetzwerk zu einem globalen Wissensspeicher und einer globalen Rechenplattform. Und dies aus gutem Grund:

- XML ist für den Einsatz im Web optimiert, es ist einfach und leicht verständlich, vor allem ist es aber über die Dokumentdefinition DTD (Document Type Definition) gut anpaßbar. Eine DTD beschreibt in einer für Menschen und Rechner lesbaren Form die Typisierung von Dokumenten, muß also branchen- bzw. anwendungsbezogen erstellt werden. Damit lassen sich spezifische Bedürfnisse umsetzen, ohne den allgemeinen Standard zu verletzen.
 - XML-Tags beschreiben den Inhalt des Dokumentes, nicht dessen Darstellung. Das Motto von HTML ist: "Ich weiß, wie es aussieht." Das Motto von XML dagegen ist: "Ich weiß, was es bedeutet – sage mir, wie es aussehen soll." In herkömmlichen Datenbanken benötigen Dateneinträge Schemata, die vom Datenbankadministrator gepflegt werden. XML-Dokumente können ohne diese Definitionen gespeichert werden, da sie Metadaten in Form von Tags und Attributen enthalten. Jeder XML-Tag kann eine unbegrenzte Anzahl an Attributen, wie Datum oder Version, enthalten. Tags, Attribute und Elementstrukturen liefern Kontextinformationen, die die Interpretation der Dokumentinhalte ermöglichen. Hier ergeben sich völlig neue Einsatzmöglichkeiten für hocheffiziente Suchmaschinen, Data Mining, maschinengestützte Dokumentenübersetzung usw. Gegenüber HTML oder reinem Text, wo Kontextinformationen schwierig oder gar nicht auszuwerten sind, ist dies ein entscheidender Vorteil.
 - Mit Hilfe der XSL Style Sheets kann das Aussehen eines XML-Dokuments (oder einer ganzen Webseite) verändert werden, ohne daß der Inhalt des Dokuments berührt wird. Verschiedene Darstellungen desselben Inhalts sind ohne weiteres möglich. Ein XML-fähiger Browser kann mit Hilfe von XML Style Sheets XML-Code sofort auf dem Bildschirm darstellen, ohne daß vorher aufwendige Masken generiert werden müssen.
 - XML unterstützt mehrsprachige Dokumente und Unicode, was für globalisierte Anwendungen von Bedeutung ist. Dies ist für Electronic Business-Anwendungen von elementarer Bedeutung.
 - Dank der Baumstruktur von XML-Dokumenten können Dokumente effizient Element für Element verglichen und gruppiert werden.
-

- XML-Dokumente können sämtliche Datentypen enthalten – von Multimedia-Daten (Bild, Ton, Video) bis zu aktiven Komponenten (Java-Applets, ActiveX). Die Erweiterbarkeit und Flexibilität von XML erlaubt die Beschreibung von Daten, die in einer breiten Palette an heterogenen Anwendungen vorhanden sind - angefangen von der Beschreibung einer Kollektion von Webseiten bis hin zu Datensätzen.
- Das Abbilden bestehender Datenstrukturen wie Dateisysteme oder relationale Datenbanken auf XML ist einfach. XML deckt alle bestehenden Datenstrukturen ab und unterstützt verschiedene Datenformate. XML-Dokumente können aus verschachtelten Elementen bestehen, die über mehrere entfernte Server verteilt sind. XML ist derzeit das am weitesten entwickelte Format für verteilte Daten, mit dem das Web zur einer riesigen Fundgrube für XML-Informationen werden kann. Da XML-basierte Daten selbstbeschreibend sind, können Daten ausgetauscht und verarbeitet werden, ohne daß eine integrierte Beschreibung der eingehenden Daten vorhanden ist. Heutzutage ist die Durchführung von Suchvorgängen in mehreren, inkompatiblen Datenbanken praktisch unmöglich. XML ermöglicht die vereinfachte Kombination von strukturierten Daten aus verschiedenen Quellen. Anders ausgedrückt: Durch XML wird das gesamte World Wide Web zu einer einzigen, allumfassenden Datenbank.
- XML bietet die Möglichkeit, vom Browser aus direkt auf XML-fähige Server zuzugreifen und Informationen zu recherchieren, die vom XML-Server aus verschiedenen Informationstypen oder auch aus vorhandenen relationalen Daten zusammengesetzt werden. Den gleichen Service kann auch eine Anwendung benutzen. Beide verwenden die URL-Adresse für die Verbindung zum gewünschten XML-Server. Klassische Anwendungen können weiterhin parallel dazu laufen und beispielsweise auf ihre SQL-Daten zugreifen.

In allen Branchen haben sich bereits Unternehmen und Organisationen zusammengefunden, um XML-Anwendungen zu entwickeln. Die einzigartig offene Struktur von XML erlaubt es, je nach Bedarf neue – selbst völlig neuartige – Elemente hinzuzufügen. Damit ist auch jederzeit eine Anpassung an spezifische Branchenvokabulare möglich, unter anderem:

- Commerce XML (cXML): ein neu eingeführter Standard zur Beschreibung von Kataloginhalten und für sichere kommerzielle Transaktionen.
- Open Buying on the Internet (OBI): ein Standard für den Business-to-Buisness-Handel übers Internet, basierend auf HTML, SSL, SET und X.509.
- Open Trading Protocol (OTP): eine Umgebung für den Verkauf an Endverbraucher übers Web einschließlich Regelung von Zahlungsmodalitäten.
- Open Financial Exchange (OFX), das von Intuit Quicken und Microsoft Money benutzte Format für die Kommunikation mit Banken.

XML liegen bereits bewährte standardbasierte Technologien zugrunde, die für das Web optimiert wurden. Das W3C arbeitet derzeit mit führenden Unternehmen einschließlich der Software AG zusammen, um Interoperabilität und Unterstützung für Entwickler, Autoren und Benutzer auf vielen Systemen und Browsern sicherzustellen sowie den XML-Standard weiter zu entwickeln.

Die XML-Initiative besteht aus drei miteinander verbundenen Strategien:

- **XML (Extensible Markup Language).** XML 1.0 ist gegenwärtig eine Empfehlung - die letzte Stufe im W3C-Prüfungsverfahren. Das bedeutet, daß der Standard stabil ist und von Web- und Tools-Entwicklern vollständig angenommen werden kann.
 - **XSL (Extensible Style Language).** XSL wird bei der Transformation von XML-basierten Daten in HTML oder in sonstige Darstellungsformate verwendet. XSL ermöglicht Entwicklern den Aufbau einer Darstellungsstruktur, die sich von der Datenstruktur unterscheidet. Beispielsweise
-

kann XSL bei der Transformation einer XML-basierten Kaufordernummer in eine Liste mit Gliederungspunkten in einer HTML-Ansicht und in eine Fußnote in einer zweiten HTML-Ansicht verwendet werden.

- **XLL (Extensible Linking Language).** XLL ist eine XML-Linksprache, die die Erstellung von Verweisen ähnlich wie in HTML sicherstellt, jedoch durch eine weitaus größere Leistungsfähigkeit gekennzeichnet ist. Zum Beispiel könnten durch den Einsatz von XLL Links multidirektional sein und die Links auf Objektebene statt lediglich auf Seitenebene existieren.

Zusätzlich folgt das XML Object Model (XML-OM) dem W3C-Standard Document Object Model (DOM), um damit einen programmatischen Zugriff auf strukturierte Daten durch Scripting zu ermöglichen, so daß Entwickler mit bzw. auf XML-basierten Daten einheitlich arbeiten können.

Die XML-Technologie steht noch am Anfang, aber schon gibt es auch anspruchsvolle Produkte, die über die anfänglich dominierenden Parser und Browser weit hinausgehen. Allmählich wird sich die IT-Welt bewußt, daß mit XML strukturierte Dokumente mehr sein können als White Papers. Der Einsatzbereich von XML ist nahezu unbegrenzt, aufgrund der Flexibilität von XML bleibt praktisch kein Anwendungsgebiet ausgeschlossen. Funktionierende, integrierte, auf XML aufbauende Lösungen sind heute keine Vision mehr. Für alle diese vielfältigen Anwendungen braucht es neben pfiffigen Konzepten natürlich auch die entsprechenden Maschinen: XML-Server unterschiedlichster Art. Weitere Informationen zu XML finden Sie im Internet unter: www.software.com

Gabriele Müller-Steinhagen
SAG Systemhaus GmbH

14.12.2001
