

# Quark XML Author erweitert Microsoft Word

## Zusammenfassung

In diesem Whitepaper wird die in Microsoft® Word integrierte XML-Unterstützung beschrieben und erklärt, wie Word mit Quark® XML Author zu einem einfach zu verwendenden Editor für XML-Dokumente wird.

## Übersicht

Microsoft hat XML-Funktionen in Office 2003 eingeführt und diese in Office 2007 verbessert. Diese Funktionen haben den Benutzern enorm geholfen, die Vorteile von XML zu nutzen. Microsoft hat die Verwendung von XML für die Softwareentwickler von Drittanbietern ausgelegt, um das Integrieren von Daten in Dokumente und von Dokumenten in Downstream-Prozesse zu ermöglichen sowie andere Produktivitätsverbesserungen für Benutzer bereitzustellen.

Die umgehende XML-Bearbeitung in Word wird von Microsoft jedoch nicht unterstützt. Microsoft empfiehlt den Benutzern, die XML-Dokumente in Word erstellen möchten, die Produkte seiner Partner, beispielsweise Quark XML Author.

Entwickler können die in Word integrierten XML-Funktionen verwenden, um XML-Bearbeitungsfunktionen für einfache Vorgänge bereitzustellen. Um jedoch die Anforderungen einer Dynamic Publishing-Anwendung zu erfüllen, sind umfassende Funktionen für die Bearbeitung von XML-Dokumenten erforderlich, die ausschließlich in Quark XML Author für Microsoft Word verfügbar sind. Quark XML Author basiert auf den XML-Funktionen von Word, um Autoren das Erstellen und Bearbeiten von XML-Dokumenten unter Verwendung der einfachen, intuitiven Benutzeroberfläche von Microsoft Word zu ermöglichen.

## Unterstützung für XML von Microsoft

**Quark XML Author basiert auf den XML-Funktionen von Word, um Autoren das Erstellen und Bearbeiten von XML-Dokumenten unter Verwendung der einfachen, intuitiven Benutzeroberfläche von Microsoft Word zu ermöglichen.**

In diesem Abschnitt sind die XML-Funktionen von Microsoft beschrieben, um Ihnen zu helfen, deren Nutzen für zahlreiche Anwendungstypen und die Einschränkungen des XML-Editors zu verstehen. Die XML-Unterstützung in Microsoft Word fällt in die folgenden beiden Bereiche: Dateiformat und Datenintegration. In diesem Abschnitt werden diese Bereiche detailliert beschrieben.

### Dateiformat

Angenommen, Ihr Unternehmen hat Tausende von Word-Dokumenten generiert, die Ihr Firmenlogo enthalten. Da das Logo geändert wurde, muss es nun in allen Dokumenten ersetzt werden. (Dieses Szenario ist in einem Microsoft-Artikel unter <http://msdn2.microsoft.com/en-us/library/aa338205.aspx> beschrieben, der sich eingehend mit dem Dateiformat von Microsoft befasst.)

Eine Methode zum Ersetzen des Logos ist, jedes Dokument zu öffnen, nach dem Logo zu suchen, dieses durch das neue Logo zu ersetzen und das Dokument zu speichern. Diese Methode ist für nur einige Dokumente ausreichend, aber für Hunderte oder Tausende von Dateien, sowohl ausgesprochen arbeits- und zeitaufwendig als auch fehleranfällig.

Eine andere Methode ist das Schreiben eines Programms, das das Öffnen der Dokumente, das Ersetzen des Logos und das Speichern automatisiert. In den Versionen vor Office 2007 hatten die Microsoft Word-Dateien jedoch das „.doc“-Format, ein proprietäres Binärdateiformat. Um das Ersetzen des Logos zu automatisieren, müssen Sie sich zuerst mit dem Dateiformat vertraut machen (an sich schon eine entmutigende Aufgabe) und anschließend die Software schreiben und testen, die die Dateien nach dem Logo durchsucht und dieses ersetzt.

Um dieses Szenario zu vereinfachen sowie viele neue Features zu aktivieren (dies ist im Artikel über den oben angegebenen Link beschrieben), hat Microsoft das .doc-Format in Office 2007 geöffnet, damit dieses zugänglicher ist. Obwohl Microsoft einfach die Spezifikationen für das vorhandene .doc-Format veröffentlichen hätte können, ist Microsoft einen Schritt weitergegangen und hat das Standardformat in eine XML-Version von .doc geändert, die .docx-Erweiterung, die als „OpenXML“ bezeichnet wird.

Der Hauptvorteil des neuen Dateiformats, das auf XML basiert, sind die zahlreichen Softwareprogramme, die bereits für die XML-Verarbeitung verfügbar sind. Das heißt, dass die Entwicklung von Software zum Verarbeiten von OpenXML einfacher und kostengünstiger ist, als das Schreiben von Software zum Verarbeiten von .doc-Dateien.

(Microsoft hat in Office 2003 ein XML-Dateiformat eingeführt, das als „Word ML“ bezeichnet wird. In Office 2007 hat Microsoft mehrere Verbesserungen an diesem Format vorgenommen, die entsprechende Dokumentation veröffentlicht und das Format als Standarddateiformat für Office übernommen. Obwohl Word 2007 .doc-Dateien öffnen kann, verwendet es standardmäßig das OpenXML-Format.)

## Datenintegration

Viele Word-Dokumente enthalten die Daten aus anderen Unternehmenssystemen. Beispielsweise kann es sich bei einem Word-Dokument um einen Katalog handeln, der Teilenummern, Beschreibungen und Preise aus einer Datenbank enthält, oder um einen Forschungsbericht mit Aktienkursen und anderen Finanzdaten von Aktiengesellschaften. Um diese Informationen in Word-Dokumente zu übertragen, müssen die Daten normalerweise in das Dokument eingegeben werden.

Aber was machen Sie, wenn die Daten geändert werden? Bis zur Freigabe von Office 2007 musste ein ähnliches Verfahren wie im Beispiel mit dem Logo ausgeführt werden: Der Benutzer musste die alten Dokumente öffnen, die veralteten Daten suchen, durch die neuen Daten ersetzen und die Dokumente speichern. Wie viele Wochen- oder Monatsberichte werden mit dieser Methode erstellt? Und wie viel Zeit wird dafür verschwendet?

Um diese fehleranfälligen und arbeitsaufwendigen Aufgaben überflüssig zu machen, hat Microsoft Funktionen für die XML-basierte Datenintegration in Office 2003 implementiert. Mit diesem Feature können Entwickler die Felder im Dokument konfigurieren, die mit Datenbanken und anderen Datenfeeds verbunden sind, um Dokumente automatisch oder mit nur einem Mausklick zu aktualisieren. Heute sind XML-basierte Datenfeeds von allen wichtigen Datenbank- und Unternehmenssystemen verfügbar.

Um diese Funktion zu implementieren, hat Microsoft in Office 2007 sogenannte Inhaltssteuerelemente integriert, mit denen der Programmierer die XML-Daten in einem Word-Dokument einbetten kann. Der Microsoft-Artikel unter <http://msdn2.microsoft.com/en-us/library/bb266218.aspx> enthält einen Screenshot, der die Verwendung eines Inhaltssteuerelements zum Einbetten von Aktienkursen in einem Dokument zeigt:



[Company Name] Stock Activity	: Last Price
Last Price	[Last price]

Der Programmierer kann das Inhaltssteuerelement so konfigurieren, dass automatisch die neuesten Aktienkurse abgerufen werden und die Tabelle aktualisiert wird, wenn der Benutzer das Word-Dokument öffnet. Außerdem kann der Programmierer in Office 2007 die Inhaltssteuerelemente sperren, damit diese von den Benutzern weder geändert noch entfernt werden können. Das Sperren von Dokumentabschnitten stellt für die Downstream-Prozesse sicher, dass die Dokumente immer die erwarteten Informationen enthalten. Dies vereinfacht das Schreiben und Testen dieser Prozesse.

## Wie Quark XML Author die XML-Unterstützung von Microsoft verbessert

**Obwohl Word einige ausgezeichnete Funktionen für XML umfasst, behauptet Microsoft nicht, dass Word ein XML-Editor ist. Tatsächlich verweist Microsoft Kunden, die einen XML-Editor benötigen, an die Produkte seiner Partner, beispielsweise Quark XML Author.**

Obwohl Word einige ausgezeichnete Funktionen für XML umfasst, behauptet Microsoft nicht, dass Word ein XML-Editor ist. Tatsächlich verweist Microsoft Kunden, die einen XML-Editor benötigen, an die Produkte seiner Partner, beispielsweise Quark XML Author.

Bevor wir beschreiben, wie Word von Quark XML Author verbessert wird, sollten Sie den wichtigsten Unterschied zwischen einem XML-Editor und Word verstehen. Mit einem XML-Editor können Sie im gesamten Dokument Ihre eigenen Tags verwenden. Für Dynamic Publishing-Anwendungen ist dies eine wichtige Anforderung, wenn Sie die gleichen Informationen mit einer jeweils anderen Formatierung für unterschiedliche Dokumenttypen und Medien veröffentlichen möchten.

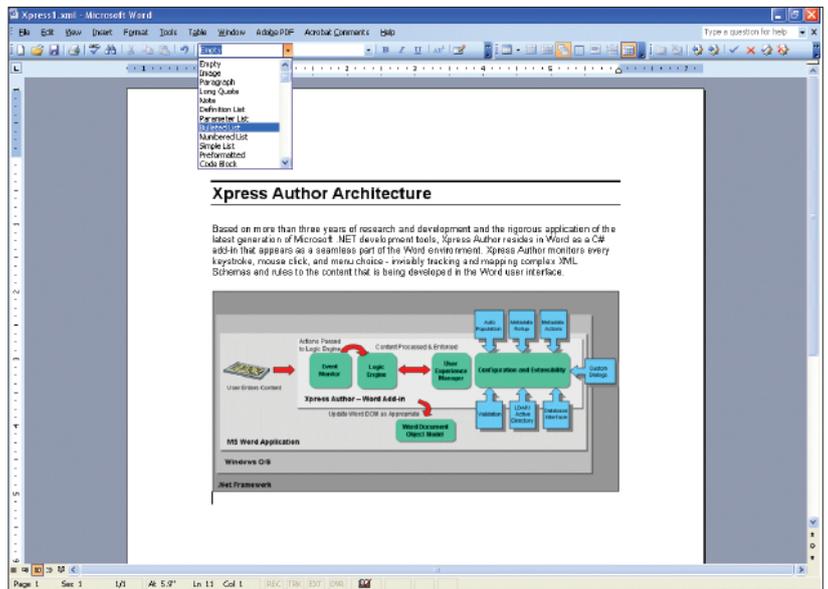
Im Gegensatz dazu hat Microsoft die Tags für das XML-Dateiformat in Word 2007 festgelegt, um die Darstellung des Inhalts in einer Word-Datei zu steuern. Das ist sinnvoll für Word, aber genau das Gegenteil der Trennung des Inhalts von der Formatierung, die das Dynamic Publishing erfordert. Beim Dynamic Publishing soll das Tag die Bedeutung des Inhalts anzeigen (beispielsweise den Titel, die Teilenummer oder eine Liste) und nicht die Darstellung. Ein separater automatisierter Publishing-Prozess legt die Formatierung basierend auf der Bedeutung fest, um den gleichen Text für unterschiedliche Zwecke jeweils anders zu formatieren.

Um die Funktionen für die Datenintegration in Word zu unterstützen, ermöglichen die Inhaltssteuerelemente in Office 2007 das Zuweisen von beliebigen XML-Tags zu bestimmten Abschnitten des Dokuments. Möglicherweise kommen Sie zu dem Schluss, dass Sie Inhaltssteuerelemente für alle Dokumentabschnitte verwenden können, um Ihren eigenen XML-Editor zu erstellen. Sie haben Recht.

Um jedoch die normalen Textbearbeitungsfunktionen durch Inhaltssteuerelemente zu ersetzen, müssen Sie den Arbeitsaufwand beurteilen, der zum Beibehalten des normalen Verhaltens von Word erforderlich ist. Wenn Sie bereit sind, das normale Verhalten von Word nachhaltig zu ändern, indem Sie ausgeklügelte und funktionelle Formate erstellen, dauert die Entwicklung wahrscheinlich nur einige Wochen. Möchten Sie aber die Störung der Benutzer möglichst gering halten, indem Sie das normale Verhalten von Word so weit wie möglich beibehalten, kann die Entwicklung mehrere Jahre dauern.

Hier kommt Quark XML Author ins Spiel: Quark XML Author erstellt eine echte XML-Bearbeitungsumgebung in Word. Um dies zu erreichen, mussten die Entwickler alle Funktionen in Word überprüfen, die über die Tastatur, Menüs, Symbolleisten oder das Windows-Betriebssystem verfügbar sind. Außerdem mussten die Entwickler Folgendes überprüfen:

- Die Möglichkeit, dass das Dokument beschädigt wird
- Die Benutzererfahrung für den normalen Word-Betrieb
- Die geeignete Implementierung für das XML-Authoring



Quark XML Author für Word wird im Hintergrund ausgeführt und greift nur minimal in den normalen Word-Betrieb ein, damit der Benutzer XML-Dokumente erstellen kann. In dieser Abbildung verwendet Quark XML Author die Word-Dropdown-Liste mit den Formatvorlagen, um alle gültigen Inhaltstypen anzuzeigen, die der Autor an der aktuellen Cursorposition einfügen kann.

Ein Beispiel der von Quark XML Author geänderten Funktionen sind die Word-Formatvorlagen. Quark XML Author stellt eine Beziehung zwischen den Formatvorlagen und XML-Elementen her, die als Smart Styles bezeichnet werden, da ausschließlich die in der Dropdown-Liste angezeigten Formate an der aktuellen Cursorposition eingefügt werden können. Quark XML Author fängt Hunderte von Funktionen auf, beispielsweise die Funktion für Formatvorlagen, und muss Tausende von Änderungen verarbeiten. Um Ihren eigenen XML-Editor in Word zu erstellen, müssen Sie alle Funktionen duplizieren (ein extrem teures Unterfangen).

Quark XML Author versucht, die normale Arbeitsumgebung bei der Verwendung von Word möglichst beizubehalten. Quark XML Author verwendet Dialogfelder und andere Steuerelementtypen nur, wenn diese die Benutzerumgebung gegenüber der Word-Benutzeroberfläche verbessern.

Um die XML-Kompatibilität zu erzwingen, geht Quark XML Author wie folgt vor:

1. Aus den Menüs und Symbolleisten werden alle unwesentlichen Funktionen entfernt
2. Ungültige Funktionen, beispielsweise das Ziehen eines Bilds an eine unzulässige Stelle im Dokument, werden deaktiviert und eine Fehlermeldung wird angezeigt

## Schlussfolgerung

Die Kombination aus Microsoft Word und Quark XML Author verbindet die Vorteile von XML mit der vertrauten, komfortablen Benutzeroberfläche von Word.

Obwohl die neuen Versionen von Microsoft Word wesentliche Verbesserungen für die Verwendung von XML umfassen, sind keine XML-Bearbeitungsfunktionen verfügbar, die umgehend verwendet werden können. Aufgrund der Word-Entwicklungsumgebung ist es jedoch möglich (mit großem Entwicklungsaufwand) einen einfach zu verwendenden XML-Editor mit umfangreichen Funktionen in Word zu erstellen. Quark XML Author für Microsoft Word macht dies dank achtjähriger und mehr als 100.000 Stunden Entwicklungsarbeit möglich. Die Kombination aus Microsoft Word und Quark XML Author verbindet die Vorteile von XML mit der vertrauten, komfortablen Benutzeroberfläche von Word.

## Über Quark

### Revolutionising Publishing. **Again™.**

Vor zwei Jahrzehnten hat Quark das Publishing mit QuarkXPress erstmals revolutioniert, einer Desktop Publishing-Software, die schnell zum Branchenstandard wurde. Heute führt QuarkXPress nicht nur Neuerungen im Desktop Publishing-Markt ein: Quark revolutioniert das Publishing erneut. Mit der Quark Dynamic Publishing-Lösung helfen wir unseren Kunden, die sich ständig ändernden Anforderungen zu erfüllen und neue Einkommensquellen zu erschließen, indem wir die Vorteile von fortschrittlichen Technologien auf den gesamten Publishing-Prozess ausweiten. Unsere Dynamic Publishing-Lösung legt einen neuen Standard für das automatisierte Multichannel-Publishing fest, da die Leistungsfähigkeit des flexiblen Layouts und Designs mit automatisierten Workflows und dem einfachen XML-Authoring für die personalisierte Kommunikation im Druckformat, im Web und in elektronischen Medien verknüpft wird.

Weitere Informationen zu Quark XML Author finden Sie unter [www.quark.com](http://www.quark.com).

**North America**  
Quark Inc.  
1800 Grant Street  
Denver, CO 80203  
[solutions@quark.com](mailto:solutions@quark.com)  
+1.800.676.4575  
+1.303.894.8888

**Europe**  
Quark Media House Sàrl  
Rue Du Monruz 2  
2000 Neuchâtel  
Switzerland  
[solutionseu@quark.com](mailto:solutionseu@quark.com)  
00800 1787 8275  
+41 (32) 720 1414

**Asia**  
Quark India Pvt Ltd.  
A 45 Industrial Area  
Phase VIII B Mohali  
160059  
Punjab, India  
[solutionsasia@quark.com](mailto:solutionsasia@quark.com)  
+91.172.2299.131

**Japan**  
Quark Japan K.K.  
West 248 bldg.  
2-4-8 Ebisu-nishi  
Shibuya-ku, Tokyo  
150-0021 Japan  
[solutionsjapan@quark.com](mailto:solutionsjapan@quark.com)  
+81.3.3476.9440

