XML – In Plain English!

Introduction

Across many business communities XML has become a hot topic, yet the misinformation and confusion surrounding XML is staggering. Though an enchanting buzzword, XML is very often misused and misunderstood. The purpose of this article is to explain, in very basic terms, what XML is and what it is not, in terms of data exchange.

What is a File Format?

A computer file contains information (text, images, video, etc.) to be used by a software application. The way this information is organized and stored within a computer file is its file format. In other words, the file format is the specification used by a software application to recognize and understand the data contained in a particular computer file.

File formats can be open or proprietary. Open (or "Free") file format specifications are publicly available and can be used and implemented by anyone. Open file formats allow multiple software applications to share the same files, and are usually maintained by a standards organization which governs the format specification and revisions to it. For instance, digital camera and scanner software can use the same open file format for saving images, allowing files created by either device to be used by a single software application. By contrast, proprietary file formats are created and maintained by private companies and software developers. Proprietary format specifications are not publicly available.

What is XML?

XML (Extensible Markup Language) is a strict, text-only "format" that has become extremely popular in recent years because of its ability to store, describe and transport data. It's important to note that the XML "language" itself doesn't "do" anything; it is simply a general set of rules for describing data. An XML file contains only plain text information; the format of a particular XML file is unique to the software application(s) that create or will use the file. Therefore, a particular XML file is of little use without an XML document definition, also known as an XML Schema.

What is an XML Schema?

An XML schema is a map for the data contained in an XML file. XML files must conform to an XML schema, and an XML schema is required to navigate through the data of an XML file.

How do Different Software Applications Communicate via XML?

Because almost every software application in the world can understand plain text, XML is ideal for passing data from one application to another. XML has enabled even very complex software applications to integrate with other applications. These transactions are not without custom programming however, since all applications using XML data must understand the data's schema. The schema acts as a contract between the applications, and the software developers create the processes for making use of the data.

A Common XML Misconception...

Though plain text is virtually a universal language among software applications across all computer operating systems, it is a common misconception that an XML file created by a particular application is automatically useful for another application that understands XML. At first glance, the contents of XML files may appear similar, but in fact, without the XML schema and a specific process developed for what to do with the data, the XML file itself is of no use.

Conclusion

XML is an excellent, flexible means of describing and transporting data, but it is by no means a "turn-key" solution for software integration. XML transactions are very specific and need to be carefully coordinated by both software applications.