

What is wrong with XML?

Author: Mihai Rotaru

Date: 13 Dec 2011

It seems somewhat inappropriate to state that there is something wrong with XML; it is, after all, just a technology designed for simple purposes: to provide an extensible markup language, which could be used on it's own, or be used to create other, specialized markup languages. The burgeoning scene of XML-based languages is a testament to XML's success, and therefore it fulfilled its purpose.

But that is not to say that XML doesn't come without drawbacks, though. One of the most important issues with XML pertains to it's verbosity - XML files can become quite large. This is inherent to XML, and is one of the trade-offs consciously made by it's designers; integrity was deemed more important than file size. This issue can be mitigated to some degree by using compression - an XML file is, after all, just text so it lends itself to compression algorithms. However, this negates one of XML's advantages, since if the archive is damaged it would hardly be possible to salvage any of the data.

Another issue consists of the relative complexity of an XML parser, as compared to alternatives available under certain conditions. For example, XML is often being used for storing configuration settings for applications. The application then has to include an XML parser, which could significantly increase the size of the application. This is less of an issue with large applications, such as Visual Studio or Microsoft Office, but simpler means of storing configuration files - such as .ini or .conf files - are available for smaller applications.

XML files can also be cumbersome to author without a specialized editor, again due to the explicit nature of XML. This is not a concern when small changes need to be made - for example, to change the resolution for a game in it's config.xml. But authoring can be an extremely tedious process when large chunks of XML need to be added, or when a document is created from scratch. This problem is compounded by the fact that XML documents should be well-formed, and in many cases valid; and small mistakes can render a document unacceptable to parsers due to well-formedness or validity issues. This issue can make XML documents 'read-only' for humans.