

LINGUAGEM XML (EXTENSIBLE MARKUP LANGUAGE)

**Cleusa Aparecida Carlos de Souza, Érika Ferreira Correia, Everton Luiz Araújo,
Henrique de Oliveira Santos, Rodrigo Ferreira da Costa e Samuel Cabral da Silva**

RESUMO

A programação para a WWW (wide world web) tem crescido de forma importante exigindo maiores recursos das linguagens de programação. A XML (extensible markup language) é uma linguagem de fácil utilização e que tem obtido ampla aceitação de mercado.

Palavra - chave: internet, XML, programação web.

ABSTRACT

The programming for WWW (wide world web) has grown in a important way requiring more resources from the programming languages. The XML (extensible markup language) is an easy to use language and has got broad acceptance from market.

Keywords: internet, XML, web programming.

pode tirar proveito da tecnologia XML, contanto que protocolos seguros (com recursos de criptografia) sejam usados para assegurar as comunicações.

Este artigo tem por objetivo estudar o padrão XML, sua definição, suas origens, características, diferenças com relação ao HTML, vantagens, desvantagens e exemplos de aplicações. Descrever como é composto um documento escrito em XML, ou seja, seus elementos, sintaxe, atributos e forma de declaração.

2. DEFINIÇÕES E ORIGENS

A XML (eXtensible Markup Language) é uma linguagem de marcação que serve basicamente para descrever dados (MACORATTI, 2000).

A primeira linguagem de marcação, a General Markup Language (GML), surgiu em 1969, através dos pesquisadores Ed Mosher, Ray Lorie e Charles F. Goldfarb. Em seguida, transformou-se em SGML – Standard Generalized Markup Language (VITAL, 2005).

Em 1986, a SGML foi adotada como um padrão internacional de troca e armazenagem de dados pela ISO – International Organization for Standardization – (VITAL, 2005).

No ano de 1996, através do Word Wide Web Consortium (W3C), que iniciou o projeto de uma linguagem de programação que combinaria a flexibilidade da SGML com a HTML. Desta forma, a XML é considerada um subconjunto da SGML (ANTONACCI, 1999).

Surge, então, a XML, que fora padronizada pelo W3C, em 1998, na sua versão 1.0 (VITAL, 2005).

2.1. Diferenças entre a HTML e a XML

Páginas que têm a intenção apenas de publicar de informações, em que a preocupação é apenas com o layout do texto e distribuição de link, são casos simples, que são resolvidos apenas com HTML (Hyper Text Markup Language), ASP (Active Server Pages) e uso de componentes (PEREIRA, 2002).

Se uma aplicação necessitar consultar uma base de dados existente em uma determinada página, em determinado certo, haverá uma dificuldade, pois as respostas reproduzidas estarão em HTML, e será necessário fazer um tratamento do texto para encontrar os dados necessários à consulta, sem contar o risco que a aplicação correrá de parar de funcionar, caso haja alguma alteração na seqüência de páginas ou no design (PEREIRA, 2002).

Segundo Pereira (2002), a HTML é considerada uma excelente linguagem para exibir informações para o usuário. Porém, quando se trata de aplicações mais complexas, em que é necessário manipular os dados, a HTML deixa muito a desejar, além do que a linguagem impõe várias limitações descritas a seguir, segundo o autor:

- É uma tecnologia para apresentações, combinando dados com a forma de apresentá-los e tornando difícil uma separação dos dois elementos;
- Tem um conjunto fixo e rígido de marcações (tags) e não permite que você use suas próprias marcações nos seus programas aplicativos;
- É uma linguagem “plana” e não permite que você especifique uma hierarquia de dados, incluindo detalhes como limites e importância;
- Dificulta a transmissão de dados para o cliente, para que sejam processados posteriormente;
- É gerada continuamente pelo servidor e o cliente é apenas um mecanismo de exibição de dados;
- Fornece apenas um modo de visualização dos dados. Se quiser fornecer diferentes modos de visualização, você precisa refazer ou gerar os dados e a página HTML completa no servidor, para exibição nos clientes;
- Não é facilmente legível, nem pelo olho humano, nem por máquina;
- Não é muito consistente. Por exemplo, algumas marcações, ou tags, exigem início e fim, como <html> e </html>, outras, só têm início, como <p> e . Desse modo, os analisadores de sintaxe HTML (existentes nos browsers) têm de saber lidar com essa formatação aleatória.

