Universidade do Estado de Minas Gerais - Campos de Frutal

Curso de Sistemas de Informações – 5º Período.

Disciplina: Banco de dados II.

Docente: Dr. Geraldo Corrêa Nunes.

Discentes: Leonardo Baldo e Leopoldo Ferreira de Paula.

**Projeto de Banco de Dados em XML**

**1. Origem da XML**

Marcação nada mais é que uma informação complementar inserida em um documento, ou seja, acrescenta marcadores nos documentos criados. Em texto, por exemplo, estes marcadores coordenam o aspecto do documento e é constituído de itens como tipo e tamanho da fonte, negrito, sublinhado, itálico etc. Podemos utilizar da marcação generalizada para uma simples designação da importância estrutural de um texto em um documento.

SGML é um tipo de linguagem de marcação com uso mais específico em medicina ou matemática devido sua dificuldade, fazendo assim com que seus usuários sejam escassos. A HTML, na verdade, é uma aplicação da SGML para a marcação de conteúdo, de modo a prepará-lo para apresentação na World Wide Web Consortium (W3C). Um projeto iniciado pela W3C em 1996 fez a junção da versatilidade da SGML e a ampla aceitação do HTML surgindo assim a XML.

XML (Linguagem Extensível de Marcação) é uma linguagem de marcação recomendada e criada pela World Wide Web (W3C) para a criação de documentos com dados organizados hierarquicamente, tais como textos, banco de dados ou desenhos vetoriais. A linguagem XML é classificada como extensível porque permite definir os elementos de marcação.

**2. Conceitos Básicos de XML e Tecnologias Relacionadas**

XML faz o uso de Tags muito similares as tags usadas em HTML, porém, a diferença é que enquanto o HTML tem um número de tags fixas, o XML define seus elementos e as funcionalidades dos mesmos.

**Exemplo 1: Trecho de documento em HTML**

<p>

<font face=”verdana” size=10>

<b>Remetente:</b> José da Silva<br>

<b>Destinatária:</b> Maria da Silva<br><br>

</font>

</p>

<p>

<font face=”arial” size=12>

Olá, Maria, bom dia!<br>

</font>

</p>

<p>

<font face=”arial” size=12>

Espero notícias suas! Até mais.<br>

</font>

</p>

A partir deste exemplo há a possiblidade de se reconhecer que se trata de uma carta, pois, temos o conhecimento prévio de que se trata de uma, já um software não tem condição de ter esse conhecimento, é preciso dar informações precisas para que uma aplicação entenda o que está sendo dito, a linguagem HTML usa as “tags” que apenas controlam como serão exibidos os dados. Interpretação das Tags: <p> diz ao interpretador, no caso, navegador de internet que deve-se quebrar uma linha e inserir uma em branco. <br> insere uma quebra de linha, <b> marca o que estará em negrito e <font> controla a fonte que será utilizada.

O problema da HTML é que ela é reconhecida como abordagem “engessada”, pois, limita o uso de tags a apenas uma função, como por exemplo o <p> que dá início e fim, é reconhecido apenas como cabeçalho de carta, e não pode ser usado posteriormente, essa limitação é reconhecida pela XML de tal forma que está exposta em seu nome, onde X significa “eXtensible”, ou extensível, significando que cada XML tem um subconjunto próprio de tags. Essas tags receberiam nome que não correspondem aos dados, apenas representam a estrutura determinada no contexto deste documento em específico. O exemplo 1 poderia ser representado da seguinte forma:

**Exemplo 2: Trecho de documento em XML**

<carta tipo="social">

<cabeçalho>

<remetente> José da Silva </remetente>

<destinatário> Maria da Silva </destinatário>

</cabeçalho>

<parágrafo>

Olá, Maria, bom dia!

</parágrafo>

<parágrafo>

Espero notícias suas! Até mais.

</parágrafo>

</carta>

Nesse exemplo a estrutura do documento está muito mais clara e é específico para uma carta, XML em si é um arquivo de texto contendo conteúdo e marcação que especifica aquele conteúdo, como no exemplo <carta> determina o início e fim de um documento e dessa forma diferente da HTML podendo existir vários documentos desde que seja marcado seus respectivos inícios e fins através das tags, em que entre as tags há a informação de o que o elementro representa. Um exemplo ilustrativo pode ser dado por: <parágrafo> Olá, Maria, bom dia! </parágrafo>. Um atributo é uma informação descritiva do elemento, e consiste de um nome e de um valor. Os atributos devem ser especificados sempre entre delimitadores, que podem ser aspas duplas ou simples. Em <carta tipo="simples">, o elemento <carta> possui um atributo chamado tipo, cujo valor é "simples". Somente uma instância de um nome de atributo é permitida dentro de uma tag inicial ou de uma tag vazia, ou seja, o exemplo a seguir seria inválido: <carta tipo="simples" tipo="comercial">.

São tecnologias relacionadas a XML:

SAX: (Simple API for XML, ou API simples para XML) é um analisador baseado em eventos, e, no caso de eventos, o analisador chama métodos ou rotinas com nomes semelhantes no programa host, que poderão tomar a ação apropriada.

DOM: (Document Object Model) é uma interface de programação de aplicação

completa para documentos XML. Essa interface permite não apenas navegar dentro de

documentos XML, mas também apanhar, acrescentar, modificar ou excluir elementos e

conteúdo.

XPath: é quem define como acessar os nós em documentos XML através de expressões de caminho que especificam um arquivo dentro de um diretório de arquivos, a expressão que identifica o atributo “tipo” da tag <carta> é dada por /carta/@tipo, com isso a localização do caminho é facilitada.

*XSL:* (Extensible Style Language) é uma linguagem para definir folhas de estilo para

XML. Ela também aceita opções de impressão e transmissão de documentos XML entre

várias plataformas.

**3. Banco de Dados em XML**

Com a expansão do uso da internet novos desafios surgiram em relação a trocas comerciais e integração de processos entre organizações, com isso foi desenvolvido um sistema de banco de dados nativo em XML.

Sistemas nativos em XML podem ser implementados sobre um modelo relacional, hierárquivo, orientado a objetos, ou mesmo utilizar um formato proprietário de armazenamento. O que difere um banco de dados nativo em XML dos demais modelos é unidade de armazenamento: em um banco de dados relacional uma linha é a menor unidade lógico já no nativos em XML a menor é o próprio documento XML. Esse documento deve ser mantido intacto, e ao ser recuperado deve estar idêntico a quando foi armazenado, é importante entender o que a XML traz de benefícios para o mundo dos bancos de dados, e qual lugar ela ocupa atualmente entre as tecnologias existentes.

A expansão do uso da internet aconteceu em cima do uso de padrões abertos, ou seja, não-proprietários: código-fonte, interfaces, sistemas operacionais etc., independentes de um fabricante específico e abertos a alterações, adequações e extensões que se fizessem necessárias, sempre mantendo o alinhamento com as regras definidas para estes padrões, com isso surgiu novas maneiras para o aproveitamento comercial da internet, destacando-se dentre elas o comércio eletrônico entre empresas (business-to-business, ou B2B), entre empresas e órgãos governamentais (business-to-government, ou B2G), e entre empresas e clientes finais (business-to-consumer, ou B2C). Segundo Renan Mendes em sua publicação O Dinheiro de Verdade está no E-Commerce B2B:

“No mundo, a expectativa é que em 2020 o mercado de E-commerce B2B dobre o tamanho do mercado online B2C — **U$6,7 trilhões vs U$3,2 trilhões** de faturamento mundial — de acordo o provedor de pesquisas Frost & Sullivan. A empresa prevê que o país da Alibaba, a China (assim como acontece no B2C), seja o maior mercado online **com U$2,1 trilhões em vendas online em 2020.**

Nos Estados Unidos, país da Amazon, o E-commerce B2B já é duas vezes maior que o B2C. A Forrester Researcher estima que o mercado B2B supere os U$780 bilhões e com um share de 9.3% das vendas de todo o mercado no fim de 2015. As previsões apontam um crescimento anual de 7,7% ao longo dos próximos 5 anos, dados resultantes da necessidade das empresas em diminuir custos.”

É indubitável que para que ocorra o comércio eletrônico deva haver troca de informações, e para facilitar a integração entre parceiros, deveria existir uma ferramenta de transmissão de dados de forma padronizada onde todos do grupo de parceiro entenderiam seu significado, até pouco tempo atrás, o que estava disponível nesse sentido era a tecnologia de troca eletrônica de dados, ou Electronic Data Interchange (EDI). Resumidamente, a EDI “define uma plataforma técnica, baseada num conjunto de padrões estruturados para troca de informações entre participantes num mercado” (ALBERTIN, 2001).

A EDI foi uma solução precária, pois, geralmente requer uma rede de conexão, privada e cara, entre os parceiros de negócios; Os parceiros precisam ter um relacionamento preexistente, para dar a eles tempo para estabelecer uma rede; Não é interativa, inibindo a possibilidade de negociação e discussão dinâmica.

Com formas de acesso baratas, preços não dependentes da quantidade de informação utilizada, padrões abertos e amplamente adotados a internet se firmou como opção a essa situação de altos custo e seletividade de usuários utilitários do e-commerce.

Como exposto por Renan, o relacionamento “um-para-um” será gradativamente substituído pelo “um-para-muitos”. Soluções fechadas como a EDI se tornaram obsoletas, a XML se enquadra nesse cenário de busca por padrões abertos que tiram o máximo de vantagem da arquitetura da internet já que redes de computadores tratam basicamente da troca de informações, e a internet nada mais é que uma rede de computadores. Com a XML, essas trocas de informações passaram a ter um padrão aberto, outra característica importante da XML é a sua portabilidade, dizer que ela é portável equivale a dizer que pode ser usada sem adaptações entre plataformas diferentes de hardware e software. A razão para tanto é que qualquer documento escrito usando o padrão XML será sempre um arquivo de texto simples, que usa as regras de marcação da XML. Assim, este arquivo texto pode ser trocado entre diferentes arquiteturas sem ser necessário sofrer alterações em seu formato: arquivos de texto podem ser reconhecidos e processados por qualquer plataforma virtual.

Não devemos levar em conta somente o uso do XML em si como ferramenta de armazenamento de dados, mas também de todas as tecnologias que o envolvem como XSL, XSD, linguagens para consulta como XPath, Xquery e XQL, e analisadores DOM e SAX, dentre outras. O uso combinado destas tecnologias, muitas das quais padrões publicados ou em fase de aceitação pelo W3C, é que torna viável o uso da XML como padrão robusto de intercâmbio e armazenamento de informações. Sem as tecnologias que o englobam, a XML é apenas mais uma linguagem de marcação, sem o poder e eficiência que apresenta quando usada em conjunto com elas. A adoção da XML vem num ritmo crescente por parte daqueles que usam a internet como meio para a troca de informações, mas é interessante observar que sua importância não se limita a troca de dados ela também invade a área das aplicações de bancos de dados, e com grande destaque.

Já existe no mercado uma incontável quantidade de sistemas de bancos de dados que lidam satisfatoriamente com dados relacionais que são predominantes no contexto empresarial. Porém, após o surgimento e fortalecimento do padrão XML para troca e armazenamento de dados, surgiram novas necessidades com as quais os desenvolvedores de sistemas de bancos de dados relacionais tiveram que lidar. Ao perceberem a importância do XML os grandes fornecedores de sistemas de bancos de dados adicionaram extensões para XML aplicadas a seus sistemas gerenciadores de bancos de dados relacionais, com o intuito de adaptar estes sistemas para que pudesse de alguma forma armazenar dados XML em suas tabelas relacionais os principais fornecedores de soluções para bancos de dados relacionais são: Oracle, Microsoft e IBM.

**4. Conclusão**

O que deve-se levar em consideração em todo e qualquer projeto, é o escopo de seu tamanho, e a real necessidade de qual tecnologia usar e qual será a melhor abordagem para o mesmo, atualmente o banco de dados do tipo relacional é o mais usado pois, com ele pode-se armazenas um imenso volume de informação com confiabilidade de processamento e segurança de armazenamento, ainda mais implementando o ACID para processamento de transações, porém, há projetos em que o uso de banco de dados XML irá melhor atender a necessidade, o que deve ser feito é sua identificação, para assim tirar melhor proveito, pois, como demonstrado ele vem se tornando cada vez mais utilizado devido suas vantagens em certos casos, o que agora deve ser trabalhado é que seja tão confiável para o armazenamento de dados e recuperação de informação quanto o modelo relacional.

Jímenez, Higasi e Masao demonstra casos em que a implementação de XML seria viável:

“[...] algumas características de sistemas que poderiam ser implementados

sob um banco de dados XML nativo:

· Sistemas de pequeno a médio porte, com pequeno volume de dados e transações;

· Sistemas para dispositivos com poucos recursos, como palmtops e hanhelds, bem como computadores mais antigos que não suportariam um grande SGBD;

· Interface de integração entre sistemas de back-end e front-end ou outros. Poderíamos ter toda a base de dados relacionada a integração entre sistemas armazenadas nesse tipo de banco de dados, ao invés de utilizar arquivos de configuração e de sincronismo;

· Implementação do próprio sistema de front-end de uma grande aplicação. Poderíamos replicar os dados necessários de um grande SGBD para o Xindice e utilizar esses dados para um sistema específico;

· Implementação de sistemas onde os dados a serem armazenados são semiestruturados, ou seja, não possuem “colunas” fixas, e essas podem ser bastante variadas. Assim, mapear esses dados em um modelo relacional poderia ser bastante trabalhoso, além de desperdiçar espaço (teríamos uma grande quantidades de colunas armazenando valores nulos). Para esses casos, o XML, e mais especificamente o Xindice poderiam ser muito mais indicados, pois a estrutura de um documento XML, apesar de ter sua estrutura, é muito mais flexível e se adaptaria a essa situação facilmente.” (2003, pag. 80)

**5. Referências**

FIGUEIREDO, Flávio J. V.; HIGASI, Hilton; MURAOKA, Lincoln M. J. “XML e Banco de Dados”. Universidade Presbiteriana Mackenzie, São Paulo/SP, 2003; Disponível em: <http://meusite.mackenzie.com.br/rogerio/tgi/2003XMLXindex.PDF> Acesso em: Setembro de 2016.

MENDES, Renan. “O dinheiro de verdade está no mercado de E-commerce B2B”. Disponível em: <http://www.profissionaldeecommerce.com.br/dinheiro-mercado-e-commerce-b2b/> Acesso em: Setembro de 2016