

# As bibliotecas na Web e vice-versa

MARCELLO PEIXOTO BAX\*

*Apresenta um panorama geral das principais tecnologias associadas à criação e disponibilização de recursos digitais na Internet. Aborda tópicos como: breve história da Web, as características de sua principal linguagem (HTML), conceitos básicos de multimídia e tecnologias mais recentes como Java e ActiveX. Apresenta conceitos fazendo referência a documentos em linha na Internet, o que possibilita consultas mais aprofundadas aos leitores interessados em obter mais detalhes sobre um determinado assunto.*

## 1 Introdução

Ninguém contesta a influência e a importância da World Wide Web, ou simplesmente Web ou WWW, para a sociedade da informação em que vivemos e aquela que está por vir. Entretanto, quem já recorreu à Internet, ou mais especificamente à Web, sabe que ainda é preciso criar meios para organizar os recursos informacionais disponíveis. Vários laboratórios e pesquisadores trabalham atualmente sobre esta questão e uma saída para amenizar o problema parece despontar: são as chamadas bibliotecas digitais. Cada vez mais presentes na Internet, estas estruturas computacionais são concebidas com o objetivo de representarem verdadeiras bibliotecas na Web.

No entanto, para que estas bibliotecas obtenham proveito de anos e anos de pesquisa em organização e tratamento da informação, arquivística e demais subáreas afins à ciência da informação, será necessário trazer a Web para o mundo das bibliotecas e centros de informação. Para isso, os pesquisadores em ciência da informação envolvidos em projetos de concepção de bibliotecas digitais devem procurar conhecer e dominar as diferentes tecnologias usadas para se construir, disponibilizar e manter recursos digitais em linha na Internet.

A Web é de importância fundamental para as bibliotecas e centros de informação. Pouco a pouco ela vai impondo uma mudança radical na maneira pela qual as bibliotecas servem à comunidade. Aquelas bibliotecas que não forem capazes de integrar estas mudanças, de forma efetiva, ainda que gradual, simplesmente desaparecerão ao longo do tempo, muito provavelmente por falta de usuários.

Hoje em dia provedores de informação comerciais conseguem chamar

<sup>1</sup>Professor da Escola de Biblioteconomia da UFMG



+  
+  
+  
+  
+  
+  
+  
+  
+  
+  
+  
+  
+  
+  
+  
+  
+  
atenção de um número cada vez maior de usuários da Internet, oferecendo mecanismos de pesquisa de informação na rede mundial. Muitos destes usuários satisfazem suas necessidades de informação exclusivamente através de mecanismos de pesquisa, como o *Surf*<sup>1</sup> no Brasil, ou o *Altavista* nos EUA, por exemplo. O fato é que tais mecanismos dificilmente integram, ou integrarão, o conhecimento desenvolvido pela área da ciência da informação durante décadas de pesquisas realizadas sobre a organização e o tratamento da informação. É obviamente desejável que esse arcabouço teórico, já desenvolvido pela ciência da informação, seja usufruído em conjunto, e com o apoio, das novas tecnologias digitais. Estas últimas devem ser consideradas como parte fundamental para a ciência da informação: não se trata mais da simples utilização de programas, ou software, mas sim de moldá-las à luz da pesquisa em ciência da informação.

+  
+  
+  
+  
+  
+  
+  
Como fruto dessa interação, estão surgindo novos software apropriados para a busca interativa e mais focalizada da informação (principalmente em se tratando de informação especializada): são as chamadas bibliotecas digitais. Tais software não têm a pretensão de indexar ou catalogar todos os recursos presentes na Internet, mas simplesmente de oferecer novos e sofisticados serviços de consulta e recuperação da informação até hoje inexistentes. Por este lado, eles não podem ser comparados aos mecanismos de pesquisas existentes atualmente.

Para que as bibliotecas digitais evoluam, levando em consideração os desenvolvimentos teóricos da área da informação, os profissionais da área precisam entender, cada vez mais, os princípios gerais da criação de recursos digitais (para a Web, por exemplo), de forma a se capacitarem e, em colaboração com analistas de sistemas e programadores, desenvolverem serviços que possam atrair e manter o interesse dos usuários.

O crescente papel da Internet, e mais especificamente da Web, nas comunicações internas de qualquer instituição pública ou privada, nas pesquisas universitárias e na educação em linha e continuada, reforçam a necessidade premente em formar profissionais capazes de entender os mecanismos de criação e disponibilização de recursos digitais na Internet. A Web pode ser utilizada por profissionais de biblioteconomia para o oferecimento de um grande número de serviços, dentre os quais: provimento de acesso a materiais informacionais, distribuição de publicações, manuscritos etc.; promoção de encontros e seminários locais; construção de repositórios de informação mais dinâmicos e personalizados. Além disso, a Web constitui um fórum ideal para a criação de projetos de colaboração institucional, onde os esforços de grupos de pesquisadores podem ser centralizados e disponibilizados para outros grupos, quebrando barreiras de tempo e espaço físico.

Este documento procura discutir a terminologia e alguns conceitos básicos associados à criação e disponibilização de recursos digitais na Internet e na Web. Algumas das mais recentes tecnologias são descritas e referidas através de documentos na Internet, possibilitando consultas mais aprofundadas. Apresentam-se tecnologias e métodos utilizados para a criação de hiperdocumentos. Embora não se tratando de um guia da linguagem HTML, relacionam-se endereços de documentos disponíveis na Internet sobre o tema. Este documento deverá dar origem à uma versão

---

<sup>1</sup>O mais recente mecanismo de pesquisa em português no Brasil, à época da redação desse artigo. Sua URL é <http://www.surf.com.br>

digital interativa e em constante evolução disponível na Web<sup>2</sup>, que trará vários outros recursos de interesse sobre o assunto na Internet.

## 2 Um pouco de história

O poder da Web reside em sua capacidade em associar um determinado elemento (palavra, frase, imagem) de um documento digital a outro localizado no mesmo computador ou afastado milhares de quilômetros. Esta conexão entre documentos é possível graças à utilização do protocolo de comunicação de dados chamado HTTP (*Hypertext Transfer Protocol*) e da linguagem HTML (*Hypertext Markup Language*). Embora o conceito de hipertexto venha sendo discutido desde que o cientista Vannevar Bush<sup>3</sup>, escreveu o artigo intitulado: *As We May Think*<sup>4</sup>, publicado em 1945, somente 45 anos depois, em 1990, é que a HTML foi concebida. HTML aparece pela primeira vez num artigo escrito por Tim BERNERS-LEE e Robert CAILLIAU do CERN<sup>5</sup> intitulado: *World Wide Web: Proposal for a HyperText Project*<sup>6</sup>. [2].

Até o desenvolvimento de programas navegadores, ou browsers, para as plataformas Windows e Macintosh, o acesso à Internet através da interface Web era restrito apenas a supercomputadores. A chegada da Web aos computadores pessoais (PCs) só ocorreu com o lançamento da versão do navegador Mosaic do NCSA (National Center for Supercomputing Applications) em 1993 [8]. Desde então, uma multidão de outras empresas, notadamente a Netscape com o seu Netscape Navigator e, com algum tempo de atraso, a Microsoft com o *Internet Explorer*, desenvolveram também programas de navegação similares.

As versões mais recentes dos programas servidores Web para Windows e Macintosh simplificou e democratizou o desenvolvimento de sites Web em PCs. Um serviço Web básico pode ser criado com o auxílio das ferramentas mais recentes, utilizando-se um PC e algumas poucas horas de trabalho. Entretanto, dependendo da extensão e da dinâmica de alteração das informações, recursos humanos dedicados ainda são essenciais para o sucesso de um *site*.

## 3 Independência de plataforma: uma questão chave para a Web

Talvez uma das principais vantagens da publicação de documentos e criação de serviços digitais usando a Web seja o fato de ser a mesma uma tecnologia completamente independente do material computacional usado e, em grande parte, até mesmo do sistema operacional do computador.

Isto significa que o mesmo recurso pode ser usado por indivíduos em computadores rodando sistemas como Macintosh, Windows ou UNIX, de qualquer ponto (cidade, estado ou país) onde estejam e durante 24 horas por dia, bastando que,

2 O documento estará disponível no endereço "cuba.eb.ufmg.br/bax".

3 Conselheiro científico do então presidente Roosevelt nos anos 40 e diretor do principal grupo de pesquisadores norte-americanos.

4 Este artigo pode ser encontrado na URL: <http://www.isg.sfu.ca/~duchier/misc/vbush/>.

5 Centre European de Recherche Nucleaire, Genebra, Suíça.

6 <http://www.w3.org/hypertextVwww/Proposal.html>





Language da *World Wide Web Organization* [9] e *A Beginners Guide to HTML da NCSA* [6]. Existem muitos outros guias de HTML disponíveis na Web. O leitor poderá verificar na rubrica Guides and Tutorials, mantida pelo site de pesquisa Yahoo!<sup>7</sup> . Para aqueles que desejam informações detalhadas, o documento que apresenta a especificação SGML completa de HTML 4.0 pode ser obtido na página da W3C: "www.w3c.org".

#### 4.1 A natureza da linguagem HTML

Talvez o ponto mais importante a ser compreendido sobre HTML é o fato de ser esta linguagem orientada à estrutura e ao conteúdo do documento, ao invés de sua aparência ou apresentação final. Dessa forma, cada navegador Web, seja ele textual ou gráfico, poderá traduzir o documento da forma apropriada à visualização local. Por exemplo, o autor deve estar consciente de que a especificação do estilo de fonte (aparência) com o qual um documento deve ser mostrado, muitas vezes não será considerada, pois caso ela exista no computador do usuário que está lendo o documento, será escolhida outra, dentre aquelas que lá estiverem presentes.

Esta natureza orientada à estrutura e ao conteúdo dos documentos traz inúmeras vantagens para HTML. Ela permite ao autor se focalizar no conteúdo de seu documento que será provavelmente mais bem escrito e fácil de entender. Sendo a HTML menos dependente de códigos de formatação de apresentação, o autor pode alterar seu documento com menos trabalho de edição. Edições podem ser feitas com o auxílio de editores de texto simples ao invés de sistemas de publicação eletrônica sofisticados. Além do mais, sendo os documentos HTML mais livres dos estilos de fontes e demais códigos de formatação, eles são, geralmente, bem menores (menos bytes) e podem ser transferidos mais rapidamente pela Internet.

Apesar de todas as aparentes vantagens citadas acima, o que realmente conta no novo paradigma de editoração chefiado pela linguagem SGML é o fato desta ser um padrão internacional reconhecido pela ISO e seguido por toda e qualquer empresa de software que queira participar do mercado. As informações se tornam, assim, independentes de uma plataforma de software específica. Elas não precisam mais, como antes, serem proprietárias e estarem aprisionadas nesse ou naquele formato particular, desse ou daquele software produzido por essa ou aquela empresa, e só por ela. A utilização de um formato internacionalmente reconhecido libera, assim, a informação de suas amarras e da inconveniente necessidade de utilização de conversores de formato. Em suma, a informação se torna muito mais perene.

Essa independência de um formato proprietário proporcionada pelo padrão SGML é devida ao fato da linguagem ser uma espécie de meta-linguagem. Dito de outra forma, ela permite a construção de tipos de documentos, os chamados DTDs (*Data Type Document*) que podem ser interpretados por qualquer software que reconheça o padrão SGML. Para isso, basta que o documento seja acompanhado da especificação do seu tipo, ou seja, do seu DTD. Como o documento é geralmente acompanhado do seu DTD, qualquer software que interprete SGML pode interpretar a sua formatação e mostrá-lo num meio qualquer, seja na tela do computador ou na impressora.

<sup>7</sup> [http://beta.yahoo.com/Computers\\\_and\\\_Internet/Internet/World\\\_Wide\\\_Web/HTML/Guides\\\_and\\\_Tutorials](http://beta.yahoo.com/Computers/_and\_Internet/Internet/World\_Wide\_Web/HTML/Guides\_and\_Tutorials).



Note que HTML nada mais é do que um tipo de documento definido em SGML específico para ser utilizado na Web. No caso específico de HTML, os documentos não precisam vir acompanhados de seus DTDs (ou especificações de tipo em SGML), pois os navegadores já são construídos previamente levando-se em consideração o DTD da última versão de HTML. Assim, os navegadores convencionais estão aptos a interpretar HTML e não SGML.

## 4.2 Ligações hipertexto

A explosão e o sucesso atual da Web deve-se à utilização de hipertextos, pois com um simples apertado do botão do mouse todo um processo de transferência de arquivos é ativado. As ligações hipertexto em documentos Web podem ser usadas para vários propósitos. Talvez o mais importante deles seja facilitar o acesso dos usuários a recursos multimídia em arquivos disponíveis em rede. Elas permitem aos usuários moverem-se entre várias páginas de um mesmo documento ou de documentos diferentes de forma livre, provendo informações adicionais e definindo melhor palavras ou clareando conceitos.

A marca HTML que cria uma ligação hipertexto é <A HREF> -A de âncora e HREF de Hipertext Reference. O texto que é associado a esta marca se torna uma ligação hipertexto para outro documento ou recurso multimídia. Esta associação entre um pedaço de texto e um recurso é, geralmente, chamada de *referência hipertexto*. Eis um exemplo de tal referência:

```
<A HREF="/beatriz/doc.html">Documento</A>
```

Neste caso, a palavra Documento será evidenciada na página mostrada no navegador geralmente sublinhada e em cor diferente identificando uma referência hipertexto. Assim, a palavra Documento é associada ao documento doc.html contido no diretório beatriz no computador onde se encontra a página que contém a referência. Quando o navegador Web localiza uma referência hipertexto ele colore as palavras e/ou as sublinha, indicando uma referência hipertexto. Clicar na palavra Documento faz com que o documento contendo a referência desapareça e seja substituído na tela do navegador pelo documento referenciado doc.html, no caso acima.

## 4.3 Tipos de caminhos e URLs

O caminho faz referência ao diretório ou nome de arquivo usado pelo navegador para localizar um determinado recurso. Existem dois tipos diferentes de caminhos usados em referências hipertexto: caminhos *relativos* ou *absolutos*.

No uso de caminhos relativos pressupõe-se que o recurso associado reside no mesmo computador servidor e no mesmo diretório ou sub-diretório relativo ao documento que está sendo visualizado. A informação sobre o protocolo de acesso e o caminho do diretório formam o chamado URL (*Uniform Resource Locator*), usado para se alcançar o recurso.



Um exemplo de caminho relativo em HTML:

```
<A HREF="url.html"> URL </A>
```

*Caminhos absolutos* são usados quando o autor referencia recursos que não são diretamente controlados por ele; geralmente localizados em outros servidores na Internet. Toda a estrutura do diretório deve, então, estar presente para que o navegador possa localizar o recurso referenciado. Um exemplo da referência acima, usando-se caminho absoluto:

```
<A HREF="http://www.eb.ufmg.br/bax/documentos/doc1/url.html"> URL </A>
```

Existem vantagens fundamentais na utilização de caminhos relativos. O seu uso não necessita do adicional

```
http://www.eb.ufmg.br/bax/documentos/doc1/url.html
```

em todas as referências hipertexto nos documentos contidos no diretório doc1. Caminhos relativos facilitam a movimentação de *sites* inteiros de um servidor para outro, tornando a operação independente da estrutura de diretórios do servidor destino.

#### 4.4 Mais sobre URLs

Como visto anteriormente, um URL especifica de que maneira um recurso pode ser acessado na Internet; informa aos navegadores sobre o tipo de protocolo de acesso a ser usado na transferência do recurso, o endereço do domínio do servidor que contém o recurso na Internet, o caminho e o nome do arquivo que contém o recurso. Todo URL, devido à sua natureza universalmente única (não podem existir dois recursos com URLs idênticos em toda a Internet), possui caminho absoluto. A sintaxe padrão de um URL é:

tipo-do-protocolo:

```
//servidor.domínio(:porta)/diretório(/subdiretório)/nome-do-arquivo.extensão
```

ou, para a futura versão em linha deste documento:

```
http://www.eb.ufmg.br/bax/documentos/WebDoc.html
```

Os protocolos, ou tipos de servidores, referidos na seção tipo-do-protocolo incluem:

- *ftp*: servidor de ftp, geralmente com acesso anônimo (o usuário não precisa se identificar);
- *gopher*: servidor Gopher;
- *http*: servidor Web;
- *news*: servidor de notícias (News);
- *telnet*: servidor Telnet;
- *mailto*: para mandar uma mensagem via correio eletrônico (e-mail).

Usando uma referência hipertexto, o seguinte código HTML criaria uma ligação para este documento:

```
<A HREF="http://www.eb.ufmg.br/bax/documentos/WebDoc.html"> Bibliotecas na Web</A>
```



## 5 Aspectos de multimídia

O uso de multimídia na Web é, desde sua criação, o foco principal de desenvolvimento. A Web é uma plataforma onde a distribuição de informações multimídia é realizada de forma bastante natural, através de redes de computadores. Além do que, a criação de recursos multimídia tem se tornado cada vez mais simples, à medida que novas ferramentas de editoração e conversão são desenvolvidas.

Os recursos multimídia referenciados em documentos Web muitas vezes não podem ser executados pelo próprio navegador. Quando o navegador encontra um formato que não consegue processar, ele procura uma aplicação, ou software específico. Os navegadores mais modernos estão sendo desenvolvidos com diferentes aplicações embutidas, tornando-se programas cada vez maiores! Entretanto, a maioria dos vários formatos de mídia podem, ainda, ser processados por software auxiliares: os chamados *plug-ins*, ou *helpers*.

Uma referência hipertexto associada com um recurso de mídia é codificada em HTML de forma similar à mostrada anteriormente para imagens. Por exemplo, o código HTML `<A HREF= "exempl03.wav">` resultará numa ligação para um arquivo de som. Ao se adicionar mídias em documentos, uma boa prática é utilizar ícones como referências aos recursos; isto ajuda o usuário a identificar o recurso como arquivo de áudio, vídeo etc..

Um exemplo usando um arquivo de áudio:

`<A HREF= "exemplo3.wav">icone.gif</a>` (*formato: WAV; tamanho: 506k*)  
será mostrado na tela assim:

(*formato: WAV; tamanho: 506k*)

Informações sobre o formato da mídia e o tamanho do arquivo também são úteis para que os usuários se informem sobre a aplicação que devem possuir para poderem interpretar o arquivo, ou decidam se estão dispostos a esperar o tempo necessário para a sua transferência.

### 5.1 Gráficos e imagens

Existe um debate considerável entre os autores de páginas Web sobre o papel das imagens em documentos HTML. Inserir imagens pode tornar um documento visualmente atrativo para o leitor; porém, o tamanho e a quantidade delas pode desviar a atenção do conteúdo do documento. Além disso, o uso de gráficos aumenta o tempo de transmissão do documento de forma drástica. O compromisso entre nem tantos e nem tão poucos, ou seja, o uso apropriado de imagens tem um efeito determinante sobre um retorno futuro do usuário ao site.

Existem duas maneiras de inserir imagens em documentos Web: podem ser *embutidas* ou podem ser *externas* ao documento. Imagens embutidas são aquelas que serão automaticamente mostradas quando da visualização do documento, independentemente de uma ação consciente do usuário em visualizá-las; elas são impostas. A marca HTML para se embutir imagens é `<IMG SRC>`. A sintaxe é idêntica àquela usada pela marca `<A HREF>`, com a localização do arquivo expressa por



caminho relativo ou absoluto. Para incluir uma linha colorida logo em seguida neste documento o código usado seria o seguinte: <IMG SRC="/images/colorline.gif">.

Imagens externas são aquelas que não são mostradas automaticamente quando do carregamento do documento no navegador. Elas precisam ser solicitadas pelo usuário para serem transmitidas pela rede. Imagens grandes não deveriam ser embutidas em documentos, mas sim acessadas como imagens externas. Geralmente imagens externas são referenciadas por cópias menores, de mais baixa resolução (chamadas em inglês de *thumbnails*), codificadas como referências hipertexto para a imagem maior. Uma cópia menor da imagem deve servir de referência hipertexto a uma imagem externa. Isso pode ser realizado com o uso de uma combinação das marcas <A HREF> e <IMG SRC>.

Por exemplo, o código HTML:

```
<A HREF="pequena.gif"><IMG SRC="grande.gif"></A>
```

resulta na imagem "pequena.gif" sendo uma referência hipertexto para a versão da mesma imagem "grande.gif", bem maior, externa ao documento. É importante notar que imagens podem servir como referência hipertexto para qualquer recurso: texto, som, filme, outra imagem etc.

Alguns formatos de imagens podem não ser reconhecidos por um determinado navegador, dependendo, muitas vezes, de qual versão é utilizada. Quando o navegador encontra uma imagem cujo formato não pode interpretar, ele tenta executar uma aplicação independente (*plug-in*) capacitada a mostrar o formato da imagem.

### 5.1.1 Formatos de arquivos

Existe uma grande variedade de formatos de imagens, mas os mais comuns são os chamados GIF (pronuncia-se "jifi" ) e JPEG (pronuncia-se "j\_peg"). Um formato mais recente é o *Portable Network Graphic* (ou PNG). O PNG foi concebido especialmente para aplicações que necessitam de visualização em linha como a Web, por exemplo. O FlashPix<sup>8</sup> é outro formato recentemente desenvolvido pela Kodak, concebido para oferecer uma espécie de resolução *à la demande*. A princípio, a imagem é mostrada com uma resolução fraca; em seguida, o próprio usuário (aquele que está visualizando a imagem), através de um menu flutuante ativado sobre a imagem, solicita qual a resolução desejada. Somente então esta parte da imagem é transmitida pela rede.

O tipo de uma imagem, assim como de outros recursos como texto, filme, som etc., pode ser identificado pela extensão do nome de seu arquivo. Definiu-se um padrão internacional para a identificação de tipos de recursos, o chamado padrão *Multipurpose Internet Mail Extension* (MIME). Extensões MIME são geralmente de três letras; para os formatos de imagens mencionados acima, tem-se: ".gif", ".jpg", ".png", ".fpx". Os navegadores usam as extensões dos arquivos para identificarem o seu formato. Assim, caso o formato não seja suportado pelo próprio navegador, este último pede ajuda a um plug-ins que será capaz de visualizar o arquivo. Essa estratégia faz com que os navegadores não cresçam de tamanho indiscriminadamente, à medida que novos formatos são criados.

---

<sup>8</sup> <http://www.kodak.com/daiHome/flashPix/flashPix.shtml>.



## 5.2 Áudio

As aplicações práticas para o uso de áudio na Web são ainda bastante limitadas. Como resultado, poucos sites utilizam este recurso. Usos potenciais dessa tecnologia incluem: música ambiente para o *site*, disponibilização de leituras ou entrevistas gravadas, histórias contadas oralmente, eventos gravados ao vivo etc. A criação de recursos de áudio requer conhecimento técnico especializado [3]. Além disso, a conversão de áudio a partir de fitas cassete em forma digital, ou a gravação direta em forma digital, requer material especializado que poucas bibliotecas podem adquirir.

Uma razão para a não inclusão de áudio é que poucos usuários possuem o material e as aplicações necessárias para seu processamento. Enquanto os Macintoshes incluem capacidades de áudio embutidas, os PCs requerem placas de som e caixas externas. A princípio, as bibliotecas não devem pressupor que seus usuários possuam tais equipamentos e configurações.

Os formatos de áudio mais comuns e suas extensões MIME são *ulaw* (.au) e *WAVE* (.wav). A principal razão da popularidade desses formatos é sua independência de plataforma e a disponibilidade de aplicações processadoras. Uma discussão profunda sobre a distribuição de áudio na Internet pode ser encontrada no documento Web, escrito por Victor LOMBARDI da New York University [3].

## 5.3 Vídeo

A inclusão do vídeo em *sites* Web tem sido também bastante limitada. Como para o áudio, a ausência de aplicações para a integração de vídeos e a falta de equipe técnica para a produção e conversão de informações visuais para a Web parece ser a principal razão do pouco uso desta mídia. Potencialmente, vídeos poderiam ser usados em bibliotecas digitais para empréstimo de vídeos educacionais, visitas remotas às instalações físicas da biblioteca, caso estas existam, e muitas outras.

Os principais formatos de vídeo e suas extensões MIME usados na Web são *MPEG* (.mpg) e *QUICKTIME* (.mov). O tamanho dos arquivos de vídeo é o maior problema enfrentado pelos autores de páginas que os disponibilizam. Um filme de trinta segundos de duração ocupa vários megabytes, mesmo após compressão. A transferência de arquivos desse tamanho através dos modems atuais, cuja taxa de transferência gira entre 28.800 e 33.600 bits por segundo (bps) parece uma tarefa sem fim. Vários produtos para visualização de filmes existem e estão atualmente disponíveis na Internet, tais como, VOSAIC, CineWeb, e InterVu. Uma tecnologia que parece ter grande potencial é a Shockwave<sup>9</sup>.

Embora esses produtos pareçam ser promissores para a disponibilização de informações multimídia, a realidade da tecnologia atual é que a largura de banda e o tráfego das redes limitam as taxas de transmissão a poucos quadros de um filme por segundo. O tamanho dos visores dos programas de visualização atuais limitam-se a um pequeno quadrado de poucos centímetros na tela. A transmissão de um vídeo que ocuparia toda a tela do computador seria hoje impensável. Esses fatores estão fazendo

<sup>9</sup> Para mais informações, procure pela palavra Shockwave em algum mecanismo de pesquisa.

com que a maioria dos autores de sites Web estejam esperando uma ou duas gerações antes de disponibilizar arquivos realmente multimídia, contendo sons e filmes.

## 6 Editores de HTML

Um documento HTML pode ser escrito em texto ASCII por intermédio de um simples editor de texto como o *Notepad* do Windows. Há poucos anos, esta era a única forma disponível para se criar um documento HTML, porém, a disponibilização de vários programas especializados na edição de páginas Web tornou esta tarefa bem mais fácil. Funcionando como verdadeiros processadores de texto, novos editores do tipo WYSIWYG (*What You See Is What You Get*) permitem a edição do documento como se este já estivesse sendo mostrado no navegador. Listas de editores HTML são mantidas por alguns *sites*, como, NSCA<sup>10</sup>, The World Wide Web Organization<sup>11</sup>, e Yahoo!<sup>12</sup>.

Todas as maiores empresas que desenvolvem processadores de texto já incluíram a possibilidade de se salvar ou se converter arquivos para o formato HTML. Programas como *Pagemaker* da *Adobe* e *Word* da *Microsoft* oferecem essa possibilidade. Estes programas permitem aos autores salvarem arquivos como documentos HTML, reduzindo a necessidade de terem de codificar todo o documento. Para aqueles que não podem ou não querem pagar pelos programas mais modernos de editoração de textos, existem ainda vários filtros conversores de textos para documentos HTML. Uma lista destes filtros é mantida pela World Wide Web Organization<sup>13</sup> e o Yahoo!<sup>14</sup>.

Programas cada vez mais especializados têm sido propostos para a editoração de documentos HTML. Com a evolução desta linguagem, que passa a ter uma estrutura mais complexa além do suporte de vários formatos adicionais (de quadros, gráficos, folhas de estilo etc.) são necessários programas mais sofisticados. Existem, também, ferramentas que se propõem à editoração de todo um *site* Web, oferecendo, por exemplo, uma janela com visão em forma de gráfico das ligações entre páginas e imagens, além de editor de páginas. O programa *FrontPage* da *Microsoft* é um exemplo.

## 7 Algumas considerações de concepção

Com a proliferação do desenvolvimento de *sites*, existe atualmente uma grande quantidade de manuais sobre estilos de concepção de páginas Web. Entretanto, não há um estilo que possa ser considerado *correto*. Cada instituição acaba desenvolvendo seu próprio estilo padrão. O importante é a continuidade e a consistência das páginas, que devem ser o objetivo principal.

Existem algumas considerações básicas sobre concepção de páginas Web que

10 <http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/faq-software.html#editors>.

11 <http://www.w3.org/hypertext/WWW/Tools/Overview.html>.

12 [http://www.yahoo.com/Computers/\\_andInternetSoftware/Internet/World\\_Wide,-Web/HTML'-Editors/](http://www.yahoo.com/Computers/_andInternetSoftware/Internet/World_Wide,-Web/HTML'-Editors/).

13 <http://www.w3.org/hypertext/WWW/Tools/Filters.html>.

14 [http://http://www.yahoo.com/Computers.-andInternetSoftware/Internet/World\\_Wide,-Web/HTML\\_converters/](http://http://www.yahoo.com/Computers.-andInternetSoftware/Internet/World_Wide,-Web/HTML_converters/).





9. Facilite a navegação através do *site* utilizando marcas de orientação nas páginas. Use elementos de navegação consistentes, tais como ligações para sua página de chegada, de forma a que os usuários naveguem sempre bem orientados.

10. Inclua ligações para os programas na Internet que seus visitantes necessitarão para navegarem no site. Sempre que fizer uso de áudio ou vídeo, aponte para os sites onde os usuários poderão descarregar os respectivos processadores destes recursos.

## 8 Alguns conceitos mais recentes

### 8.1 As tecnologias *Java* e *ActiveX*

Em sua maioria, os documentos atualmente criados com HTML são estáticos: embora contenham ligações hipertexto e imagens animadas, as páginas não são muito interativas. Isto está mudando com o surgimento de novas linguagens, como *Java* e *ActiveX*, de uso (quase) específico para a Internet.

A linguagem *Java*, criada pela Sun Microsystems, permite colocar dentro de documentos pequenos programas, ou *applets* (microaplicações), para a realização de várias tarefas. Tal como HTML, a linguagem *Java* foi concebida de forma independente de hardware. Suas aplicações podem ser executadas numa grande variedade de CPUs e sistemas operacionais. Aplicações escritas em *Java* podem ser transportadas pelas redes e são acessíveis a qualquer um que possua um interpretador *Java* (*HotJava* da Sun ou Internet Explorer e Netscape, a partir da versão 2.0.).

Atualmente, as aplicações *Java* para bibliotecas e centros de informação estão apenas começando a ser pesquisadas. Sendo *Java* uma linguagem concebida com base na sintaxe de C++, a maioria dos bibliotecários responsáveis por *sites* Web a acharão de difícil manipulação. Para essas pessoas foi criada a linguagem de script chamada *JavaScript* que estende as capacidades de programação de *Java* a um número maior de autores. *JavaScript* é simples o suficiente para qualquer um habituado ao uso de HTML. Enquanto *Java* é usado por programadores para criação de novos programas ou aplicações complexas, *JavaScript* foi concebida para os autores de páginas HTML criarem pequenos scripts que podem ser executados tanto no lado do servidor quanto no do cliente. O site da Sun Microsystem disponibiliza maiores informações sobre *Java* e *JavaScript*.

*ActiveX* parece ser a resposta da Microsoft à linguagem *Java*. Trata-se de um conjunto de tecnologias também independentes de plataforma (pelo menos segundo seus criadores); uma versão moderna da tecnologia OLE, que possibilita a comunicação entre aplicações (mesmo através de redes, incluindo a Internet) que seguem o mesmo padrão. Chamadas de componentes na terminologia *ActiveX*, essas aplicações estão sendo promovidas pela Microsoft como uma tecnologia aberta, compatível com um grande número de outros padrões. A idéia, que aliás é originária da programação orientada a objetos, é de poder desenvolver pequenos blocos (componentes) que podem ser reutilizados várias vezes em diferentes aplicações. Os componentes *ActiveX* podem tanto ser usados por páginas Web, como também por







agrupa o resultado da pesquisa, e o mostra na tela do navegador como uma página HTML. O serviço EbscoDOC<sup>17</sup> é um exemplo deste tipo de construção dinâmica de documentos.

A concretização de tais sistemas necessita de pessoas especializadas e algum tempo de desenvolvimento, mas uma vez criado, será necessária pouca manutenção técnica. Por questões técnicas, o gerenciamento de *sites* inteiros de forma dinâmica pode não ser a melhor estratégia para a maioria das bibliotecas. Entretanto, a criação dinâmica de certas páginas pode ser uma forma útil de se delegar a responsabilidade de manutenção de partes do site para outros membros da equipe de trabalho. Com efeito, as informações são adicionadas diretamente na base de dados, sem passar por qualquer edição de arquivos HTML. Por exemplo, bases de dados podem ser criadas com os horários de funcionamento da biblioteca, incluindo informações sobre a equipe, agenda de encontros e workshops etc. As páginas contendo tais informações podem ser criadas de forma dinâmica a partir da base de dados.

Existem muitos programas que podem ser usados para a criação dinâmica de documentos Web. A Oracle lançou, recentemente, seu WebServer que pode ser integrado ao gerenciador de base de dados Oracle. Ao invés de escrever programas CGI ou scripts PERL, o WebServer cria páginas HTML de forma dinâmica através de requisições ao programa que é chamado Web Request Broker; uma espécie de negociador de requisições que se coloca entre o banco de dados e o servidor Web. Informix é outra empresa de banco de dados que cria produtos para autoria e gerenciamento remoto de sites Web.

Atualmente, uma ferramenta popular para a concepção de intranets<sup>18</sup> é o *IntraBuilder*<sup>19</sup>, da empresa de software americana Borland. O IntraBuilder oferece uma série de facilidades para a criação de sites dinâmicos possibilitando programação em JavaScript de forma visual (no estilo Delphi ou Visual Basic). Tais ferramentas são geralmente acompanhadas de estruturas de sites padrões, além de bibliotecas de elementos gráficos mais comuns, que facilitam a criação de páginas.

## 9 Considerações futuras

A Internet está passando de uma mídia antes baseada em texto para uma mídia altamente visual. A crescente integração de recursos multimídia está mudando o formato de transmissão de informações e novos protocolos de comunicação são criados para englobarem tais mudanças. O efeito dessas mudanças para as bibliotecas está apenas começando a ser estudado. O interesse atual sobre o tema *bibliotecas digitais*, demonstrado pelo grande número de congressos e simpósios na área, atesta o grau de relevância das pesquisas na área para uma sociedade cada vez mais informatizada.

<sup>17</sup> <http://www.ebsco.com>.

<sup>18</sup> Intranet é uma rede do tipo Internet interna à uma empresa ou organização. A disponibilização da informação, as ferramentas de software (clientes e servidores) e os protocolos de transferência de dados são os mesmos utilizados na Internet e WWW. Porém, a rede pode não estar conectada à Internet ou conectada através de um computador que serve de ponte segura, evitando ataques externos. São os chamados Firewalls.

<sup>19</sup> <http://www.borland.com/intrabuilder>.



Conforme se viu, apoiando-se sobre uma rede física de computadores -a Internet-, a própria Web é um conceito em constante mudança. Os avanços tecnológicos não param, ocorrem simultaneamente, alterando sua estrutura e seu conteúdo. Novos conceitos como documentos dinâmicos, Java e arquivos PDF mostram seu impacto, mudando a maneira pela qual os recursos Web são criados.

Um importante desenvolvimento que afetará, sem dúvida, o futuro das aplicações Web é a linguagem de programação chamada VRML (*Virtual Reality Modeling Language*). Esta linguagem permite aos autores criarem mundos virtuais em três dimensões (3D). Usando navegadores como *Cosmo*<sup>20</sup>, usuários interagem com *sites* Web através de uma interface realista em três dimensões, reproduzindo o mundo físico com o qual o usuário está habituado. Embora ainda considerada experimental, esta linguagem mostra sinais de amadurecimento e vários *sites* 3D começam a despontar na Internet.

Usando VRML as bibliotecas poderão, dentro em breve, oferecer *sites* tridimensionais que reproduzam seu espaço físico, oferecendo interfaces mais interativas para seus usuários. A viabilização destas idéias no seio das bibliotecas atuais e futuras trará, sem dúvida, um enorme avanço para a sociedade de informação do próximo século e dependerá somente da habilidade e capacidade técnica dos profissionais da informação e bibliotecários para conceber, gerenciar e manter tais recursos.

### *The libraries in the Web and vice-versa*

*Presents an overview of the main technologies for development and provision of Internet digital resources. Oriented towards information professionals, the article approaches topics such as a brief history of the Web, characteristics of its main language HTML, basic concepts of multimedia and more recent technologies like Java and ActiveX. Throughout the text, citations to Internet on-line documents are provided to readers interested about more details on specific topics.*

## Referências Bibliográficas

- AMDUR, D. Dynamic Media for the Web. *New Media*, v. 5, n.11, p.42-52, Nov.1995.
- LEE Berners-Lee T. J., CAILLIAU R, GROFF J.-F. World Wide Web: The Information Universe. *Electronic Networking*, v. 2, n.1, p.52-59, Spring 1992.
- LOMBARDI, V. Audio on the Internet.  
<http://www.nyu.edu/atg/library/papers/internet.audio>
- MAGID, J., JONES, P., MATTHEWS, R. D., MCCONVILLE, D. *The Web Server Book: Tools and Techniques for Building Your Own Internet Information Site. Online Companion*. Ventana Press. 1995.  
[Http://www.vmedia.com/wc/onlcomp/wsb](http://www.vmedia.com/wc/onlcomp/wsb)
- METCALFE, E.S., FRISSE, M. E., HASSAN, S.W., SCHNASE J.L. *Academic Networks: Mosaic and the World Wide Web. Academic Medicine*, v. 69, n. 4,p. 270-273, Apr. 1994.
- NATIONAL Center for Supercomputing Applications, *A Beginners Guide to HTML*. <http://www.ncsa.uiuc.edu/general/internet/www/HTMLPrimer.html>
- POWELL J. *Spinning the World Wide Web. Database* v. 18, n.1, p. 54-59, Feb./Mar. 1995.
- SCHATZ, B. R., HARDIN, J. B. *NCSA Mosaic and the World Wide Web: Global Hypermedia Protocols for the Internet. Science*, v. 265, n. 5174, p. 895-901, Aug.121994.
- The WORLD Wide Web Organization. *HyperText Markup Language*.  
<http://www.w3.org/hypertext/WWW/MarkUp>

Recebido em 19/05/1997

<sup>20</sup> <http://webspacesgi.com/cosmoplayer>.