



Universidade do Estado do Rio de Janeiro
Centro de Tecnologia e Ciências
Escola Superior de Desenho Industrial

Bruno Sérgio Coelho de Oliveira

**Aplicando técnicas de hipermídia adaptativa
em jornais online**

Rio de Janeiro
2007

Bruno Sérgio Coelho de Oliveira

**Aplicando técnicas de hipermídia adaptativa
em jornais online**



Dissertação apresentada como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Design da Universidade do Estado do Rio de Janeiro. Área de concentração: Design.

Orientador Prof. Dr. André Soares Monat

Rio de Janeiro
2007

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / CTC/G

O482 Oliveira, Bruno Sérgio Coelho de.
Aplicando técnicas de hipermídia adaptativa em jornais
online / Bruno Sérgio Coelho de Oliveira. – Rio de Janeiro, 2007.
92 f.

Orientador: Prof. Dr. André Soares Monat.
Dissertação (mestrado) – Universidade do Estado do Rio de
Janeiro, Escola Superior de Desenho Industrial.

1. Sistemas Hipermídia – Teses. 2. Hipermídia adaptativa
– Teses. 3. Jornais online – Teses I. Monat, André Soares. II.
Universidade do Estado do Rio de Janeiro. Escola Superior de
Desenho Industrial. III. Título.

CDU 004.032.6

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial
desta tese / dissertação.

Assinatura

Data

Bruno Sérgio Coelho de Oliveira

**Aplicando técnicas de hipermídia adaptativa
em jornais online.**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Design da Universidade do Estado do Rio de Janeiro. Área de concentração: Design.

Aprovada em 20 de agosto de 2007

Banca examinadora:

Prof. Dr. André Soares Monat (Orientador)
ESDI – UERJ

Prof^a. Dr^a. Mônica Moura
Universidade Anhembi Morumbi

Prof. Dr. Washington Dias Lessa
EBA – UFRJ

Rio de Janeiro
2007

AGRADECIMENTOS

Ao Senhor Deus por todas as coisas que tem feito e ainda que fará. Pelo Seu amor incondicional e sustento milagroso durante este período de muito trabalho. Obrigado Senhor!

Ao meu orientador André Monat, principal responsável por guiar este trabalho e por ter se tornado mais do que um orientador, um amigo que me recebeu e acompanhou durante todos estes 2 anos.

Aos meus pais, Paulo e Daura por todo zelo, apoio e força, desde quando nasci até hoje, e sem os quais não seria possível aqui estar, além do Robson, meu grande irmão e amigo. Amo vocês!

À minha linda, querida, amada varoa e logo esposa, Carla, sempre presente com seu apoio e compreensiva desde o começo deste trabalho, apesar do tempo destinado para tal. Te amo, minha linda!

Ao professor Roberto Eppinghaus, sem dúvida um grande responsável por meus primeiros e atuais passos no mundo acadêmico, com quem muito aprendi e certamente aprendo.

Aos meus gerentes e coordenadores da Globo.com, pela compreensão e cessões de tempo durante este período: André Braz, Tatiana Roza, Márcio Tristão, Eduardo de La Rocque e Diogo Cabral.

Aos meus colegas de trabalho que, durante este período, torceram e deram apoio para este trabalho: Anderson Frisone, André Cardoso, Felipe Memória, Eduardo Brandão e Rodrigo Estevam.

Um agradecimento especial para os professores Frank Barral, João Bezerra, Luis Cláudio Portugal e Rafael Cardoso, pelo aprendizado, além dos exemplos de dedicação e disciplina a serem seguidos.

Aos meus grandes amigos João Marcelo, Guilherme Tardin, Rafael Carlos, Ricardo Barrocas e Vitor Carlos que sempre me apoiaram, mesmo me tendo um tanto quanto ausente durante este período.

Aos irmãos da Comunidade Cristã Semente do Amanhã, pelas orações e, em especial, aos meus líderes Anderson e Marcus Paulo pelo zelo durante este tempo. Somos mais que vencedores!

Aos alunos da FESO, em Teresópolis, pela disponibilidade para participação na pesquisa e, especialmente, ao Leonardo Teixeira pelo desenvolvimento do sistema idealizado nesta dissertação.

Aos meus alunos do primeiro ano da ESDI em 2006 e 2007, com quem também muito aprendo, inclusive àqueles que também participaram da pesquisa aqui presente.

A todos que de alguma forma contribuíram para este trabalho. Desde os autores citados na bibliografia até àqueles que contribuíram com uma simples idéia ou palavra de apoio: obrigado!

RESUMO

OLIVEIRA, Bruno Sérgio Coelho de. *Aplicando técnicas de hipermídia adaptativa em jornais online*. 2007. 92 f. Dissertação (Mestrado em Design) - Escola Superior de Desenho Industrial, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2007.

Esta pesquisa corresponde ao estudo de técnicas de hipermídia adaptativa (HA), visando melhorias nos meios para organização e disponibilização de conteúdo jornalístico online de acordo com o comportamento e interesse individual de um usuário em busca de notícias e informação. Do ponto de vista do design, a abordagem sobre o assunto corresponde, basicamente, à busca por métodos inteligentes de alimentação de uma interface gráfica com conteúdo noticioso, baseando-se em estudos do comportamento de navegação do usuário. A idéia é que as estatísticas matemáticas deste comportamento sejam traduzidas na exibição do conteúdo para atender de forma antecipada possíveis interesses do usuário. Concentrou-se como área de interesse focal, as matérias que serão incluídas na página principal e a forma de como serão distribuídas. Para isto foram criados alguns modelos de layout para a página principal. Estes modelos de layout possuem slots (espaços) onde as notícias são inseridas. Estes slots, por sua vez, estão distribuídos hierarquicamente de tal forma que os considerados mais importantes são aqueles com fotos e localizados nas melhores posições da página. A metodologia aqui criada foi aplicada como estudo na página principal do portal “www.globo.com” de conteúdo predominantemente noticioso e pertencente a um dos principais grupos de empresas de comunicação do Brasil.

Palavras-chave: Hipermídia adaptativa. Hipermídia. Modelagem do perfil do usuário. Design. Design da informação. Jornalismo online. Jornalismo.

ABSTRACT

This research concerns a methodology to organize the homepage of online news in order to adapt it according to the main interests of the users. The designer's approach to this issue involves the usage of intelligent procedures based on the navigational behavior of the user and statistical analysis of such behaviour. Our main interest concerns the selection of the news to be inserted into this home page and how to distribute them. We deal with some templates for such first page. These templates have slots where news is inserted. These slots are hierarchically distributed in such a way that most important slots are those with photos and larger texts and are located in the best positions. These best slots are filled out with the most interesting news according to each user's interests. This whole methodology has been applied to the journalistic site www.globo.com which belongs to a major Brazilian communication company.

Keywords: Adaptive hypermedia. Hypermedia. User modeling. Design. Information Design. Journalism. Online news.

Lista de Ilustrações

1 - "A Cauda Longa" (Área "B"). (Adaptado da <i>Wikipedia</i>)	3
2 - Hipermídia adaptativa como combinação entre hipermídia e modelagem de usuário.	7
3 - Classificação taxonômica dos métodos e técnicas de hipermídia adaptativa.	11
4 - Anotação de link no sistema <i>InterBook</i> .	13
5 - Exemplo do comportamento de um fragmento de texto condicional no "AHA!".	14
6 - Sugestão de compra (abaixo do produto) baseada em outros usuários. (www.amazon.com)	16
7 - O "My Yahoo" pode ser personalizado manualmente.	17
8 - Exibição do sistema com as editorias acessíveis: economia e política. (SeAN)	19
9 - Exibição da página de notícias personalizada. (Imagem de uma tela utilizando o SeAN.)	20
10 - Arquitetura do sistema SeAN.	20
11 - Estereótipo de interesse usuário. (SeAN)	22
12 - Página principal do "Chicago Online", versão online do "Chicago Tribune". (iml.jou.ufl.edu/carlson)	27
13 - Página principal do "Mercury Center", versão online do "San Jose Mercury News". (iml.jou.ufl.edu/carlson)	28
14 - O "Nando Land", uma versão online via BBS para o "News & Observer". (iml.jou.ufl.edu/carlson)	28

15 - O “@ Times” do “News York Times” na AOL. (iml.jou.ufl.edu/carlson) _____	29
16 - O jornalista Sérgio Charlab e a primeira página do JB Online. (www.jbonline.com.br) _____	30
17 - Primeira página principal do “UOL”. (www.uol.com.br) _____	30
18 - O “Globo On” em uma de suas primeiras aparições na Internet. (www.oglobo.com.br) _____	31
19 - O portal “IG” em 2000. (www.ig.com.br) _____	32
20 - A primeira versão do portal “Globo.com” (esquerda, 2000) e a atual (direita, 2007). (www.globo.com) _____	33
21 - O “Jornal do Terra” em 2007. (www.jornaldoterra.com.br) _____	34
22 - O “Jornal Interativo” em 2007. (www.alltv.com.br) _____	35
23 - G1: o <i>site</i> de notícias da Globo. (www.g1.com.br) _____	35
24 - As versões <i>online</i> e impressa atuais do “The New York Times”. (www.nytimes.com) _____	37
25 - O papel eletrônico da “Fujitsu” (retirado de diginfo.tv) _____	38
26 - Área inicial visível do <i>site</i> “Wikipedia” em um computador na resolução de 1024 x 768 <i>pixels</i> . _____	42
27 - O Diagrama de Gutenberg. (retirado do livro “ <i>Universal Principles of Design</i> ”) _____	43
28 - Leitura ordenada ocidental para GOLDSMITH apud CHUEKE (2004) _____	43
29 - Um gravador <i>EyeTrack</i> e as imagens por ele geradas. (retirado de poynterextra.org) _____	44
30 - As áreas de prioridade nas interfaces gráficas digitais. (poynterextra.org) _____	44
31 - O caminho em forma de “F” verificado na visualização de três interfaces. (www.useit.com) _____	45

32 - As etapas de construção da interface gráfica de uma página para a Web. (retirado do livro “ <i>The Elements of User Experience</i> ”) _____	48
33 - Exemplo de uma “arquitetura da informação”. (Globo.com) _____	49
34 - Exemplo de um “esqueleto de interface e navegação”. (Globo.com) _____	50
35 - A superfície visível de uma “interface gráfica digital”. (www.globo.com/tvglobo) _____	50
36 - Expectativa de posicionamento dos elementos “voltar para a página principal”, “links internos” e “busca no site”. (Retirado de uma apresentação de Mônica Fernandes feita para a Globo.com) _____	51
37 - As interfaces gráficas digitais dos sites de comércio eletrônico “Submarino” e “Americanas.com”. (www.submarino.com.br e www.americanas.com) _____	52
38 - As interfaces gráficas dos sites “Último Segundo”, “Folha Online” e do “ <i>The New York Times</i> ”. (www.ultimosegundo.ig.com.br, www.folha.uol.com.br e www.nytimes.com) _____	52
39 - As versões da página principal do site G1 para um computador pessoal, um telefone celular avançado e um celular com menos recursos. _____	54
40 - A versão atual (1024 x 768 pixels) do Globo Online. _____	55
41 - A relação entre as questões éticas, sociais e políticas na sociedade da informação. _____	58
42 - Página de política de privacidade da Globo.com. _____	61
43 - Os três modelos de layout. _____	64
44 - Hierarquia dos slots na página principal. _____	65
45 - Formulário para avaliação do interesse do usuário pelos assuntos das notícias publicadas. _____	66
46 - A classificação de acordo com a taxonomia dos assuntos. _____	67
47 - Data mart de cliques. _____	67

48 - Notas para as notícias da versão adaptada (esquerda) e não-adaptada (direita). _____ **70**

49 – Distribuição de pontos de acordo com a “nobreza” do *slot*. _____ **70**

50 - Distribuição de pontos de acordo com a “nobreza” do *slot* de um usuário. _____ **71**

Sumário

1.	Introdução	01
1.1.	Justificativa	01
1.2.	Descrição do trabalho	05
2.	Hipermídia adaptativa	07
2.1.	Definição e evolução	07
2.2.	Taxonomia de Brusilovsky	10
2.3.	Aplicações e principais sistemas	12
2.4.	Aplicações em jornalismo <i>online</i>	18
3.	Histórico e características do jornalismo <i>online</i>	24
3.1.	Evolução do jornalismo <i>online</i>	24
3.2.	Comportamento do leitor	39
3.3.	Diagramação no jornalismo <i>online</i>	48
4.	Questões éticas do monitoramento do usuário	56
4.1.	Ética em sistemas da informação	56
4.2.	Política de privacidade	61
5.	Metodologia para uso de hipermídia adaptativa em jornalismo <i>online</i>	63
5.1.	Apresentação geral da metodologia	63
5.2.	Aplicação da metodologia	66

5.3.	Avaliação dos resultados	69
6.	Conclusões e futuros trabalhos	73
	Bibliografia	76
	Anexos	80

1. Introdução

1.1. Justificativa

A Internet tornou-se um dos mais importantes veículos de informação e tem mostrado uma crescente importância no jornalismo, onde ocupa um mercado anteriormente restrito a mídias convencionais como jornais, revistas, rádio e canais de TV.

Diferentemente de um jornal ou revista, o conteúdo de um jornal *online* é virtualmente ilimitado. Um jornal *online* pode facilmente fornecer uma enorme quantidade de informações para os seus usuários, embora isto possa causar algum tipo de desorientação entre os mesmos. Conseqüentemente, existe uma grande necessidade de organizar o acesso a esta informação. Os usuários devem estar aptos a localizar seus principais interesses e serem diretamente direcionados para as principais informações disponíveis de forma rápida e eficiente. Adaptar a primeira página de um jornal *online* serve exatamente para ajudar a resolver este problema de desorientação, fazendo com que o usuário encontre o que deseja da forma mais rápida. Esta é, portanto, uma típica aplicação do conceito de design da informação.

Em mídias convencionais, é muito difícil adaptar o conteúdo de, por exemplo, um material impresso de acordo com a personalidade de cada usuário. Mesmo na Internet, a maioria dos *sites* jornalísticos ainda funciona sem nenhum tipo de consideração específica com as individualidades.

Encurtar o caminho entre o buscar e o achar, não é somente um problema de arquitetura da informação ou organização de conteúdo. Saber o que o usuário quer e saber o que disponibilizar, melhora em muito a experiência com o produto. Em um mundo onde a personalização de serviços ganha cada vez mais valor, toda a possibilidade de adequar melhor o serviço ao usuário é bem-vinda.

Personalização é uma palavra-chave de grande importância para um mundo de produtos e empresas que cada vez mais buscam atender as individualidades e consumidores que querem, e já podem, buscar produtos e serviços mais personalizados.

Pierre Lévy, em seu livro “Cibercultura”, trata sobre os fenômenos sociais e culturais que estão por trás desta nova era técnica que vivemos e de como as pessoas se relacionam com ela. Entre outros aspectos, Lévy comenta sobre como são, em geral, utilizados os meios de comunicação tradicionais (redes de TV abertas, jornais impressos, rádio, etc.), onde se procura utilizar soluções que visam uma “totalização”, com conteúdos ou serviços

que são disponibilizados para todos da mesma forma e de acordo com um “denominador comum” mental. Ele afirma que na Internet, esta realidade não é mais viável:

Por meio dos computadores e das redes, as pessoas mais diversas podem entrar em contato, dar as mãos ao redor do mundo. Em vez de se construir com base na identidade do sentido, o novo universal se realiza por imersão. Estamos todos no mesmo banho, no mesmo dilúvio de comunicação. Não pode mais haver, portanto, um fechamento semântico ou uma totalização.

Uma nova ecologia das mídias vai se organizando ao redor das bordas do ciberespaço. Posso agora enunciar seu paradoxo central: Quanto mais universal (extenso, interconectado, interativo), menos totalizável. Cada conexão suplementar acrescenta ainda mais heterogeneidade, novas fontes de informação, novas linhas de fuga, a tal ponto que o sentido global encontra-se cada vez menos perceptível, cada vez mais difícil de circunscrever, de fechar, de dominar. (Lévy, 1999)

Lévy usa a palavra “universal” como contraste a palavra “totalização”. Ele afirma o “universal” como uma outra maneira na qual se instaura a presença virtual dentro da Internet:

A cibercultura, por outro lado, mostra precisamente que existe uma outra forma de instaurar a presença virtual da humanidade em si mesma (o universal) que não seja por meio da identidade do sentido (a totalidade). (Lévy, 1999)

Podemos afirmar, portanto, que a realidade dos produtos na “grande rede” não deve se ater mais na tentativa de se “totalizar”, ou tentar atender a todos com a mesma solução, mas sim atender as diversidades deste universo sendo flexível e aproveitando o potencial interativo desta nova mídia. Os produtos desenvolvidos para este novo meio devem ser os mais adequados possíveis a esta realidade. As técnicas de hipermídia adaptativa visam exatamente permitir atender melhor às individualidades deste universo.

David Siegel, grande estrategista de negócios para Internet, utiliza uma analogia interessante que mostra o quão importante é o papel do cliente na definição de metas de uma empresa. Ele afirma que por volta de 1850, Charles Darwin encontrou uma orquídea rara com o bulbo de 37 centímetros e concluiu que deveria existir então uma mariposa com a língua gigante de 37 cm para polinizar a planta. Apesar de muitos rirem de sua conclusão, 21 anos depois de Darwin, biólogos descobriram tal mariposa. Baseado nisto ele afirma que: “Na natureza o alimento molda o organismo. Nos negócios o cliente molda a empresa”. Em semelhante afirmação, Siegel também declara:

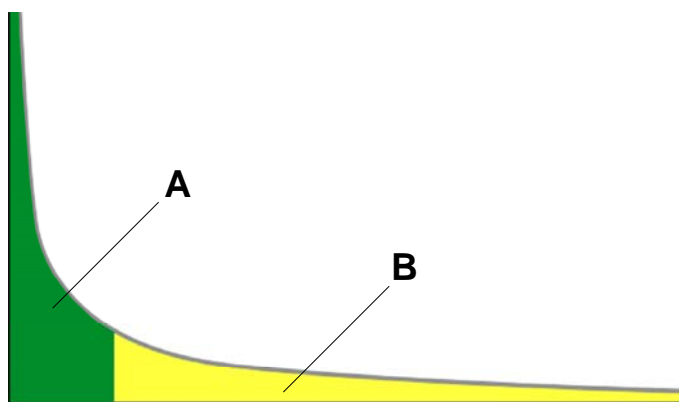
A transição do velho mundo da gerência de negócios para o novo mundo de clientes *Web* é um acontecimento único e bastante singular. Antes do ano 2000, as empresas empurravam os produtos para os clientes que estavam à espera deles. No ano 2000 os clientes “puxam” os produtos e serviços de acordo com a demanda. Em 2010, os compradores, em sua maioria, estarão conectados, completando a revolução conduzida pelo cliente. (Siegel, 1999)

A partir da proposição anterior, Siegel também conclui que: “No Velho Mundo o cliente não tinha escolha. No novo mundo, os clientes têm todo o poder”. Isto mostra que o caminho é realmente deixar o produto ser moldado pelo próprio o usuário. Quanto mais isto for uma realidade, melhor.

A frase “*Content is king*” (Conteúdo é rei), de Jakob Nielsen, afirma que em qualquer solução de design para Internet, o conteúdo é que manda e deve ter prioridade nas decisões de interface. Em um *site* com conteúdo jornalístico, certamente o conteúdo é a estrela principal, pois é isto que o usuário espera encontrar de interessante. O ponto crucial é que, aquilo que é interessante para um usuário, pode não ser para outro. Enquanto para um usuário, o mais importante é o resultado do campeonato carioca de futebol, para outro pode ser o paulista, já para outro o mineiro e para outro ainda, o gaúcho e assim por diante. Ainda que esteja em destaque o campeonato carioca, para determinado usuário o mais interessante pode ser o jogo do time A, para outro o time B e assim por diante novamente. É impossível, pois, com a mesma página, atingir uma qualidade máxima de interesse para cada um dos milhares de usuários.

O termo “*The Long Tail*” (A Cauda Longa), foi publicado pela primeira vez por Chris Anderson em um artigo da revista “*Wired*” em outubro de 2004 para descrever certos tipos de negócios tais como a livraria “*Amazon*” (www.amazon.com), que trabalha vendendo não somente “campeões de venda”, mas também produtos de nichos bem específicos. “*Long Tail*” é também uma expressão comumente utilizada em estatística para identificar distribuições de dados da curva de Pareto, curva esta na qual os volumes de dados são classificados de forma decrescente.

É bastante comum encontrar curvas do tipo “Cauda Longa” para ilustrar a procura dos consumidores no mercado (figura 1). A curva demonstra geralmente uma grande busca por um grupo pequeno de produtos (área “A” na figura 1) em contraste com uma procura bastante reduzida para um grande número de produtos, que em muitos casos soma um



1 - “A Cauda longa” (Área “B”). (Adaptado da Wikipedia)

montante de demanda maior do que os dos produtos mais buscados (área “B” na figura 1). Na economia tradicional, esta curva é importante, pois auxilia a decidir quais são os produtos que valem a pena serem mantidos em estoque em função do custo de manutenção dos mesmos enquanto não são comprados, para que não haja prejuízo.

É possível observar, portanto, que a área em amarelo pode ser tão significativa ou até mesmo maior que a área em verde. Na verdade, em muitos casos, o volume da área amarela será muito maior do que o da área verde. Isto significa que o volume de demanda

dos poucos produtos que muito vendem pode ser bem menor do que a demanda dos muitos produtos que pouco vendem.

Por isso, com relação à nova economia, decidir quais são os produtos que valem a pena serem mantidos em estoque, já não se faz tão necessário. Com a “virtualização”, e a falta de necessidade do mundo físico para algumas das operações de uma loja virtual, o custo de um produto que vende muito e outro que vende menos, pode ser o mesmo.

Tratando-se de um produto digital, que entre outras possibilidades, pode ser um software, um serviço de informação, uma revista ou um jornal *online*, certamente, a diferença entre a manutenção de um produto que vende muito e um produto que vende menos é praticamente nula.

A seguir, um texto presente na “*Wikipedia*” (www.wikipedia.com), enciclopédia digital mantida pelos próprios usuários, sobre o conceito de “Cauda Longa” aplicado a produtos digitais:

Importante destacar que o conceito de Cauda Longa se aplica a praticamente todo mercado, inclusive o mercado de mídia. Com a convergência digital, é muito provável uma reorganização na distribuição da audiência, não apenas por conta de alterações nas características dos meios e na maneira como são consumidos, mas principalmente por alterações no próprio comportamento do consumidor.

A afirmação anterior vem confirmar exatamente o valor que a personalização pode ter em um jornal *online*. Neste caso a personalização seria uma forma de se adequar ao comportamento de um consumidor cada vez mais exigente e específico em sua busca por informações de interesse pessoal.

Neste sentido, outra grande vantagem existente na personalização de um jornal *online* está na venda de áreas destinadas à publicidade para os anunciantes. A partir do conhecimento do perfil do usuário, as publicidades podem ser vinculadas especificamente para os usuários de acordo com seus interesses. Por exemplo, enquanto um usuário visualiza um *banner*, na página principal do portal, com um anúncio de venda da camisa do seu time, outro vê o anúncio com a camisa do seu próprio time e, enquanto isto, um usuário que nunca visita as páginas de futebol, pode ver o anúncio de um outro produto.

Podemos afirmar que a página principal de um jornal *online* é a principal “vitrine” deste, e que as notícias da vitrine são os principais “produtos” do mesmo. No entanto, com centenas de notícias publicadas em todo o *site* simultaneamente, com milhares de usuários interessados em diversos assuntos, mas com poucas dezenas de espaço nesta vitrine, fica bem difícil resolver esta equação, atendendo bem a todos com uma só solução.

No entanto, apenas o fato do conteúdo estar na página, não é o suficiente como ideal. Um outro aspecto importante é que a hierarquia de valor da disposição das notícias na página também corresponda, o mais aproximadamente possível, ao modelo mental de valores que o usuário tem dos seus assuntos favoritos.

A dissertação e pesquisa aqui presentes têm como objetivo, através do estudo do comportamento de navegação do usuário, combinado com o uso de técnicas de hipermídia adaptativa (HA), estudar, descobrir e propor meios para disponibilizar da forma mais antecipada e correta possível, na página principal do jornal *online*, a informação procurada por este usuário e como apresentar outros dados afins que possam ser de seu interesse de modo a enriquecer a experiência incrementando o interesse pelo produto.

1.2. Descrição do trabalho

Para melhor compreensão do conteúdo desta dissertação, faz-se necessário apresentar, nos tópicos seguintes, os principais pontos tratados em cada um dos capítulos da mesma:

Capítulo 2 - Hipermídia adaptativa:

Definição e evolução: São definidos aqui os principais termos e conceitos necessários que são utilizados neste trabalho afim de uma melhor compreensão das partes subseqüentes do mesmo.

Taxonomia de Brusilovsky: Este item do capítulo é dedicado à classificação taxonômica mais referenciada, importante e conhecida, dos métodos e técnicas de HA, feita por Peter Brusilovski.

Aplicações e principais sistemas: Neste tópico são demonstradas as principais aplicações e possibilidades de uso das técnicas de hipermídia adaptativa. Uma maior ênfase é destinada aos três sistemas que têm recebido maior foco de atenção como objetos de pesquisa: hipermídia educacional, sistemas de informação on-line e os sistemas recuperação e busca de informação.

Aplicações em jornalismo *online*: Neste item são apresentadas as principais possibilidades de uso das tecnologias de HA em *sites* de conteúdo jornalístico.

Capítulo 3 - Histórico e características do jornalismo *online*:

Evolução do jornalismo *online*: Uma breve história do jornalismo *online*, contando as primeiras experiências na área até sua chegada e desenvolvimento no Brasil. São descritos, aqui também, alguns casos e projetos que trouxeram um crescimento qualitativo a esta mídia relativamente nova.

Comportamento do leitor: São demonstrados aqui algumas conclusões e estudos sobre o comportamento dos usuários ao ler um jornal *online*.

Diagramação no jornalismo *online*: Apresentam-se aqui as características específicas relativas ao projeto de diagramação e interface da página em um jornal *online*, inclusive comparando-as com aspectos da diagramação do jornal impresso.

Capítulo 4 - Questões éticas do monitoramento do usuário:

Ética em sistemas da informação: São abordadas aqui, as questões relativas à ética necessária para o uso de tecnologias que lidam com dados pessoais monitorados e armazenados em um banco de dados eletrônico.

Política de privacidade: Neste item explica-se o que é política de privacidade de um *site* e são demonstrados os principais pontos sobre o assunto e que devem estar presentes no sistema proposto neste trabalho.

Capítulo 5 - Metodologia para uso de hipermídia adaptativa em jornalismo *online*:

Apresentação geral da metodologia: Apresenta-se aqui, baseando-se nas possibilidades de uso das técnicas de HA, uma metodologia para a adaptação de um *site* com conteúdo predominantemente jornalístico.

Aplicação da metodologia: A forma como foi aplicada a metodologia é demonstrada e justificada neste tópico. As páginas geradas pelo algoritmo criado são testadas com os usuários do *site*.

Avaliação dos resultados: A avaliação aqui demonstrada procura medir o grau de eficiência da versão adaptada em relação à não adaptada (padrão do *site*).

Capítulo 6 - Conclusões e futuros trabalhos:

Baseando-se nos resultados alcançados pela metodologia testada, novos passos são sugeridos como possibilidades de aprimoramento do sistema e são demonstrados outros caminhos a serem estudados para o uso das técnicas de HA no jornalismo *online*.

Bibliografia: Listagem com as referências bibliográficas utilizadas neste trabalho.

Anexos: Formulários e material utilizados na fase de aplicação da metodologia.

2. Hipermídia Adaptativa

2.1. Definição e evolução

Segundo Brusilovsky (2001), a hipermídia adaptativa (HA) é uma linha de pesquisa resultante da combinação entre as áreas de conhecimento de hipermídia e da modelagem de perfil do usuário (Figura 2).



2 - Hipermídia adaptativa como combinação entre hipermídia e modelagem de usuário.

Para melhor entender o significado da expressão “hipermídia adaptativa”, faz-se necessário, definir primeiramente o conceito de “hipermídia” (hypermedia), cuja criação é atribuída a Theodor Holm Nelson (Ted Nelson), que também é criador do termo “hipertexto” (hypertext), ambos criados por volta de 1965.

De acordo com Laufer & Scavetta, Hipermídia é a união de diversos tipos de mídias em um suporte computacional eletrônico. Conforme também a definição utilizada pela W3C (*World Wide Web Consortium*), podemos concluir que o conceito “hipermídia” contém o conceito “hipertexto”, textos contendo vínculos internos a outros textos. Porém o conceito “hipermídia” vai além do “hipertexto” pois, não se restringindo ao texto, abrange também diversos tipos de mídia como gráficos, vídeo e áudio.

Uma outra breve definição para compreensão do que vem a ser hipermídia adaptativa é a do termo “modelagem do perfil do usuário”, ou modelagem do usuário (user modeling).

A modelagem de perfil do usuário é um processo usado em sistemas eletrônicos no qual informações não conhecidas sobre o usuário são deduzidas baseadas nas informações já conhecidas sobre o mesmo, por exemplo, durante o uso do sistema. (ZUKERMAN, ALBRECHT & NICHOLSON, 1999). Modelos de usuários (user models) podem ser criados usando-se as próprias informações recebidas pelo usuário. A semelhança do comportamento de um usuário quando ele começa a usar um determinado sistema, pode ser comparada ao uso de outros usuários e enquadrá-lo em um determinado perfil. O sistema pode então, baseado em experiências anteriores de usuários com perfil semelhante, ou do próprio, formatar o seu conteúdo de acordo com o perfil deduzido. Este perfil pode ser

modificado constantemente de acordo com os usos futuros deste usuário. A modelagem do perfil do usuário pode ser baseada no comportamento deste (behavior-based) ou no seu conhecimento (knowledge-based) (MIDDLETON et AL, 2001).

Quanto à definição do termo hipermídia adaptativa, além da definição um pouco mais geral do pioneiro Brusilovsky (2001), apresentada no início deste capítulo, há ainda outras definições com ênfases que complementam mutuamente o conceito.

Devido ao fato de ser uma nova área de pesquisa, não há uma definição hermeticamente fechada sobre sistemas de HA (PALAZZO, 2000). Este fato abre espaço para algumas tentativas de definição.

Neste contexto, ainda de acordo com Brusilovsky (1996), sistemas de HA são todos os sistemas de hipertexto e hipermídia que refletem algumas características dos usuários em modelos de usuário. São sistemas que aplicam este modelo para adaptar diversos aspectos visíveis do sistema para o usuário.

Segundo De Bra (2000), sistemas de HA constroem modelos de acordo com os objetivos, preferências e conhecimento de cada usuário e modificam este modelo durante a interação, com o objetivo de adaptá-lo de acordo com as necessidades desse mesmo usuário.

Conforme Koch (2000), os sistemas de HA procuram resolver problemas relacionados à desorientação e sobrecarga cognitiva, adotando uma abordagem centrada no usuário. A partir da observação do usuário, o sistema constrói um modelo do usuário e adapta aspectos visíveis do sistema ao mesmo. Isto que evita a sobrecarga cognitiva, pois mostra a informação apropriada muitas vezes com *layout* adequado para cada usuário. Segundo Palazzo:

O objetivo geral dos sistemas e modelos de HA é, portanto, prover seus usuários com informação atualizada, subjetivamente interessante, com a ilustração multimídia pertinentes, num tamanho e profundidade adequados ao contexto e em correspondência direta com o modelo do usuário. Este funciona como uma referência para o sistema, que busca adaptar seu ambiente - um hiperespaço, por vezes caótico - às expectativas particulares de seus usuários. (PALAZZO, 2000, p. 30)

Palazzo também afirma, a partir da definição acima, que os sistemas de HA devem satisfazer a três critérios básicos: ser um sistema hipertexto ou hipermídia, possuir um modelo do usuário, e ser capaz de adaptar a hipermídia do sistema usando tal modelo. Esta afirmação mostra exatamente os elementos mencionados por Brusilovsky: hipermídia, modelagem de usuário e a relação entre os dois, que é o cerne de um sistema de HA.

Dentre as principais desvantagens da hipermídia tradicional (estática) em relação a sistemas de hipermídia adaptativos, está o fato de apresentar o mesmo conteúdo e navegação para toda a diversidade de usuários, não considerando a diferença de seus perfis.

Para um produto de hipermissão tradicional é realmente difícil ser “tudo para todos” (BRUSILOVSKY, 2001), principalmente quando o mesmo deve ser projetado para usuários com diferentes graus de conhecimento, cultura, sexo, faixa etária, entre outras características. Além disso, ele não se adequa de forma ideal para pessoas com diferentes objetivos em um sistema de hipermissão, podendo demandar um maior esforço para conclusão dos objetivos do usuário.

Do ponto de vista do conteúdo podemos citar alguns exemplos de benefícios que sistemas de HA podem trazer a *sites* convencionais (não adaptativos). Enquanto uma livraria on-line baseada em hipermissão tradicional oferece as mesmas ofertas de best-sellers para usuários com diferentes gostos de leitura, uma livraria on-line adaptativa poderia oferecer os livros mais vendidos dos assuntos preferidos do usuário. De forma similar, um museu virtual que oferece a mesma “visita guiada” com a mesma narração para visitantes com diferentes metas e conhecimentos culturais, poderia a partir de dados recolhidos durante a visita, deduzir qual o tipo de assunto que mais interessa ao visitante e apresentar ao mesmo outras opções afins. Isto seria possível se, por exemplo, o sistema verificasse que o visitante sempre fica mais tempo vendo pinturas renascentistas de Da Vinci do que esculturas de Rodin e, em função disto, o sistema poderia sugerir uma visita a uma seção do *site* somente com as obras renascentistas de Da Vinci e até mesmo outros pintores do renascimento.

A linha de evolução histórica de pesquisa sobre sistemas de hipermissão adaptativos apresenta, uma evolução mais perceptível durante os anos 90, principalmente devido ao crescimento da Internet.

No início dos anos 90 as áreas de hipertexto e a de modelagem de usuário alcançaram um nível maturidade, que alimentou pesquisas entre estas áreas na mesma época na qual a Internet começava a chegar às residências familiares.

No entanto, estes primeiros esforços neste sentido eram de poucos grupos independentes, que não atentavam para outros trabalhos. Foi a estabelecida comunidade de pesquisa em modelagem de usuário que auxiliou o encontro destes grupos de pesquisa e promoveu a hipermissão adaptativa como uma linha de pesquisa independente dentro da área de modelagem de perfil do usuário.

De acordo com Brusilovsky, após o ano de 1996, cresce o número de grupos de pesquisa, teses de doutorado e livros sobre hipermissão adaptativa. Paralelamente a isso, o crescimento do uso da Internet com sua diversidade de tipos de usuários faz o interesse pela HA crescer. Enquanto isto, no campo acadêmico, o próprio acúmulo de pesquisas na área faz com que pesquisadores comecem a referir-se a trabalhos anteriores de outros grupos e propor a elaboração ou a extensão de técnicas usadas anteriormente.

Antes de 1996, os sistemas de HA gerados pelos pesquisadores, consideravam exclusivamente as características do usuário para a modelagem de seu perfil. Como exemplos de características do usuário, podemos citar: objetivos, conhecimento, cultura, nível de experiência como usuário de Internet, preferências, Interesses pessoais, características de personalidade.

Após 1996, outros pontos, tais como os dados de uso e do ambiente, começaram a serem considerados importantes para a modelagem de perfil do usuário. Os dados de uso podem ser obtidos a partir de dados armazenados em um banco de dados específico, denominado *data mart* de cliques, que corresponde ao conjunto de dados sobre a seqüência de cliques do usuário, que pode ser alimentado por dimensões como tipo de notícias, data do clique, dados do usuário e hora do clique.

Quanto aos dados ambientais mais significativos para adaptação, podemos citar: hardware (Tipo de Máquina), tamanho da tela do dispositivo, software (Tipo de plataforma ou programa), velocidade de conexão, além da localização física do usuário (onde ele está, para onde ele está olhando e seus movimentos).

2.2. Taxonomia de Brusilovsky

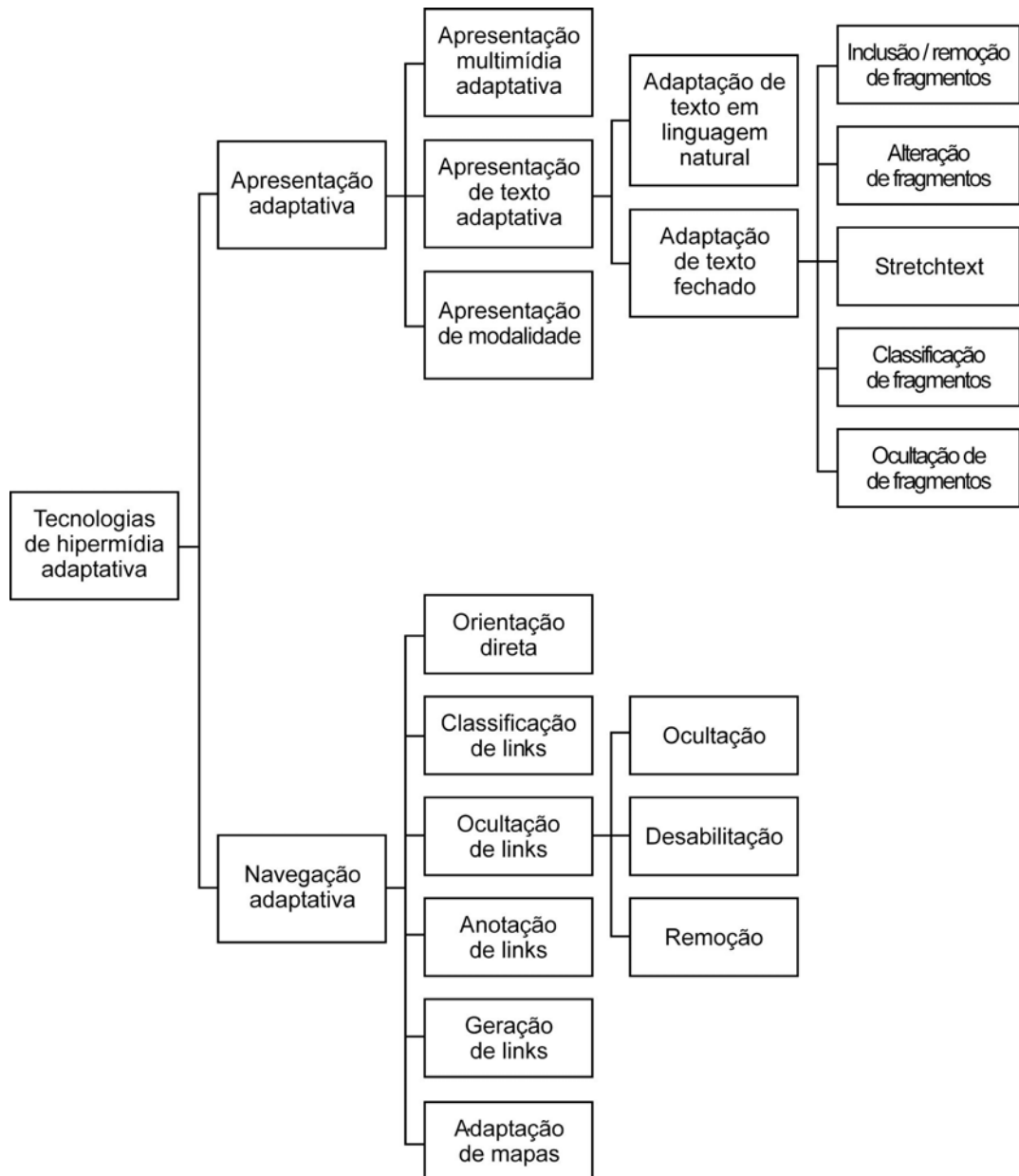
Brusilovsky (2001) propôs uma taxonomia para demonstrar e classificar hierarquicamente métodos e técnicas de hiperímia adaptativa. A classificação divide-se em duas grandes áreas: os métodos de apresentação adaptativa e os de navegação adaptativa (figura 3).

Classificam-se como métodos de apresentação adaptativa aqueles que implicam em alteração do conteúdo das páginas para melhor atender o perfil de determinado usuário. Já os métodos de navegação adaptativa são os métodos que, com o objetivo de melhor guiar o usuário dentro do *site*, modificam as opções de busca dos *links* de forma a apresentar, de acordo com o perfil do usuário, outras opções para continuar a navegação em outras páginas do sistema.

Em relação a esta classificação taxonômica, Brusilovski afirmou que outros tipos de métodos e técnicas podem ser agregados a essa classificação demonstrada, já que, naturalmente, surgiriam novos tipos e aplicações de adaptação.

Palazzo menciona especificamente a “adaptação de *layout*”, que apesar de não estar explicitamente citada na classificação de Brusilovski, estaria implicitamente incluída dentro de “apresentação adaptativa”, pois pode ser considerada como um tipo de adaptação de conteúdo do *site*. Neste trabalho, de busca de formas de criação de um jornalismo *online* individualizado, consideramos o uso da adaptação de *layout* como principal foco.

Na adaptação do *layout*, a personalização pode interferir tanto o posicionamento do conteúdo da página como também a diagramação dos elementos. A proposta de adaptação presente no capítulo 5 desta dissertação trabalha com essas as duas modalidades de adaptação. A intenção é exatamente exibir para o usuário o conteúdo que ele tem interesse e também destacá-lo no layout de acordo com o interesse pelo mesmo.



3 - Classificação taxonômica dos métodos e técnicas de hiperímia adaptativa.

2.3. Aplicações e principais sistemas

São amplas e ainda não totalmente exploradas as possibilidades de aplicação de hipermissão adaptativa. Conforme já mencionado anteriormente, as técnicas de HA podem trazer diversos benefícios aos chamados *sites* convencionais (não-adaptados ao perfil específico do usuário). As principais áreas de utilização de HÁ são:

- **Hipermissão educacional;**
- **Sistemas de informação on-line**
- **Sistemas de ajuda on-line**
- **Sistemas para busca da informação**
- **Hipermissão institucional**
- **Sistemas para gerenciamento de visão personalizados em espaços para informação**

Dos sistemas acima citados, três deles têm recebido maior foco de atenção como objetos de pesquisa. São eles: hipermissão educacional, sistemas de informação on-line e os sistemas recuperação e busca de informação. Um dos principais motivos para este foco específico, é porque os estudos sobre estes sistemas crescem à medida que a Internet, local onde a presença deles é bastante comum, também cresce.

Os sistemas de hipermissão educacional crescem especificamente devido ao interesse pela educação à distância através da Internet. Novas ferramentas de publicação, destinadas a esse conteúdo, surgem paralelamente. Quase a totalidade dos sistemas de HA educacionais desde 1996 foram desenvolvidos em plataforma *Web*.

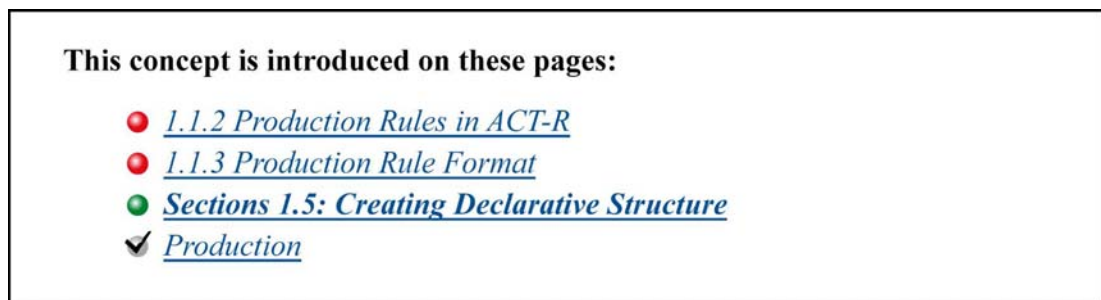
De acordo com De Bra (2002), em um livro impresso convencional, um estudante segue a organização e a ordem específica dos caminhos estabelecidos pelo autor do livro. Para livros publicados na Internet, o leitor espera poder navegar livremente através da navegação provida pelo autor, mas sem esbarrar em dificuldades como, por exemplo, termos desconhecidos para ele.

Devido à necessidade de conhecimento de determinados tópicos ou páginas como pré-requisito para se entender outras partes do livro, não é factível a criação de um livro educacional baseado em um sistema *Web* que possa ser lido aleatoriamente. Exatamente para isso que os métodos de HÁ servem, pois podem ser usados para sinalizar ao usuário que ele ainda não tem o conhecimento necessário para acessar determinado *link*, ou ainda compensar a falta de conhecimento sobre determinado termo, adicionando explicações às páginas visitadas pelo usuário.

Ainda de acordo com De Bra, podemos afirmar que um sistema de hipermissão educacional adaptativo rastreia o comportamento do usuário durante o uso para tentar determinar quais são suas experiências, conhecimentos e interesses. Estes aspectos podem mudar à medida que o usuário navega pelo sistema, que funciona como uma espécie de armazém da hierarquia de conhecimentos adquiridos. Neste caso, o modelo de perfil do usuário

corresponde à armazenagem sobreposta do nível de conhecimento adquirido pelo usuário para cada um dos conceitos existentes. Quando o usuário lê determinada página, o sistema entende que ele está adquirindo conhecimentos sobre os conceitos ali presentes, o que pode, inclusive, ser aferido através de um teste *online*.

O InterBook (BRUSILOVSKY et al., 1998) é um sistema de hipermedia adaptativo educacional que, baseando-se no que o usuário já demonstrou conhecer, sinaliza cada um dos *links* apresentados na página de acordo com esse conhecimento adquirido. Para isto o sistema utiliza uma das mais comuns formas de navegação adaptativa: a anotação de *link* (figura 4).



4 - Anotação de *link* no sistema *InterBook*.

Neste sistema, os *links* marcados com um ponto vermelho direcionam o usuário para páginas as quais o usuário não tem alguns dos conhecimentos previamente requeridos para os tópicos abordados. Os pontos verdes indicam *links* para páginas recomendadas de acordo com o conhecimento já adquirido pelo usuário. O ponto branco indica que o usuário já domina os conhecimentos existentes na página para a qual o *link* aponta. A marcação de “visto”, sobreposta ao ponto branco, mostra que o usuário já leu a página.

É interessante observar que no exemplo da figura 3 acima, a seção 1.5 recomendada para o usuário não corresponde à ordem estabelecida na hierarquia do livro, pois os itens da seção 1.1 não são recomendados ainda. Isto mostra que a navegação pelo conteúdo do livro não precisa obedecer a uma única ordem linear estabelecida para o aprendizado de determinado conceito.

Dentre os recursos apresentados pelo InterBook, está uma lista dos conceitos requisitados para a página que está sendo utilizada, uma lista de conceitos que poderão ser explorados após a leitura da mesma e um botão “teach me” (ensina-me), para gerar uma listagem das páginas necessárias para adquirir o conhecimento da página em questão.

O sistema AHA!, Adaptive Hypermedia Architecture (DE BRA et al., 1999), utiliza a técnica de ocultação de *link*. Nesta técnica os *links* coloridos em azul significam que ainda não foram visitados e estão habilitados para tal, enquanto a cor lilás indica que o *link* está habilitado, mas já foi visitado. Já a cor preta, indica que o usuário não está habilitado a visitar o *link*.

O interessante desta técnica é que ela faz uma analogia cromática com o comportamento padrão das cores dos *links* nos principais softwares para navegação na Internet (*browsers*), que utiliza azul para os *links* não visitados e lilás para os *links* já visitados. De acordo com Owen Conlan (2003), sistemas de hipermídia adaptativo podem estender esta idéia dos *browsers* para mostrar os *links* em três estados. São eles: “já aprendido”, “bem-aprendido” e “desconhecido”. Outros autores sugerem outros possíveis estados para o *links*. Como exemplos: “já visitado”, “não visitado”, “atualmente exibido”, “sugerido” (Eklund, Sawers 98).

No AHA!, existe ainda uma outra técnica adaptativa que pode ser utilizada para suprir a necessidade de uma “explicação adicional” sobre algum conceito básico para a aprendizagem de determinado texto em um curso *online*. Na figura 5 é possível ver uma “inclusão de fragmento de texto”, entre parênteses, para explicar a palavra “Xanadu” na primeira imagem e a remoção do mesmo texto na segunda.

In Xanadu (a fully distributed hypertext system, developed by Ted Nelson at Brown University, from 1965 on) there was only one protocol, so that part could be missing. Within a node every possible (contiguous) subpart could be the destination of a link.

In Xanadu there was only one protocol, so that part could be missing. Within a node every possible (contiguous) subpart could be the destination of a link.

5 - Exemplo do comportamento de um fragmento de texto condicional no “AHA!”.

Diferentemente dos sistemas de HA com objetivos educacionais, um segundo grande grupo, os sistemas de informação on-line têm como objetivo levar o usuário à informações de natureza um pouco mais pontuais e deduzir opções que possam ser de interesse do usuário. No grupo de possíveis aplicações para estes sistemas temos: enciclopédias eletrônicas, quiosques de informação, museus virtuais, guias, comércio eletrônico (E-commerce) e sistemas de auxílio a diagnóstico.

Em enciclopédias eletrônicas o sistema pode deduzir, através da navegação, o interesse do usuário sobre determinado objeto e apresentar comparações entre objetos similares ou ainda deduzir outros interesses através da navegação do usuário e oferecer artigos relevantes. Por exemplo: ao perceber que todos os últimos *links* clicados pelo usuário tinham alguma relação comum com o assunto “dinossauros do período cretáceo”, o sistema pode automaticamente sugerir a visita a uma página especial sobre esse assunto ou ainda sobre assuntos afins. Uma enciclopédia eletrônica pode personalizar um artigo para incrementar ainda mais o conhecimento e interesse do usuário sobre o seu próprio conteúdo.

A personalização em um sistema adaptativo pode ser feita tanto através da criação de perfis individualizados para cada usuário, quanto na criação de perfis pré-estabelecidos (arquetipos de usuário) e adequação do usuário a um destes perfis. Personas é um termo

que tem sido utilizado para definir arquétipos de usuários, quando um sistema qualquer é desenvolvido para identificar e atender grupos de usuário com características próprias.

Um exemplo da aplicação de um sistema adaptativo aplicado em quiosques de informação pode ser encontrado no trabalho de MONAT (1997). O trabalho é um sistema adaptativo desenvolvido para quiosques da cidade de Nova Friburgo - RJ, que tem como objetivo atender aos diferentes tipos de turistas que procuram a cidade. Basicamente, existem três grupos predominantes de objetivos turísticos dentro da cidade.

Enquanto alguns turistas buscam cachoeiras, esportes radicais e caminhadas (“mochileiros”), um segundo grupo busca o patrimônio histórico da cidade e um terceiro grupo, também conhecidos como “sacoleiros”, ficam somente um dia na cidade e vão às compras em busca de roupas.

Durante o uso do sistema, de acordo com a navegação utilizada pelo usuário (número de cliques, páginas visitadas, etc.), o sistema atribui uma pontuação ao mesmo e procura reconhecer em qual dos três arquétipos o usuário se enquadra melhor. O tempo de visitação destinado a cada página do sistema contribui para esta dedução.

Após o sistema deduzir por similaridade o arquétipo no qual o usuário melhor se enquadra, o mesmo sugere um roteiro turístico. Caso o sistema tome uma conclusão errônea, a qual não foi de interesse do usuário, os pesos das variáveis (número de cliques, tempo gasto nas páginas visitadas, etc) se alteram para evitar a repetição do erro. Por exemplo, caso o sistema venha a sugerir um mapa mais apropriado a “sacoleiros” para um “mochileiro” (dado este que pode ser deduzido caso haja uma rejeição), ocorre uma reavaliação e alteração dos pesos que conduziram a uma seleção errada do perfil, para que um novo erro deste tipo seja evitado.

Outro tipo inovador de sistema de informação *online* são os museus virtuais. Com o uso de um sistema adaptativo, existe a possibilidade de oferecer visitas adaptadas e guiadas de acordo com a demonstração do interesse do usuário durante seu caminho virtual. O sistema pode, por exemplo, verificar uma tendência do usuário em buscar e passar a maior parte do tempo vendo “esculturas” e sugerir uma visita à uma galeria especial de esculturas em 3D da arte renascentista.

Este mesmo conceito também pode ser utilizado em um museu real (físico). O tempo que um observador empreende vendo, ou se dedica ao ouvir uma narração (áudio por fone de ouvido) sobre determinada obra de arte pode ser considerado um sinal sobre seus interesses. O sistema pode, por exemplo, verificar uma tendência do usuário em buscar e passar a maior parte do tempo vendo as esculturas abstratas e levá-lo a um caminho virtual especial que sobre o abstracionismo.

No comércio eletrônico, a grande possibilidade de uso de sistemas adaptativos se dá, basicamente, para incrementar as vendas. O sistema, entre outras possibilidades, pode verificar a ocorrências de compras de produtos que os usuários costumam fazer paralelamente e sugerir vendas casadas entre produtos. Por exemplo, ao verificar que muitos usuários que compraram o filme “Sr. & Sra. Smith” compraram também o filme

“Sahara”, entre outros. A figura 6 mostra este exemplo no *site* da Amazon. Na verdade, esta uma aplicação tipicamente utilizada em gestão de relacionamento com o cliente, CRM (Customer Relationship Management), que pode ser considerada como um exemplo de adaptação de sistema relativamente simples se comparada com diversas possibilidades visionárias aqui citadas.

Uma destas possibilidades visionárias é o uso em sistemas de auxílio a diagnóstico. Nesta categoria, podemos enquadrar sistemas que, de forma adaptativa, o auxiliam no reparo técnico de algum produto com defeito, onde o usuário pode responder um questionário com os sintomas e receber como resposta um diagnóstico com a solução, ou as possíveis soluções, destes defeitos.

The screenshot shows the Amazon.com interface for a DVD product. At the top, there are navigation links for 'Your Store', 'DVD', and 'See All 32 Product Categories'. A search bar contains the text 'DVD'. Below the search bar is a promotional banner for a Tanita scale. The main product section is for 'Mr. & Mrs. Smith (Widescreen Edition) (2005)'. It features a product image, a list price of \$29.98, a current price of \$15.98, and a 47% off discount. The page also includes a 'Rate This DVD' section, a 'Favorite Magazines!' section, and a 'Customers who bought this DVD also bought:' section with links to other products like 'Sin City DVD', 'Sahara (Widescreen Edition) DVD', 'Kingdom of Heaven (Widescreen Edition) DVD', and 'National Treasure (Widescreen Edition) DVD'. There are also buttons for 'Add to Wish List' and 'Add to Wedding Registry'.

6 - Sugestão de compra (abaixo do produto) baseada em outros usuários. (www.amazon.com)

De forma um pouco mais ousada e até polêmica, estuda-se também a possibilidade de prover tratamento médico utilizando tais técnicas para o diagnóstico de doenças. Da mesma forma o usuário preencheria um formulário com seus sintomas e o sistema forneceria um diagnóstico adaptado ao perfil do usuário. Neste caso, combinam-se a inteligência da máquina (artificial) e a humana para resolver problemas. Um sistema como este pode auxiliar, por exemplo, a um médico novato.

Um terceiro grande grupo de sistemas que têm recebido grande foco de atenção em pesquisas sobre adaptação são os sistemas auxiliam ao usuário a encontrar informações dentro do infinito espaço da Internet: os sistemas para busca da informação. O uso de hipermedia adaptativa, neste caso, tenta traçar a meta de busca de informação e

recomendar *links* mais relevantes para o usuário sobre o assunto; ordenando, acrescentando ou suprimindo *links* de maior ou menor interesse.

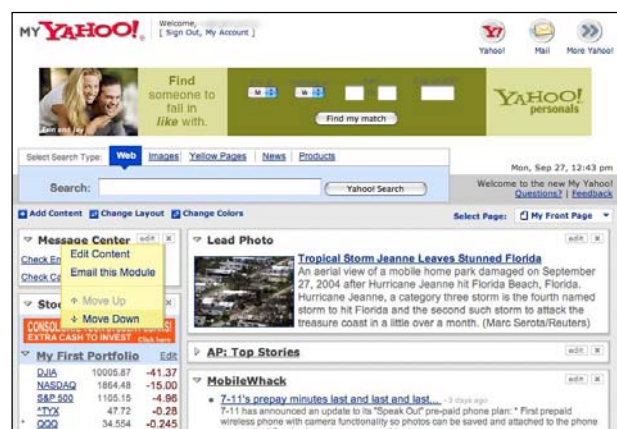
Estes sistemas também podem ser utilizados em um espaço de busca restrito, como um único *Website*, por exemplo. Mas a maioria destes sistemas tem sido desenvolvida para explorar toda a grande rede. Brusilovsky (2001) diferencia dois grandes grupos de sistemas de busca adaptativos: os orientados por busca e os orientados pela navegação.

Os sistemas adaptativos de busca orientados por busca (propriamente dita) se diferenciam dos sistemas tradicionais, pois não levam em consideração somente o conjunto de palavras digitadas para exibir o resultado da busca. Eles também consideram o perfil do atual usuário (interesses e preferências) para melhor hierarquizar os *links* exibidos. Um exemplo deste tipo de sistema está em uma tese apresentada por Mirco Spereta em 2000 (*Personalizing Search based on User Search Histories*), onde se combina os interesses do usuário com os resultados de busca do *site* de buscas “Google” (www.google.com), para um resultado final de busca com maior probabilidade de ser mais interessante para este usuário.

De forma um pouco diferente, os sistemas adaptativos de busca orientados por navegação auxiliam o usuário utilizando técnicas de suporte a navegação adaptativa. Como exemplo, é possível marcar cada *link* do resultado da busca com um sistema de sinalização que indica a relevância daqueles *links* para os objetivos do usuário. Uma outra possibilidade é marcar os *links* através da técnica de anotação de *links*, onde é possível criar e classificar diferentes tipos de marcações para os *links*.

Há ainda outros campos e estudos para adaptação que merecem atenção. Entre eles estão os sistemas para gerenciamento de visão personalizados em espaços para informação. A idéia básica destes sistemas é que cada usuário tenha uma versão personalizada do mesmo para melhor servi-lo. O “My Yahoo” (my.yahoo.com) é um exemplo do tipo. Ele é considerado um sistema adaptável, porém não adaptativo (figura 7).

O exemplo do *site* “My Yahoo” e seu uso atual por uma grande audiência comprovam e também demonstram uma necessidade que já é há tempos percebida tanto pelos usuários quanto pelas empresas de tecnologia: a personalização, ou a adaptação ao usuário. A vantagem das principais técnicas de HA é a de poder fazer isto automaticamente, sem precisar da interferência do usuário, o que diminui seu trabalho e pode melhorar a experiência do mesmo no uso do *site*.



7 - O “My Yahoo” pode ser personalizado manualmente.

2.4. Aplicações em jornalismo *online*

Dentre os sistemas onde são possíveis as aplicações das técnicas de HA descritas por Brusilovski (2001), podemos dizer que os *sites* de jornalismo *online* contêm tanto características de um sistema de informação *online* (enciclopédias eletrônicas, quiosques de informação, museus virtuais, etc), quanto às de um sistema de busca da informação (busca específica dentro de um *Website* ou em toda a Internet). Portanto, as técnicas estudadas para aplicação de adaptação nestes dois tipos de sistema são as mais interessantes para uso em jornalismo *online*. A necessidade do usuário é o que define a técnica a ser utilizada.

O usuário pode entrar no sistema para “saber das notícias”, desejando conhecer as informações que estão disponíveis, o que é um comportamento típico da função de um sistema de informação *online*, ou para “buscar a notícia”, procurando uma notícia sobre um assunto específico, o que é uma função típica de um sistema de busca da informação.

Podemos então afirmar que um bom jornal adaptativo deveria tanto selecionar as informações mais interessantes a serem exibidas (função de um sistema de informação *online*), quanto hierarquizar (ordenar) esta informação dentro da página de acordo com os interesses do usuário. Há ainda sistemas que prevêm a possibilidade de personalizar o detalhamento e tamanho da notícia.

A intenção de aplicação de tais técnicas em jornais eletrônicos é atender interesses específicos de um usuário da forma mais acertada possível dentro do infinito universo de produção de notícias que ganham os banco de dados dos servidores dos *Websites*. Classificar, tratar e exibir estas notícias nos *sites* é o grande objetivo desses sistemas.

Uma outra aplicação possível que não tem relação direta com o tratamento do banco de dados das notícias é a personalização da publicidade no *site*. Através do reconhecimento do perfil do usuário, é possível exibir no espaço publicitário do *site*, um anúncio especificamente feito para o mesmo. Isto significa que ao mesmo está sendo exibida a publicidade de um creme hidratante para um usuário do sexo feminino, é possível exibir o anúncio de um aparelho de barbear para os usuários do sexo masculino. Como outro exemplo, pode ocorrer que, ao mesmo tempo em que a área publicitária exibe a oferta de uma camisa de um clube de futebol para um torcedor deste clube, exibe a venda de camisa de outros clubes para seus respectivos torcedores.

Apesar de serem extremamente promissoras tais possibilidades, ter uma página especialmente feita para cada usuário, utilizando uma inteligência de máquina, ainda não é uma realidade para os grandes jornais *online* do mundo. No entanto, mesmo sendo uma área de pesquisa em crescimento, existem alguns esforços no sentido de se chegar a um modelo de adaptação para jornais *online*.

Neste sentido e baseando-se na afirmação que a personalização é uma das chaves para o sucesso nos serviços disponibilizados via Internet, Liliana Ardissono, Luca Console e Ilaria Torre da universidade de Turim apresentam em 2001 o SeAN (*Server for Adaptive News*),

Servidor para notícias adaptativo. Trata-se de um sistema adaptativo para acesso personalizado a servidores de notícias na Internet. O SeAN concentra-se em três formas de adaptação:

1º tipo: O sistema pode selecionar apenas os tópicos (assuntos ou editorias) e as notícias que sejam de interesse do usuário, mas permitindo também o acesso a outras notícias.

2º tipo: O sistema pode adaptar o nível de detalhamento da notícia publicada de acordo com a experiência e interesse do usuário.

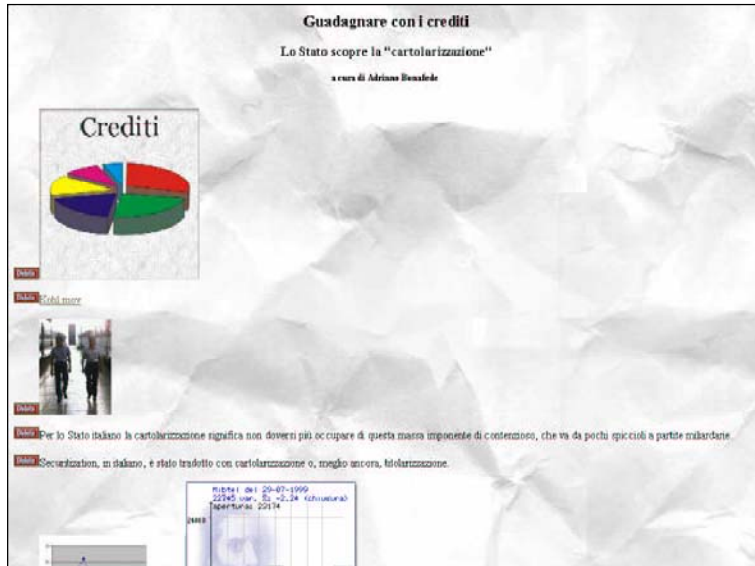
3º tipo: O anúncio publicitário exibido na página deve ser apropriado ao usuário e também ao contexto da página onde o mesmo está inserido.

Como premissa, o sistema define que a adaptação deve atender tanto os usuários que usam o sistema pela primeira vez quanto aos usuários que o usam frequentemente. Por isso, o SeAN gera um modelo inicial do usuário, quando este se conecta pela primeira vez ao *site* e preenche um formulário com perguntas específicas e gerais. Durante o uso, o sistema deve então ser capaz de rever o modelo deste usuário e modificar este modelo, percebendo suas possíveis mudanças de interesses. Porém, o sistema não deve impor suas escolhas ao usuário, permitindo a este modificar suas próprias configurações com relação às editorias e detalhamento das notícias desejadas, o que também auxilia o sistema a ter acesso a suas características e interesses mais recentes. Nenhum dos assuntos (editorias) fica inacessível, o que acontece é que eles podem ser omitidos ou ativados a qualquer momento pelo usuário. Na figura 8, temos um exemplo de um experimento do sistema com as editorias de economia e política ativadas e outras, como a de esportes, desativadas na página principal. As editorias podem ser retiradas ou adicionadas na página através de um sistema de seleção por “pull down” ao final da mesma com as funções de retirar (*Elimina*) ou adicionar (*Aggiungi*).



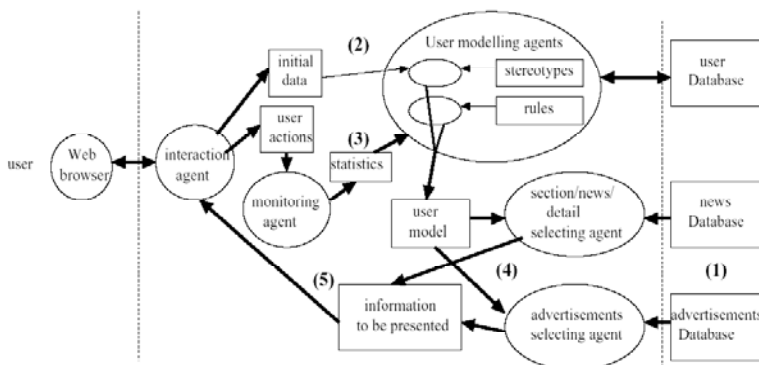
8 - Exibição do sistema com as editorias acessíveis: economia e política. (SeAN.)

A figura 9 mostra a página de notícia personalizada. Apenas itens relacionados com a notícia são demonstrados. Estes itens podem ser removidos através de um botão (*delete*) ao lado esquerdo da notícia. Os itens também podem ser explorados através de um simples clique sobre o título.



9 - Exibição da página de notícias personalizada.
(Página que utiliza o SeAN)

A organização do conteúdo deste sistema no servidor se dá através de uma arquitetura do mesmo (figura 10). O banco de dados de notícias contém todas as matérias que podem ser exibidas para o usuário. Nele as notícias armazenadas são classificadas por assunto e nível de detalhamento. Diversos agentes trabalham para que, a partir dos bancos de dados existentes (notícia, anúncios publicitários e usuário) seja exibida a página personalizada.



10 - Arquitetura do sistema SeAN.

Para definir a ordem em que são exibidas as editorias, sub-editorias e notícias, o sistema considera a mesma ordem que o jornal original as exhibe, não exibindo apenas os assuntos rejeitados pelo usuário.

A modelagem do usuário no SeAN é baseada em quatro características:

Interesses: Esta dimensão armazena os interesses do usuário em relação aos tópicos correspondentes as editorias no banco de dados de notícias.

Experiência: Esta dimensão guarda a experiência (conhecimento) do usuário sobre os tópicos relativos as editorias no banco de dados de notícias.

Características de cognição (retenção da informação): Esta dimensão define os parâmetros correspondentes às capacidades cognitivas, ou de retenção, do usuário. Isto permitirá ao sistema determinar a quantidade de detalhes que o usuário pode assimilar ou reter.

Estilo de vida: Esta dimensão captura as preferências correspondentes aos diferentes estilos de vida dos usuários.

Para criar o modelo inicial do usuário, o SeAN trabalha com estereótipos (*personas*) que representam tipos de consumidores. Os dados para definição de qual estereótipo é o mais próximo do usuário são recolhidos através de um breve questionário de registro, no qual o usuário responde perguntas sobre seus dados pessoais básicos (idade, sexo, nível de escolaridade,...), profissão, se acessa as notícias do trabalho ou não, as coisas que mais gosta de fazer (viagens, esportes, ir as compras,...) e afins.

Apesar do formulário não ser extenso, o sistema não obriga o usuário a preenchê-lo totalmente, podendo até mesmo não ser preenchido por desistência do mesmo. Caso não tenha a quantidade de dados suficientes para classificá-lo em um dos estereótipos disponíveis, o sistema poderá defini-lo como um “usuário padrão”.

O sistema considera diferentes estereótipos para cada uma das quatro características (dimensões) nas quais o usuário possui uma classificação (interesses, experiência, características de cognição e estilo de vida).

Cada uma destas quatro dimensões é utilizada para definir o estereótipo do usuário da seguinte forma:

Interesses: A partir de dados como idade, sexo, profissão, motivos principais de acesso a Internet e das coisas que gosta de fazer, o usuário será enquadrado em um estereótipo que define qual é o seu nível de interesse em relação às editorias e subeditorias das notícias disponíveis nos servidores.

Experiência: Dados como o grau e tipo de educação e a profissão, auxiliam o sistema a definir o estereótipo que melhor combina com o conhecimento do usuário sobre os tópicos correspondentes as notícias disponíveis.

Características cognitivas: Este grupo de estereótipo deduz a receptividade do usuário em relação à quantidade de informação que o mesmo tem interesse em receber. Os dados de nível de educação, idade e frequência de acesso a Internet são utilizados para definir o nível de detalhamento das notícias que serão exibidas para este usuário.

Estilo de Vida: Os estereótipos aqui disponíveis classificam o usuário de acordo com sua posição social e prioridades. A classificação do usuário dentro de um dos estereótipos irá, neste caso, definir que anúncios publicitários serão exibidos para o mesmo.

Na figura 11 podemos ver com é definido o modelo de perfil do usuário, quando este é enquadrado no estereótipo de interesse como um “profissional do setor financeiro”. Este modelo é dividido em duas partes: o perfil do usuário em si e as predições de interesse. O sistema armazena dados estatísticos. É possível, por exemplo, observar ao ler a linha - *reason of connection: work: 0.9; personal: 0.1* - que o usuário se conecta mais vezes para tarefas relacionadas ao trabalho (90%) do que por motivos pessoais (10%). Já na área de predição de interesses podemos ver dados, gerados pelo sistema, que podem ser utilizados como estatísticas para exibir o conteúdo do jornal *online*. É possível ver, por exemplo, que o sistema considera o interesse deste usuário por economia sempre alto (*economy: high:1*). Já em política, a probabilidade deste interesse ser alto é de 70% (*politics: high: 0,7*).

FINANCIAL PROFESSIONAL:
profile:
age: <20: 0; 20-25: 0.1; 26-35: 0.2; 36-45: 0.3; 46-65: 0.3; >65: 0.1
gender: Male: 0.8; Female: 0.2
job: manager: 0.57; self-trader: 0.3; self-employed: 0.05; ...; student: 0.01
job field: {financial, banking insurance}: 0.8; {politics, law, civil services}: 0.14; ...
reason of connection: work: 0.9; personal: 0.1
hobbies - theatre: a lot: 0.1; some: 0.3; a little: 0.4; not at all: 0.2;
hobbies - watching sports: a lot: 0.4; some: 0.3; a little: 0.2; not at all: 0.1;
 ...
predictions on interests:
economy: high: 1; medium: 0; low: 0; null: 0
politics: high: 0.7; medium: 0.3; low: 0; null: 0
sport: high: 0.2; medium: 0.4; low: 0.3; null: 0.1
culture: high: 0; medium: 0.2; low: 0.5; null: 0.3
technology: high: 0; medium: 0.3; low: 0.6; null: 0.1
 ...

11 - Estereótipo de interesse usuário. (SeAN)

O SeAN, apesar de seu protótipo visualmente simples, tem um sistema que usa as estatísticas de forma bem objetiva para alcançar a personalização. Ao serem testados alguns de seus conceitos, obteve resultados encorajadores. O projeto tem também como grande interesse o fato de ter sido pensado desde o início como um *site* adaptativo, o que é um fato ainda raro.

Na proposta de adaptação presente no capítulo 5 desta dissertação propusemos-nos a experimentar uma metodologia para adaptação de um *site* jornalístico que possui algumas diferenças em relação ao SeAN. A primeira é que, no caso de nossa proposta, a adaptação é feita em cima de um *site* já existente: a página principal do portal Globo.com (www.globo.com), de conteúdo predominantemente jornalístico. O portal Globo.com possui uma classificação de notícias que permite um detalhamento de assuntos mais profundo, de forma que podemos não apenas levantar o interesse do usuário nas macro-editorias (esportes, entretenimento e notícias), mas também descobrir e utilizar interesses mais específicos (assuntos mais lidos, clube de futebol favorito, novela preferida, séries de TV