



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ADMINISTRAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO
MESTRADO EM ADMINISTRAÇÃO - UFRGS / Urcamp



Um estudo sobre a demanda de informações em *sites Web*:
o caso de uma unidade de pesquisa de uma empresa de P&D
para o agronegócio brasileiro

Elaborado por
Ricardo Martins Bernardes

Projeto de Dissertação apresentado ao
Curso de Mestrado em Administração da
UFRGS, Escola de Administração,
Programa de Pós Graduação em
Administração, em convênio com a
Urcamp.

Orientador: Professor Henrique Freitas, Dr

Bagé-RS, dezembro de 1999.

SUMÁRIO

GLOSSÁRIO	
3	
RESUMO	
5	
1 INTRODUÇÃO	
6	
2 OBJETIVOS	
9	
2.1 Objetivo Geral.....	9
2.2 Objetivos Específicos	9
3 REVISÃO DE LITERATURA	10
3.1 Rede, Estratégia e Sobrevivência.....	10
3.2 <i>Internet & Web</i>	11
3.3 Números que dobram	13
3.4 A resposta das Empresas	16
3.5 Usuários, suas trilhas e comportamento	19

4	MÉTODO DE PESQUISA	22
4.1	Nível e <i>design</i> da pesquisa	22
4.2	Investigações a realizar na busca do objetivo de pesquisa	22
4.3	Descrição das fontes de dados, coleta, análise e interpretação dos dados	22
4.3.1	Registro de acessos redirecionados por outros <i>sites</i>	22
4.3.2	Registro de acessos de páginas eletrônicas	23
4.3.3	Registro de expressões e termos de busca utilizados pelos usuários do <i>site</i>	23
4.3.4	Registro de expressões e termos de busca utilizados pelos leitores do informativo externo	24
4.3.5	<i>E-mails</i> contendo solicitação de informações ou documentos	24
4.4	Esquema da pesquisa e considerações sobre seu enquadramento metodológico.....	25
5	O CONTEXTO DO ESTUDO	27
6	CONTRIBUIÇÕES POTENCIAIS	33
7	CRONOGRAMA DE EXECUÇÃO	35
	REFERÊNCIAS BIBLIOGRÁFICAS	36

GLOSSÁRIO

Backbone - estrutura de comunicações - física e lógica - que possibilita a interligação de redes de comunicações de diferentes organizações/países/continentes.

Banners - publicidade exibida nas páginas *Web*, geralmente utilizando-se de recursos gráficos animados ou não.

Browser - é o programa que “localiza” e exibe uma determinada página eletrônica no computador requisitante. É a parte *cliente* do processo (Netscape, MS-Explorer, HotJava, etc) .. O HTTP é a parte *servidor*.

Cookies - registro de preferências dos usuários ao navegar por um *site Web*, que reside no computador do próprio usuário.

Extranet - rede que utiliza a mesma tecnologia da *Internet*, porém com acesso restrito

a algumas organizações que se relacionam (fornecedores, principalmente).

Gateway - qualquer computador ou roteador intermediário entre o computador acessado e o computador que acessa.

Hits - unidade de contagem de acessos a páginas eletrônicas. Muito utilizada para se referir a popularidade do *site*.

HTML - Hyper Text Markup Language - linguagem utilizada para a elaboração de páginas eletrônicas.

HTTP - Hyper Text Transfer Protocol – protocolo que gerencia os acessos às páginas de um *site*. é a parte *servidor* do processo. O *cliente* é representado pelo *browser*.

Hypertexto - é uma metáfora para a apresentação de informações nas quais textos, imagens, sons e ações fiquem interligados em uma teia complexa e não linear de associações que permitem ao usuário percorrer assuntos interrelacionados independentemente da ordem que os tópicos são apresentados. O termo foi criado por Ted Nelson em 1965. Atualmente o termo *hypermedia* tem sido usado em razão dos componentes não textuais do hypertexto, como som, vídeo, imagens e etc. (Microsoft Press dicionário de informática)

Intranet – rede que utiliza a mesma tecnologia da *Internet* (TCP/IP), porém com acesso restrito aos membros de uma mesma organização.

ISDN - *Integrated Services Digital Network*. uma rede mundial de comunicações eletrônicas que evoluiu a partir dos serviços telefônicos existentes. O objetivo da ISDN é a total digitalização do sistema de comunicações, sem a necessidade qualquer conversão de sinais.

Kbps - *Kilo bits per second*. Milhares de bits por segundo. Unidade utilizada para medir a taxa de transferência de dados por segundo entre dispositivos e redes. Cada kbps equivale a 1.024 bits (128 caracteres).

Links - elementos de ligação que compõe o hypertexto, permitindo a navegação entre diferentes documentos.

Log - Registro de transações efetuadas por sistemas de computador. Estes registros podem dizer respeito aos processos internos do sistema, bem como de processos de interação do sistema com o mundo exterior.

Mbps - *Mega bits per second*. Milhões de bits por segundo. Unidade utilizada para medir a taxa de transferência de dados por segundo entre dispositivos e redes. Cada mbps equivale a 1.024.000 bits (128.000 caracteres).

MS-Windows - Sistema Operacional da Microsoft, para computadores pessoais (PC).

OpenWindows – Interface gráfica para sistemas Unix.

Proxies - semelhante aos *gateways* os *proxies* possuem funções específicas entre o cliente e o servidor. É utilizado normalmente para manter uma memória de todo o conteúdo que circula nas linhas de comunicações a fim de diminuir o tráfego de dados (*cache*).

SQL - Structured Query Language. Linguagem padrão para operações em sistemas gerenciadores de banco de dados (SGBD).

TCP/IP - Acrônimo de *Transport Control Protocol/Interface Program*, um protocolo de software base da *Internet*, desenvolvido pelo Departamento de Defesa dos Estados Unidos para a comunicação de computadores.

URL - Uniform Resource Locator. Como são chamados os endereços *Web* (<http://>, <ftp://>, etc).

Wireless - tecnologias para comunicação sem a utilização de cabos físicos.

X-Windows – Interface gráfica para sistemas Unix.

RESUMO

Será proposto um método para verificar a existência de conhecimentos úteis para a organização nos registros de acessos realizados (*logs* de transações) em um site na rede *Web*. Para isso, será realizado um estudo de caso longitudinal onde serão analisados os *logs* de acessos de um *site* de uma organização de P&D para o agronegócio, coletados ao longo de 1 ano. Considerando o contexto atual, onde cada vez mais as organizações tenderão a buscar e interagir com clientes e parceiros através da *Web*, pode-se supor que estudo daqueles dados pode revelar aspectos úteis relacionados a oportunidades de negócios e as atividades de *marketing* da organização em estudo. É possível que este conhecimento possa ser transformado em vantagem competitiva.

1 INTRODUÇÃO

Nos últimos anos temos assistido um crescimento extremamente acelerado da capilaridade da rede *Internet*. Esta capilaridade está dependente da estrutura de comunicações (*backbone*), da proliferação de provedores de acesso (*gateways*) e na facilidade de aquisição de microcomputadores pessoais, principalmente.

É o crescimento da capilaridade rede *Internet* que possibilita a expansão da rede *World Wide Web*. Enquanto a primeira estabelece os conceitos de como a informação trafega pela rede, a segunda estabelece os padrões de apresentação da informação ao usuário. O protocolo base da *Internet* foi criado em 1974 por Vinton Cerf e Bob Kahn. Já o protocolo base da *Web* foi criado em 1990 por Tim Bernes-Lee. Uma distância de 16 anos as separam. Em resumo, a *Web* navega pela a *Internet*.

A rede *Web* tirou a *Internet* do meio técnico-científico para torná-la útil também às pessoas comuns, às empresas e organizações, aos governos, enfim, para a economia mundial. E isto tudo a uma velocidade espantosa. Uma pesquisa recente estima a existência de oito milhões de *sites* em todo o mundo¹. Poderia ser pouco, não fosse o fato disto ter se tornado possível para o cidadão comum a apenas pouco mais de 5 anos.

Pode-se afirmar, também, que a *Web* contribui para o fenômeno da globalização uma vez que ela pressiona as barreiras econômicas, sociais e políticas que impedem a livre

¹ <http://www.netcraft.com/survey/>, consulta em 10/1999.

troca de tecnologias, serviços e produtos entre países, homogeneizando as regras de mercado em âmbito global. Por sua vez, é razoável admitir que a globalização represente um dos fatores que acirram a competitividade.

Com a palavra competitividade inserida firmemente no contexto empresarial dos últimos anos, estamos assistindo uma atenção crescente das empresas para com sua presença na *Web*. Pesquisas do IDC Columbia, estimam que 90% dos negócios da América Latina tinham ou estavam desenvolvendo um *Web site* em finais de 1998; 80% tinham ou estavam desenvolvendo sua *Intranet*; 45% tinham ou estavam desenvolvendo sua *extranet* e 44% tinham ou estavam desenvolvendo *sites* com comércio eletrônico².

O comércio eletrônico, em particular, sinaliza para a mudança do conceito tradicional de “horário de expediente” das empresas. Os espaços virtuais são concebidos para funcionarem independentes da presença de pessoas no ambiente da organização. Desta maneira, nesta nova mídia as empresas nunca fecham, ficando disponíveis e acessíveis para que clientes obtenham informações de seus produtos e serviços, efetuem compras, transações, reclamações e sugestões durante as 24 horas por dia, sete dias por semana. A presença das empresas na *Web* tornou-se obrigatória, deslocando a competitividade também para aquele espaço.

Na esteira do movimento da qualidade total vivenciado pelo País desde inícios da década de 90, as empresas procuram agora atingir e satisfazer o consumidor/cliente também na *Web*. Tornou-se comum a prática das empresas divulgarem em sua publicidade também seus endereços na *Internet*. Observa-se que esta tendência abrange todo espectro de organizações, sejam elas privadas ou públicas, filantrópicas, comerciais, industriais, de serviços, de ensino, de pesquisa e etc.

Diante desta tendência, os *sites Web* têm experimentado significativas melhorias no sentido de tornarem-se adaptativos, agradáveis e chamativos às pessoas. Também tornaram-se obrigatórios o uso de mecanismos de captura de informações dos visitantes, sejam dados cadastrais, sócio-econômicos, preferências, etc, sejam suas impressões a cerca dos produtos e serviços da empresa. Ambientes sensíveis a escolhas realizadas pelos usuários durante a navegação pelos *sites* tornam-se cada dia mais sofisticados. Isto inclui *banners* dirigidos, utilização de *cookies*. Agentes e *robots* representam, também, um dos focos da pesquisa acadêmica relacionada a *Web* e sua interação com os usuários.

Neste sentido, muitas pesquisas têm sido realizadas com o objetivo de investigar como recuperar informações relevantes, qual o comportamento do usuário da *Web* e quais suas estratégias de busca de informação. Os fundamentos destas pesquisas se baseiam, fortemente, nas linhas de pesquisa da multimídia e da inteligência artificial, tendo como principais interessados as áreas de informações e de *marketing* das organizações. Neste

² <http://www.linkset.co.uk/pstats.htm>, consulta em 13/10/1999.

estudo, as premissas gerais são:

a) pelo fato de ser a rede *Web* um espaço universal onde habitam todos os tipos de interesses, pode-se admitir que existe demanda por informações de qualquer natureza;

b) ao procurar informações na rede o usuário deixa “rastros” de sua navegação, que são capturados e registrados de forma automática pelos servidores de páginas dos *sites* acessados (W3C, 1999). Os dados deixados pelos usuários podem incluir qual mecanismo de busca ou *site* e qual a palavra-chave o levou a determinada página da organização. Uma vez dentro do *site* da organização, o usuário pode seguir *links* ou utilizar-se de novos mecanismos de busca - quase sempre restritos ao âmbito do *site* - para obter a informação desejada. Nos dois casos ele, novamente, deixa registros que podem incluir sua “trilha” pelo *site* e as palavras-chave sobre as quais ele julga que encontrará a informação desejada. Além disso, a origem do acesso, ou seja, a organização ou localização geográfica do usuário, assim como dados de data e hora de acesso a cada página podem ficar registradas no *site* da organização acessada;

c) considerando o contexto atual, onde cada vez mais as organizações tenderão a buscar e interagir com clientes e parceiros através da *Web*, pode-se supor que estudo daqueles dados pode revelar aspectos úteis relacionados a oportunidades de negócios e as atividades de *marketing* das empresas. É possível que este conhecimento possa ser transformado em vantagem competitiva.

Esta pesquisa tem por objetivo analisar os registros de navegação de usuários através de um *site* com a finalidade de extrair conhecimento útil para a organização que o mantém. O estudo será efetuado em uma unidade de negócios de uma organização governamental de pesquisa e desenvolvimento para o setor agropecuário, localizada em Bagé, no Estado do Rio Grande do Sul, Brasil. O centro de pesquisa em estudo tem a missão de “Atender com soluções tecnológicas as necessidades dos sistemas produtivos integrados ao agronegócio de bovinos e ovinos na Região Sul em benefício da sociedade”. Para isto, serão analisados registros automáticos de acessos ao *site* da organização (*logs*), capturados ao longo de um ano. Serão também consideradas as requisições de informações, via mensagens eletrônicas (*e-mail*) recebidas pelo *site*.

Adicionalmente, pretende-se discutir a aderência entre as informações demandadas através do *site Web* analisado e seu alinhamento com a proposta estratégica da organização.

É possível que a análise dos dados revele algum conhecimento útil para a organização, não somente para a configuração do *site*, mas também como subsídio para suas estratégias de negócios e para sua efetiva orientação ao mercado. Caso a investigação revele conhecimentos relevantes, estes poderão ser utilizados em seus processos de planejamento das atividades de *marketing* e P&D.

Procurar-se-á mostrar que a *Web* é uma poderosa ferramenta para promover a imagem da organização, bem como para a estabelecer um canal de comunicação eficiente e um ambiente propício para a satisfação do cliente. Para que a ferramenta possa apresentar

resultados eficazes, entretanto, é necessário que ela esteja alinhada com o negócio da organização. Este alinhamento é possível, desde que se observem os preceitos de *marketing*, efetuando pesquisas junto aos clientes e usuários a fim de investigar os processos utilizados para a busca e a recuperação de informações, assim como suas expectativas e suas necessidades de informações. Para isto, serão propostos métodos e técnicas de análise que podem ser aplicadas aos dados, de maneira a se obter uma visão agregada dos processos de recuperação de informações em um servidor de páginas para a rede *Web*.

Para isto, o capítulo 2 apresentará os objetivos gerais e específicos deste estudo. Na revisão de literatura - capítulo 3 - o tema será desenvolvido em uma abordagem *top-down*, ou seja, serão apresentadas impressões de pessoas expoentes a respeito do fenômeno “rede *Web*”. A seguir serão exibidos dados históricos de seu surgimento, algumas projeções sobre seu crescimento e a reação das empresas. Finalmente, será abordado o tema principal que diz respeito a análise de registros de transações em servidores de informação (*sites*). O capítulo 4 abordará aspectos metodológicos do estudo e, no capítulo 5, será abordado seu contexto de aplicação. Na seqüência, o capítulo 6 apresentará os benefícios esperados com o estudo e, para finalizar, será apresentado o cronograma de execução da pesquisa.

2 OBJETIVOS

2.1 Geral

- propor um método para verificar a existência de conhecimentos úteis para a organização nos registros de acessos realizados (*logs de transações*) em um *site* na rede *Web*.

2.2 Objetivos Específicos

- coletar, analisar e discutir os registros de navegação de usuários através das páginas de um *site*, procurando definir suas necessidades e expectativas de consumo de informação;
- coletar, analisar e discutir os registros externos que levaram o usuário ao *site*, verificando sua adequabilidade à estratégia da organização em estudo;
- aplicar e validar procedimentos para extrair conhecimento útil dos registros de

navegação de usuários através de um *site*.

- examinar, analisar e discutir solicitações de informações encaminhadas por sistema de *e-mail* às áreas de comunicação e negócios e biblioteca;
- aplicar e validar métodos e técnicas estatísticas que permitam a elaboração de uma síntese dos dados coletados.

3 REVISÃO DE LITERATURA

Neste capítulo serão expostos alguns conceitos e estudos relacionados com o objetivo desta pesquisa. Procurar-se-á caracterizar a presença na rede *Web* como crucial para a sobrevivência das organizações atuais. Para isto serão expostos alguns conceitos básicos que tornam a existência da rede possível, dados de seu crescimento, suas tendências e suas ligações com *marketing*. Serão também apresentadas impressões recentes de pessoas atuantes nas áreas de administração e informação.

3.1 Rede, Estratégia e Sobrevivência

Nota-se que cada vez mais a presença na rede *Web* é colocada como questão de sobrevivência para as organizações atuais. Esta tendência é corroborada pela visão de teóricos e práticos, especialistas e empresários. Os cenários previstos por diferentes visões partilham um aspecto em comum: a presença na rede.

Especialistas em estratégia, como OHMAE (1998), consideram que estamos vivendo na

Era da Informação, e que a capacidade da empresa em atingir os consumidores através da *Internet* é essencial. Ohmae coloca ainda que com o comércio eletrônico tornando-se realidade, a empresa não precisa estar no mercado para atingir seus clientes e consumidores. Adverte, entretanto que “...para que uma companhia sobreviva a era digital deverá entender a tecnologia, as redes e, acima de tudo, a psicologia de seus principais clientes mundiais...” .

Na mesma lógica, PENZIAS (1998) coloca o advento das redes como a mudança mais importante ocorrida nas últimas décadas. Em suas palestras afirma que a utilização das redes evoluirá além das fronteiras corporativas, criando o que chama de “empresa expandida”. Sobre sua importância para as empresas atuais, Penzias diz: “...cada conjunto significativo de dados, cada pessoa e provavelmente quase todas as máquinas estão ligadas em alguma forma de rede. Portanto, as empresas que explorarem essas redes para encontrar pessoas e informações serão certamente as vencedoras...”. Esta percepção o levou a trocar o título de suas palestras de “Administração em um mundo de alta tecnologia” para “Administração em um mundo ligado em rede”.

Para CASTELS (1999) a tecnologia da informação é uma nova revolução tecnológica, um novo paradigma que possui algumas características tidas como aspectos centrais, que no conjunto formam a base material da sociedade da informação. A primeira característica do novo paradigma é que a informação é sua matéria-prima: são tecnologias para agir sobre a informação, não apenas informação para agir sobre a tecnologia, como foi o caso das revoluções tecnológicas anteriores. As outras características são a lógica de redes, a flexibilidade exigida às organizações em razão da lógica de redes, e a crescente convergência de tecnologias específicas para um sistema altamente integrado.

Ainda na mesma linha, TAPSCOTT (1999) afirma que as empresas e os países que não conseguirem administrar a transição para uma nova economia e uma nova tecnologia estarão em perigo. Especificamente no caso das empresas, Tapscott considera que “...as que não conseguirem se transformar em organizações em rede e forem incapazes de criar comunidades de comércio eletrônico deixarão de ser competitivas e definirão até desaparecer...”. Prevê ainda que no ano de 2005 cerca de um bilhão de pessoas terão acesso a *Internet* . TAPSCOTT & CASTON (1995) já haviam formulado uma escala evolutiva para empresas, onde um dos estágios seria a Empresa Expandida, resultado da computação interempresa obscurecendo as linhas divisórias entre organizações e viabilizando novas formas de relacionamentos comerciais. Segundo os autores, “...as redes estão sendo ampliadas... transformando a natureza das interações comerciais e levantando questões de extrema relevância sobre as estratégias de negócios”.

Por último, GATES (1999), um dos ícones empresariais deste final de século, prevê que na próxima década a velocidade com que uma empresa reúne, administra e usa a informação determinará se ela será vencedora ou perdedora. Afirma que os “... os vencedores serão aqueles que desenvolverem um “sistema nervoso digital”... de tal forma que a informação possa fluir com facilidade através de suas empresas para um máximo e constante aprendizado”.

3.2 *Internet & Web*

Neste tópico serão abordados alguns detalhes, números e aspectos históricos da criação e evolução da *Internet* e da *Web*. Embora o assunto esteja amplamente disseminado pela literatura acadêmica e pelos veículos de comunicação populares (jornais, revistas, televisão e na própria *Internet*), alguns aspectos considerados mais importantes, entretanto serão enfocados.

Sabe-se que a expansão da *Internet* tem se dado de forma bastante acelerada sendo, atualmente, um fenômeno global. A proliferação de provedores de acesso, as possibilidades de acesso via satélite, via estrutura de TV a cabo, as melhorias nas malhas físicas tradicionais de comunicação e o barateamento dos microcomputadores estão contribuindo para o aumento da sua capilaridade e, conseqüentemente, para que cada vez mais pessoas se liguem aquela rede. Para efeito desta proposta, definiremos a *Internet* como uma grande malha de comunicação digital interligada por computadores (*gateways*), que para se comunicarem partilham um protocolo comum, chamado *TCP/IP*. Seu nascimento é atribuído à necessidades de aplicações militares, mais tarde expandindo-se para comunidades de ensino e pesquisa americanas. Atualmente experimenta um crescimento acelerado em razão das aplicações comerciais³.

O surgimento da *Web* (também conhecida como *Word Wide Web*, *WWW* ou *W3*) deu-se aproximadamente 20 anos após ao advento da *Internet*. A *Web*, da forma como a conhecemos agora, foi implementada em 1990 por Tim Berners-Lee e Robert Gaillian, pesquisadores do Laboratório Europeu de Física da Partícula - CERN. A idéia inicial era elaborar uma interface amigável, que possibilitasse a qualquer usuário o acesso a informações de seu interesse, utilizando-se do conceito de *hipertexto*. Para isto desenvolveram os conceitos básicos de *URL*, *HTTP* e *HTML*⁴.

A *URL* - *Uniform Resource Locator* - é o endereço do computador onde está armazenado a página, ou recurso. O *HTTP* - *Hypertext Transfer Protocol* - é o protocolo que gerencia o envio das páginas ao requisitante e *HTML* - *Hypertext Markup Language* - é a linguagem que define como as páginas devem ser remontadas pelo requisitante. A remontagem é realizada por um *browser* instalado no computador requisitante. O primeiro *browser* gráfico (*Mosaic*) surgiu em 1993, criado do Marc Andreessen, na época programador do *National Center for Supercomputing Applications* - *NCSA* e, mais tarde, um dos sócios da *Netscape Communication*, uma das empresas que, juntamente com a *Microsoft*, dominam o mercado daqueles programas.

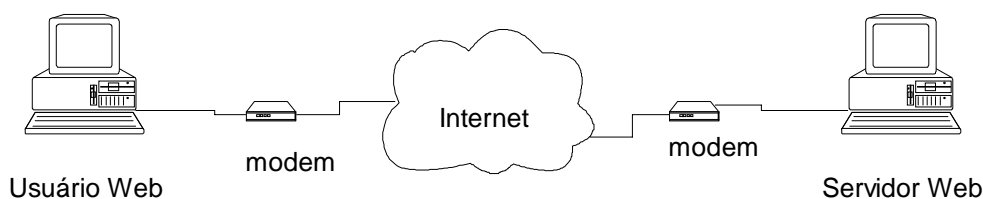
Resumidamente, a forma tradicional de acesso à *Web* pode ser assim descrita: um computador, com algum sistema operacional, com uma interface gráfica, um *browser* instalado, uma interface de rede ou modem conectados a uma rede que acesse a *Internet*, pode acessar páginas eletrônicas localizadas em qualquer outro computador,

³ <http://www.internet.org> e <http://www.isoc.org/internet/history/brief.html>

⁴ <http://www.w3.org/History/>

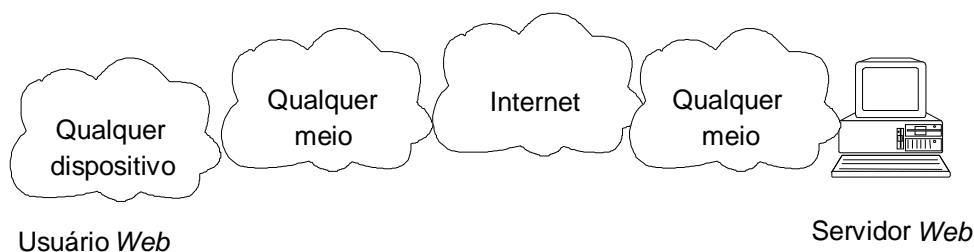
também em rede, que tenha um programa servidor de páginas baseado no protocolo *HTTP*. As páginas são transmitidas pelo servidor em forma de comandos (*tags*) em *HTML* e, na medida em que chegam ao computador requisitante, vão sendo remontadas pelo *browser* que, para isto, obedece as formas descritas pelos tags enviados pelo servidor. Conexões separadas são abertas para outros elementos como imagens e sons.

Figura 1: Forma de acesso tradicional na rede *Web*



No entanto, a descrição acima merece algumas considerações: computadores de diferentes arquiteturas de hardware e sistemas operacionais podem acessar a rede e, atualmente, aparelhos de televisão, telefone (*Web Video Phone*), refrigeradores, fornos de microondas e outros aparelhos também podem acessar a *Web*. As formas de acesso estão evoluindo rapidamente, destacando-se as tecnologias *Wireless*, *ISDN*, *ADSL*, *cable modem*, fibra ótica e, recentemente, viabilização do acesso via rede elétrica. A interface gráfica pode ser o *MS-Windows*, *OpenWindows*, *X-Windows* ou outra. O *browser* pode ser o *Netscape Communicator* ou *MS-Explorer*, ou outro menos conhecido. O que torna possível que equipamentos diferentes, com sistemas operacionais diferentes e conectados através de diferentes meios físicos troquem informações é o protocolo *TCP/IP*, o protocolo padrão de comunicação da rede *Internet*. Isto leva-nos a pensar em um novo modelo para representar do processo de acesso a rede *Web*:

Figura 2: Uma descrição mais atual do acesso a rede *Web*



Apesar da importância da *Internet*, BERNES-LEE (1999) - um dos criadores da *Web* - diferencia a *Internet* da *Web* da seguinte forma:

“... a *Web* é um espaço abstrato (imaginário) de informação.
Na [Inter]Net, as conexões são cabos entre computadores; na

Web, conexões são links de hipertexto. A Web existe por causa dos programas os quais se comunicam entre computadores na [Inter]Net... A Web fez a [Inter]Net útil por que as pessoas estão realmente interessadas em informação e não querem, na verdade, ter que saber sobre computadores e cabos...”.

3.3 Números que dobram

Como vimos no tópico anterior, é o crescimento da capilaridade rede *Internet* que possibilita a expansão da rede *World Wide Web*. Enquanto a primeira estabelece os conceitos de como a informação trafega pela rede - tendo como base o protocolo *TCP/IP* -, a segunda estabelece os padrões de apresentação da informação ao usuário - tendo como base o protocolo *HTTP*. O protocolo base da *Internet* foi criado em 1974 por Vinton Cerf e Bob Kahn. Já o protocolo base da *Web* foi criado em 1990 por Tim Bernes-Lee. Na seqüência, serão apresentados alguns dados a fim dar uma idéia de sua expansão.

Em 1980, a espinha dorsal do que hoje é a *Internet* funcionava a 56 Kbps. Esta velocidade já era de 45Mbps, em 1990 o que representa um crescimento de mais de 800 vezes em 10 anos. Com a implementação da *Internet2*, chegou-se a velocidades superiores a 100 Mbps, havendo perspectivas de gigabits/s até o fim do século (SIMON et. al., 1997). O microcomputador, lançado em 1976 espalhou-se no mundo de forma muito rápida. Além do preço, o fator utilidade está sendo determinante para sua disseminação em massa.

Com a estrutura de comunicações em expansão e a crescente massificação de microcomputadores pessoais, em meados de 1997 o número de equipamentos ligados na *Internet* dobrava a cada 15 meses, o número de servidores da teia mundial *World Wide Web* duplicava a cada 14 semanas. Como consequência, o tráfego de dados dobrava a cada 100 dias.

Atualmente, o número de pessoas conectadas na *Internet* no mundo continua a crescer aceleradamente. Pesquisas elaboradas por órgãos especializados estimavam em 201 milhões de usuários em todo o mundo, em setembro de 1999. Eram 16 milhões em dezembro de 1995. No Canadá e Estados Unidos se concentra a grande maioria (112,4 milhões), enquanto na América do Sul apenas 5,29 milhões de pessoas estão conectadas. Entretanto, previsões apontam 24,3 milhões de usuários da *Internet* na América Latina em 2003⁵.

⁵ http://www.nua.ie/surveys/how_many_online/index.html, consulta em 13/10/1999.



Gráfico 1: Estimativas de usuários da *Internet* no mundo no período 1995-2005
(em milhões)

Fonte: Nua Surveys

No Brasil, pesquisa conduzida pelo IBOPE em nove regiões metropolitanas (36,9 milhões de habitantes) em junho de 1999, estimava em 3,3 milhões de usuários. Segundo a mesma pesquisa, entre dezembro de 1998 e junho deste ano, 750 mil pessoas aderiram a *Internet* – um crescimento de 30% -, confirmando o Brasil com um dos países que mais usam a rede no mundo. Apesar da crise cambial, o que elevou o preço dos microcomputadores, os fatores do crescimento estão sendo creditados à expansão da oferta de linhas telefônicas, a tendência de barateamento dos preços das assinaturas de provedores e a notoriedade da rede em função da divulgação nas mídias tradicionais. O aumento do acesso da classe C à *Internet* subiu 33%, enquanto das classes A e B foi de 20%. Em janeiro de 1996 apenas 170,4 mil pessoas tinham acesso à *Internet* no Brasil⁶.

Entretanto, os motivos que estão fazendo a rede crescer no momento não são os mesmos motivos que levaram a sua criação. De aplicações militares, no início, passando para aplicações acadêmicas de cunho técnico e científico num determinado momento, atualmente a rede tornou-se a vitrine do comércio mundial. O advento da *Web* representou o surgimento de um novo canal para as organizações anunciarem e venderem seus produtos e serviços, difundir sua imagem e ter contato com seus clientes, usuários e consumidores.

Conforme um recente relatório do IDC, a economia em torno da *Internet* movimentará US\$ 1 trilhão em 2001, devendo alcançar US\$ 2,8 trilhões no final de final de 2003.

⁶ <http://www.ibope.com.br>, consulta em 13/10/1999.

Para o ano corrente, o mesmo relatório estima a movimentação de US\$ 507 bilhões⁷.

Outro relatório estima que, no mundo todo, o comércio online entre as empresas e seus consumidores (*business-to-consumer*) deve movimentar US\$ 380 bilhões em 2003. Só nos Estados Unidos serão US\$ 108 bilhões. Já as transações entre empresas (*business-to-business*) movimentarão US\$ 1,331 trilhão no mesmo período. Em 1998 estes números foram de US\$ oito e US\$ 43 bilhões, respectivamente. Uma pesquisa da *Forrester Research* estima que serão gastos US\$ 30 milhões em publicidade *online* no Brasil em 2004. No mundo todo o total de gastos em publicidade na *Web* será de US\$ 33,1 bilhões⁸.

É esta força que faz a *Internet* crescer atualmente. A participação de domínios comerciais na *Internet* cresceu de 1,5% em meados de 1993 para 62,6% no início de 1997. Em outubro de 1999 esta participação representava mais de 90% no Brasil. Entre setembro de 1998 e outubro de 1999, o número de domínios .COM.BR cresceu de 107,6 para 115,3 mil. Em janeiro de 1996 eram apenas 479. Até outubro de 1999 a FAPESP, órgão administrador dos domínios da *Internet* no Brasil contabilizava mais de 130 mil *sites* registrados no País⁹.

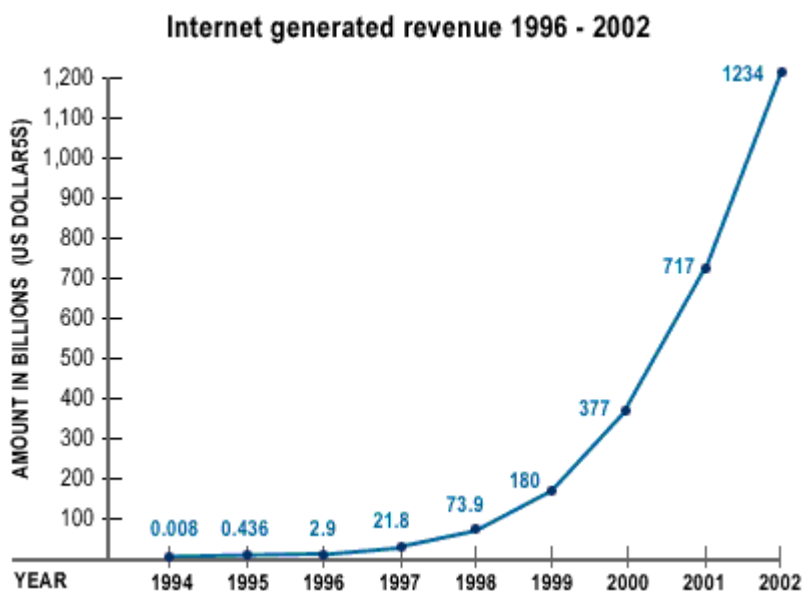


Gráfico 2: Estimativas de renda geradas pela *Internet* no período 1994-2002

Fonte: www.activmidia.com

⁷ <http://www.idc.com>, consulta em 11/1999.

⁸ http://www.nua.ie/surveys/graphs_charts/comparisons/ecommerce_us.html, consulta em 13/10/1999.

⁹ <http://www.registro.fapesp.br>, consulta em 13/10/1999.

Apesar do crescimento e da importância cada vez para o público urbano, um estudo realizado em 1999 pela Research International Brasil, a pedido da Associação Brasileira de Marketing Rural, mostra que o acesso a informática pelo produtor rural brasileiro ainda é modesto. De acordo com o estudo, 18% dos produtores rurais possuem computadores e apenas 4% acessam a Internet. Apesar de tímidos, os resultados mostram o potencial de crescimento para a informática no meio rural¹⁰.

A revista *Internet Word*, em seu número de maio de 1998, informava o seguinte “...enquanto o rádio levou 38 anos para atingir 50 milhões de ouvintes, a televisão levou 13 anos para alcançar a mesma marca. A *Internet*, no entanto, atingiu 100 milhões em pouco mais de quatro anos”.

3.4 A resposta das empresas

Naturalmente, a medida em que a era da informação e do conhecimento se materializa, a rede se torna um canal efetivo de realização de negócios (*e-commerce*), com o elemento competitividade se deslocando para o espaço virtual proporcionado pelas redes. Face a isto, as tecnologias e estudos sobre suas aplicações se sucedem rapidamente.

O emprego intensivo da tecnologia da informação, como ferramenta indispensável para as organizações se adaptarem de forma dinâmica ao seu mercado, já é considerado um pressuposto básico. Entretanto, estes recursos devem ser usados não somente para cobrir os níveis operacional e tático da organização, mas também, serem utilizados efetivamente pelo nível estratégico.

Quanto a isto, FREITAS (1993) tece várias considerações sobre o que uma empresa deve fazer para que seja competitiva na era da informação e do conhecimento. O autor considera que a empresa deve adaptar-se ao cliente e antecipar suas expectativas, desenvolver sua capacidade de evolução para assegurar sua perenidade, desenvolver sua capacidade de escuta do exterior e de seu próprio futuro para melhor se adaptar, gerindo a informática como um meio integrado na estratégia.

Nesta linha, GONÇALVES & FILHO (1995) previram que, com a redução dos custos dos equipamentos, a expansão das redes e desenvolvimento do *software* e *hardware*, as possibilidades de se aplicar a tecnologia da informação no *marketing* aumentarão a cada dia. Neste contexto, ganham as empresas inovadoras que souberem utilizar a TI para melhorar o relacionamento com os clientes, na busca da vantagem competitiva. Segundo os autores, as inovações na TI estão voltadas para os sistemas de apoio ao cliente (SAC). Esta inovação pode se dar através de pesquisas de mercado e serviços de informações para inteligência competitiva, entre outras.

¹⁰ Technonews - On Line (<http://www.technovet.com.br/tnews>, ano IV - n.46 - Out. 1999).

Nos últimos cinco anos a *Web* transformou-se em uma forma básica de relacionamento Empresa X Cliente. Analisando pesquisas relacionadas, observa-se que o assunto tem se tornado objeto de estudo de pesquisadores e profissionais.

SALAM et. al. (1996) realizaram pesquisa em 45 *sites* de empresas americanas incluídas no *Fortune 1000*, utilizando quatorze critérios, ou sugestões elaborados por RESNIK & STERN (1977)¹¹, citado pelos autores. Os critérios avaliados foram existência de preço ou valor, qualidade, performance, componentes ou conteúdo, disponibilidade, oferta especial, gosto (estética), forma, garantia, segurança, nutrição, pesquisa independente, pesquisa apoiada pela empresa e novas idéias. Os resultados evidenciados pela pesquisa mostraram que 85,71% das páginas eram informativas, contendo no mínimo um dos critérios avaliados. Somente 64,29% das páginas continham no mínimo dois dos critérios e 57,14% continham no mínimo três critérios. A pesquisa mostrou ainda que, para os usuários, o critério mais frequentemente buscado era a performance, seguida de perto por preço e qualidade. Entretanto, apenas 48.9% das páginas examinadas continham estes três critérios.

Em pesquisa conduzida entre 101 das 500 maiores e melhores de 1997 da Revista Exame, SOARES e HOPPEN (1998) concluíram que os *sites* estudados refletiam uma tendência predominante com a imagem corporativa, o que revelava estarem em estágios primários em termos de evolução do uso da *Web*. Esta conclusão está apoiada no modelo evolucionário para as grandes empresas do ramo tradicional que iniciam um *Web site*, proposto por QUELCH & KLEIN (1996)¹², citado pelos autores, a qual passa por cinco estágios: 1) imagem da empresa e informações sobre produtos, 2) coleta de informações e pesquisa de mercado, 3) suporte ao consumidor e serviços, 4) suporte interno à empresa e serviços e 5) transações via rede.

Outros dados da mesma pesquisa merecem ser observados: apenas 19% das empresas possuíam um mecanismo para receber *feedbacks* de clientes em seus *sites*. Ou seja, três entre cada cinco empresas da amostra não estavam interessadas da opinião do visitante *Web*, ou não viam na *Web* um meio potencial para transformar dados daqueles que acessam seus *sites* em oportunidades de negócio; 55% das empresas não possuíam qualquer estrutura para recolhimento de dados do visitante do *site*; 77% não explicitaram seus clientes alvos; 22% da comunicação era destinada à clientes institucionais e 18% à clientes individuais; dos que tinham algum mecanismo para coleta de dados de visitantes, 98% pediam dados de identificação, 35% dados sócio-econômicos e 15% dados de hábitos ou preferências de consumo. Ainda, 89% não ofereciam nenhum incentivo para os visitantes deixarem seus dados.

Outra pesquisa, efetuada por SILVA (1997), na qual foi avaliada a presença empresarial das empresas brasileiras na *Web*, constatou que a grande maioria das empresas

¹¹ RESNIK, A. & STERN, B. Na analysis of information content in television adverting. *Journal of Marketing*, January 1977, pp. 50-53.

¹² QUELCH, J. & KLEIN, L. *The internet and international marketing*. Boston: Sloan Magagement Review, Spring 1996.

incentivava o *feedback*, respondendo sempre as solicitações de seus clientes. Entretanto, poucas estavam se utilizando do *marketing* de relacionamento. Apenas 17% utilizavam *database*, 8% utilizavam estudos de *prospects* e 29% efetuavam pesquisas com clientes. Os valores foram considerados inexpressivos pelo autor.

Na época (1997), a pesquisa de Silva constatou que estávamos na etapa inicial do comércio eletrônico, com as empresas fornecendo prioritariamente informações sobre a empresa (85%) e seus produtos (92%). Apenas 21% das empresas faziam suporte aos clientes e 13% efetuavam suporte aos produtos. As promoções eram efetuadas em apenas 19% das empresas pesquisadas. Confirmando a pesquisa de SOARES e HOPPEN (1997) o autor constatou também que a maioria das empresas mantinha sua presença na *Internet* como forma de solidificar sua imagem institucional e também como fonte de negócios. Interessante que 71% das organizações já realizavam estudos de frequência dos usuários as suas páginas. Para isto, utilizavam relatórios de frequência e contagem de *hits*.

Certamente a situação relatada apresenta-se diferente na atualidade, mais de dois anos após as pesquisas. Entretanto, já estarão as empresas utilizando dados capturados em seus *sites* como ferramenta de inteligência de mercado ?

De fato, se conseguirmos entender as intenções das pessoas através das suas escolhas ao navegar por um *site* na rede *Web*, poderemos utilizar esta informação para obtenção de vantagem competitiva. MONTGOMERY & WEINBERG (1998) consideram que a análise de consumidores e não consumidores é, possivelmente, a mais valiosa e a mais negligenciada área da inteligência estratégica. Análises de consumidores e não consumidores podem revelar tecnologias emergentes, vantagens e desvantagens competitivas e novas idéias e produtos. Relatam um estudo sobre inovação tecnológica, no qual 74% de 137 inovações estudadas foram originadas pelos consumidores, três vezes mais que os departamentos de pesquisa das empresas. Segundo os autores, seis funções de transformação pelas quais os dados podem se tornar informação – transmissão, acumulação, agregação, análise, reconhecimento de padrões e mistura - variam consideravelmente em complexidade

3.5 Usuários, suas Trilhas & Comportamento

Em um nível mais baixo, muitos estudos têm sido realizados no sentido de investigar como o usuário se comporta frente a determina aplicação/sistema de consulta e quais suas estratégias para a busca e recuperação de informações de seu interesse. Alguns resultados destas pesquisas servem para realimentar a interface das aplicações, tornando-as mais eficientes no processo de comunicação usuário X sistema de informações. Ultimamente, com o advento do comércio eletrônico, o foco dos estudos está sendo deslocado para a interação usuário X *Web*, tendo a área de *marketing* como uma das principais interessadas.

Em um trabalho anterior a popularização da rede *Web*, FREITAS (1993) propôs um modelo para a avaliação de um sistema de apoio a decisão (SAD). O modelo se

baseava em um método implícito e automático de coleta e armazenamento de todas as ações de um usuário frente a um sistema teleinformatizado de *marketing*. Os resultados do estudo conduzido pelo autor permitiram a elaboração de uma tipologia de usuários finais para o sistema utilizado. Para seu estudo, Freitas utilizou uma aplicação customizada, que permitia o controle total das ações dos usuários na interação com o sistema. Este não é o caso dos sistemas acessados via *Web*, na qual, muitas ações dos usuários não podem ser capturadas. Além disso, a estrutura de navegação de uma aplicação customizada difere da estrutura possível em páginas de hipertexto. Enquanto na primeira é possível elaborar uma estrutura de menus pré-determinada e finita, a modelo de uma “árvore”, na segunda a estrutura se assemelha a uma “malha”, que possibilita ao usuário ampla escolha dos caminhos a seguir.

Um trabalho semelhante ao de FREITAS (1993) foi efetuado por SAKAMOTO (1998). O autor elaborou uma biblioteca (*library*) que, inserida em um *browser*, permite total controle das ações do usuário frente a uma sessão de navegação pela *Web*. Entre as medidas possíveis com a ferramenta, estão a rota tomada para acessar uma página (via *hyperlink*, *bookmark*, ou entrada direta da URL), o tipo de ação em uma página (impressão ou registro no *bookmark*), o tempo gasto na leitura de uma página e o tempo que uma página leva para ser mostrada em um terminal. O estudo de Sakamoto é útil em diversos aspectos. Para provedores de conteúdo e publicitários provê meios para medidas de audiência e auditagem. *Webmasters* podem entender melhor as tendências de uso como o número de páginas vistas, bem como obter elementos para a estratégia futura do *site*. Para os usuários, fornece elementos para implementar serviços personalizados para recuperação de informações, recomendações e customização.

Uma das diferenças entre os dois trabalhos é que no de Freitas os dados dos usuários ficavam no servidor e, no de Sakamoto, os dados ficam no cliente, muito embora nada impeça de serem enviados a um servidor remoto para arquivamento. Isto, no entanto, teria sérias restrições em relação a privacidade, sendo aconselhável seu uso apenas em situações controladas e com o consentimento dos usuários.

Uma semelhança, é que as duas aplicações geravam registros de transações entre usuário X sistema (*logs*) para posterior análise e utilização.

Neste aspecto específico, ABDULLA et. al. (1997) examinaram os *logs* de transações de um servidor de *proxy* visando verificar similaridades entre acessos realizados por instituições educacionais versus acessos da indústria, governo e provedores de acesso comercial. Os autores não encontraram diferenças entre os acessos, considerando nove fatores comuns que eles denominam *invariants*. Estes fatores, foram estabelecidos por ARLITT & WILLIAMSON (1996)¹³, citado pelos autores. A não-diferença encontrada nas medidas observadas é importante, uma vez que contribui para a generalização e a obtenção modelos estatísticos que podem ser utilizados para simulações e estudos de

¹³ ARLITT, M.F. & WILLIAMSON, C.L. Web server workload characterization: the search for invariants. Proc. SIGMETRICS, Philadelphia, PA, April 1996. ACM, 160-169.

modelagem.

Uma pesquisa mais voltada ao comportamento de usuários foi realizada por JANSEN et. al. (1998) usando os *logs* do mecanismo de busca *Excite*. Foi analisado um *subset* de 51.453 consultas efetuadas por 18.113 usuários escolhidas randomicamente entre as consultas efetuadas em 10/03/1997. Os dados revelados pela pesquisa mostraram uma média de 2,8 consultas por usuário, o que denota ações de refinamento nas buscas. Na média, as consultas possuíam 2,21 termos. Mais de 80% das consultas possuíam de um a três termos, sendo que 31,46% utilizavam dois termos e 30,81% , um termo. Em relação a construção das consultas, a pesquisa revela que o uso de operadores booleanos foi baixo. Apenas 9,32% usaram o operador AND. O uso do operador OR foi de 0,26%, o operador AND NOT foi usado 0,23% e parênteses foram usados em 0,53% das consultas. Mesmo assim, o percentual de incorreções na montagem das consultas com os operadores AND, OR, AND NOT e parênteses foi de 26,3%, 34,85%, 65,83% e 32,23% respectivamente.

Não somente descrevendo os dados obtidos através da análise de *logs*, mas agora aplicando a informação obtida, JOHN & DIMITRIS (1998) propuseram um algoritmo para reorganizar a estrutura de um *site Web* a partir da popularidade das diferentes páginas (*Relative Page Popularity – RPP*). Este algoritmo baseia-se, entre outras medidas, em estatísticas armazenadas nos *logs* de transações coletados e armazenados automaticamente pelo servidor de páginas do *site*. Através de um estudo de caso, os autores relatam que o algoritmo contribuiu para o aumento de número de acessos ao *site* em que foi aplicado.

O tema começa a despertar, também, o interesse em especialistas da área de *datamining*. SPILIOPOULOU & FAULSTICH (1998) e SPILIOPOULOU et. al. (1999) consideram que a análise do comportamento de usuários possuem dois aspectos, um relativo aos interesses do usuário e a informação que eles acessam, e outro relativo ao caminho ou maneira de como a informação é acessada. O primeiro aspecto consiste em estabelecer perfis de usuário, não sendo peculiar a *Web*. Já o segundo possui seu foco nas técnicas de analisar *logs* de servidores *Web*. Para este último, os autores propuseram um algoritmo que foi implementado através de um software (WUM - *Web Utilization Miner*), o qual trilha os caminhos dos usuários através de um *site* utilizando pré-processamento dos *logs*, técnicas sofisticadas de heurística e uma linguagem adicional para livre consulta, a modelo da SQL (*Structured Query Language*). Entretanto, os autores argumentam que o usuário *Web* é caracterizado pelos seus interesses e pelo seu comportamento navegacional sendo portanto, os dois aspectos, complementares.

Outros trabalhos, com algum nível de relacionamento foram desenvolvidos nos últimos anos. A preocupação destes, entretanto, foca a organização da informação nos *sites* visando tornar sua recuperação mais eficiente. Na verdade, estes estudos podem ser beneficiados a partir das pesquisas que investigam o comportamento dos usuários, relatadas anteriormente.

DOAN et. al. (1998) introduzem um novo conceito de documentos *Web*, o qual se

baseia fortemente em metadados incluídos nas páginas HTML e na proposição de dois tipos de *robots*: os “generalistas” e os “especialistas” que coletam e organizam aqueles metadados a fim de localizar recursos na *Web*. Entre os benefícios do trabalho efetuado pelos autores, estão o menor consumo do *bandwidth* pelos *robots* e a obtenção de respostas mais relevantes nas consultas efetuadas.

Em linha semelhante, ATTARDI et. al. (1998) descreve a técnica de “categorização por contexto”, o qual explora o conteúdo percebível da estrutura de documentos HTML para extrair informação útil para a sua classificação. A categorização proposta é realizada de forma automática e, segundo os autores, experimentos preliminares mostraram bons resultados.

SPINK et. al. (1998) observaram que usuários com um problema para resolver (*problem-at-hand*), tendem, ao longo do tempo, a procurar no mesmo ou possivelmente em diferentes sistemas interativos (bibliotecas digitais, sistemas de recuperação de informações, serviços na *Web*) por respostas para o mesmo ou problema de informação, ou a ele relacionado. Segundo os autores, este processo é chamado de *successive search phenomenon*. Observaram também que a grande maioria dos usuários tendem a empregar estratégias simples de busca. Na mesma pesquisa, os autores relatam que apenas 5.24% das consultas continham operadores booleanos, tidos como chaves para o refinamento de buscas. Neste levantamento, o número médio de termos informados por usuário/consulta foi de 3.34. Salienta-se que estes dados foram fornecidos pelos usuários, não envolvendo nenhuma técnica de coleta automática.

Os resultados do presente trabalho poderão contribuir para a melhor especificação de documentos HTML através da definição mais precisa sobre sua natureza e utilidade para quem os procura. Neste sentido, pode ser de grande valia as técnicas aplicadas por FREITAS et. al. (1996) e FREITAS & MOSCAROLA (1998), as quais expõe a importância de ferramentas específicas para análises de conteúdo empregando a análise fatorial de correspondências (AFC).

4 MÉTODO DE PESQUISA

Através da revisão de literatura, pode-se inferir que o tema abordado é emergente, carecendo ainda de métodos e técnicas apropriadas para sua interpretação. No final deste capítulo se argumentará a favor do delineamento escolhido. A seguir serão definidos os pressupostos metodológicos e as técnicas empregadas na condução desta pesquisa.

4.1 Nível e *design* da pesquisa:

Exploratório e Descritivo (Gil, 1999:44; Yin, 1994:4; Triviños, 1987) Estudo de caso (Gil, 1999:72, 1996:58; Trivinos, 1987:133; Yin, 1994) multimétodo, de forma longitudinal.

4.2 Investigações a realizar na busca do objetivo de pesquisa

Será buscado essencialmente definir quais são as características dos acessos realizados em um servidor de informações para a *Web* em uma empresa de P&D. Tal realização permitirá propor um método para verificar a existência de conhecimentos úteis para a organização nos registros de acessos realizados de forma automática (*logs de transações*) em um *site* na rede *Web*. Neste intuito, diversas outras questões de trabalho surgem, como por exemplo:

Quais termos ou expressões estão trazendo o visitante ao *site*?

Quais são as características gerais dos acessos (origem do acesso, horários, dias, semanas, páginas mais acessadas, domínios que mais acessaram, origem, distribuição espacial)?

O que busca o visitante que chega ao *site*?

O que busca o leitor do informativo externo da organização (Folha do Produtor)?

O que o visitante do *site* procura em cada mecanismo ?

Quais são as “trilhas” predominantes do usuário através do *site* ?

Como está classificada a oferta de informações no *site*?

Como medir a adequabilidade entre oferta e a demanda de informações?

4.3 Descrição das fontes de dados, coleta, análise e interpretação dos dados

4.3.1 Registro de acessos redirecionados por outros *sites*

O redirecionamento do acesso é prática comum na *Web*. Isto acontece quando um usuário, acessando um diretório de busca (ou indexador) - como o Yahoo, Altavista, Cade, etc -, insere palavras-chaves para uma pesquisa e acessa algum *link* retornado, ou quando um determinado *site* é acessado através de outro que publica o endereço de *sites* relacionados. Este tipo acesso é registrado em arquivo (*referer_log*). A análise deste arquivo revela quais termos/expressões estão levando o visitante ao *site*, bem como quais *sites* estão fazendo este redirecionamento.

Uma linha típica do *log* de acessos redirecionados:

```
http://www.altavista.com/cgi-
bin/query?pg=q&kl=pt&q=embrapa+trifolium+repens+bage&search=Search ->
/index.html
```


4.3.2 Registro de acessos das páginas eletrônicas

Cada acesso realizado no *site* deixa gravado em arquivo o nome ou endereço *IP* do computador que está requisitando a página, a data e hora do acesso e o nome da página transferida, entre outras informações de controle que se referem ao status da transferência em si (W3C, 1999). Os servidores *NCSA* (*National Center for Supercomputer Application*) e *Apache*, para *Unix* (ambiente deste estudo), tipicamente armazenam estas informações em `/usr/local/etc/httpd/logs/access_log`¹⁴.

Muitas ferramentas para análise de *logs* têm sido produzidas ultimamente, dado o crescente interesse em tornar os *sites* adaptativos a partir das escolhas dos usuários. Enquanto algumas ferramentas se limitam a produzir relatórios e gráficos descritivos dos acessos, sem empregarem nenhum método para descoberta de conhecimento e padrões, outras começam a empregar sofisticadas técnicas de *datamining*.

Os arquivos de dados passarão por sucessivas fases de preparação para serem submetidos a diferentes *softwares* genéricos para análise estatística e específicos para análise de *logs*. A fase de preparação envolverá o desenvolvimento de programas específicos a serem desenvolvidos em linguagens de programação como C ou Basic.

Pretende-se classificar as páginas segundo a natureza de seu conteúdo e do seu requisitante. Isto será ser feito diretamente no arquivo de acessos.

Uma linha típica do *log* de acesso de páginas:

```
scm619.ufrgs.br - - [30/Jun/1999:20:44:09 -0300] "GET /unidade/equipe.html
HTTP/1.0" 200 37285
```

4.3.3 Registro de expressões e termos de busca utilizados pelos usuários do *site*

O mecanismo de busca do *site* é uma ferramenta de domínio público (`search.cgi`) que possibilita a execução de pesquisas através de termos chaves utilizando-se operadores booleanos *AND*, *OR* e *NOT*. A expressão de busca utilizada fica armazenada em arquivo, contendo o nome do computador requisitante, a expressão de busca utilizada e o número de documentos recuperados pela consulta. Será necessário classificar os termos de busca segundo o mesmo critério utilizado para a contextualização das páginas. Nestes dados serão aplicados ferramentas e procedimentos estatísticos a serem discutidos com o orientador. Ferramentas capazes de manipular texto e de *datamining* poderão ser desejáveis.

¹⁴ <http://www.w3.org/Daemon/User/Config/Logging.html#common-logfile-format>

As três linhas típicas do *log* do mecanismo de busca do *site*:

Termos Opcionais: gado, de, corte
 Search by kobe.inf.ufrgs.br:
 Resultados da Busca: 14 Documentos Foram Encontrados

4.3.4 Registro de expressões e termos de busca utilizados pelos leitores do informativo externo

Este mecanismo, a modelo do mecanismo de busca do *site* também aceita a formação de expressões booleanas. Entretanto, ele se restringe aos usuários que acessam o informativo externo *online* do *site* (Folha do Produtor). Este informativo é orientado aos produtores rurais, estudantes, técnicos da extensão rural, professores, etc. A publicação *online* acompanha a versão em papel e atualmente está em seu número 14.

Uma linha típica do *log* do mecanismo de busca do informativo da organização:

gw2.ulbra.tche.br - - [09/Sep/1999:16:19:03 -0300] "GET /cgi-bin/sfolha.cgi?terms=preservação+and+campo+and+nativo HTTP/1.0" 200 4214

4.3.5 *E-mails* contendo solicitação de informações ou documentos

Nos últimos anos ficaram cada vez mais frequentes os pedidos de informações e consultas encaminhadas para a Empresa através do sistema de *e-mail*. Esta demanda determinou a estruturação de um fluxo padrão de encaminhamento de mensagens com pedidos de informações e publicações (*Serviço de atendimento ao Cidadão - SAC*). Cada unidade da empresa possui uma conta de *e-mail* chamada *sac@nome-da-unidade.embrapa.br*.

Geralmente é a Área de Comunicação e Negócios Tecnológicos de cada unidade quem encaminha a solução para o usuário. Entretanto, muitas mensagens ainda são encaminhadas ao *Postmaster* ou ao *Webmaster* do Centro.

O conteúdo é extremamente variado. O único dado que as mensagens, com certeza, é a informação requerida pelo usuário. A utilização do nome do domínio do provedor do usuário neste caso não é confiável pelo fato de que muitos endereços são de provedores gratuitos de *e-mail*, que podem estar localizados em qualquer parte do mundo (bigfoot, zipmail, hotmail, etc).

Uma possibilidade seria aplicar os mesmos procedimentos aplicados para análise dos *logs* do mecanismo de busca.

4.4 Esquema da pesquisa e considerações sobre seu enquadramento metodológico

O diagrama abaixo ilustra onde se situa a análise de *logs* no processo em investigação e quais seriam as suas aplicações práticas.

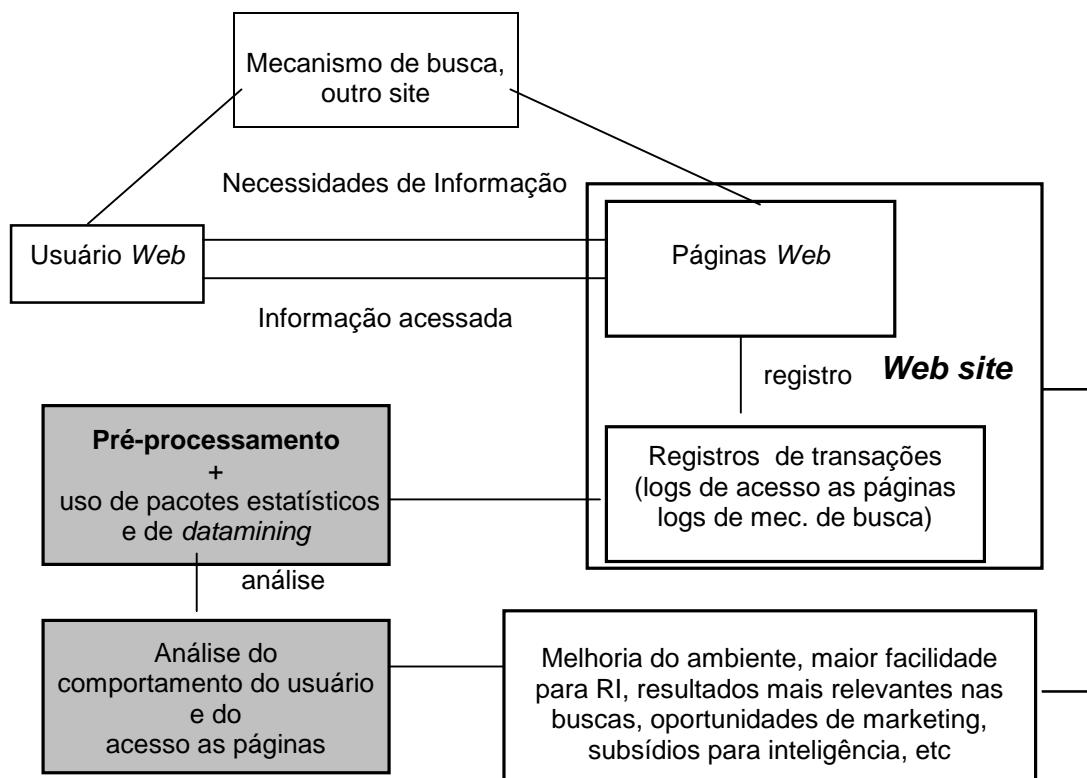


Figura 3: Diagrama do escopo da pesquisa

Estudos preliminares utilizando os dados coletados até o momento - cerca de 90% do total -, mostram a necessidade e a importância da etapa de pré-processamento antes da utilização dos procedimentos estatísticos e de *datamining*.

Estes estudos mostraram a necessidade de classificar os dados, segundo parâmetros relacionados a atividade fim da organização estudada, seu contexto de aplicação. Dado esta particularidade, optou-se pelo enquadramento da pesquisa como estudo do caso. O conjunto de atributos que serão usados para a classificação dos dados será objeto de consenso entre o orientador e especialistas da organização estudada.

Além disso, conforme pode ser observado na Tabela 1, os dados encontram-se em diferentes fontes e formatos e terão sua discussão final realizada individualmente. Entretanto, seria interessante analisa-los conjuntamente a fim de obter-se uma visão geral e permitir comparações entre as diferentes fontes.

Tabela 1: Resumo das fontes de dados, suas características e procedimentos de análise.

Fonte de Dados	Características	Procedimentos
Registro de termos redirecionados pelos diretórios de busca	Período: 1 ano, coleta automática	Cluster, Frequência (SAS); Classificação, Geração de regras (C4.5);
Registro de acesso à páginas do <i>site</i>	Período: 1 ano, coleta automática	Cluster, Frequência (SAS); Classificação, Geração de regras (C4.5); Comportamento do usuário (WUM)
Registro de termos no mecanismo de busca do <i>site</i>	Período: 1 ano, coleta automática	Frequência (SAS); Classificação, Geração de regras (C4.5);
Registro de termos no mecanismo de busca do informativo <i>online</i> da organização	Período: 1 ano, coleta automática	Frequência (SAS); Classificação, Geração de regras (C4.5); Comportamento do usuário (WUM)
Mensagens eletrônicas	Período: 1 ano, coleta automática	Cluster, Frequência (SAS); Análise de Conteúdo (Sphinx)
Páginas eletrônicas do <i>site</i>	Período: 1995/1999	Análise de Conteúdo (Sphinx)

Obs.: SAS é uma abreviatura de *Statistical Analysis System*. Pacote estatístico para análise de dados.
WUM é uma abreviatura de *Web Usage Miner*. Software para análise de *logs* de *sites* criado por pesquisadores da Humboldt University, Berlin, Alemanha.
C4.5 é um software de geração de regras criado pelo Dr. J. R. Quinlan, da Universidade de Sydney, Austrália, cujo algoritmo é amplamente incorporado em outros softwares do gênero.
Sphinx é um software para análise de dados de propriedade da Le Sphinx Développement & Société Ergole, França.

5. O CONTEXTO DE APLICAÇÃO

Neste capítulo procurar-se-á não somente fazer uma narrativa do contexto de aplicação,

em uma perspectiva cronológica, mas também ressaltar aspectos que tornam esta pesquisa útil para a organização estudada, inserindo-a em uma perspectiva estratégica. Espera-se, através da análise de documentos estratégicos da organização, proporcionar uma visão do “por quê” e do “para quê” esta investigação será conduzida.

A Empresa Brasileira de Pesquisa Agropecuária - Embrapa, ambiente deste estudo, é uma empresa estatal, fundada em 1973, que vem passando por rápida transformação em sua maneira de atuar, especialmente a partir do início da década de 1990. Sua estrutura consiste de 37 centros de pesquisas e 24 serviços especiais espalhados pelo território nacional que, ao longo dos últimos dez anos vem passando por mudanças contínuas visando sua adequação as reformas do Estado.

Uma das mudanças mais significativas foi a dimensão conceitual assumida de que “a pesquisa começa e termina na sociedade”, em contraposição ao paradigma anterior que pregava que “a pesquisa começa e termina no produtor [rural]”. Documento estratégico orientador da atuação da empresa para a década de 1990 justificava a mudança de enfoque afirmando que “...diferentes segmentos sociais têm demandas diferentes para serem atendidas por diferentes conjuntos de informações técnico-científicas geradas pela Embrapa e ofertadas sobre diferentes formas...” (FLORES, 1991). O mesmo documento colocava como diretrizes a ampliação e fortalecimento das atividades de difusão de tecnologia e a integração com o setor privado, entre outras, pondo como prioridade “...formular e implantar políticas e diretrizes gerais para as áreas estratégicas tais como cooperação nacional e internacional, marketing e comercialização, comunicação social, P&D e captação de recursos...”.

Dentre algumas mudanças implantadas desde o início desta década, rumo a uma orientação ao mercado estão a implantação do SEP (Sistema Embrapa de Planejamento), que alterou profundamente sua maneira de levantar demandas ambientais para delinear suas atividades de P&D, a política de comunicação para delinear o relacionamento com o ecossistema em que atua, a política de negócios tecnológicos, para consolidar a nova orientação voltada ao mercado, a implantação do *SIGER* (Sistema de informações Gerenciais) para gerenciar projetos de P&D, mudanças estruturais em sua cúpula administrativa, adoção de estrutura funcional por projeto, gerência por processos, redução de vagas, redução de custos, redução de níveis hierárquicos e etc.

A introdução do programa de qualidade na empresa remonta ao ano de 1992. Em 1994, publicação interna da empresa, baseada em diversos pressupostos ambientais que pressionavam as instituições de C&T a mudarem a forma de se relacionar com o seu ecossistema, assinalava que “A mudança básica consiste em estabelecer como alvo principal de sua missão e de seus objetivos, os clientes, beneficiários e usuários...” (POPINIGIS et Alii, 1994). Este pressuposto traz em seu bojo, a mudança da orientação ao produto para uma orientação ao mercado, refletindo políticas estratégicas baseadas na tendência de diminuição do papel do Estado em determinadas atividades e a necessidade crescente de geração de recursos próprios.

FLORES & SILVA (1992) observaram a que as atividades de *marketing* eram

incipientes em nossas instituições de pesquisa, creditando isto a forte orientação ao produto e não ao mercado existente na atividade. Salientavam que era necessário romper com tal paradigma.

KOTLER (1996) considera que existem cinco conceitos diferentes sob os quais as organizações conduzem suas atividades no mercado: produção, produto, vendas, marketing e marketing societal. Conforme o autor, o conceito de produto assume que os consumidores favorecerão aqueles produtos que oferecem mais qualidade, desempenho ou características inovadoras. Kotler adverte que esta orientação pode levar os administradores a ter um “caso amoroso” com seus produtos deixando de observar que o mercado pode estar menos preocupado com a qualidade oferecida. Já o conceito de *marketing* assume que a chave para atingir as metas organizacionais consiste em determinar as necessidades e desejos dos mercados-alvo e oferecer as satisfações desejadas de forma mais eficaz e eficiente do que os concorrentes.

A Embrapa orientou suas atividades baseada no conceito de produto por quase 20 anos, tendo assumido explicitamente a orientação ao mercado a partir de meados da década de 1990. Além da mudança de orientação, veio a clara definição de que a Empresa deveria atuar, como agente do sistema técnico-científico nacional, em pesquisa e desenvolvimento (P&D). Neste aspecto, ZAWISLACK (1996) caracteriza o objetivo da pesquisa tecnológica dos agentes desta classe como “...à procura de aplicação e soluções a problemas econômicos reais, ou seja, diz respeito ao momento da inovação...”.

A Política de Negócios Tecnológicos (1998) , em particular, consolida direções já expressas em documentos estratégicos da Embrapa, as quais apoiam-se fundamentalmente em *marketing*. Este direcionamento já havia ficado claro no documento Estratégia Gerencial da Embrapa - Gestão 95/98 o qual define, como política global de administração, que

“A Embrapa adota o Marketing, na sua acepção mais ampla de “filosofia de relacionamento com o macro-ambiente”...Isto equívale dizer que Marketing não deve ser apenas preocupação ou domínio de um departamento da sede ou de um setor numa unidade descentralizada...é uma atitude nova, um compromisso de todos os empregados da Embrapa para com a sociedade brasileira”.

Para amparar esta orientação o documento estabelece três políticas visando atender os quatro compostos de *marketing* (4P’S). Estas políticas eram a Política de Pesquisa e Desenvolvimento (Produto), a Política de Vendas ou Distribuição (Preços e Pontos de Venda) e uma Política de Comunicação (Promoção). O mesmo advertia que

“adoção desta nova postura não pode ser retardada por visões segundo as quais marketing não se aplicaria à Embrapa por ser uma empresa que gera, em grande monta, tecnologias de cunho social que não teriam preços e não seriam vendáveis”. Para

fortalecer esta nova orientação, o documento estabelecia que “... tudo o que se gerar na Embrapa será objeto de rigoroso escrutínio segundo a visão de Marketing ...”.

A orientação ao mercado se apóia em quatro pilares básicos (KOTLER, 1996). Estes pilares são mercado-alvo, necessidades dos consumidores [clientes e usuários], marketing coordenado e rentabilidade. A identificação de segmentos de mercado e seleção de mercados-alvo é, portanto, uma etapa obrigatória para a adoção de uma orientação ao mercado. Segundo o autor, as empresas atualmente estão abandonando a prática de mercado de massa por não valer a pena, pois estes estão sendo pulverizados e sendo transformado em micro-mercados com compradores diferentes à procura de produtos diferentes em canais de distribuição diferentes e respondendo a diferentes canais de distribuição. A segmentação de mercado é seguida da escolha do mercado-alvo e o posterior posicionamento, quando da adoção de *marketing* de mercados-alvo.

Entre os 12 objetivos específicos estabelecidos pelo documento em discussão, para auxiliar a empresa no cumprimento de sua missão, estava o de “valorizar as ações de desenvolvimento de produtos e processos, de difusão de informação e de comercialização de tecnologias, serviços e produtos...”.

Mais adiante, o referido documento estabelecia vários projetos gerenciais para auxiliar na consecução dos objetivos específicos já enunciados. Dentre os projetos, destacam-se três projetos da categoria Informação e Comunicação, os quais são: Comunicação na Embrapa (projeto 20), Sistema Embrapa de Informação (Projeto 2) e *Internet* (projeto 30). Embora seja possível estabelecer relações da *Web* com os três projetos, os dois últimos possuem ligações mais estreitas com ela. O projeto 2 - Sistema Embrapa de Informação - tinha como objetivo implementar um sistema informatizado e em rede que tornasse disponíveis, para a própria Embrapa e para o público em geral, todas as informações geradas pela Empresa. O projeto 30 - *Internet* - tinha como objetivo estabelecer uma política clara que regulasse a disponibilidade de informações institucionais e comerciais, padronizando formatos e estabelecendo mecanismos de proteção, em coordenação com o projeto Sistema Embrapa de Informação, com o projeto Comunicação na Embrapa, e outros projetos ligados às relações comerciais e de cooperação da Empresa.

Já o documento Estratégia Gerencial da Embrapa, Macroprioridades/1997, estabelecia Inovar Métodos e Meios de transferência como prioridade institucional. Esta macroprioridade preconizava o seguinte:

“... os avanços na área de telecomunicações e informática estão trazendo profundas modificações nos hábitos de informação do meio rural. As oportunidades de reciclagem dos produtores se ampliaram muito e se tornaram mais rápidas com a proliferação das antenas parabólicas e da comunicação via computador. Tais circunstâncias exigem que novos canais e processos de transferência de tecnologias sejam imaginados de sorte a melhorar a eficiência da Embrapa neste processo, aumentando a

oferta de informações e o número de técnicos e produtores atendidos, a custos menores...”.

Em 1998, após a etapa de redefinição de planos da Empresa, o documento Plano Diretor da Embrapa - Realinhamento Estratégico 1999-2003, reafirma todos os conceitos expostos anteriormente, apontando como grandes transformações a globalização com abertura de mercado, a importância do meio ambiente, a reforma do Estado, a força do consumidor e a revolução tecnológica, caracterizando o agronegócio brasileiro do futuro como um setor competitivo, com qualidade e produtividade, tecnologicamente avançado, demandante de informação técnico-gerencial e promotor de emprego e renda.

A missão, agora redefinida, consiste em “...viabilizar soluções para o desenvolvimento sustentável¹⁵ do agronegócio¹⁶ brasileiro por meio de geração, adaptação e transferência de conhecimentos e tecnologias, em benefício da sociedade...”. Importante caracterizar a visão do que seja cliente: “...a Embrapa considera como cliente todo o indivíduo, grupo ou entidade, pública ou privada, cujo sucesso em suas atividades dependa dos produtos e serviços, de natureza econômica ou social, oferecidos pela empresa e seus parceiros...”.

Além disto, a Empresa coloca como objetivo que colaborará para a consecução de sua missão, o de viabilizar soluções tecnológicas para o desenvolvimento de um agronegócio competitivo em uma economia global, estabelecendo para isto que será necessário monitorar e analisar o comportamento dos mercados afins do agronegócio, identificando tendências e oportunidades.

Visando o fortalecimento de sua posição como instituição viabilizadora de soluções para o agronegócio e para consecução de políticas governamentais em bases sustentáveis e competitivas, o mesmo documento estabelece algumas diretrizes estratégicas para as atividades de pesquisa e desenvolvimento (P&D) e de transferência de tecnologia, entre outras. Reafirma uma política geral de administração embasada nos conceitos de *marketing* e qualidade total, que assegure a) a disponibilidade de conhecimentos e tecnologias que tenham real interesse para a sociedade, b) a negociação para a distribuição de conhecimentos e das tecnologias gerados pela Empresa e por seus parceiros e c) a promoção dos conhecimentos e tecnologias entre os diversos públicos.

Especificamente para as atividades de pesquisa e desenvolvimento a Empresa pretende “...adotar mecanismos de levantamento e priorização de demandas e do grau de

¹⁵ Entende-se por desenvolvimento sustentável o arranjo político, sócio-econômico, cultural, ambiental e tecnológico que permite satisfazer as aspirações e necessidades das gerações atuais e futuras.

¹⁶ O conceito de agronegócio engloba os fornecedores de bens e serviços à agricultura, os produtores agrícolas, os processadores, os transformadores e os distribuidores envolvidos na geração e no fluxo dos produtos agrícolas até o consumidor final. Participam também do agronegócio os agentes que coordenam o fluxo dos produtos, tais como o governo, os mercados, as entidades comerciais, financeiras e de serviços.

satisfação dos clientes com os conhecimentos e as tecnologias disponíveis...”. Já a transferência de conhecimentos e tecnologias está apoiada nas políticas de negócios tecnológicos e de comunicação. Para isto, a Empresa deverá, entre outras ações, “...criar meios para a transferência de conhecimentos e tecnologias desenvolvidos ou mobilizados pela Embrapa para o maior número de clientes...”. Além disso, a Embrapa “...implementará estratégias de comunicação empresarial que sejam orientadas à melhoria da interação interna e à potencialização do relacionamento da empresa com o ambiente externo, especialmente quanto a transferência de tecnologia...”. Para atingir este objetivo, a Empresa deverá criar, manter e ampliar fluxos de comunicação de modo a estimular a interação entre a empresa e seus públicos interno e externo além de inovar e modernizar os métodos e instrumentos de comunicação com os diversos públicos.

A ênfase de utilizar redes de computadores e redes de comunicação para transferir tecnologias fica mais clara quando, no mesmo documento, a empresa define projetos chamados estruturantes, em particular o projeto "Transferência de Tecnologia", cuja justificativa é a seguinte:

“...os dias atuais exigem procedimentos mais ágeis de transferência tecnológica. Novos sistemas de comunicação e de informática permitem imprimir maior velocidade ao processo de transferência tecnológica. A empresa intensificará e institucionalizará novos mecanismos de transferência tecnológica, caracterizados pela rapidez com que a informação transita entre o pesquisador, as bases de dados e o usuário, pela rápida atualização das informações disponibilizadas, pelo estímulo à interação entre as equipes de pesquisadores, os agentes de assistência técnica e o usuário e pela facilidade de acesso às informações tecnológicas. Para tanto, as redes de computadores serão usadas para atendimento do público em geral e dos profissionais de assistência técnica que serão credenciados pela Embrapa...”.

Todas estas políticas, diretrizes e objetivos apresentados possuem em comum o fato de atribuir grande importância a tecnologia da informação, especialmente os recursos de rede, como ferramenta indispensável no processo de transferência de tecnologia, de soluções tecnológicas e da consolidação da imagem da Empresa junto aos seus públicos. Assim, são necessários estudos que identifiquem os mercados-alvo que possam alavancar o desenvolvimento tecnológico do setor rural a partir de uma maior - e mais direcionada - oferta de informações.

Além das intenções apresentadas, a empresa investiu pesadamente em recursos de informática na década de 90, implantando 37 redes locais, formadas por cerca de 3000 microcomputadores e mais de 100 workstations. Recentemente, concluiu projeto interligando suas 37 unidades descentralizadas em um sistema de comunicações via satélite, que possibilita o tráfego de som, dados e imagem (videoconferência).

A entrada da empresa na rede *Web* deu-se em 1995, quando foi implantada a estrutura

de redes locais de computadores em todas as suas unidades e iniciou-se o processo de ligação destas redes locais na *Internet*. A interligação da totalidade da empresa foi um processo lento em razão das limitações impostas pela localização das unidades descentralizadas - geralmente em áreas rurais. Apesar de ter adotado o protocolo padrão da *Internet* (TCP/IP), as soluções adotadas para elaboração da malha física da rede apresentam-se muito heterogêneas até da implantação e consolidação do *backbone* próprio, via satélite, em 1999.

No período de implantação e consolidação da rede, as unidades descentralizadas foram gradativa e isoladamente desenvolvendo suas páginas, valendo-se da maior ou menor capacidade técnica do pessoal da área de informática. Atualmente, todas as suas 37 unidades de pesquisa possuem páginas eletrônicas individualizada, que são acessadas através de uma página principal, localizada na sede da empresa, em Brasília (<http://www.embrapa.br>), ou acessadas diretamente através de suas URLs individuais (www.sigla-da-unidade.embrapa.br). Por terem evoluído sob diferentes capacidades e visões, as páginas atualmente encontram-se muito diferenciadas, nos aspectos de opções temáticas e de apresentação.

Não existe um padrão rigoroso entre as páginas das unidades, apesar de existir um documento oficial que estabelece cores, fontes, *links* padrões (conteúdo) e outros elementos a serem utilizados no seu processo de elaboração. Existem, também, diversas orientações quanto a responsabilidade pela elaboração e manutenção das páginas eletrônicas entre as unidades. Estima-se que a maioria das páginas foram criadas e sejam mantidas por técnicos da área de informática, muitas delas apresentando apenas informações institucionais

As páginas eletrônicas da unidade de pesquisa considerada começaram a ser disponibilizadas a partir de finais de 1995 e atualmente publicam informações variadas sobre as atividades desenvolvidas, como projetos de P&D em andamento, equipe de pesquisadores, publicações *on-line*, serviços, produtos e etc. O *site* está acessível 24 horas/dia, sete dias/semana, sendo possível localizá-lo através dos mecanismos de busca mais conhecidos da *Web* como *altavista* (<http://www.altavista.com>), *yahoo* (<http://www.yahoo.com>) e *cade* (<http://www.cade.com.br>). Consulta efetuada no site do Altavista no mês de julho do ano corrente, retornou um total de 55 páginas apontando para o site considerado.

6 CONTRIBUIÇÕES POTENCIAIS

A Embrapa Pecuária Sul é uma unidade de pesquisa e desenvolvimento (P&D) descentralizada da Empresa Brasileira de Pesquisa Agropecuária - Embrapa, localizada em Bagé-RS, que tem a missão institucional de “Atender com soluções tecnológicas as necessidades de dos sistemas produtivos integrados ao agronegócio de bovinos e ovinos na Região Sul em benefício da sociedade”.

Seus usuários/clientes podem ser caracterizados por produtores rurais mais ligados à pecuária, estudantes de agronomia, veterinária e escolas técnicas de agropecuária, sindicatos rurais, associações de produtores, extensionistas rurais, cooperativas, órgãos de fomento, outras instituições de pesquisa agropecuária, universidades e empresas públicas ou privadas envolvidas com as cadeias produtivas da carne (ovina e bovina) e do leite. Sua região de abrangência é entendida como sendo os estados da Região Sul (PR, SC, e RS).

Uma agropecuária que pretende ser competitiva não pode abrir mão da atividade de pesquisa tecnológica. Esta atividade entretanto, deve possuir mecanismos eficientes e eficazes de transferência da tecnologia gerada. Um destes mecanismos poderá ser a *Web*.

Assim, o benefício que os resultados deste estudo pode trazer para a sociedade está apoiado na premissa de que uma maior e mais direcionada oferta de informações tecnológicas pode contribuir para a aceleração do processo de geração/aplicação de conhecimentos advindos da pesquisa tecnológica agropecuária. Esta aceleração, por sua vez, poderá contribuir para o aumento dos índices de produtividade agropecuária impactando toda a cadeia produtiva da carne e do leite, com reflexos positivos na renda do produtor rural e da cadeia como um todo.

Neste caso, saber o que os usuários, atuais e potenciais, de tecnologias desejam ou desejariam encontrar sobre tecnologias agropecuárias na *Web* poderá ter grande utilidade. Isto, além de possibilitar o alinhamento das estratégias da organização às necessidades da matriz produtiva, possibilita também que a organização identifique e monitore áreas de crescente interesse que possam ser utilizadas para impactar a eficiência e eficácia na transferência de seus resultados. As informações disponibilizadas poderão, em algum nível, contribuir para o processo de tomada de decisões que permeia todos os segmentos da cadeia produtiva dos produtos considerados.

Já para a área de SI, a análise do ecossistema virtual da organização, poderá apontar recomendações sobre a forma e o conteúdo das páginas do *site*, de forma a torna-lo adaptativo aos usuários. A pesquisa poderá contribuir, também, para estudos que visam alinhar esta TI emergente (*Web*) com as estratégias e objetivos de uma organização, no

sentido de que pode apontar oportunidades de negócios, de parcerias, áreas de interesse crescente, etc.

Além de recomendações, o projeto poderá contribuir para a especificação ou melhoria de uma ferramenta - ou metodologia - de análise de *logs*, que monitore acessos ao *site* a partir de um enfoque de maior utilidade para a área de *marketing* estratégico. Poderá, também, apontar a necessidade de bases de dados estruturadas, modeladas a partir das necessidades dos clientes e usuários do Centro.

Para o aluno, a identificação de elos entre a sua prática diária e o corpo teórico técnico-científico da área de administração de SI contribui, de um lado, para o alargamento da sua base teórica, e de outro, para aumentar suas qualificações como profissional da área de SI. Uma melhor compreensão de métodos de pesquisa quantitativa e qualitativa também é um fator que contribui para o aperfeiçoamento de seus métodos e técnicas no ensino universitário.

Em relação ao contexto imediato, as recomendações que os resultados obtidos poderão apontar, serão implementadas nas páginas do site em estudo no sentido de torna-lo alinhado com as expectativas dos usuários, clientes, parceiros e visitantes ocasionais. Apontar possíveis oportunidades estratégicas e nichos de consumo para informações tecnológicas no âmbito da missão da unidade considerada são, também, possibilidades. Os resultados poderão ajudar também na identificação de oportunidades para promoção de negócios tecnológicos, na identificação de oportunidades para a promoção da imagem da empresa, na transferência de tecnologias e divulgação das atividades da empresa junto ao seu mercado-alvo, em futuros estudos visando implementar transações comerciais eletrônicas entre a Embrapa e seus clientes e parceiros e na elaboração de *FAQs* (*Frequently Asked Questions*) sobre assuntos de interesse de clientes, parceiros e visitantes ocasionais;

REFERÊNCIAS BIBLIOGRÁFICAS:

- ABDULLA, G.; FOX, E.A. & ABRAMS, M. Shared User Behavior on the World Wide Web. Proceedings of the WebNet 97 – World Conference of WWW, Internet and Intranet. Toronto-Canada, Nov., 1-5, 1997.
- ATTARDI, G.; DI MARCO, S. & SALVI, D. Categorization by context. Proceedings of the WebNet 98 – World Conference of WWW, Internet and Intranet. Orlando-Flórida, Nov., 1-12, 1998.
- BERNES-LEE, T. Frequently Asked Questions. Available from World Wide Web: <<http://www.w3.org/People/Bernes-Lee/FAQ.html>>, consulta em abr./1999.
- CASTELS, M. A sociedade em rede. São Paulo: Paz e Terra, 1999. 617p. (A era da informação: economia, sociedade e cultura, 1).
- DOAN, B.; BEIGBEDER, M.; GIRARDOT, J. & JAILLON, P. Using metadata to improve organization and information retrieval on the WWW. Proceedings of the WebNet 98 – World Conference of WWW, Internet and Intranet. Orlando-Flórida, Nov., 1-12, 1998.
- ESTRATÉGIA GERENCIAL DA EMBRAPA: GESTÃO 95/98. Empresa Brasileira de Pesquisa Agropecuária. D.E., Brasília, 1995. 27p.
- ESTRATÉGIA GERENCIAL DA EMBRAPA: MACROPRIORIDADES/1997. Empresa Brasileira de Pesquisa agropecuária. D.E., Brasília, 1997, 27p.
- FLORES, M.X. Projeto EMBRAPA: a pesquisa agropecuária rumo ao século XXI. Brasília: EMBRAPA-SEA, 1991. 38p. (EMBRAPA-SEA. Documentos, 4).
- FLORES, M.X. & SILVA, J. de S. Projeto EMBRAPA II: do projeto de pesquisa ao desenvolvimento sócio-econômico no contexto do mercado. Brasília: EMBRAPA-SEA, 1992. 55p. (EMBRAPA-SEA. Documentos, 8).
- FRANÇA, J.L.; BORGES, S.M.; VASCONCELLOS, A.C.de; MAGALHAES, M.H.A. Manual para normalização de publicações técnico-científicas. 4.ed. rev. amp. Belo Horizonte: Ed. UFMG, 1998. 213p.
- FREITAS, H. A Informação como ferramenta gerencial. Porto Alegre: Ortiz, 1993. 355p. Cap.4: Um método implícito e automático para a avaliação de um sistema de

apoio a decisão, p.161-199.

FREITAS, H. A Informação como ferramenta gerencial. Porto Alegre: Ortiz, 1993. 355p. Cap.5: Uma avaliação do SADAD. p. 203-290.

FREITAS, H. A Informação como ferramenta gerencial. Porto Alegre: Ortiz, 1993. 355p. Cap.6: Uma tipologia comportamental dos usuários finais. p.295-319.

FREITAS, H.; CUNHA JR; M.V.M. & MOSCAROLA, J. Pelo resgate de alguns princípios da análise de conteúdo: uma aplicação prática qualitativa em marketing. Porto Alegre: UFRGS/FCE/PPGA, 1996. 17p. (Série Documentos pra Estudo, PPGA/UFRGS, 05/96).

FREITAS, H. & MOSCAROLA, J. Content analysing qualitative data on information systems. Americas Conference on Information Systems. Baltimore, Maryland.. Aug., 14-16, 1998.

GATES, B. A empresa na velocidade do pensamento. São Paulo: Companhia das Letras, 1999. 444p.

GIL, A.C. Como elaborar projetos de pesquisa. 3.ed. São Paulo: Atlas, 1996. 159p.

GIL, A.C. Métodos e técnicas de pesquisa social. 5. ed. São Paulo: Atlas, 1999. 206p.

GONÇALVES, C. A. & FILHO, C. G. Tecnologia da Informação e *Marketing*. Como obter clientes e mercados. Revista de Administração de Empresas – RAE. São Paulo, v.35, n.4, p.21-32. Jul./Ago, 1995.

JANSEN, B.J.; SPINK, A.; BATERMAN, J. & SARACEVIC, T. Searchers, the subject they search, and sufficiency: a study of a large sample of Excite searchers. Proceedings of the WebNet 98 – World Conference of WWW, Internet and Intranet. Orlando-Flórida, Nov., 1-12, 1998.

JOHN, G., PANAGIOTIS, M.D. How to use HTML page popularity to improve a web site's structure. Proceedings of the WebNet 98 – World Conference of WWW, Internet and Intranet. Orlando-Flórida, Nov., 1-12, 1998.

KOTLER, P. Administração de Marketing; análise, planejamento, implementação e controle. 8. ed. São Paulo: Atlas, 1996. ?p.

LYRA, J. de; MANDEL, A. e SIMON, I. Informação: Computação e Comunicação. Universidade de São Paulo. Julho de 1997. 42p. Disponível na World Wide Web <<http://www.usp.br/...>>, consulta em jan./98.

MONTGOMERY, D. B. & WEINBERG, C. B. Toward Strategic Intelligence Systems. Marketing Management, Winter, 1998.

- OHMAE, K. Começando de novo. Revista HSM Management, n.11, ano 2, nov/dez 1998. p.6. (Entrevista).
- PENZIAS, A. Lições de um Prêmio Nobel. Revista HSM Management. n.11, ano 2 – nov/dez 1998. p.30. (Entrevista).
- PLANO DIRETOR DA EMBRAPA: REALINHAMENTO ESTRATÉGICO 1999-2003. Brasília: Embrapa-SPI, 1998.36p.
- PLANO DIRETOR DO CENTRO DE PESQUISA DE PECUÁRIA DOS CAMPOS SULBRASILEIROS. Brasília: EMBRAPA-SPI, 1993, 42p.
- POLÍTICA DE COMUNICAÇÃO. Embrapa, 1996. 57p. (documento para uso interno).
- POLÍTICA DE NEGÓCIOS TECNOLÓGICOS. BRASÍLIA: Embrapa-SPI, 1998. 44p. (documento para uso interno).
- POPINIGIS, F.; BRANDINI A.; LIMA, S.M.V. & MENDONÇA, S.J.B de Gestão pela Qualidade Total. In: GOEDERT, W.J.; PAEZ, M.L.D'A. & CASTRO, A.M.G. de Gestão em ciência e tecnologia: pesquisa agropecuária. Brasília, EMBRAPA-SPI, 1994.
- SAKAMOTO, Y. Tracking web user behavior using event hooks. Proceedings of the WebNet 98 – World Conference of WWW, Internet and Intranet. Orlando-Flórida, Nov., 1-12, 1998.
- SALAM, A.F.; RAO, H.R. & PEGELS, C.C. Na exploratory analysis of information content of corporate web pages as adversing media.17th International Conference on Information Systems. Dec. 16-18, 1996. Cleveland-Ohio.
- SILVA, J. A. R. e O Marketing na Internet.br – Uma avaliação da presença empresarial na *World Wide Web*. Anais ANPAD 1997. (Edição em CD)
- SOARES, R. & HOPPEN, N. Aspectos do uso da *Internet* nos negócios pelas grandes empresas no Brasil: um estudo exploratório baseado em *Sites Web*. Anais ANPAD 1998. (edição em CD)
- SPILIOPOULOU, M. & FAULSTICH L.C. WUM: A tool for web utilization analysis. In EDBT Workshop. WebDB'98, Valencia, Spain, Mar. 1998. Available from Word Wide Web: <<http://wum.wiwi.hu-berlin.de/wumDescription.html#Publications>>
- SPILIOPOULOU, M.; FAULSTICH L.C. & WINKLER, K. A data miner analysing the navigational behavior of web users. In Proc. of the Workshop on Machine Learning in user modeling. ACAI'99, Int. Conf., Creta, Greece, July 1999. Available from Word Wide Web: <<http://wum.wiwi.hu-berlin.de/wumDescription.html#Publications>>

- SPINK, A.; BATERMAN, J. & JANSEN, B.J. User's searching behavior on the Excite web search engine. Proceedings of the WebNet 98 – World Conference of WWW, Internet and Intranet. Orlando-Flórida, Nov., 1-12, 1998.
- TAPSCOTT, D. O que esperar do mundo digital. Revista HSM Management n.12, ano 2, jan/fev 1999. p.6. (Entrevista).
- TAPSCOTT, D. & CASTON, A. Mudança de paradigma: a nova promessa da tecnologia da informação. São Paulo: Makron-McGraw-Hill, 433p.
- TRIVIÑOS, A.N.S. Introdução a pesquisa em ciências sociais: a pesquisa qualitativa na educação. São Paulo: Atlas, 1987, 175p.
- ZAWISLACK, P.A. Uma proposta de estrutura analítica para sistemas técnicos-científicos: o caso do Brasil. Economia & Empresa, São Paulo, v.3, n.2, p4-29, abr./jun. 1996.
- W3C - World Wide Web Consortium. Logging Control In W3C httpd. Available from Word Wide Web <[http://www.w3.org/Daemon /User/Config/ logging.html](http://www.w3.org/Daemon/ /User/Config/ logging.html)>, consulta em 07/10/1999.
- YIN, R.K. Case Study Research: design and methods. 2nd ed. London: Sage Publications, 1994. 171p. (Applied social research methods series,.5)