

ORLINDO EUGENIO

**BASE DE DADOS GENÉRICA DE PERGUNTAS E RESPOSTAS
ALIMENTADA POR ENQUETES E PESQUISAS
VIA WEB**

Monografia apresentada à disciplina de Projeto Final I, do curso de Bacharelado em Sistemas de Informação da Faculdade Expoente

Professora Elaine Simoni Angelotti

**CURITIBA
2008**

1 INTRODUÇÃO

Para compreender a importância que a tecnologia da informação tem para as organizações atualmente e, principalmente, para entender que o grande diferencial do fator de desequilíbrio entre elas, está nas pessoas e na forma como elas criam, sistematizam, utilizam e compartilham o conhecimento, vale atentar para a evolução de nossa sociedade que passa pela transição da sociedade industrial para a sociedade pós-industrial, ou anteriormente chamada de sociedade da informação e, agora, de sociedade do conhecimento (MÜLLER, 2003).

E nesta sociedade do conhecimento, fruto do desenvolvimento das tecnologias da informação, todo dia novas tecnologias ou novas aplicações para as tecnologias já existentes trazem algum avanço, facilidade ou agilidade aos processos do cotidiano das pessoas e das organizações. E a internet tornou-se o principal meio para a disseminação da informação nesta nova sociedade.

Direta ou indiretamente, todas as atividades humanas podem se beneficiar dessa rede de comunicações que cobre nosso planeta. Com isso, todas as áreas do conhecimento humano têm tido um ganho significativo de resultados em pesquisas e conseqüentemente, novas descobertas, novas tecnologias e mais agilidade na manipulação do conhecimento, criando assim um ciclo de evolução exponencial.

A evolução para um modelo de aprendizagem mediada pela tecnologia traz, em si, um novo leque de possibilidades, permitindo a formação de comunidades virtuais construídas sobre as afinidades de interesses, de conhecimentos, sobre projetos mútuos, em processos de cooperação ou de troca, tudo isso independentemente das proximidades geográficas e das filiações institucionais.

A Internet e as Intranets, ou seus sucedâneos, certamente terão um papel fundamental nesse processo. O aspecto mais evidente está no fato de que a Internet caminha rapidamente para se tornar o grande repositório que armazenará todo tipo de informação que for tornada pública no mundo daqui para frente (MÜLLER, 2003).

Vê-se com isso que a sociedade humana como um todo foi definitivamente afetada e as pessoas e empresas não podem mais agir como se fosse possível manter velhos paradigmas¹ de um mundo desconectado.

Uma informação publicada hoje corre o mundo em segundos e mesmo tendo consciência deste fato, às vezes as pessoas se sentem seguras em acreditar que um fato negativo publicado a seu respeito não vai ter grande repercussão. Isso torna-se um risco proporcionalmente maior quanto mais popular for a pessoa ou empresa, como no caso da Intel, em 1994.

A história da economia digital vai lembrar o mês de novembro de 1994 como divisor de águas. As dolorosas experiências da Intel Corporation foram seminais. No final de outubro de 1994, as coisas iam bem na Intel. O chip Pentium — produto estratégico daquela organização — recebia altos elogios da crítica e era um enorme sucesso de venda. As receitas e os lucros estavam sólidos. Em 30 de outubro, o professor Thomas Nicely, do Lynchburg College, postou uma mensagem na Internet reportando uma falha no chip. Em poucos dias a mensagem já havia circulado no ciberespaço entre milhares de pessoas e estava formada uma multidão furiosa que rivalizava com os maiores protestos da década de 1960. A Intel tinha uma crise em mãos — mais perigosa do que qualquer outra em sua história.

A Intel demorou a compreender a dinâmica dos mercados eletrônicos, menosprezando o problema com declarações clássicas e unidirecionais do seu departamento de relações públicas. Em 27 de novembro, a declaração do CEO Andy Grove na Net fez aumentar a indignação, o mesmo ocorrendo com a declaração do diretor do laboratório de software, Richard Wirt, no dia 8 de dezembro. No dia 12 de dezembro, a IBM, participante secundária na venda do Pentium e concorrente da Intel com seu chip PowerPC, anunciou na Net e à mídia que ela estava interrompendo as remessas de IBM PCs com chips Pentium. Surgiu um servidor da Web com piadas sobre o Pentium. A comoção espalhou-se para toda a mídia e, em 20 de dezembro, a Intel admitiu seu erro e concordou em fazer um recall generalizado do chip Pentium (Tapscott, 1997, p. 102).

¹ Segundo Fergusson (1980, p. 26), um paradigma é uma estrutura de pensamento, [...] um esquema para a compreensão e a explicação de certos aspectos da realidade.

Neste vasto cenário, os profissionais de informática tem um verdadeiro paraíso de possibilidades a serem estudadas, implementadas, experimentadas e oferecidas ao mercado. E ninguém sabe exatamente o que virá após esse ciclo vertiginoso da evolução humana que aparentemente está apenas começando.

1.1 CONTEXTUALIZAÇÃO

Na vida acadêmica, estudantes e professores não estão diretamente envolvidos com a administração de uma entidade de ensino, entretanto, a vivência nesse meio revela uma peculiaridade desse ramo de atividade. Uma característica inerente a uma empresa moderna, em qualquer área de atuação, é a necessidade de feedback do negócio e informações gerenciais que possibilitem a rápida tomada de decisões e o melhor impacto possível ao menor custo para estas decisões; e diferentemente de empresas de outras áreas, numa instituição de ensino o relacionamento com seus alunos e pais de alunos é muito próximo. O convívio é diário. Os alunos passam muito tempo nas instalações da instituição, estudando e usufruindo de sua infra-estrutura. Já os pais, ou responsáveis, têm grande interesse na satisfação dos filhos com a instituição e também no conteúdo do ensino aplicado.

Mas, para se obter um feedback desses pais e alunos, um processo manual de pesquisa de satisfação é trabalhoso e inadequado. Uma quantidade elevada de avaliadores com preenchimento de formulários em papel torna uma avaliação demorada e o resultado da avaliação que poderia apontar anseios dos clientes e soluções para o negócio, não sai em tempo adequado, pois depende de transcrição para um sistema informatizado.

Quando o resultado chega a ser apresentado, pode ser tarde para pequenas correções estratégicas. As mudanças necessárias exigem então maiores investimentos, ou aquele anseio do cliente já evoluiu e a solução quando apresentada atende apenas parcialmente, refletindo uma realidade de meses atrás. Certamente este problema também acomete empresas de muitas outras áreas de atividade.

A criação de uma ferramenta genérica com base dinâmica de perguntas e respostas, alimentada por enquetes e pesquisas, realizadas por usuários através da Web, pode ser uma solução viável para todo esse mercado. E ao mesmo tempo, com a difusão do uso da ferramenta, a formação de uma base de dados ao longo do

tempo, pode gerar um novo produto, um repositório de questões com possibilidades de gerar opiniões e estatísticas diversas. Essa base será construída em dois pilares principais: perguntas e respostas. A relação simples entre perguntas e respostas permitirá ao sistema adaptar-se a praticamente qualquer assunto, ampliando a sua utilização para todas as áreas em que se tome conhecimento de sua existência e queira-se fazer pesquisas on-line.

A idéia de montar, com o passar do tempo, uma base grande o suficiente de perguntas e respostas inter-relacionadas, nas mais variadas configurações e tratando dos mais diversos assuntos e ainda de um público multicultural e representativo de boa parte da sociedade, que possibilite extrair conhecimento e que permita tecer opiniões culturais e estatísticas de mercado é forte estímulo motivacional para este projeto.

A difusão de uso e alimentação da base, apoiada na idéia de uma boa ferramenta de consultas, pesquisas e enquetes, de uso inteiramente grátis, disponível pela internet, altamente customizável, com resultados on-line e downloads em XML e CSV, fortalece a viabilidade do projeto.

1.2 OBJETIVO GERAL

Desenvolver uma ferramenta genérica para criação de enquetes e armazenamento de respostas de forma a se obter uma base dinâmica de perguntas e respostas. Esta base será alimentada por usuários e ordenada por assuntos, com três tipos de usuários padrão, sendo: Administrador, Pesquisador e Avaliador, em ambiente web.

1.3 OBJETIVOS ESPECÍFICOS

Para que se possa alcançar o objetivo geral supracitado será necessário o cumprimento dos seguintes objetivos específicos:

- Criar uma base de dados que suporte a taxa de crescimento ao longo do tempo e que seja estruturada de tal forma a permitir a extração de conhecimento;
- Criar um sistema Web de enquetes e pesquisas customizável, onde o usuário Pesquisador poderá organizar as perguntas como lhe convier, fazer

quantas perguntas quiser até o limite de 200 perguntas, escolher o tipo de resposta que deseja e terá opção de ter Avaliadores anônimos ou identificados, neste caso com o consentimento destes;

- Criar um sistema de relatórios e conversão de dados para envio ao Pesquisador, o resultado da pesquisa poderá ser enviado ao pesquisador em formato CSV ou XML;
- Criar manual de instalação;
- Criar manual de ajuda on-line;
- Criar política de privacidade e Termo de uso.

1.4 JUSTIFICATIVA

A falta de uma ferramenta com estas características faz com que as empresas improvisem pesquisas, na maioria das vezes ineficientes e incompletas, demonstrando a existência de um mercado interessante e fácil de ser cativado com uma boa ferramenta.

O desenvolvimento de um projeto de banco de dados com alimentação via Web e com a preocupação em estruturar a base de forma a facilitar futuras aplicações, além do fato de imaginar uma base que suporte pesquisas em qualquer assunto com interface em formato atrativo aos usuários; e mais, a concepção de relatórios genéricos que sejam realmente úteis e levem conteúdo ao Pesquisador, mais que interessante, isso tudo forma um estimulante desafio para aplicação do preparo recebido no curso de Bacharelado em Sistemas de Informação.

1.5 DELIMITAÇÃO

O sistema Enquetes&Pesquisas não permitirá gravação de imagens, tratando apenas de dados nos formatos texto e número.

A customização da apresentação dos resultados on-line das pesquisas não será possível. Os gráficos e relatórios terão formatação estática e apenas os dados serão alterados. O pesquisador terá opções de relatórios por pergunta, por bloco de perguntas, por pesquisa e comparativo entre pesquisas republicadas.

O sistema Enquetes&Pesquisas servirá apenas para a realização de enquetes, pesquisas e visualização de relatórios de percentuais e comparativos de enquetes republicadas, tendo como principal finalidade alimentar a base de dados.

O sistema não vai tratar essas informações ou aplicar qualquer técnica de IA (como Data Mining).

1.6 SUMÁRIO DOS CAPÍTULOS

No próximo capítulo, intitulado revisão da literatura, faz-se um breve estudo do surgimento dos sistemas gerenciadores de bancos de dados, também da estrutura de um banco de dados e suas instâncias e esquemas, além da questão de armazenamento e consulta. Em seguida, explica-se o desenvolvimento de software para Web e faz-se um comparativo entre desenvolvimento Web e desenvolvimento Desktop. A gestão da informação é o próximo assunto. Demonstra-se a importância da informação, apresenta-se a hierarquia da informação e esclarece-se de forma sintética a diferença entre gestão da informação e gestão do conhecimento. Um tópico importante deste capítulo é a informação coletada por formulários, explicando algumas dificuldades desta tarefa. Por fim comenta-se algumas modalidades de sistemas com bases dinâmicas.

O capítulo 3 apresenta alguns trabalhos relacionados com o tema deste projeto, tanto trabalhos acadêmicos como trabalhos comerciais.

O capítulo 4 trata da metodologia, explicando como foi feita a coleta de dados e a análise. Mostra também algumas restrições de acesso, apresenta o ambiente de desenvolvimento e execução escolhidos. Dá-se uma visão do que é desenvolvimento Web em plataforma Microsoft .NET, das características do banco de dados IBM DB2 Express-C 9, comenta-se a geração de relatórios e as configurações mínimas para execução do sistema. Também passa-se uma descrição das funcionalidades do sistema e o cronograma de desenvolvimento. Termina-se este capítulo com uma análise de riscos e uma descrição dos testes que serão realizados com o sistema.

O quinto e último capítulo apresenta os diagramas utilizados para a modelagem do sistema.

2 REVISÃO DA LITERATURA

2.1 SURGIMENTO DOS SISTEMAS GERENCIADORES DE BANCO DE DADOS

Um banco de dados é um repositório de informações. Segundo Date (2000),

Em essência, um sistema de bancos de dados é apenas um sistema computadorizado de armazenamento de registros. O banco de dados pode, ele próprio, ser visto como o equivalente eletrônico de um armário de arquivamento. Em outras palavras, é um repositório ou recipiente para uma coleção de arquivos de dados computadorizados.

Elmasri (2005) afirma que, para uma coleção de dados ser considerada como um banco de dados, deve possuir as seguintes propriedades implícitas:

- Um banco de dados representa alguns aspectos do mundo real, sendo chamado, às vezes, de minimundo ou de universo de discurso. As mudanças no minimundo são refletidas em um banco de dados.
- Um banco de dados é uma coleção lógica e coerente de dados com algum significado inerente. Uma organização de dados ao acaso (randômica) não pode ser corretamente interpretada como um banco de dados.
- Um banco de dados é projetado, construído e povoado por dados, atendendo a uma proposta específica. Possui um grupo de usuários definido e algumas aplicações preconcebidas, de acordo com o interesse desse grupo de usuários.

O banco de dados, com o passar do tempo, deixou de ser mais um componente de sistemas computacionais, com a finalidade de guardar as informações para simples consultas. A descoberta de que “a coleção de dados, normalmente chamada de banco de dados, contém informações relevantes a uma empresa” (Silberschatz, 2006, p. 1) trouxe nova importância ao banco de dados.

A revolução da Internet no final da década de 90 aumentou muito o acesso de usuários a bancos de dados (Silberschatz, 2006, p. 2). Com isso, o crescimento dos bancos de dados ganhou um impulso que não perdeu mais força, muito pelo contrário, há cada vez mais e maiores bancos de dados dando suporte aos mais variados assuntos e negócios.

Os sistemas mais antigos de gerenciamento computadorizado de dados, com a finalidade de permitir o acesso e a manipulação dos dados, contava com uma

coleção de programas de aplicação que trabalhavam com os arquivos. Era um “sistema de processamento de arquivos”. Os sistemas de bancos de dados surgiram em resposta a esses métodos mais antigos de gerenciamento computadorizado de dados comerciais (Silberschatz, 2006, p. 2).

Com o tempo, esses sistemas de processamento de arquivos mostraram-se deficitários em controlar os dados. Inúmeras desvantagens surgiram. Conforme Silberschatz (2006, p. 2) as mais importantes foram:

- Redundância e inconsistência de dados: informações duplicadas em vários arquivos elevando desnecessariamente o custo de armazenamento e causando problemas de inconsistência de dados por alterações parciais nas várias cópias do mesmo registro;
- Dificuldade de acesso aos dados: a dinâmica das necessidades de diferentes visões de um determinado universo de registros exige aplicações específicas para cada foco que se deseja, ou a manipulação das informações a partir de algum relatório ou planilha padrão que se tenha disponível causando, por vezes, excessivo trabalho de filtragem e seleção manual dos dados;
- Isolamento de dados: a dispersão dos dados em vários arquivos, muitas vezes de diferentes formatos dificulta a elaboração de novas aplicações para acessar esses dados;
- Problemas de integridade: são causados pela necessidade de se aplicar restrições de consistência, como este tipo de restrição é comum, aplicá-las a um grupo grande de itens de dados espalhados em diversos arquivos, gera um grande trabalho;
- Problemas de atomicidade: muitas vezes, na ocorrência de falha numa transação ou operação, é de crucial importância que os dados retornem ao estado inicial da transação para evitar inconsistências. Garantir atomicidade em uma transação num sistema de processamento de arquivos é uma tarefa muito difícil;
- Anomalias de acesso concorrente: para se garantir a correta manipulação dos dados por usuários ou aplicações em acessos simultâneos é necessário que haja uma coordenação rigorosa entre os acessos, o que também é difícil de se conseguir em sistemas de processamento de arquivos;

- Problemas de segurança: impor restrições de segurança em um ambiente em que a toda hora novas aplicações estão sendo integradas e estão tendo acesso aos arquivos é praticamente impossível.

Os sistemas gerenciadores de bancos de dados (SGBD) surgiram para sanar estes problemas e dar mais confiança às informações das instituições.

Um sistema de gerenciamento de banco de dados é uma coleção de dados inter-relacionados e um conjunto de programas para acessar esses dados. A coleção de dados normalmente chamado de banco de dados, contém informações relevantes a uma empresa ou domínio. O principal objetivo de um SGBD é fornecer uma maneira de recuperar informações de um banco de dados que seja tanto conveniente quanto eficiente (Silberschatz, 2006, p. 1).

Elmasri (2005, p. 4) diz que o SGBD é um sistema de software de propósito geral que facilita os processos de definição, construção, manipulação e compartilhamento de bancos de dados entre vários usuários e aplicações.

Sendo inegável a importância que os bancos de dados possuem em nossos dias, a adoção de um banco de dados não gerenciado na atualidade só pode ser pensada para pequenas aplicações domésticas, como por exemplo, pequenos arquivos pessoais de fotos ou agenda de contatos de clientes.

2.1.1 Estrutura de um banco de dados

Um banco de dados pode ser desenvolvido em diversas modelos. Hoje predomina no mercado o modelo relacional e começa a ser utilizado um modelo orientado a objetos, o qual ainda é muito pouco utilizado. Antes de haver o modelo relacional, os mais comuns eram o modelo hierárquico e o modelo de rede. Entretanto, conforme Date (2001, p. 21), os produtos de SGBDs baseados no modelo relacional de dados (“sistemas relacionais”) vieram para dominar o mercado de bancos de dados. Mais que isso, a vasta maioria das pesquisas sobre bancos de dados ao longo dos últimos trinta anos também se baseou — embora um tanto indiretamente, em alguns casos — nesse modelo. De fato, a introdução do modelo relacional em 1969—1970 foi sem dúvida o evento mais importante em toda história da área de bancos de dados. Assim sendo, apenas o modelo relacional será objeto deste trabalho.

Neste modelo, a fim de atingir altos níveis de eficiência e desempenho em bancos de dados, os projetistas utilizam complexas estruturas de dados. Entretanto, para facilitar a interação com os usuários, os projetistas ocultam esta complexidade através do uso de camadas de abstração.

Segundo Machado (2006, p. 21), toda realidade é sempre, em princípio, bastante nebulosa e informal. Através da observação podemos extrair dela (realidade) fatos que nos levam a conhecê-la de uma forma mais organizada.

Para que se possa elaborar um projeto de banco de dados, faz-se necessário também o uso de abstração. Temos então, duas situações de abstração em um projeto de banco de dados: uma na análise inicial e outra na apresentação para o usuário.

Também conforme Machado (2006, p. 22), para registrarmos as necessidades de informação de uma realidade, precisamos fazer uso de um modelo, ou seja, algo que nos mostre como as informações estão relacionadas (fatos). E com base no modelo criado, os analistas podem interagir com os usuários validando seus objetivos e metas, permitindo a construção de um sistema de informações cada vez mais próximo da realidade do usuário.

Na visão de Silberschatz, o processo de construção do projeto desenvolve-se da seguinte forma:

A fase inicial do projeto de banco de dados é caracterizar completamente as necessidades de dados dos prováveis usuários do banco de dados. O projetista de banco de dados precisa interagir extensivamente com especialistas de domínio e usuários para realizar essa tarefa. A saída dessa fase é uma especificação das necessidades do usuário.

Em seguida, o projetista escolhe um modelo de dados e, aplicando os conceitos do modelo de dados escolhido, traduz essas necessidades para um esquema conceitual do banco de dados.

Um esquema conceitual completamente desenvolvido também indica as necessidades funcionais da empresa. Em uma especificação das necessidades funcionais, os usuários descrevem os tipos de operações (ou transações) que serão realizadas nos dados. Operações de exemplo incluem modificar ou atualizar dados, pesquisar e recuperar dados específicos e excluir dados. Nessa fase

do projeto conceitual, o projetista pode revisar o esquema para garantir que ele atenda as necessidades funcionais.

O processo de transição de um modelo de dados abstrato para a implementação do banco de dados ocorre nas duas fases finais do projeto.

Na fase de projeto lógico, o projetista mapeia o esquema conceitual de alto nível para o modelo de dados de implementação do sistema de banco de dados que será usado. O modelo de dados de implementação normalmente é o modelo de dados relacional, e essa etapa geralmente consiste em mapear o esquema conceitual definido usando o modelo entidade-relacionamento em um esquema de relação.

Finalmente, o projetista usa o esquema de banco de dados específico do sistema resultante na fase do projeto físico, em que os recursos físicos do banco de dados são especificados. Esses recursos incluem a forma de organização de arquivo e as estruturas de armazenamento internas (2006, p. 133).

Para simplificar as interações do usuário com o sistema, Silberschatz (2006) apresenta os níveis de abstração da seguinte forma:

- Nível físico: é o nível mais baixo e descreve em detalhes como os dados são realmente armazenados.
- Nível lógico: um nível acima do físico. Descreve que dados estão armazenados no banco de dados e que relações existem entre eles. De forma simples, descreve o banco de dados inteiro em poucas estruturas.
- Nível de visão: o nível de abstração mais alto descreve apenas parte do banco de dados. Existe para simplificar a interação com o sistema. Muitos usuários precisam acessar apenas uma parte do banco de dados e o sistema pode fornecer muitas visões para o mesmo banco de dados.

2.1.2 Instâncias e Esquemas

É usual referir-se aos dados em um banco de dados como “persistentes” (mesmo que eles na verdade não persistam por muito tempo!). Por persistente queremos sugerir intuitivamente que os dados desse banco de dados diferem em espécie de outros dados mais

efêmeros, como dados de entrada, dados de saída, instruções de controle, filas de trabalho, blocos de controle de software, resultados intermediários e, em geral, quaisquer dados que tenham natureza transiente. De modo mais preciso, dizemos que os dados do banco de dados “persistem” porque, uma vez aceitos pelo SGBD para entrada inicial no banco de dados, eles só podem ser removidos subsequentemente do banco de dados por alguma solicitação explícita ao SGBD, não como um mero efeito colateral de (por exemplo) conclusão da execução de um programa. Essa noção de persistência nos permite portanto dar uma definição ligeiramente mais precisa do termo banco de dados”: Um banco de dados é uma coleção de dados persistentes utilizada pelos sistemas de aplicação de uma determinada organização (Date, 2000, p. 9).

Os bancos de dados mudam com o tempo, à medida que as informações são inseridas e excluídas. A coleção das informações armazenadas no banco de dados em um determinado momento é uma instância do banco de dados. O projeto geral do banco de dados é o esquema do banco de dados (Silberschatz, 2006, p. 5).

O esquema é estático, uma vez definido não pode sofrer alterações, as instâncias, ao contrário, são constantemente sobrepostas e cada instância é diferente das anteriores.

No desenvolvimento de um sistema com base dinâmica, onde se espera coletar dados estatísticos por intervalos de tempo, por exemplo, deve-se ter especial atenção à característica de instância da base de dados. Ou seja, quais informações, eram válidas em determinada instância.

Uma solução para recuperar dados em diversas instâncias ao longo do tempo é a modelagem de dados temporais.

De acordo com Silberschatz (2006, p. 201), em geral, dados temporais são dados que possuem um intervalo de tempo associado, durante o qual eles são válidos. Existe ainda o modelo onde o atributo temporal indica quando um fato foi registrado no banco de dados.

Quando se pretende utilizar uma base de dados em consultas temporais é importante pensar nas possibilidades que se pretende alcançar com relação às pesquisas que poderão ser feitas na base futuramente.

A inclusão de um atributo temporal em uma chave pode levar à necessidade de alguma mudança de projeto. Por exemplo, vamos supor de modo bastante

artificial, que numa determinada empresa, a quantidade de cada remessa de um produto específico é determinada pelo mês em que a remessa ocorre. Veja na Figura 1 a tabela demonstrando esse exemplo, onde IDM é o identificador do mês e QTD é quantidade do produto. Então, a cada repetição de um mês qualquer, a quantidade também se repete causando redundância desnecessária e desnormalização², pois existe dependência transitiva entre os atributos IDM → QTD; assim, ela deve ser ainda mais normalizada, como indica a Figura 2. Infelizmente, os projetistas de sistemas de apoio à decisão raramente se preocupam em considerar tais dependências induzidas (Date, 2000, p. 603).

F#	P#	IDM	QDE
F1	P1	3	300
F1	P1	5	100
F1	P2	1	200
F1	P3	7	400
F1	P4	1	200
F1	P5	5	100
F1	P6	4	100
F2	P1	3	300
F2	P2	9	400
F3	P2	6	200
F3	P2	8	200
F4	P2	1	200
F4	P4	8	200
F4	P5	7	400
F4	P5	11	400

**Figura 1 – Amostra de valores para a tabela FP, incluindo IDs de meses
(Fonte: adaptado de Date, 2006, p. 603)**

² O processo de normalização pode ser visto como o processo no qual são eliminados esquemas de relações (tabelas) não satisfatórios, decompondo-os, através da separação de seus atributos em esquemas de relações menos complexas mas que satisfaçam as propriedades desejadas.

O processo de normalização como foi proposto inicialmente por Codd conduz um esquema de relação através de um bateria de testes para certificar se o mesmo está na 1a, 2a e 3ª Formas Normais. Estas três Formas Normais são baseadas em dependências funcionais dos atributos do esquema de relação (Marília, p. 27).

F#	P#	IDM
F1	P1	3
F1	P1	5
F1	P2	1
F1	P3	7
F1	P4	1
F1	P5	5
F1	P6	4
F2	P1	3
F2	P2	9
F3	P2	6
F3	P2	8
F4	P2	1
F4	P4	8
F4	P5	7
F4	P5	11

IDM	QDE
1	300
2	100
3	200
4	400
5	200
6	100
7	100
8	300
9	400
10	200
11	200
12	200

**Figura 2 – Equivalente normalizada da Figura 1
(Fonte: Adaptado de Date, 2006, p. 604)**

A falta de um atributo importante ou até mesmo uma tabela que, a princípio, não esteja no escopo do projeto e não faça parte dos requisitos dos usuários iniciais do sistema, pode derrubar qualquer possibilidade de aproveitamento da base para um futuro trabalho de mineração para estatísticas temporais, por exemplo, caso o sistema seja colocado em produção sem a percepção desta falha.

Obviamente, nem tudo pode ser previsto e mudanças num projeto de software são inevitáveis ao longo da sua elaboração. No entanto, quanto menos mudanças precisarem ser feitas ou quanto menor seu impacto, melhor, porque todas as dificuldades e riscos do projeto devem ter sido medidos e avaliados na fase de análise. Alterações ao longo da elaboração do projeto podem gerar novas dificuldades na execução e no cumprimento do cronograma.

2.1.3 Armazenamento e consulta

As principais funcionalidades de um sistema de banco de dados podem ser divididas em dois componentes básicos: gerenciador de armazenamento e processador de consulta.

O gerenciador de armazenamento é o componente que faz a interface entre as aplicações que acessam o banco de dados e os dados armazenados fisicamente em disco; ele faz isso através de conversões das instruções DML³ em comandos do sistema de arquivos do sistema operacional.

O gerenciador de armazenamento é importante porque os bancos de dados normalmente exigem uma grande quantidade de espaço de armazenamento e é fundamental que o sistema de banco de dados estruture os dados de modo a minimizar a necessidade de mover dados entre disco e memória principal (Silberschatz, 2006, p. 13).

Silberschatz (2006, p. 13) também escreveu que o processador de consulta é importante porque ajuda o sistema de banco de dados a simplificar o acesso aos dados. As visões de alto nível ajudam a alcançar esse objetivo; com elas, os usuários do sistema não são desnecessariamente afligidos com detalhes físicos da implementação do sistema.

O processador de consulta tem a finalidade de interpretar e registrar as instruções DDL⁴, além de otimizar as consultas através de um compilador DML que traduz as instruções DML em instruções de baixo nível avaliando e selecionando a abordagem de menor custo para a consulta dentre as opções possíveis.

2.2 DESENVOLVIMENTO WEB

Nos dias de hoje, muitas empresas têm um vasto parque de equipamentos interconectados, entre computadores, impressoras, PDA's, smartphones, equipamentos diversos de automação, etc. Isto é possível porque houve um grande desenvolvimento em tecnologias de rede para realizar esta interconexão.

Porém, conforme Tanenbaum (2003), talvez mais importante que compartilhar recursos físicos, como scanners, impressoras e gravadores de CDs, seja compartilhar informações. Toda empresa de grande e médio porte e muitas empresas pequenas têm uma dependência vital de informações computadorizadas.

³ DML (Data Manipulation Language) é uma linguagem de manipulação de dados e está incluída na SQL (Structured Query Language) que é a linguagem mais utilizada em bancos de dados.

⁴ DDL (Data Definition Language) é uma linguagem de definição de dados e também está incluída na SQL.

A própria informação, como veremos mais adiante, evoluiu com o tempo, e passou de um patrimônio a ser guardado ao nível de um recurso a ser distribuído e trabalhado para promover o crescimento da organização que a detém.

Tanenbaum também nos alerta para o fato de que nem a Internet e nem a World Wide Web é uma rede de computadores. Pois, a Internet é uma rede de redes e a Web é um sistema distribuído⁵ que funciona na Internet (2003).

A Internet foi vislumbrada pelo cientista J. C. R. Licklider, do MIT, em agosto de 1962. O conselheiro de ciência do presidente Roosevelt durante a Segunda Guerra Mundial. [...] Licklider falou de uma “rede galáctica”, uma rede globalmente interconectada em que qualquer computador podia acessar diretamente quaisquer outros e trocar dados (Meyer, 2000, p. 280).

Mas, a Internet vai além de meras trocas de dados entre computadores. Ela permite que um computador conectado opere outro computador remotamente, enviando comandos pela rede. Uma chave para o sucesso da Internet é que esses comandos funcionam mesmo se o computador remoto for de marca e modelo diferentes (Meyer, 2000, p. 281). Esta característica, chamada interoperabilidade, garante a comunicação entre computadores de qualquer plataforma.

A interoperabilidade possibilitou a implementação de diversos serviços na Internet, o que impulsionou sua popularidade.

O serviço de correio eletrônico (E-Mail) é sem dúvida o mais popular e consiste de um servidor de caixas postais que permanece 24 horas on-line e recebe suas mensagens para você, guardando-as para que você possa acessá-las quando se conectar; além disso, o serviço também se encarrega de conectar-se com outros servidores de E-Mail e encaminhar corretamente as mensagens que você envia.

Outro serviço bastante utilizado é o World Wide Web (www) que, segundo Meyer (2000, p. 285), é um sistema global de hipertexto implementado na Internet e funciona através de servidores de DNS (*domain name service*) que localizam endereços de computadores na rede por associação de URL's e números de endereço IP⁶. Os servidores WWW (ou Web) disponibilizam o serviço de Home Page's.

⁵ Na prática, um sistema distribuído é um sistema de software instalado em uma rede. O software dá ao sistema um alto grau de coesão e transparência (Tanenbaum).

⁶ Endereço IP (Internet Protocol) é um número em um formato padrão que identifica um ponto de acesso na rede, que pode ser um computador, uma impressora ou outro dispositivo.

O FTP (*file transfer protocol*) é um serviço de transferência de arquivos pela Internet e pode transferir arquivos de um servidor para a sua máquina ou vice e versa.

Existem ainda, muitos outros serviços disponíveis na Internet, como por exemplo: Usenet, IRC, Voip, Pager, MUDs, Telnet, Ping e tantos outros. Mas, não é o foco deste trabalho explicar todos esses serviços.

É fácil perceber que todo esse arsenal de serviços funciona porque existem máquinas conectadas à Internet oferecendo-os. São os servidores. Parte de uma arquitetura conhecida como cliente/servidor e que consiste de um computador servidor que disponibiliza os serviços na Internet e vários computadores clientes que acessam o servidor à distância e fazem uso desses serviços. Essa arquitetura tornou-se um padrão para sistemas Web. Mas como pode um servidor atender a um grande número de clientes ao mesmo tempo, como ocorre com frequência nas redes? Tanenbaum nos explica que, se examinarmos o modelo cliente/servidor em detalhes, veremos que há dois processos envolvidos, um na máquina cliente e um na máquina servidora. A comunicação toma a forma do processo cliente enviando uma mensagem pela rede ao processo servidor. Então, o processo cliente espera por uma mensagem em resposta. Quando o processo servidor recebe a solicitação, ele executa o trabalho solicitado ou procura pelos dados solicitados e envia de volta uma resposta (2003) (Figura 3).

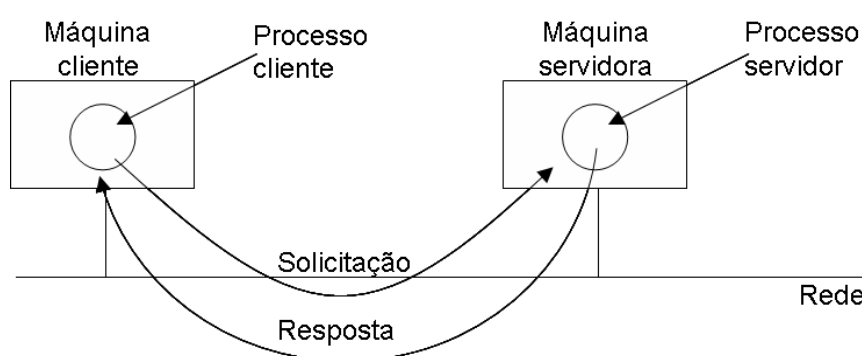


Figura 3 – Modelo cliente/servidor

Desenvolver aplicações para este tipo de ambiente difere das aplicações desktop em muitos aspectos, principalmente na questão da segurança das informações. Para atender uma solicitação de um cliente o servidor precisa identificá-lo ou poderá enviar a resposta ao cliente errado. Outros problemas são a

interceptação de informações, a disseminação de vírus, programas espiões e as invasões de sistemas.

Entretanto, a razão de todos estes problemas é a mesma do sucesso e da popularidade da plataforma Web: o livre acesso à rede.

Com tanta facilidade de comunicação a baixo custo e passados oito anos da explosão de popularidade da Internet⁷, já não há mais dúvidas de que a plataforma Web é o grande canal de comunicação do século XXI. Conforme Ramalho (2005), o grande atrativo da Internet e da Web é a possibilidade de trazer o mundo para a sua área de trabalho com um clique do mouse.

2.2.1 Desenvolvimento Web X desenvolvimento Desktop

Mesmo com os problemas de segurança e todas as dificuldades de conectividade e perda de dados que fazem com que um sistema Web necessite de muitos controles e componentes para garantir seu correto funcionamento, as aplicações nesta plataforma crescem em número e funcionalidades. A necessidade de levar as informações a quem precise de forma rápida, e também tê-las sempre atualizadas, tem levado as empresas a migrarem suas aplicações e processos para esse ambiente.

Aplicações desktop podem se comunicar por redes, mas precisam ser instaladas em cada máquina onde serão utilizadas e muitas vezes têm problemas de portabilidade, ou seja, não são compatíveis com todos os sistemas operacionais em uso nas diversas máquinas.

Além dos problemas de portabilidade, nas aplicações Desktop destinadas a trabalhar em rede, cada máquina irá precisar do driver para conexão com o banco de dados. Também é necessário atualizar o software em cada máquina cada vez que entra em produção uma nova versão.

Já nas aplicações Web não existe instalação nas máquinas do usuário; é necessário apenas um servidor para armazenar o sistema, e todos os usuários o acessam através de um browser, como o Internet Explorer ou Mozilla Firefox, e podem acessar o sistema em qualquer computador, de qualquer lugar, em qualquer browser e em qualquer sistema operacional.

⁷ Ano 2000

Para alguns aplicativos é interessante ou até mesmo necessário que se execute localmente, como antivírus, firewall, player's de vídeo e áudio e outros de uso específico.

As aplicações Web visam, principalmente, o compartilhamento de bases de dados, o trabalho em grupo, a disseminação da informação, comércio eletrônico, trocas de informação on-line e outras.

2.3 GESTÃO DA INFORMAÇÃO

As eras econômicas anteriores foram marcadas por longos períodos de estabilidade, seguidos por curtos períodos de mudança. Os evolucionistas chamariam esse fenômeno de "equilíbrio pontuado". Hoje, as forças da informação digital estão criando um ambiente empresarial em constante mudança. Os evolucionistas chamariam isso de caos pontuado: agitação constante marcada por breves repousos. O ritmo da mudança é por vezes perturbador (Gates, 1999, p. 389).

A gestão da informação em nossa sociedade, hoje, é inegavelmente, fator decisivo para o sucesso, tanto das pessoas quanto das organizações. Não se pode conceber uma organização funcionando de forma estática ao longo dos anos, alheia às mudanças, ou uma pessoa de sucesso que não esteja conectada ao mundo ao seu redor.

Rezende (2003, p. 97) afirma que a informação nos dias de hoje tem um valor altamente significativo e pode representar grande poder para quem a possui, seja pessoa, seja instituição. Atualmente a informação deixou de ser algo secundário ao patrimônio físico e tornou-se parte importante do patrimônio total de uma pessoa ou instituição, um patrimônio intelectual.

A informação e seus respectivos sistemas desempenham funções fundamentais e estratégicas nas organizações em sua totalidade. A informação apresenta-se como recurso estratégico sob a ótica da vantagem competitiva. O enorme salto tecnológico ocorrido na última década transformou, de maneira decisiva, a execução do trabalho. Cada vez mais, as empresas passam a ver na informação o principal recurso estratégico (Rezende, 2003, p. 107).

A informação que antes era vista como um recurso secundário, guardada para quando fosse necessária, e tendo seu conhecimento e disseminação

controlados, passou agora a ser considerada como um recurso a ser utilizado e compartilhado diariamente. Uma ferramenta para a evolução da organização.

2.3.1 A importância da informação

A importância da informação para as organizações é universalmente aceita, constituindo, senão o mais importante, pelo menos um dos recursos cuja gestão e aproveitamento estão diretamente relacionados ao sucesso desejado. A informação também é considerada e utilizada em muitas organizações como um fator estruturante e um instrumento de gestão. Portanto, a gestão efetiva de uma organização requer a percepção objetiva e precisa dos valores da informação e do sistema de informação (Tarapanoff, 2001, p. 111).

Quanto à relevância da informação, Rezende (2003, p. 108) afirma que a informação não se limita a dados coletados. Para que estes dados se transformem em informações, é necessário que sejam organizados e ordenados de forma que se tornem úteis.

A informação também está diretamente ligada à competitividade nas empresas. Com relação a este fato, Rezende (2003, p. 109) explica que ao atuar na definição da estratégia sobre o ambiente competitivo e sobre a organização, a informação auxilia os executivos a identificar tanto as ameaças quanto as oportunidades para a empresa e cria o cenário para uma resposta competitiva mais eficaz.

Os processos de interface com o cliente talvez sejam os mais importantes para o sucesso de uma empresa. A satisfação do cliente e a eficiência dos processos são essenciais para o fluxo de caixa de uma organização. Contudo, são múltiplas as possibilidades de inovação nos processos de contato com o cliente, no gerenciamento de marketing e vendas, e no gerenciamento de processos. (Rezende, 2003, p. 109).

A informação, segundo sua finalidade para uma organização, pode ser classificada em quatro categorias, conforme Tarapanoff (2001, p. 112):

- Informação crítica: crucial para a sobrevivência da empresa;
- Informação mínima: importante para a gestão da organização;
- Informação potencial: adiciona vantagem competitiva;
- Informação sem interesse: lixo.

Devido aos custos operacionais para obtenção da informação, os esforços principais de uma organização devem priorizar a busca e a manutenção da informação crítica, mínima e potencial, respectivamente. Quanto à informação sem interesse, deve-se evitar o desperdício de recursos em sua busca e armazenamento.

Entretanto, classificar a informação não é tarefa trivial. É preciso definir parâmetros para que se possa levar a efeito uma classificação. A informação crítica para uma organização pode ser lixo para outra. Uma empresa que atua no mercado de *commodities*⁸ terá grande interesse em informações relativas ao agronegócio, por exemplo. Alterações climáticas fortes, que prejudiquem as lavouras em determinadas regiões afetarão diretamente o modo de a empresa conduzir suas operações com produtos dessas regiões. As informações advindas dos locais atingidos fundamentarão decisões negociais importantes e terão uma classificação crítica para esta empresa. Informações que para muitas empresas não têm a menor importância, são críticas para o negócio de *commodities* e o custo para a obtenção da informação mais atual possível tem seu retorno numa decisão negocial mais à frente, onde quem tiver a informação mais atual, terá a vantagem de uma visão mais ampla.

A informação deve ser classificada dependendo do contexto da organização, o que implica a própria empresa fazer uma análise minuciosa de suas informações, fontes, custos, etc.. A Figura 4 demonstra como a informação crítica para a organização 1 pode ser informação sem interesse para a organização 2, e vice e versa dependendo do contexto da área de atuação de cada uma.

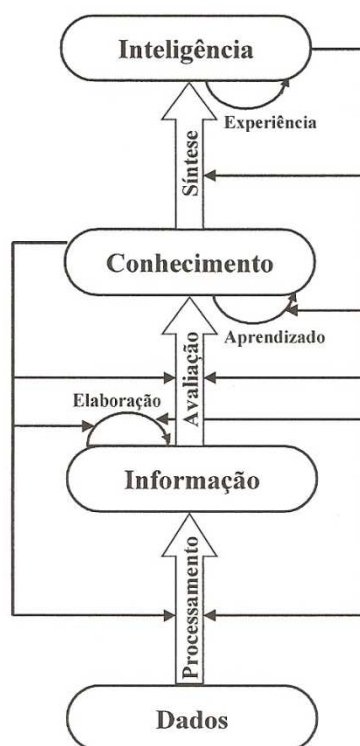


Figura 4 – A informação no contexto das organizações
(Fonte: Adaptado de Tarapanoff, 2001, p. 115)

⁸ Commodities são produtos de consumo, homogêneos, que são produzidos e negociados por uma grande variedade de empresas. Podem ser produtos agropecuários, como gado, milho, café; minerais, principalmente ouro, prata e petróleo; industriais, como tecido de algodão, poliéster, etc.; e financeiros, dólar, euro, ações, títulos, etc.

2.3.2 A hierarquia da informação

Conforme Tarapanoff (2001, p. 117), genericamente, o termo informação é usado para referir-se a todas as maneiras de descrição ou representação de sinais ou dados. Mas é importante reconhecer que existem, de fato, quatro classes diferentes de informação, que são as seguintes: dados, informação, conhecimento e inteligência (Figura 5). Assim sendo, é importante saber as diferenças entre essas classes porque elas têm diferentes valores no contexto de uma organização.



**Figura 5 – Os níveis hierárquicos da informação
(Fonte: Tarapanoff, 2001, p. 118)**

Os dados são itens que representam fatos, textos, números, gráficos, imagens estáticas, sons, segmentos de vídeo, etc. e compreendem a classe mais baixa da hierarquia da informação. Representam a matéria-prima da informação.

A informação é a próxima classe dessa hierarquia. Obtêm-se informação aplicando algum tipo de processamento sobre os dados, como a exibição de arquivos numa tela em formato de texto ou gráfico, por exemplo, ou a elaboração de um relatório consolidado. Segundo Tarapanoff (2001, p. 117), o processo de transformação envolve a aplicação de procedimentos que incluem formatação, tradução, fusão, impressão e assim por diante.

A análise das informações obtidas através do processamento dos dados resulta em uma nova classe de informação. Essa nova informação pode ser uma hipótese para a solução de um problema, uma justificativa para um acontecimento, a descoberta de um novo padrão de relacionamento entre os dados, etc.

O processamento dos dados, então, gerou informações que definiram um quadro de uma determinada situação, que por sua vez levou ao conhecimento de uma realidade num determinado universo de coisas. Assim sendo, o conhecimento é a nova classe da informação.

Tarapanoff (2001, p. 119) diz que o conhecimento não é estático: modifica-se por meio da interação com o ambiente e é denominado processo de aprendizado [...], além disso, conhecimentos novos podem resultar de um processo de inferência na própria estrutura do conhecimento.

A classe suprema dessa hierarquia é a inteligência, que Tarapanoff (2001) diz poder ser entendida como a informação como oportunidade, ou seja, o conhecimento contextualmente relevante que permite atuar com vantagem no ambiente considerado. Numa outra visão, pode-se dizer que inteligência é um conhecimento aplicado numa determinada situação para aprofundar-se no entendimento da mesma. Tarapanoff (2001, p. 119) define a inteligência como o resultado da síntese de corpos de conhecimentos, com o uso do julgamento e da intuição daquele que toma decisões.

Com a crescente concorrência global, as empresas terão cada vez mais necessidade de informação, tanto do mercado, quanto a respeito dos planos e intenções de consumidores e dos competidores. A informação correta e bem trabalhada norteia as diretrizes da empresa, embasa suas estratégias e garante seu sucesso operacional.

2.3.3 Gestão da informação X gestão do conhecimento

Segundo Moraes (2006, p. 125), alguns autores classificam o processo de gestão da informação em um ciclo de 5 etapas relacionadas da seguinte maneira:

- Etapa 1 – Determinação da necessidade de informação: envolve compreender as fontes e os tipos de informações necessárias para um bom desempenho do negócio, bem como suas características, fluxos e necessidades;

- Etapa 2 – Obtenção: inclui as atividades relacionadas à coleta dos dados;
- Etapa 3 – Processamento compreende atividades de classificação (define o melhor modo de acessar as informações necessárias) e de armazenamento (seleciona o melhor lugar e os recursos para o arquivamento) das informações obtidas;
- Etapa 4 – Distribuição e apresentação: envolve escolher, entre diferentes metodologias, qual pode ser mais adequada para se apresentar a informação, disponibilizando-a aos usuários por diferentes formas e fontes e estilos;
- Etapa 5 – Utilização: após a apresentação da informação, segue-se a etapa de utilização da mesma pelas pessoas da empresa, que as incorporarão às etapas de elaboração, execução e avaliação da estratégia empresarial, auxiliando, assim, o processo de gestão estratégica.

Depois da última etapa, quando a informação é utilizada, uma nova demanda torna necessária a busca de informação, e o processo de gerenciamento da informação é reiniciado, já que esse processo é estratégico e contínuo. A figura 6 demonstra esse ciclo.

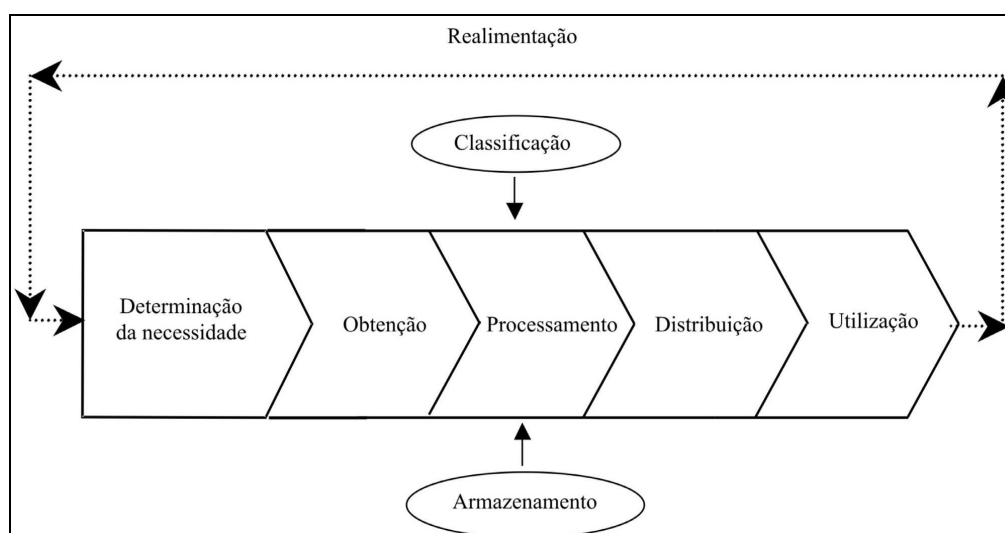


Figura 6 – Processo de gerenciamento da informação
(Fonte: Moraes, 2006, p. 126)

Então, gestão da informação é um processo que se baseia nas atividades de obtenção, processamento, distribuição e utilização de informações de um modo cíclico. Em qualquer formato e em qualquer meio em que se encontre (arquivos digitais, documentos, textos, imagens, filmes obras de arte, etc.). Seu objetivo é

fazer com que as informações cheguem às pessoas que delas necessitam e no momento certo para apoiar a tomada de decisões.

Já para a gestão do conhecimento, inúmeros são os modelos apresentados. Conforme Tarapanoff (2001, p. 143), as abordagens variam e podem ser culturais, com ênfase nas relações humanas, de reengenharia, com ênfase nas tecnologias de informação e processos de negócio, e as voltadas para o pensamento sistêmico, com ênfase na compreensão de todo o sistema e das interdependências dos fatores internos e externos que influenciam a organização.

Tarapanoff (2001, p. 147) cita que em 1998, no Workshop Brasileiro de Inteligência Competitiva, na cidade do Rio de Janeiro, Maria Fátima L. Stollenwerk apresentou uma análise comparativa dos diversos modelos de gestão do conhecimento com o intuito de identificar semelhanças que levassem a um modelo genérico de gestão do conhecimento.

Com base na análise comparativa dos modelos, pode-se depreender que existem idéias básicas que permeiam todos eles, apesar das especificidades e das contribuições individuais de cada modelo (Tarapanoff, 2001, p. 147).

Essa análise levou à composição do modelo genérico (Figura 7) composto por sete processos que devem ser considerados na gestão do conhecimento: identificação; captura; seleção e validação; organização e armazenagem; compartilhamento; aplicação; criação. Foram incluídos no modelo genérico, ainda, quatro fatores facilitadores dessa gestão: liderança; cultura organizacional; medição e recompensa; tecnologia de informação e comunicação (Tarapanoff, 2001, p. 148).

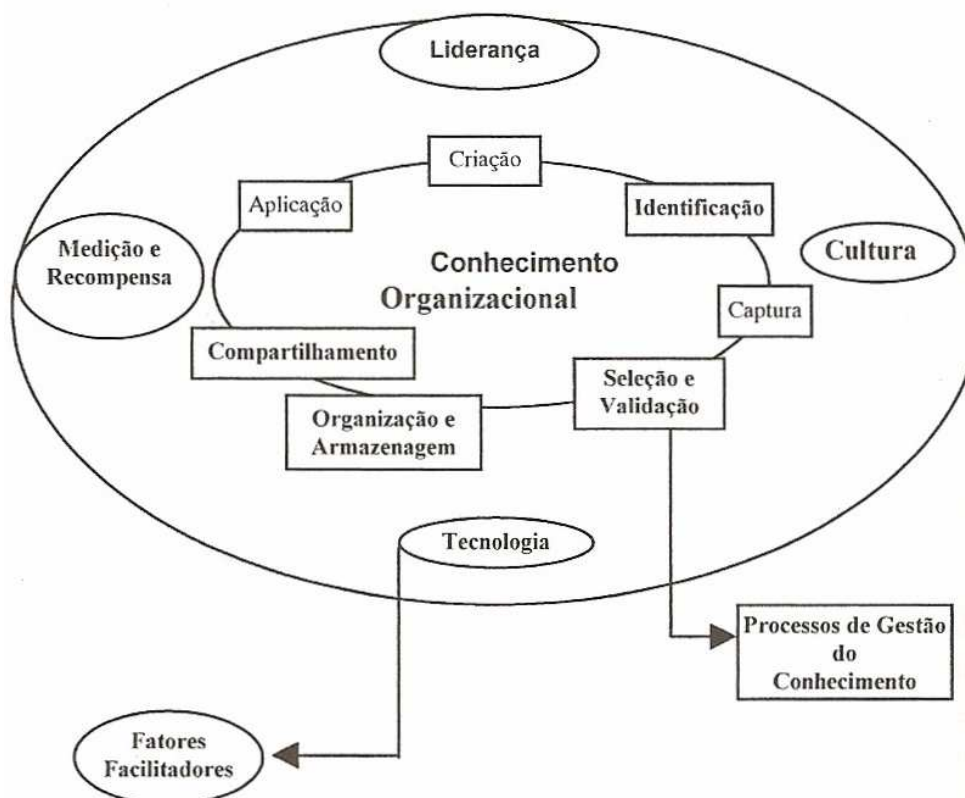


Figura 7 – Modelo genérico de gestão do conhecimento
(Fonte: Tarapanoff, 2001, p. 148)

Identificação do conhecimento é o primeiro processo do modelo genérico. Ele visa, entre outras coisas, identificar que competências são críticas para o sucesso da organização. Cada competência essencial está ligada a áreas específicas do conhecimento que à sustentam.

Captura é o processo de aquisição de conhecimento, habilidades e experiências necessárias para criar e manter as competências essenciais e áreas de conhecimento.

Seleção e validação são processos que visam filtrar o conhecimento, avaliar sua qualidade e sintetiza-lo para fins de aplicação futura. Nem todo o conhecimento gerado, recuperado ou desenvolvido deve ser armazenado na organização (Tarapanoff, 2001, p. 151).

Organização e armazenagem garantem a recuperação rápida, fácil e correta do conhecimento. Segundo Tarapanoff, quanto mais se formalizar o conhecimento, mais eficaz será o processo de organização e armazenagem (2001, p. 152).

Compartilhamento é o processo que mais depende da tecnologia da informação. Afim de que o conhecimento seja disponibilizado e compartilhado em

larga escala, de forma rápida e eficaz, é necessário que haja um mecanismo automático de distribuição das informações aos interessados.

Aplicação diz respeito à aplicação do conhecimento em processos decisórios, soluções de problemas, aprendizagem e registro dos ganhos obtidos; tem a finalidade de aprimorar os processos da organização.

Criação é o processo voltado a gerar novos conhecimentos com aproveitamento de aprendizagem, pensamento criativo, pesquisa, experimentação, descoberta, inovação, etc.

O fator facilitador da liderança é primordial no processo de gestão do conhecimento, sem o qual a eficácia da gestão fica altamente prejudicada.

O fator facilitador da cultura organizacional visa dar suporte às inúmeras características da organização, como desempenho, foco no cliente, foco em melhoria, foco em excelência, flexibilidade, etc. e criar um ambiente de confiança e cooperação para que haja o compartilhamento efetivo do conhecimento.

O fator da medição e avaliação acompanha o desempenho, o comportamento e as atitudes para poder efetuar-se o devido reconhecimento e a recompensa dos colaboradores.

Finalmente, o fator da tecnologia de informação é o que dá suporte a toda a estrutura, disponibilização e controles do fluxo do conhecimento na organização.

Entende-se por conhecimento, então, a informação interpretada, ou seja, o conjunto dos significados de cada informação e que impactos cada informação pode causar no domínio tratado, de modo que esse conjunto possa ser utilizado para importantes ações e tomadas de decisões.

2.3.4 A informação coletada por formulários

Quanto aos formulários, nos quais você preenche dados que no fim das contas irão para o computador, a hora de abolir o papel é agora. Caso contrário, você terá que manter grupos de trabalho cuja função é pegar os formulários em papel e digitar a informação no sistema informatizado.[...] O uso de formulários eletrônicos desde o início dos processos não só elimina trabalho extra como assegura consistência, mensuração fácil e autorização apropriada.

Se alguém julgar alguma coisa confusa num formulário eletrônico, ou achar que um dos campos deveria ser preenchido automaticamente, sempre pode clicar na linha do “e-mail para” e digitar algumas palavras dizendo: “Ei, isso deveria funcionar melhor”. E não leva um ano para fazer essas alterações, o que pode acontecer com formulários em papel. Com um formulário eletrônico, a maioria das modificações é feita em alguns dias, ou no máximo em poucos meses (Gates, 1999, p. 68).

O uso de formulários eletrônicos permitiu, entre outras coisas, a realização de enquetes e pesquisas on-line abrindo um grande canal para a coleta de informações. Entretanto, a Internet ainda representa uma novidade nos meios de pesquisa. Sem dúvida, a Internet revolucionou a distribuição e troca de informação. A partir dela, todos podem informar a todos. Mas, se ela pode facilitar a busca e a coleta de dados, ao mesmo tempo oferece alguns perigos; na verdade, as informações passadas por essa rede não têm critérios de manutenção de qualidade da informação. Devemos levar em conta que toda e qualquer informação colhida na Internet deverá ser confirmada antes de divulgada. Por conta disso, todo trabalho realizado em ambiente Web deve ponderar sobre a fonte das informações coletadas. Sempre que possível deve-se buscar garantias, como identificação por login/senha, links de único acesso, confirmação por E-Mail, etc., dependendo da aplicação e do fim a que se destinam as informações coletadas.

Mas a informação não se limita a dados coletados. Para que esses dados se transformem em informações, é necessário que sejam organizados e ordenados de forma que se tornem úteis (Rezende, 2003, p. 108). Para tanto, faz-se necessário um sistema de coleta e uma base de dados bem elaborados, a fim de abraçar o maior leque possível de possibilidades dentro do domínio a que o sistema se propõe.

As vantagens de um sistema de bancos de dados sobre os métodos tradicionais baseados em papel para manutenção de registros são inúmeras. Aqui estão algumas delas:

- Densidade: não há necessidade de arquivos de papel, possivelmente volumosos.
- Velocidade: a máquina pode obter e atualizar dados com rapidez muito maior que o ser humano.

- Menor trabalho monótono: grande parte do tédio de manter arquivos à mão é eliminada. As tarefas mecânicas são sempre feitas com melhor qualidade por máquinas.
- Atualidade: informações precisas e atualizadas estão disponíveis a qualquer momento sob consulta.

É claro que as vantagens anteriores se aplicam com intensidade ainda maior a um ambiente multiusuário, no qual o banco de dados provavelmente será muito maior e mais complexo que no caso do ambiente de usuário único. Porém, há uma vantagem adicional predominante em tal ambiente, qual seja: o sistema de banco de dados proporciona à empresa o controle centralizado de seus dados (que é um de seus bens mais valiosos). Tal situação contrasta nitidamente com aquela encontrada em uma empresa sem um sistema de banco de dados, onde em geral cada aplicação tem seus próprios arquivos privativos — muitas vezes também suas próprias fitas e seus discos particulares — de modo que os dados ficam amplamente dispersos e são difíceis de controlar de forma sistemática (Date, 2000).

O sistema Enquetes&Pesquisas será, com certeza, uma excelente ferramenta para as organizações que desejarem efetuar a automação de formulários, não só de pesquisas, mas também de cadastros diversos, porque através de perguntas e respostas elas poderão montar cadastros e depois fazer o download destes em arquivos XML ou CSV.

2.4 SISTEMAS COM BASES DINÂMICAS

O termo base dinâmica, normalmente, relacionado à inteligência artificial, na verdade é muito mais abrangente. Qualquer base que responda com resultados diferentes em diferentes instâncias a uma mesma busca é uma base dinâmica.

A Internet vista como um grande repositório de informações pode ser considerada uma gigantesca base de dados hipermídia e dinâmica. O termo hipermídia, conforme Jonassen (1996, p. 78), é derivado do hipertexto, que é um método não-sequencial, não-linear para organização e exibição de texto. A hipermídia foi projetada para permitir ao usuário o acesso direto à informação desejada à partir de links que consistem em pontos centrais (trechos ou fragmentos de texto, figuras, animações, sons ou documentos) que interligam arquivos na

internet ou pontos dentro de um mesmo arquivo. Ao estudar uma base de conhecimento hipermídia, os usuários podem acessar pontos centrais em qualquer ordem que os satisfaça. Em muitos sistemas de hipermídia, os pontos centrais podem ser emendados ou modificados pelo usuário, de modo que o sistema possa ser uma base dinâmica de conhecimento que continue a crescer, representando novos e diferentes pontos de vista (Jonassen, 1996, p. 78).

O ensino à distância é uma variante deste conceito de internet como base dinâmica que agrega facilidade de acesso às pesquisas, aos contatos, ao conteúdo da disciplina estudada e às diferentes visões de cada tema através das variações de links visitados por cada participante.

Em empresas de comunicação, como rádio, televisão e propaganda, a hipermídia facilitou o acesso às imagens e sons arquivados em suas bases digitalizadas.

Além de melhorar a qualidade de seus conteúdos, a base de dados dinâmica permite disponibilizar novas formas de interação com o seu público e criar outros produtos e serviços. As empresas de comunicação devem saber que esse tipo de armazenamento de conteúdos pode ser utilizado não apenas como instrumento pelos mais diversos veículos, mas também para a criação de novos produtos digitais mais criativos e estabelecer uma relação mais estreita com o usuário, que encontrará inúmeras possibilidades nessa imersão digital (Quadros, 2005, p. 414).

3 TRABALHOS RELACIONADOS

A idéia proposta neste projeto de um modo global não foi identificada em nenhum dos trabalhos consultados. Os trabalhos acadêmicos de um modo geral tendem à pesquisas no formato tradicional e não genérico ou ao estudo do processo de pesquisa e não à produção de uma ferramenta. Os trabalhos comerciais têm uma maior semelhança com a ferramenta de pesquisas, porém, nenhum deles parece ter uma estrutura genérica capaz de criar uma base de dados crescente de perguntas e respostas com pretensões de futura mineração.

3.1 TRABALHOS ACADÊMICOS

Como dito anteriormente, não foram localizados, na pesquisa realizada, trabalhos acadêmicos similares ao aqui proposto. Entretanto, alguns trabalhos analisados possuem alguma relação com o tema “pesquisas em formulários on-line”.

São eles:

- Dinâmica do processo de coleta e análise de dados via web: trata-se de uma análise da Internet como ferramenta de pesquisas. Freitas (2004) apresenta a Web como um universo novo de funcionalidades e possibilidades a ser explorado. Cita estudos e abordagens que dão um novo sentido ao uso da Internet como fonte de informação. Também explana sobre diversos aspectos, como rapidez, economia, qualidade do suporte multimídia, interatividade e acessibilidade, seletividade e representatividade. Analisa as etapas do processo de pesquisa on-line, as tecnologias e os fluxos de informação.
- Relacionamento entre o webjornal Campo Grande e os seus usuários: Esta dissertação trata do uso das tecnologias no relacionamento entre o jornal e seus leitores. Utiliza uma enquete através da Internet para fazer uma análise do perfil e dos hábitos dos usuários e entrevistas com os produtores do webjornal para obter dados qualitativos da produção e dos próprios produtores. Apresenta os resultados das pesquisas, propõe mais interação entre os usuários e o jornal e por fim aponta sugestões para a realização de novas pesquisas.

3.2 TRABALHOS COMERCIAIS

Inúmeros são os portais comerciais que oferecem pesquisas on-line pagas. Alguns até tem opções de pacotes grátis, porém, nestes casos o usuário fica muito limitado e é induzido a um pacote pago para conseguir fazer uma pesquisa com bons resultados. Os melhores exemplos encontrados foram:

- enquetes.com.br: site de enquetes, onde pode-se fazer enquetes grátis, mas apenas com uma pergunta por enquete. As pesquisas mais avançadas são pagas. Tem também um serviço chamado Rotator que agrupa várias enquetes. Os resultados estão disponíveis on-line, mas não são agrupados num relatório ou gráfico único. Talvez nas opções pagas isso seja possível.
- equestiona.com: site de pesquisas on-line pago com vários pacotes de assinatura. Existe uma opção grátis por 30 dias, com quantidade ilimitada de pesquisas e perguntas por pesquisa. Com limitação de 100 convites por e-mail e 50 respostas completas. Pode-se acompanhar os resultados em tempo real com alguns gráficos em pizza ou barras, mas o download do resultado é em arquivo MS-Excel e distribuído em várias planilhas com os dados dispostos de formas diferentes em cada uma delas, tornando a análise um pouco confusa.
- formfacil.com: O Formfácil é um serviço gratuito de criação de formulários online. Dentre os serviços pesquisados é o que mais se assemelha ao projeto Enquetes&Pesquisas. O serviço é indicado para pessoas e empresas que precisam cadastrar informações em formulários de forma prática e fácil, permite visualizar os dados preenchidos em tempo real pela Internet. Ele permite o cadastro de produtos e serviços, adicionar formulários criados ao seu site ou blog, relatórios dinâmicos que permitem escolher as colunas e ordenar os dados. Este site tem o problema de ser muito pesado para a abertura. No navegador Mozilla Firefox 2.0.0.14 e numa conexão de 2 MB o site só abriu depois de várias tentativas e o problema persistiu em repetidos testes, mesmo com os cookies do acesso anterior no disco rígido. No Microsoft Internet Explorer 7.0.5730.13 e mesma conexão de 2 MB, o site não abriu, mesmo depois de sete tentativas.
- Supertráfego.com: O SuperTráfego é um portal baseado em três pilares: ferramentas, geração de tráfego e conteúdo para, principalmente, donos de

sites, blogs e fotologs. Na área de enquetes possui até 12 opções por enquete. Criação de um fórum associado a enquete. Resultados públicos ou privados. Possibilidade de alterar a quantidade de votos de cada opção. Área administrativa para gerenciamento das enquetes, onde é possível alterar, apagar ou adicionar opções à enquete. Resultados mostrados em gráfico de barras e números/porcentagem. Também é possível colocar uma cor para cada opção na barra de resultados. Em todas as enquetes existe um link para o usuário enviá-la a um amigo. Estatísticas de uso, para conferir se os usuários estão usando ou não a ferramenta. O mesmo usuário pode votar mais de uma vez por dia ou não. Personalização das telas. Entretanto, o site não permite a criação de pesquisas com mais de uma pergunta.

4 METODOLOGIA

Após uma visão geral do problema das pesquisas por formulário e a apresentação da idéia de criar uma base de conhecimento alimentada por pesquisas realizadas pela Internet, faz-se necessário uma metodologia adequada para a construção do sistema de forma a garantir a eficiência do mesmo.

4.1 COLETA DE DADOS

Para elaboração deste trabalho, fez-se uso de material disponível na biblioteca da Faculdade Expoente, alguns livros de acervo próprio e também da Orientadora, além de material disponível na Internet.

Todo o material consultado está relacionado na seção de referências.

4.2 ANÁLISE

O modelo mais apropriado para a modelagem do sistema Enquetes&pesquisas é o modelo orientado a objetos (OO) com linguagem UML (*Unified Modeling Language*). Para tanto se apresenta três razões. Primeiro, por ser este um sistema Web em arquitetura cliente/servidor onde esse padrão de modelagem mostra-se extremamente eficaz. Segundo, porque a modelagem OO com UML é o padrão de mercado e isso facilita a interpretação do projeto. E terceiro, porque a linguagem adotada para a codificação do sistema é a C#, uma linguagem totalmente orientada a objeto.

4.2.1 Restrições de acesso

Algumas restrições de acesso são cruciais para o funcionamento do sistema em ambiente Web e também devem ficar bem claras para o entendimento da estrutura do sistema, são elas:

- Apenas o usuário Administrador terá acesso ilimitado ao sistema.
- O usuário Avaliador só terá acesso aos resultados on-line da pesquisa que estiver respondendo, caso o Pesquisador libere esta funcionalidade.

- O download de arquivo com o resultado da pesquisa só estará disponível para o usuário Pesquisador.
- O usuário Pesquisador terá acesso aos resultados somente das suas pesquisas.
- O usuário Avaliador só poderá responder uma única vez uma pesquisa a qual tenha recebido acesso, entretanto, lhe será permitido responder em multi-sessão.

4.3 AMBIENTE DE DESENVOLVIMENTO E EXECUÇÃO

Na construção do projeto está previsto para a base de dados a utilização do IBM DB-2 Express-C 9 por ser de livre uso para desenvolvimento e produção e com capacidade ilimitada de armazenamento.

As camadas de negócio e de apresentação serão desenvolvidas no MS Visual Studio 2008 com linguagens ASP.NET e C#, sobre o .NET Framework 3.5. A escolha se deve por se tratar de ferramenta com excelente tecnologia e oferecida gratuitamente aos alunos da Faculdade Expoente pela própria Microsoft.

O sistema de relatórios e dados para o Pesquisador será criado com Crystal Reports, ferramenta disponível no pacote do MS Visual Studio.

Para o manual de instalação e ajuda serão criados arquivos no formato HTML, assim também para a política de privacidade e Termo de uso.

Para efetuar os testes necessários à validação do sistema será aplicada simulação de carga em vários níveis e, se necessário, será demandado um espaço em servidor da Faculdade e disponibilização de porta neste mesmo servidor para atender ao serviço em caráter experimental/temporário.

4.3.1 Desenvolvimento Web em plataforma Microsoft .NET

Em junho de 2000, a Microsoft anunciou sua iniciativa .NET, uma visão nova para abranger a Internet e a World Wide Web no desenvolvimento, engenharia e uso de software (Deitel, 2003, p. 13).

A idéia básica da plataforma .NET é melhorar a produtividade do desenvolvedor com reutilização de software e aproveitamento de componentes

prontos pré-empacotados, como botões, caixas de texto e barras de rolagem. Analogamente, os desenvolvedores podem criar um aplicativo usando serviços da Web para bancos de dados, segurança, autenticação, armazenamento de dados e tradução de idiomas, sem ter de conhecer os detalhes internos desses componentes (Deitel, 2003, p. 14).

O centro da estratégia .NET é o Microsoft .NET Framework, uma estrutura que gerencia e executa aplicativos e serviços Web. O .NET Framework possui uma vasta biblioteca de classes, a FCL (*Framework Class Library*), que garante a segurança e também fornece muitos recursos de programação.

Para desenvolver modernas aplicações Web, a Microsoft desenvolveu para o projeto .NET uma linguagem ágil e de alto nível para desenvolvimento de aplicações comerciais com excelente produtividade. Esta linguagem, derivada da linguagem C e C++, foi batizada de C# (C sharp), e é a linguagem nativa da plataforma Microsoft .NET.

4.3.2 Banco de dados IBM DB2 Express-C 9

Como foi dito no início deste sub-item, está previsto para a base de dados a utilização do IBM DB-2 Express-C 9 por ser de livre uso para desenvolvimento e produção e com capacidade ilimitada de armazenamento. Suas principais características são (IBM):

- Instalação simplificada – A instalação do DB2 Express-C é feita de forma a torná-lo um componente transparente dentro da solução de análise de dados.
- Flexibilidade – Roda em Linux ou Windows, provê scripts para simplificar a configuração da base de dados e dá liberdade de escolha de hardware e sistema operacional.
- Funcionalidades de Gerenciamento Autônomo – Oferece funcionalidades avançadas de gerenciamento para otimizar a performance e monitorar a operação. Possui alertas para possíveis erros em processos operacionais e providencia avisos quando os mesmos são solucionados.
- Projetado para operar 24x7 – Não existe a necessidade de desativar o banco de dados para manutenção. Alterações no banco de dados podem

ser realizadas ao mesmo tempo em que o banco está online, obtendo-se resultados imediatamente.

- Suporte para o desenvolvimento – Suporta uma gama de interfaces de aplicações de programação, incluindo ODBC, OLE DB, ADO, ADO.NET, JDBC, SQLJ, DB2 CLI, e SQL embutido, além de interfaces emergentes, como a XML e Serviços da Web.
- O DB2 UDB provê add-ins para IDEs populares (como WebSphere Studio Application Developer e Microsoft Visual Studio .NET). Isto facilita uma integração transparente na utilização do aplicativo, e proporciona aos desenvolvedores a habilidade para reduzir o *time-to-market*.

4.3.3 Plug In para relatórios

Para a geração de relatórios ao Pesquisador será utilizado o componente Crystal Reports embutido no MS Visual Studio 2008, devido principalmente ao fato de ter-se integrados ao ambiente de desenvolvimento .net os principais recursos do Crystal Reports, facilitando o desenvolvimento de relatórios, bem como a sua utilização.

4.3.4 Configurações mínimas para execução do sistema

Pelo fato de o Enquetes&Pesquisas ser um sistema Web, para sua perfeita execução serão necessários dois tipos de configuração em computadores ligados em rede e devido ao objetivo de alimentar a base de dados, é primordial a disponibilização do serviço na Internet. Assim sendo, sugere-se para o sistema a seguinte configuração mínima de hardware e software:

Para o servidor:

- Computador PC Pentium IV 3 Ghz;
- 2 GB de memória RAM;
- Disco rígido com 40 GB de espaço e velocidade de 7.200 RPM;
- Sistema operacional Windows 2003 Server;
- Framework Microsoft .net 3.5;
- Servidor Microsoft Internet Information Services (IIS) versão 6.
- Link de Internet de 2 MB.

Para o cliente:

- Computador com processador de 800 Mhz;
- 512 MB de memória RAM;
- Conexão com a Internet de 56 Kb;
- Navegador Internet Explorer 7, Firefox 2 ou outro compatível.

4.4 DESCRIÇÃO DAS FUNCIONALIDADES DO SISTEMA

De forma geral o sistema Enquetes&Pesquisas terá 18 funcionalidades, conforme descritas na tabela 1.

Tabela 1: Funcionalidades do sistema

Nº	Descrição	Usuário com acesso
01	Manter base de dados	Administrador
02	Criar pesquisa	Administrador Pesquisador
03	Cadastrar assunto	Administrador Pesquisador
04	Cadastrar bloco	Administrador Pesquisador
05	Cadastrar pergunta	Administrador Pesquisador
06	Cadastrar alternativa de resposta	Administrador Pesquisador
07	Descadastrar pergunta	Administrador Pesquisador
08	Publicar pesquisa	Administrador Pesquisador
09	Encerrar pesquisa	Administrador Pesquisador
10	Convidar Avaliadores	Administrador Pesquisador
11	Cadastrar tipo de usuário	Administrador
12	Incluir usuário	Administrador

		Pesquisador Avaliador ⁹
13	Alterar usuário	Administrador Pesquisador ¹⁰
14	Desativar usuário	Pesquisador Avaliador
15	Responder pesquisa	Avaliador
16	Finalizar participação	Avaliador
17	Visualizar estatísticas	Pesquisador Avaliador ¹¹
18	Baixar arquivo de respostas	Pesquisador

4.5 DEFINIÇÃO DOS DIAGRAMAS

Entre os diagramas utilizados pela UML, os mais comuns são: diagrama de caso de uso, diagrama de classes, diagrama de objeto, diagrama de estado, diagrama de seqüência, diagrama de colaboração, diagrama de atividade, diagrama de componente e diagrama de execução.

Dentre eles, serão usados neste projeto os seguintes diagramas:

- Diagrama de caso de uso: por ser de importância primordial para a visualização das funcionalidades do sistema e do papel dos atores;
- Diagrama de classes: também é fundamental para se visualizar o sistema do ponto de vista de sua forma estática, demonstrando o lugar de cada objeto e seus relacionamentos;
- Diagrama de execução: assim como o diagrama de componente mostra o lado funcional do sistema, neste caso exhibe os componentes físicos de hardware e software e sua arquitetura. Apresenta uma macro visão interessante do sistema.
- Diagrama de entidade/relacionamento (DER): este diagrama, apesar de não fazer parte dos diagramas da UML, originalmente, ainda é muito

⁹ Avaliador só poderá incluir a si próprio.

¹⁰ Pesquisador só poderá alterar cadastro de usuário que ele próprio tenha incluído.

¹¹ Avaliador só poderá visualizar estatísticas, caso o Pesquisador tenha liberado esta funcionalidade.

utilizado devido à clareza com que retrata o banco de dados. Exatamente por este motivo é indispensável neste projeto.

Os seguintes diagramas não serão utilizados no presente projeto:

- Diagrama de estados: por ser importante apenas em alguns casos, para a análise do comportamento de alguma classe específica durante sua utilização pelo sistema.
- Diagrama de seqüência: também é utilizado apenas para análise específica de alguns objetos, principalmente para se analisar a execução concorrente dos mesmos.
- Diagrama de objeto: por ser uma variação do diagrama de classes e utilizar quase a mesma notação este diagrama é dispensável neste caso. A diferença reside no fato de que o diagrama de objetos mostra os objetos que foram instanciados das classes. Ele mostra o perfil do sistema em um dado momento de sua execução, não sendo necessária sua utilização no presente projeto;
- Diagrama de colaboração: semelhante ao diagrama de seqüência, podendo-se optar pela utilização de um ou de outro;
- Diagrama de atividade: este diagrama foca a execução de implementação de uma operação (método) e suas atividades numa determinada instância de um objeto qualquer, um nível de detalhamento dispensável neste projeto;
- Diagrama de componente: mostra o lado funcional do sistema, exibindo os componentes como tipos e suas dependências com outros componentes. Também mostra um nível de detalhamento dispensável neste projeto;

4.6 CRONOGRAMA DE DESENVOLVIMENTO

Para que se tenha êxito na execução de qualquer projeto é de suma importância que se atente para os prazos definidos no cronograma. Um acompanhamento diário do cronograma permite que se perceba com antecedência um possível atraso na implantação para que se tome as medidas para os ajustes necessários ao cumprimento dos prazos. Segue nas figuras 8 e 9 o cronograma para o desenvolvimento da monografia e do sistema Enquetes&Pesquisas.

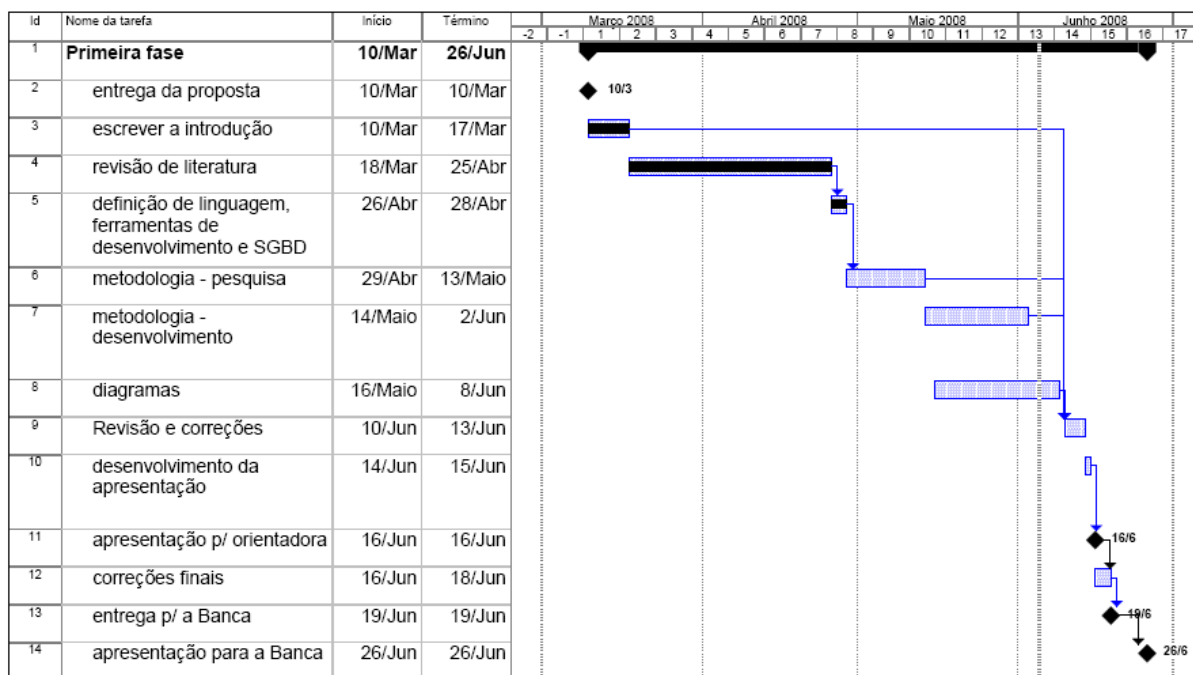


Figura 8 – cronograma para o primeiro semestre de 2008

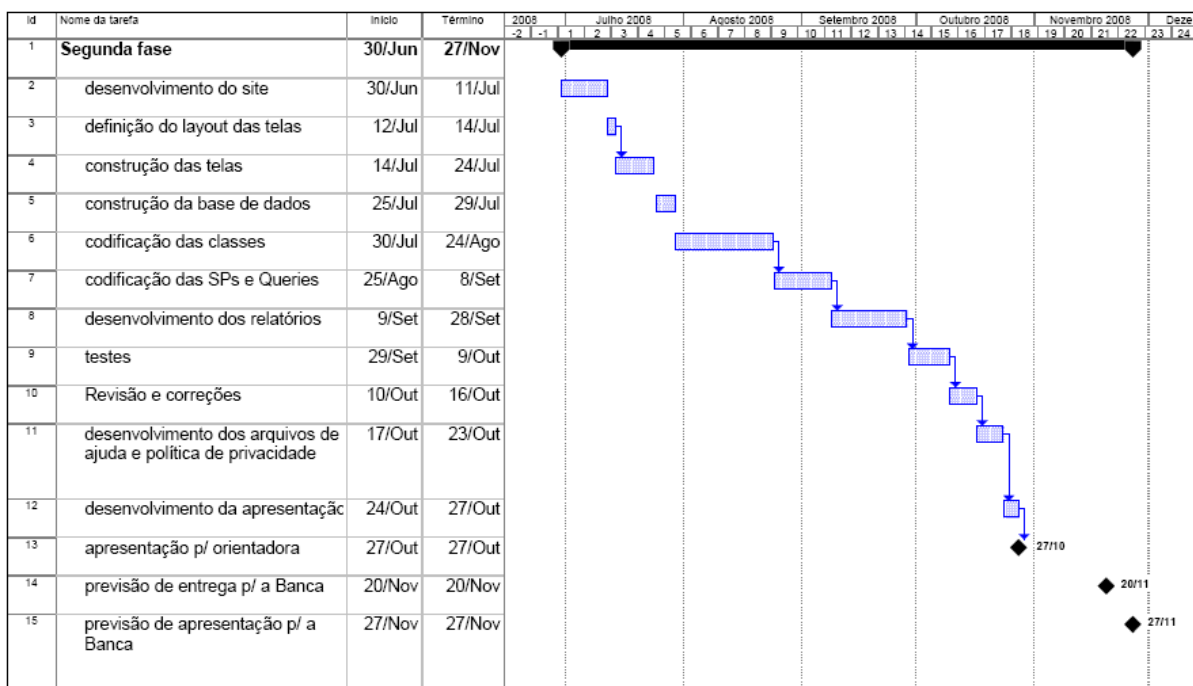


Figura 9 – cronograma para o segundo semestre de 2008

4.7 ANÁLISE DE RISCOS

Estimar a produção de um software é uma tarefa que carrega riscos inerentes. Segundo Sanches (2003), alguns fatores, como complexidade, tamanho e grau de definição de estruturas, aumentam os riscos (Figura 10). Outros fatores,

porém, podem minimizar os riscos, como reuso, dados históricos e uma boa definição de estruturas.

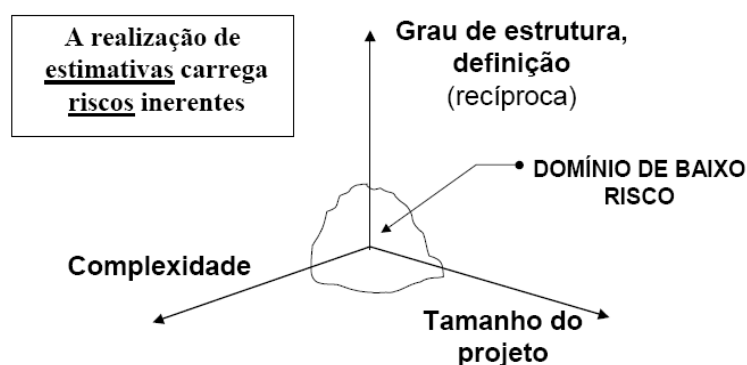


Figura 10 – Fatores que aumentam o risco
(Fonte: Sanches, 2003)

Pressman (1995, p. 131) diz que a análise dos riscos é, de fato, composta por quatro atividades distintas: identificação, projeção, avaliação e administração dos riscos.

A identificação dos riscos envolve a análise de um leque de categorias, tão amplas e complexas quanto maior e mais complexo for o projeto.

Como este projeto é individual e acadêmico, não necessita de análise de riscos de muitas das categorias possíveis, como riscos de pessoal e de custos, por exemplo. Também não se pretende neste trabalho, enumerar as diversas categorias de riscos existentes, para tanto se indica um estudo aprofundado de engenharia de software.

Dentro da categoria de riscos de projeto, tamanho e complexidade são riscos possíveis e foram, portanto, analisados. Também foram analisadas duas restrições do IBM DB2 que dizem respeito ao servidor do banco de dados e tiveram especial atenção, uma impondo limite máximo de dois processadores para rodar o banco de dados e a outra é o limite máximo de 4 GB para uso em memória RAM .

Por fim, na categoria de riscos de negócio, foram analisados riscos de utilidade e divulgação, focando a importância do sistema para o mercado e as possibilidades de crescimento do número de usuários por divulgação na própria rede.

A projeção, ou estimativa, está relacionada à probabilidade de que um fato tido como risco, ocorra realmente. Pressman (1995, p. 133) indica o estabelecimento de uma escala que reflita a probabilidade percebida de ocorrência de um risco, o

delineamento das conseqüências do risco, a estimativa de impacto sobre o projeto e anotação da precisão global dessas estimativas para evitar-se mal-entendidos, como sendo as quatro atividades necessárias à projeção dos riscos.

A escala citada no parágrafo anterior pode ser definida de diversas formas. Neste projeto adotou-se uma escala de probabilidades qualitativa, indicada como uma boa abordagem por Pressman (1995), a qual terá os seguintes níveis: 1, 2, 3, 4 e 5, indicando respectivamente: altamente improvável, improvável, moderado, provável e altamente provável.

A avaliação, segundo Pressman (1995) deve ser feita sob a ótica de um nível definido de risco. Este nível baseia-se geralmente em três pilares, custo, prazo e desempenho. Como o presente projeto não tem custo financeiro, foram estabelecidos apenas níveis de prazo e desempenho. A tabela 3 apresenta os riscos identificados para o sistema e a figura 11 mostra o resultado da avaliação desses riscos. Observe-se que, embora os níveis críticos para as limitações de memória e processadores tenham sido atingidos, considerou-se neste caso uma situação hipotética de sucesso de uso da ferramenta, com muitos acessos simultâneos e um forte crescimento da base de dados, o que é esperado ao longo do tempo, mas não num primeiro momento que será de afirmação do serviço no mercado. Também neste caso, aceitou-se o nível das estimativas à tolerância máxima devido ao fato de o banco de dados trabalhar apenas com dados nos formatos texto e números, e os registros textuais não serem extensos, o que teoricamente permite um grande número de registros sem afetar significativamente o desempenho.

Tabela 3: Riscos identificados para o sistema Enquetes&Pesquisas

Risco	Nível	Impacto no projeto
Tamanho do sistema	2	Atraso na entrega
Complexidade do sistema	3	Atraso na entrega
Utilidade do sistema	2	Fracasso na intenção de crescimento da base de dados
Divulgação da ferramenta	3	Povoamento insuficiente da base de dados
Limite de processadores para rodar o DB2	4	Queda no desempenho em caso de excesso de conexões simultâneas
Limite de memória RAM para rodar o DB2	4	Queda no desempenho em caso de excesso de conexões simultâneas

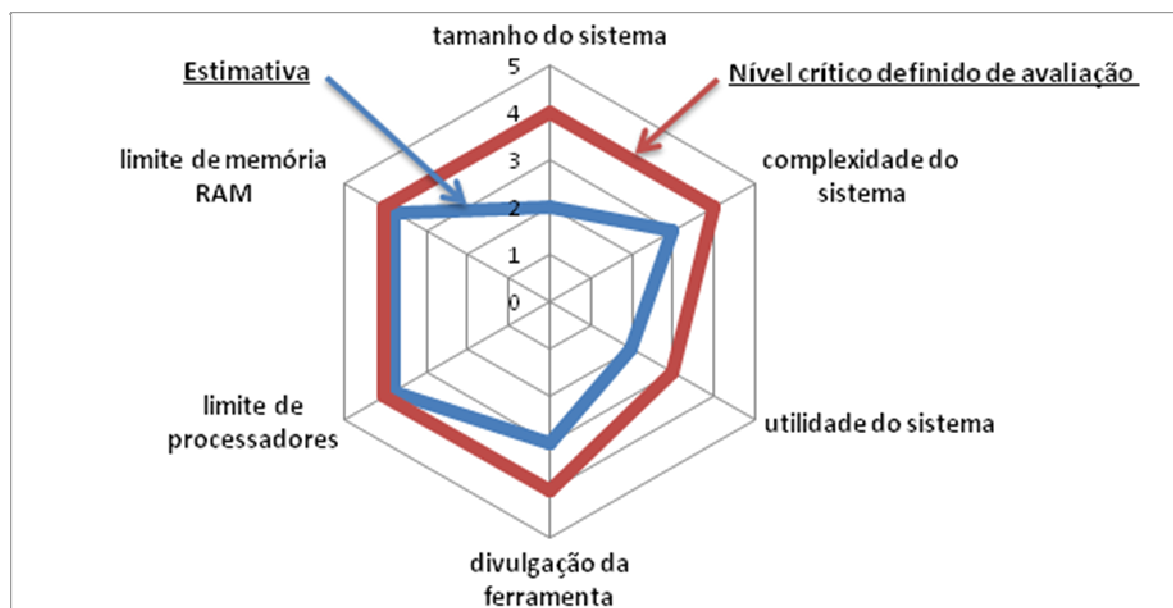


Figura 11 – Níveis de risco para o sistema Enquetes&Pesquisas

A administração dos riscos identificados envolve uma análise que vai de medidas para evitarem-se os riscos, até os custos para implantação dessas medidas preventivas ou outras corretivas caso o risco se torne real. A tabela 3 demonstra os procedimentos a serem adotados, caso um risco venha a ocorrer.

Tabela 3: Procedimentos para ocorrências de riscos

Risco	Nível	Impacto no projeto	Medida	Tipo
Tamanho do sistema	2	Atraso na entrega	Previsão de término antecipada no cronograma	Preventiva
Complexidade do sistema	3	Atraso na entrega	Previsão de término antecipada no cronograma	Preventiva
Utilidade do sistema	2	Fracasso na intenção de crescimento da base de dados	Criação de uma ferramenta atraente e intuitiva para o usuário Pesquisador	Preventiva
Divulgação da ferramenta	3	Povoamento insuficiente da base de dados	Divulgação em sites de relacionamento	Corretiva
Limite de processadores para rodar o DB2	4	Queda no desempenho por excesso de conexões simultâneas	Migração da base de dados para outro SGBD que não imponha este limite	Corretiva
Limite de memória RAM para rodar o DB2	4	Queda no desempenho por excesso de conexões simultâneas	Migração da base de dados para outro SGBD que não imponha este limite	Corretiva

4.8 TESTES

Para garantir um bom funcionamento do sistema após a implantação, é fundamental a execução de uma série de testes que verifiquem se o sistema vai suportar as situações a que será submetido quando em produção.

A tabela 3 mostra a seqüência de testes a qual o sistema Enquetes&Pesquisas será submetido para validar seu bom funcionamento..

Tabela 3: Testes a serem realizados

Seqüência	Teste	Objetivo
1	Teste de navegação das telas.	Verificar a consistência dos links de navegação.
2	Teste de povoamento da base de dados.	Verificar a consistência da estrutura do banco de dados e a variação do tempo de resposta.
3	Teste de validação das regras de negócio	Verificar se o sistema responde adequadamente a dados incorretos e tentativas de burlar as regras.
4	Teste de carga de acessos simultâneos com crescimento exponencial, iniciando com um usuário, depois dez usuários, cem usuários, mil usuários e assim por diante.	Verificar o comportamento do sistema em situações de acessos simultâneos.

Para a realização dos testes de carga de acessos simultâneos será utilizada a ferramenta Microsoft Visual Studio Team System Test Edition, uma ferramenta direcionada para criação de testes provendo uma poderosa infra-estrutura para controle de qualidade. Será utilizada uma versão com licença de 90 dias, mais que suficiente para os testes necessários.

5 DIAGRAMAS DO SISTEMA

Para a realização da análise do sistema Enquetes&Pesquisas utilizou-se dos diagramas da UML e, apesar de o projeto do sistema ser OO, para a modelagem da base de dados será utilizado o diagrama de Entidade Relacionamento (DER) devido ao cuidado especial que se pretende com o projeto em relação a base de dados visando futuro aproveitamento em *Data Mining*.

O diagrama de execução (figura 12) demonstra de forma simples, uma configuração possível para a arquitetura global do sistema.

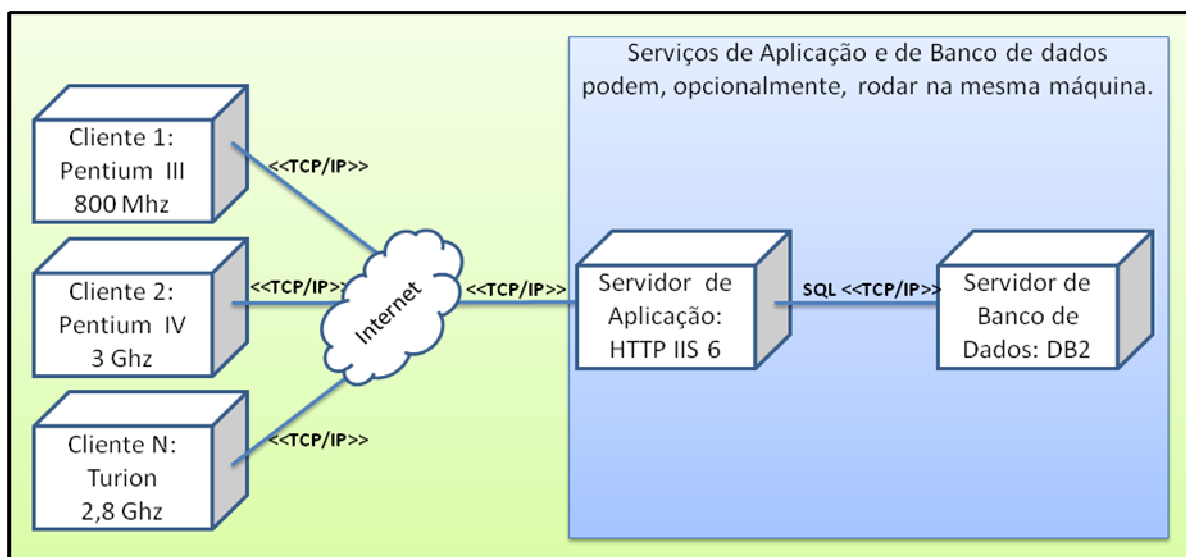


Figura 12 – Diagrama de execução do sistema Enquetes&Pesquisas

A figura 13 mostra o diagrama de casos de uso. Foram identificados 25 casos de uso para o sistema, sendo que os principais casos de uso são “Identificar usuário” e “Criar pesquisa” que estão relacionados mais diretamente com as classes “Usuário” e “Pesquisa”, as quais também são as classes mais importantes conforme pode ser visto no diagrama de classes, mais adiante.

A figura 14 apresenta o diagrama de classes. As classes “Usuário” e “Pesquisa” são as principais classes do sistema.

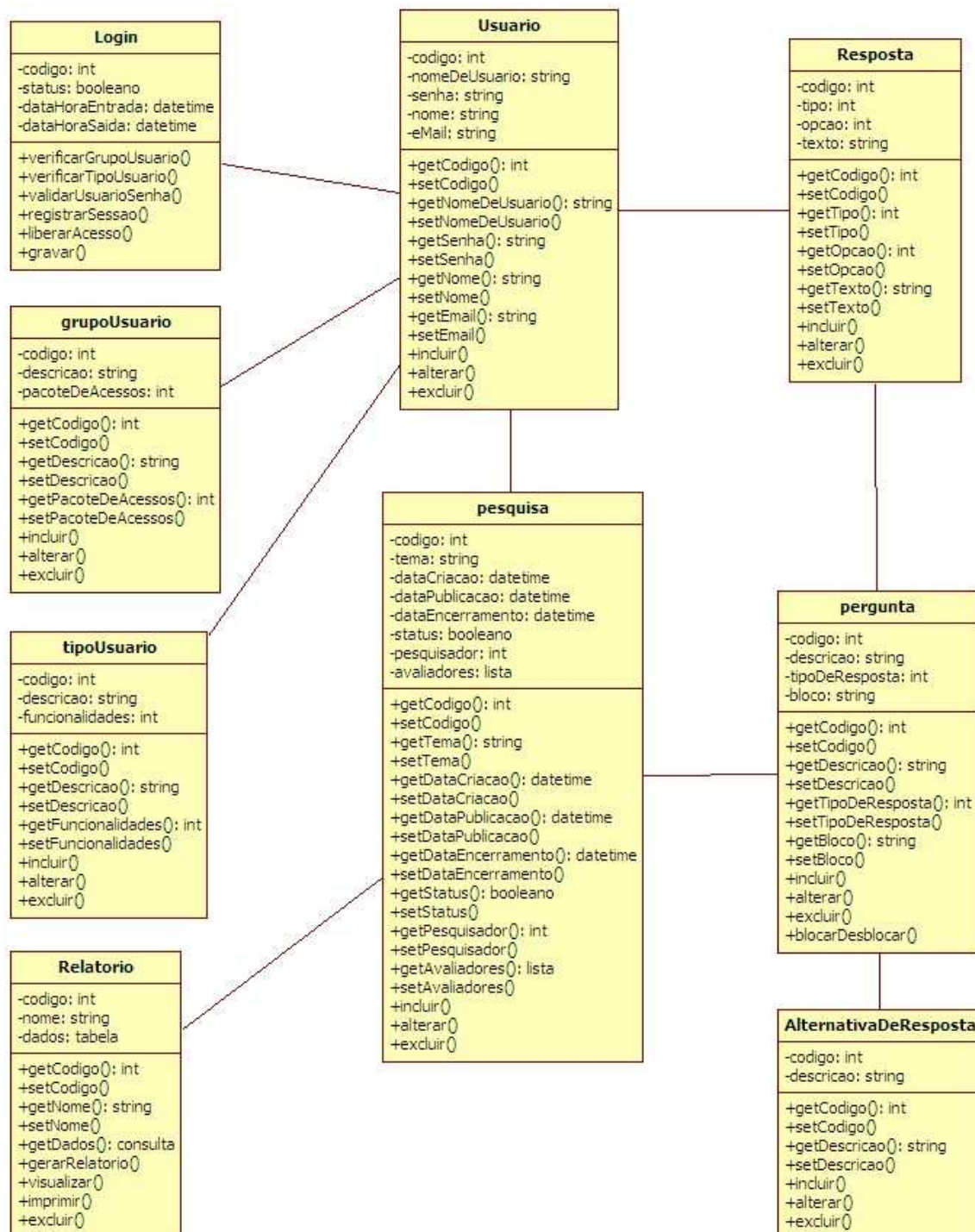


Figura 14 – Diagrama de classes do sistema Enquetes&Pesquisas

O diagrama de entidade relacionamentos (DER) (Figura 15) mostra a estrutura do banco de dados. Ele divide-se basicamente em duas partes principais, controle de usuários e controle de pesquisas.

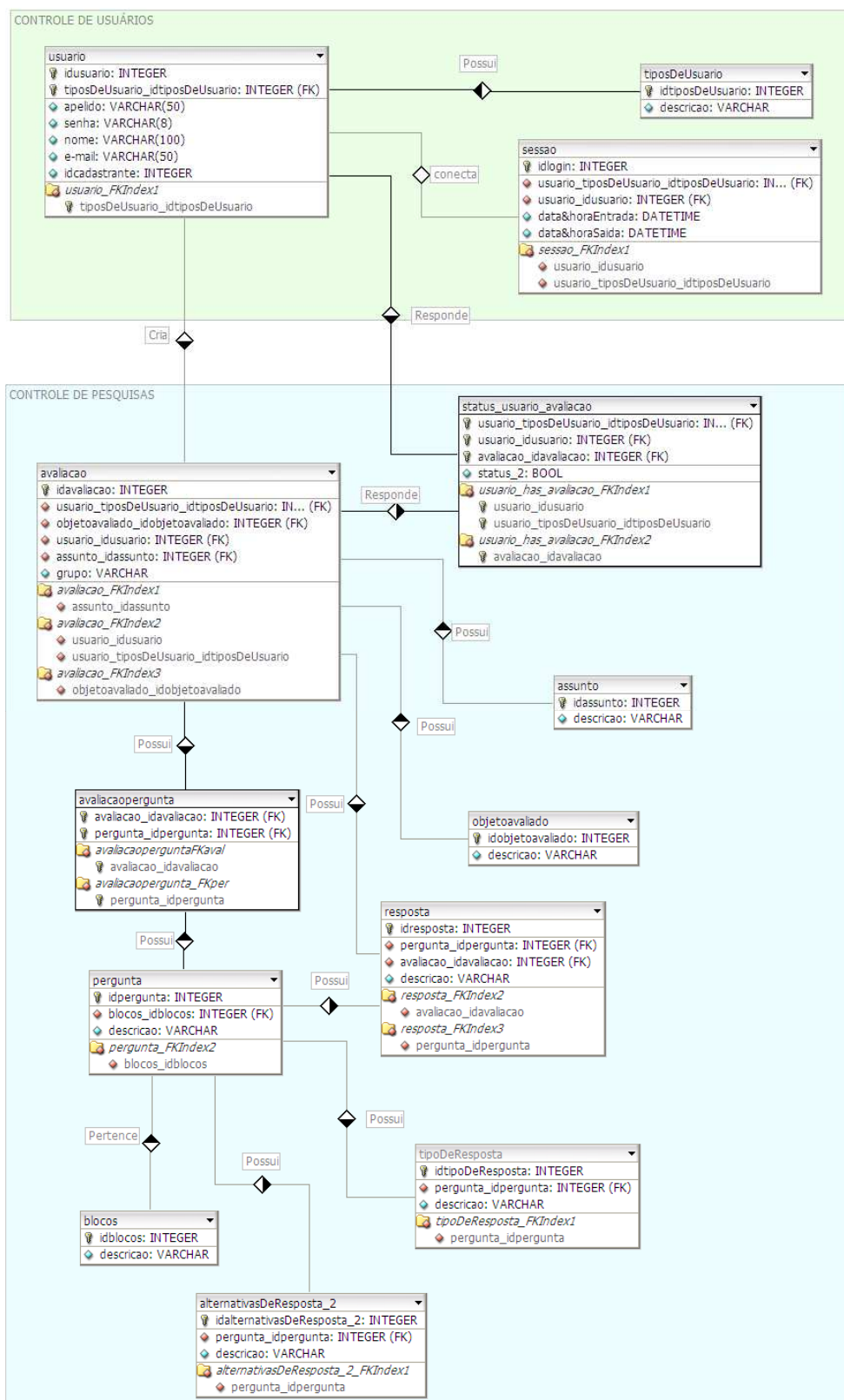


Figura 15 - DER do sistema Enquetes&Pesquisas

6 CONCLUSÃO

A história tem demonstrado que na evolução das tecnologias de informação nada é definitivo. Se hoje temos esse grande leque de possibilidades à frente dos profissionais da área, temos também uma grande incógnita do que será importante ou que tendências se afirmaram no mercado. Ferramentas e tecnologias cruciais ao gestor de informação de hoje, podem não ter mais valor ou utilidade daqui a cinco anos.

A pesquisa realizada demonstrou a necessidade de uma ferramenta ágil e segura para a realização de pesquisas rápidas e eficientes com baixíssimo custo, como a ora apresentada. Além disso, o fato de não ter-se encontrado nenhum trabalho acadêmico neste tema indicou que este pode ser referência para futuros trabalhos relacionados. Também buscou-se ao máximo a projeção de uma ferramenta genérica com muitas possibilidades de aplicação, como pesquisas de mercado, pesquisas de satisfação, enquetes, cadastro de pessoas e informações diversas, pesquisas pessoais em círculos de amigos, etc., visando com isso uma maior vida útil para a ferramenta e com isto uma garantia de alimentação para a base de dados por um período de tempo maior.

Acima de tudo, espera-se com este trabalho, contribuir para a melhoria da qualidade de vida das pessoas e da sociedade em geral, com um canal de comunicação que possa levar ao entendimento das necessidades e desejos, das dúvidas e curiosidades, das reclamações e diferenças entre fornecedor e cliente, sejam eles quem forem, e seja qual for a relação existente entre eles.

REFERÊNCIAS

DATE, C.J. Introdução a Sistemas de banco de dados. 7.ed. Rio de Janeiro: Campus, 2000.

DEITEL, Harvey M. C#: Como programar. 1.ed. São Paulo: Makron Books, 2003.

ELMASRI, Ramez; NAVATHE, Shamkant B. Sistemas de banco de dados. 4.ed. São Paulo: Addison Wesley, 2005.

ENQUETES.COM.BR. Disponível em: <http://www.enquetes.com.br>. Visitado em: 18/05/2008.

EQUESTIONA. Disponível em: <http://www.equestiona.com>. Visitado em: 11/05/2008.

FERGUSSON, Marilyn. A conspiração aquariana: transformações pessoais e sociais nos anos 80. 4.ed. Rio de Janeiro: Record, 1980.

FORMFÁCIL. Disponível em: <http://www.formfacil.com>. Visitado em: 17/05/2008.

FREITAS (H.), JANISSEK (R.) e MOSCAROLA (J.). Dinâmica do processo de coleta e análise de dados via web. CIBRAPEQ - Congresso Internacional de Pesquisa Qualitativa, 24 a 27 de março, Taubaté/SP, 2004. 12 p..

GATES, Bill. A empresa na velocidade do pensamento: Com um sistema nervoso digital. 1.ed. São Paulo: Companhia das Letras, 1999.

IBM. Página do DB2 Express-C para o Brasil. Disponível em: http://www.ibm.com/expressadvantage/br/catalogo/db2_express-c.phtml. Visitado em: 13/05/2008.

JONASSEN, David. Chefe do Departamento de Educação Continuada da Pennsylvania State University. O uso das novas tecnologias na educação a distância e a aprendizagem construtivista. Em Aberto, Brasília, ano 16, n.70, abr./jun.1996

MACHADO, Felipe Nery Rodrigues. Projeto de banco de dados: uma visão prática. 13.ed. São Paulo: Érica, 2006.

MARILIA, Universidade de. Apostila de banco de dados. São Paulo.

MEYER, Marilyn. Nosso futuro e o computador. 3.ed. Porto Alegre: Bookman, 2000.

MORAES, Giseli Diniz de Almeida. A gestão da informação diante das especificidades das pequenas empresas. Ci. Inf., Brasília, v. 35, n. 3, p. 124-132, set./dez. 2006.

MÜLLER, Felipe Martins. Gestão do Conhecimento Mediada por Tecnologia da Informação (UFSM / CT / PPGEP). Artigo apresentado no XXIII Encontro Nac. de Eng. de Produção - Ouro Preto, MG, Brasil, 21 a 24 de out de 2003.

PRESSMAN, Roger S. Engenharia de software. 3.ed. São Paulo: Makron Books, 1995.

QUADROS, Claudia Irene de. Base de dados: a memória extensiva do jornalismo. Em Questão, Porto Alegre, v. 11, n. 2, p. 409-423, jul./dez. 2005.

RAMALHO, José Antônio. Curso completo para desenvolvedores Web. 1.ed. Rio de Janeiro: Elsevier, 2005.

REINO, Lucas Santiago A. - Relacionamento entre o webjornal Campo Grande e os seus usuários. Dissertação submetida à Comissão Examinadora designada para obtenção do título de Mestre em Ciência da Informação na Universidade de Brasília, 2006, 103 p..

REZENDE, Denis Alcides. Tecnologia da informação aplicada a sistemas de informação empresariais: o papel estratégico da informação e dos sistemas de informação nas empresas. 3.ed. São Paulo: Atlas, 2003.

SANCHES, Roseli. Engenharia de software: planejamento do projeto. Apostila da Universidade de São Paulo. São Paulo: USP, 2003.

SILBERSCHATZ, Abraham. Sistemas de banco de dados. 5.ed. Rio de Janeiro: Elsevier, 2006.

SUPERTRÁFEGO. Disponível em: http://www.supertrafego.com/enquetes_home.asp. Visitado em: 11/05/2008.

TANENBAUM, Andrew S. Redes de computadores. 4.ed. São Paulo: Rio de Janeiro, 2003.

TAPSCOTT, Don. Economia digital: promessa e perigo na era da inteligência em rede. 1.ed. São Paulo: Makron Books, 1997.

TARAPANOFF, Kira(organizadora). Inteligência organizacional e competitiva. 1.ed. Brasília: Universidade de Brasília, 2001.

SUMÁRIO

1 INTRODUÇÃO	1
1.1 CONTEXTUALIZAÇÃO	3
1.2 OBJETIVO GERAL	4
1.3 OBJETIVOS ESPECÍFICOS	4
1.4 JUSTIFICATIVA	5
1.5 DELIMITAÇÃO	5
1.6 SUMÁRIO DOS CAPÍTULOS	6
2 REVISÃO DA LITERATURA	7
2.1 SURGIMENTO DOS SISTEMAS GERENCIADORES DE BANCO DE DADOS ...	7
2.1.1 Estrutura de um banco de dados	9
2.1.2 Instâncias e Esquemas	11
2.1.3 Armazenamento e consulta	14
2.2 DESENVOLVIMENTO WEB	15
2.2.1 Desenvolvimento Web X desenvolvimento Desktop	18
2.3 GESTÃO DA INFORMAÇÃO	19
2.3.1 A importância da informação	20
2.3.2 A hierarquia da informação	22
2.3.3 Gestão da informação X gestão do conhecimento	23
2.3.4 A informação coletada por formulários	27
2.4 SISTEMAS COM BASES DINÂMICAS	29
3 TRABALHOS RELACIONADOS	31
3.1 TRABALHOS ACADÊMICOS	31
3.2 TRABALHOS COMERCIAIS	32
4 METODOLOGIA	34
4.1 COLETA DE DADOS	34
4.2 ANÁLISE	34
4.2.1 Restrições de acesso	34
4.3 AMBIENTE DE DESENVOLVIMENTO E EXECUÇÃO	35
4.3.1 Desenvolvimento Web em plataforma Microsoft .NET	35
4.3.2 Banco de dados IBM DB2 Express-C 9	36
4.3.3 Plug In para relatórios	37
4.3.4 Configurações mínimas para execução do sistema	37
4.4 DESCRIÇÃO DAS FUNCIONALIDADES DO SISTEMA	38
4.5 DEFINIÇÃO DOS DIAGRAMAS	39

4.6 CRONOGRAMA DE DESENVOLVIMENTO	40
4.7 ANÁLISE DE RISCOS	41
4.8 TESTES	45
5 DIAGRAMAS DO SISTEMA	46
6 CONCLUSÃO	50
REFERÊNCIAS	51