

PRÓ-REITORIA DE GRADUAÇÃO
TRABALHO DE CONCLUSÃO DE CURSO

Bacharelado em Sistemas de Informação

Sistema de Apoio à Decisão aplicado ao Ambiente
Acadêmico da Universidade Católica de Brasília, com
ênfase no Corpo Discente

Autoras: Paula Souza dos Santos

Sybelles Nogueira Batista Gomes

Orientador: Prof. MSc. Milton Pombo da Paz

2008

Paula Souza dos Santos
Sybelles Nogueira Batista Gomes

**Sistema de Apoio à Decisão aplicado ao Ambiente
Acadêmico da Universidade Católica de Brasília,
com ênfase no Corpo Discente**

Projeto Final de Graduação apresentado ao Programa de Graduação em Sistemas de Informação da Universidade Católica de Brasília, como requisito para obtenção do título de Bacharel em Sistemas de Informação.

Orientador: Prof. MSc. Milton Pombo da Paz

Brasília

2008

TERMO DE APROVAÇÃO

Projeto Final de Graduação defendido e aprovado como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação, em 18 de junho de 2008, pela banca examinadora constituída por:

Prof. MSc. Milton Pombo da Paz

Prof. MSc. Vilson Carlos Hartman

Brasília

UCB

Dedicatória

À nossa família, pelo apoio e incentivo
em todas as horas.

“Uma mente que se abre a uma nova idéia
jamais volta ao seu tamanho original.”

Albert Einstein

AGRADECIMENTOS

Primeiramente a Deus, por permitir e tornar possível a realização e conclusão de mais um objetivo em nossas vidas.

Aos meus pais, Daniel Bravo dos Santos (*in memorian*) e Amasile de Fátima Souza dos Santos, incentivadores incondicionais em minha formação e em todo meu percurso de vida, responsáveis pela formação do meu caráter e pela construção de quem sou hoje e que não mediram demonstrações de amor e esforços para que alcançasse mais este objetivo.

Aos meus irmãos, Daniela, Fernanda e Daniel Júnior, pelo apoio, carinho, compreensão e por saber que, em todos os momentos, posso contar com vocês.

Aos meus amigos, pela ajuda na conclusão deste objetivo, pela parceria em vários outros projetos pessoais e profissionais, pelos agradáveis momentos vividos e pelo grande elo de amizade formado.

Particularmente a amiga Sybelle, pela parceria em mais um projeto, pela solidariedade, ânimo, alegria e amizade compartilhados.

Ao professor e orientador Milton Pombo da Paz, pela paciência na orientação e inspiração no amadurecimento dos meus conhecimentos.

A todos os professores que contribuíram decisivamente para a minha formação acadêmica, profissional e pessoal.

A Universidade Católica de Brasília, por ter nos permitido uma formação de qualidade.

Paula Souza dos Santos.

AGRADECIMENTOS

Primeiramente a Deus, pois, sem a presença Dele em nossas vidas, nada disso seria possível.

Aos meus pais, Fabiano Batista Gomes e Valdice Nogueira Batista Gomes, por acreditarem e apostarem em minha educação, que com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida.

Ao meu marido Rodrigo Vitor do Reis, que acompanhou toda minha caminhada, incentivando-me em cada pequena conquista e descoberta, que trouxe alegria, amor e ternura para meus dias, não me deixando desanimar e me fazendo acreditar na realização de sonhos.

Em especial ao meu bem mais precioso, meu filho, que mesmo ainda em meu ventre, transborda nossas vidas de alegria.

As minhas queridas irmãs, Fabiana e Isabella, pelo apoio emocional para enfrentar os desafios dessa longa caminhada, e por saber que, em todos os momentos, posso contar com vocês.

Aos meus amigos, pela ajuda na conclusão deste objetivo, pela parceria em vários outros projetos pessoais e profissionais.

Particularmente a amiga Paula Souza dos Santos, pelo seu ombro amigo em mais um projeto, pela solidariedade, ânimo, alegria e amizade compartilhados.

Ao professor e orientador Milton Pombo da Paz, pela paciência na orientação e maior fonte de estímulo nos momentos de dispersão e insegurança.

A todos os professores que contribuíram decisivamente para a minha formação acadêmica, profissional e pessoal.

A Universidade Católica de Brasília, por ter permitido uma formação de qualidade.

Sybelle Nogueira Batista.

RESUMO

Este projeto apresenta a descrição de conceitos relacionados a um Sistema de Apoio à Decisão e objetiva orientar sua utilização na coleta de informações essenciais à tomada de decisões por parte da Universidade Católica de Brasília, em função dos registros solicitados e disponibilizados ao Ministério de Educação e Cultura, por meio do Censo da Educação Superior, com ênfase no Corpo Discente da Instituição.

Palavras-chaves: Educação Superior, Censo, Sistemas de Apoio à Decisão, *Business Intelligence*, *Data Warehouse*, *Data Mart*, *Data Mining*, Ferramentas OLAP e Ferramentas OLTP.

ABSTRACT

This project presents the description of concepts related to a Decision Support System and objective orient their use in the collect of essential information for decision making on the part of the University Catholic of Brasília, according to the registers requested and provided to Ministry of Education and Culture, through the Census of Superior Education, with emphasis on Corps Discente of the Institution.

Keywords: *Superior Education, Census, Decision Support System, Business Intelligence, Data Warehouse, Data Mart, Data Mining, OLAP Tools and OLTP Tools.*

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO.....	1
1.1 - Motivação	2
1.2 – Breve Histórico	3
1.3 – Problemas Diagnosticados	5
1.4 – Surgimento da necessidade do trabalho.....	5
1.5 – Usuários do trabalho.....	6
1.6 – Organograma.....	6
CAPÍTULO 2 - OBJETIVOS	8
2.1- Objetivo Geral.....	8
2.2 - Objetivos Específicos	8
CAPÍTULO 3 – PROPOSTA DA PESQUISA	9
3.1 - Descrição da Pesquisa	9
3.2 - Resultados esperados	10
3.3 - Relação Custo x Benefício.....	11
3.4 – Áreas afetadas pela pesquisa	11
CAPÍTULO 4 – METODOLOGIA DO TRABALHO.....	12
CAPÍTULO 5 – DETALHAMENTO DO TRABALHO – ESTUDO TEÓRICO	13
5.1 – Censo	14
5.1.1 – Histórico	14
5.1.2 – Censo da Educação Superior	15
5.2 – Sistemas de Apoio à Decisão	16
5.2.1 - <i>Business Intelligence</i>	23
5.2.2 - <i>Data Warehouse</i>	28
5.2.3 - <i>Data Mart</i>	33
5.2.4 – OLTP.....	35
5.2.5 – OLAP.....	36
5.2.6 – OLAP x OLTP.....	42
5.2.7 – <i>Data Mining</i>	43
CAPÍTULO 6 – ESTUDO DE CASO	46
6.1 – Sistema Existente	46
6.2 – Detalhamento dos Problemas Diagnosticados	47
6.3 – Proposta de solução	48
6.4 – Modelo Diagrama de Classes.....	50
6.5 – Modelo Entidade Relacionamento	51
6.6 – Modelagem do Data Warehouse	52
6.6.1 – Descrição das tabelas do Modelo <i>Data Warehouse</i>	53
6.6.2 – Exemplos das tabelas do Modelo <i>Data Warehouse</i>	56
6.7 – Ferramentas utilizadas	58
CAPÍTULO 7 – CONCLUSÃO	65
7.1. Trabalhos Futuros.....	67
REFERÊNCIAS BIBLIOGRÁFICAS.....	68

Lista de Figuras

Figura 1 - Organograma da Universidade Católica de Brasília - UCB.....	6
Figura 2 - Organograma da Secretaria Acadêmica – SA	7
Figura 3 - Modelo Conceitual – SAD [Binder, 1994].....	22
Figura 4 - Visão geral das relações entre BI e Gerência do Conhecimento [Barbieri, 2001]	24
Figura 7 - Funcionamento BI [SOL7 - BI]	27
Figura 8 - Camadas da arquitetura de um ambiente de DW (www.sqlmagazine.com.br)	32
Figura 9 - Relatório Bidimensional (www.microsoft.com)	36
Figura 10 - Relatório Multidimensional (www.microsoft.com).....	37
Figura 11 - Arquitetura de um sistema OLAP (www.microsoft.com)	39
Figura 12 - Opções de armazenamento/ implementação de estruturas dimensionais (Barbieri, 2001)	42
Figura 13 - Passos do Data Mining (Navega, 2002).....	45
Figura 16 - Modelagem do Data Warehouse	52
Figura 17 - Cursos oferecidos (Visão na Ferramenta Qlik View).....	60
Figura 18 - Informações dos discentes (Visão na Ferramenta Qlik View).....	61
Figura 19 - Identificação de público alvo - Gráfico: Faixa Etária (Visão na Ferramenta Qlik View)	61
Figura 20 - Identificação de público alvo – Tabela e Filtro: Faixa Etária (Visão na Ferramenta Qlik View).....	62
Figura 21 - Identificação de Perfil – Gráfico e Tabela com Seleção Múltipla: Sexo (Visão na Ferramenta Qlik View)	62
Figura 22 - Aproveitamento do índice VIDA – Gráfico: Curso (Visão na Ferramenta Qlik View)	63
Figura 23 - Aproveitamento do índice VIDA – Gráfico: Ano (Visão na Ferramenta Qlik View)	63
Figura 24 - Desempenho de candidatos por Processo Seletivo – Gráfico (Visão na Ferramenta Qlik View).....	64
Figura 25 - Desempenho de candidatos por Processo Seletivo – Tabelas e filtros (Visão na Ferramenta Qlik View)	65

Lista de Tabelas

Tabela 1 - Sistemas de Apoio à decisão	20
Tabela 2 - Tabela comparativa entre Data Mart e Data Warehouse	34
Tabela 3 - Diferenças entre OLAP e OLTP	42
Tabela 4- Descrição da Tabela Fato: Histórico	53
Tabela 5 - Descrição da Tabela Dimensão: Processo Seletivo.....	53
Tabela 6 - Descrição da Tabela Dimensão: Candidato	53
Tabela 7 - Descrição da Tabela Dimensão: Aluno	54
Tabela 8 - Descrição da Tabela Dimensão: Curso	54
Tabela 9 - Descrição da Tabela Dimensão: Tempo	54
Tabela 10 - Descrição da Tabela Dimensão: Situação	54
Tabela 11 - Exemplo da Tabela Processo Seletivo	56
Tabela 12 - Exemplo da Tabela Candidato	56
Tabela 13 - Exemplo da Tabela Aluno	57
Tabela 14 - Exemplo da Tabela Curso.....	57
Tabela 15 - Exemplo da Tabela Tempo	57
Tabela 16 - Exemplo da Tabela Situação	57

Lista de Abreviaturas

ABNT: Associação Brasileira de Normas Técnicas

AQL: *Associative Query Logic*

BI: *Business Intelligence*

DOLAP: *Desktop On-Line Analytical Processing*

DW: *Data Warehouse*

ETL: *Extract, Transform, Load*

HOLAP: *Hybrid On-Line Analytical Processing*

IBGE: Fundação Instituto Brasileiro de Geografia e Estatística

IES: Instituição de Ensino Superior

MEC: Ministério da Educação e Cultura

MOLAP: *Multidimensional On-Line Analytical Processing*

OLAP: *On-Line Analytical Processing*

OLTP: *On-Line Transaction Processing*

PI: Pesquisador Institucional

ROLAP: *Relational On-Line Analytical Processing*

SAD: Sistema de Apoio à Decisão

SEEC: Serviços de Estatística de Educação e Cultura

SIEDSUP: Sistema Integrado de Informações da Educação Superior

TI: Tecnologia da Informação

UCB: Universidade Católica de Brasília

CAPÍTULO 1 – INTRODUÇÃO

Aos trinta e cinco anos de Educação Superior de Qualidade e doze anos como universidade, a Universidade Católica de Brasília conta com quarenta cursos de Graduação Presencial, nove cursos de Graduação à Distância, trinta e dois cursos de Lato Sensu (MBA e Especialização), doze Cursos de Stricto Sensu (Mestrado e Doutorado) e treze cursos de Lato Sensu à Distância (Especialização).

Em mais de três décadas de ensino superior, a Universidade Católica de Brasília já formou mais de trinta mil alunos nas suas diversas áreas de conhecimento. Atualmente, a Universidade Católica de Brasília possui mais de quatorze mil alunos de Graduação e aproximadamente seiscentos alunos de Pós-Graduação.

A Universidade Católica de Brasília tem se tornado referência de ensino superior de qualidade na região centro oeste, fato comprovado pela publicação do Guia do estudante, 2006.

A demanda do Ministério da Educação para as Instituições Ensino Superior tem aumentado consideravelmente nos últimos anos, em especial as informações relativas ao Censo da Educação Superior.

Preocupada com a não utilização interna dos dados enviados ao Ministério da Educação e Cultura, pretende-se utilizar na Universidade Católica de Brasília uma ferramenta que proporcione análises, por meio dos dados do censo de ensino superior, encaminhados anualmente pela Universidade Católica de Brasília ao Ministério da Educação e Cultura, que servirá de suporte à tomada de decisão da alta gerência da UCB.

Pretende-se, inclusive, identificar elementos que possam tornar-se informações capazes de gerar subsídios à tomada de decisões.

O trabalho está estruturado em sete capítulos:

No primeiro capítulo será apresentado o foco deste trabalho, bem como as informações mais relevantes a cerca das instituições envolvidas e o papel chave do Pesquisador Institucional.

No segundo capítulo apresentam-se os objetivos do trabalho, de forma a elucidar as expectativas gerais e específicas.

No terceiro capítulo será descrito o trabalho em seus aspectos gerais.

No quarto capítulo será apresentada a metodologia utilizada para a realização do trabalho.

No quinto capítulo será descrito o histórico do censo, a coleta de dados e o censo sob o ponto de vista da Universidade Católica de Brasília. Será descrito também o detalhamento das ferramentas de apoio que contribuíram para o desenvolvimento do trabalho, tais como: Sistemas de Apoio à Decisão, *Data Warehouse*, *Data Mart*, *Data Mining*, Ferramentas OLAP (*On-Line Analytical Processing*) e Ferramentas OLTP (*On-Line Transaction Processing*).

No sexto capítulo será descrito o estudo de caso do Censo da Educação Superior no âmbito da Universidade Católica de Brasília, com ênfase no corpo discente, onde serão destacados os benefícios da aplicação de uma ferramenta OLAP com os dados do censo, de forma a oferecer subsídios à tomada de decisão por parte da alta gerência da Universidade Católica de Brasília.

No sétimo capítulo será apresentada a conclusão do trabalho e as perspectivas futuras.

1.1 - Motivação

O Censo da Educação Superior é uma demanda anual do Ministério da Educação para as Instituições de Ensino Superior do País, sendo pré-requisito para diversos processos junto ao MEC. A Universidade Católica de Brasília tem contribuído de forma efetiva na realização deste trabalho, no entanto, não existem ferramentas na IES que se proponham a manipular os dados do censo, de forma a agregar informações que gerem conhecimentos na tomada de decisão para alta gerência.

A coleta de dados compreende as informações dos cursos de graduação, dados sobre docente e pessoal administrativo, financeiros, extensão, infra-estrutura,

compreendendo bibliotecas, instalações, equipamentos e outros recursos institucionais.

O trabalho da coleta de dados é gerido pelo Pesquisador Institucional, através da consulta ao sistema acadêmico e setores estratégicos da UCB, que insere os dados no sistema do Ministério da Educação e Cultura e não agregam qualquer valor informacional para alta gerência da Universidade Católica de Brasília.

Acredita-se neste projeto, que a utilização de uma ferramenta focada em *Business Intelligence* possibilitará a Universidade Católica de Brasília uma micro visão do ambiente acadêmico, resultando na identificação de informações capazes de gerar subsídios à tomada de decisões da alta gerência.

1.2 – Breve Histórico

A Secretaria Acadêmica é o órgão de apoio à administração acadêmica da Universidade Católica de Brasília – UCB, e está diretamente ligada a Reitoria. É responsável por receber, processar e distribuir informações e dados sobre a vida acadêmica dos alunos, desde o seu ingresso na instituição até a colação de grau, expedição e registro do diploma. Tem, também, a responsabilidade de controlar, efetuar, corrigir e atualizar os registros acadêmicos, de modo a garantir a segurança e a preservação dos documentos escolares, nos termos da legislação e normas vigentes, em prol da preservação da memória da Universidade e da contínua melhoria da qualidade dos serviços prestados.

Desde o ano de 2003, o diretor da Secretaria Acadêmica assumiu a figura do pesquisador institucional da UCB, sendo o interlocutor e responsável pelas informações da Universidade Católica de Brasília junto ao Ministério da Educação. A nomeação foi realizada pelo Reitor da Universidade Católica de Brasília, pois cabe ao dirigente da IES nomear o Pesquisador Institucional.

O período para lançamento do Censo da Educação Superior para as IES do país é definido por meio de portaria ministerial, que, pelo Sistema Integrado de

Informações da Educação Superior, torna disponível o questionário eletrônico do Censo.

Após a publicação desta portaria, a Universidade Católica por meio de seu pesquisador institucional avalia as informações requeridas e inicia o processo de coleta de dados em todos os setores envolvidos, que envolve dados de todos os cursos de graduação, tecnológicos e seqüências, como a modalidade presencial e a distância.

O pesquisador institucional, com o auxílio da diretoria de tecnologia da informação, coleta no sistema *Stratus*¹ os dados necessários dos cursos com modalidade presencial e a distância, de acordo com o formato exigido pelo Ministério da Educação e Cultura. As demais informações são coletadas por meio de solicitações aos setores chaves da IES, tais como: financeiro, extensão, biblioteca, pós-graduação e recursos humanos.

Cada setor viabiliza sua coleta de dados de forma independente, com auxílio ou não de um sistema, e os módulos de sistemas existentes não tem interação, o que dificulta o cruzamento de informações.

Após o recebimento dos dados dos setores em forma impressa, o pesquisador institucional acessa o questionário do censo pela página <http://www.ensinosuperior.inep.gov.br/> com sua senha máster e inicia o lançamento dos dados para cada curso. O Censo demanda tempo, pesquisa e atenção, dado que o sistema SiedSup do Ministério da Educação e Cultura realiza cruzamento de informações, de forma a minimizar erros de compreensão. No decorrer do lançamento de informações, é possível visualizar críticas que impendem o avanço de questões.

¹ Stratus – Sistema de Controle Acadêmico da Universidade Católica de Brasília – UCB.

1.3 – Problemas Diagnosticados

A alta gerência da Universidade Católica de Brasília não tem conhecimento da riqueza de informações que o censo da educação superior dispõe, dado que a coleta é realizada pelo pesquisador institucional, que alimenta a base de dados do Ministério da Educação e Cultura, que o utiliza para definições de políticas públicas para o desenvolvimento de ações acadêmicas.

A aplicação de uma ferramenta que proporcione análises em tempo real, capacitará à definição de ações acadêmicas na Universidade Católica de Brasília e a geração de informações facilitadoras para a tomada de decisão da alta gerência.

De acordo com estas informações, os principais problemas diagnosticados são:

- Os dados informados no Censo não são utilizados nas estratégias para a tomada de decisão;
- Inexistência de acompanhamento dos dados informados no decorrer do ano que antecede ao censo;
- Impossibilidade de gerar relatórios ou gráficos com as informações censitárias; e
- Macro visão do MEC x Micro visão da IES.

1.4 – Surgimento da necessidade do trabalho

A necessidade da utilização de uma ferramenta focada em *Business Intelligence* para manipular os dados censitários enviados ao Ministério da Educação e Cultura surgiu devido as análises e identificações relevantes para o esclarecimento da situação de evasão, como por exemplo, onde e porque a evasão na instituição tem sido tão presente nos últimos anos.

Na preocupação em abstrair os dados inseridos no censo de forma produtiva, verificou-se a oportunidade de utilizá-los como suporte na obtenção da visão interna da Instituição, gerando subsídios para a tomada de decisões, a serem disponibilizadas para a alta gerência.

1.5 – Usuários do trabalho

Os usuários de fato das informações geridas pela ferramenta disponibilizada por este projeto serão os dirigentes da Universidade Católica de Brasília, capacitando a compreensão dos dados por detrás do Censo da Educação Superior, para a tomada de decisão.

Utilizando-se dos benefícios proporcionados por ferramentas gráficas e interativas para análise de dados do censo da UCB, os dirigentes poderão idealizar e aplicar políticas de melhoria contínua na qualidade do ensino.

1.6 – Organograma

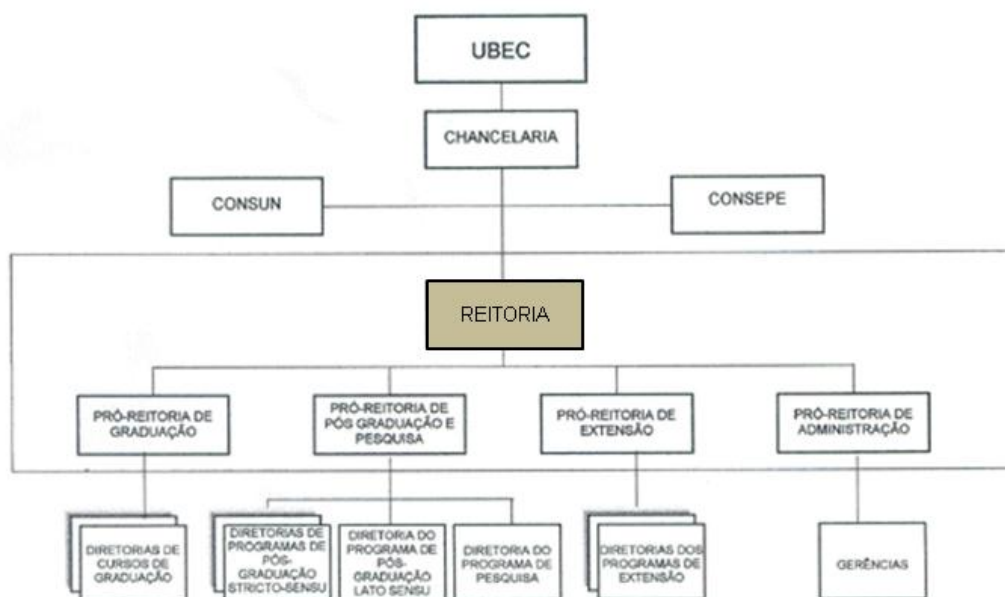


Figura 1 - Organograma da Universidade Católica de Brasília - UCB

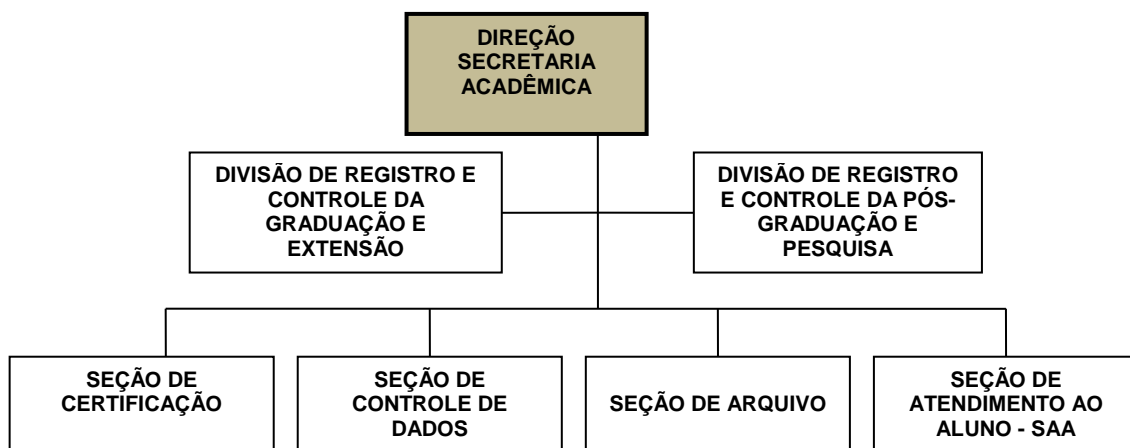


Figura 2 - Organograma da Secretaria Acadêmica – SA

CAPÍTULO 2 - OBJETIVOS

2.1 - Objetivo Geral

Utilizar ferramentas para o apoio à decisão da alta gerência da Universidade Católica de Brasília, baseado no Censo Educacional de Ensino Superior, utilizando navegações intuitivas nos dados dos discentes da IES.

Serão utilizadas informações no âmbito discente, proporcionando a radiografia das dimensões que compõe o Censo realizado pelo Ministério da Educação e Cultura, possibilitando a identificação de elementos que possam se tornar chaves nos aspectos estratégicos da Universidade Católica de Brasília.

2.2 - Objetivos Específicos

Com foco no auxílio à alta gerência da Universidade Católica de Brasília na tomada de decisões e na compreensão de fatores por detrás das informações censitárias do ensino superior, foram identificados os seguintes objetivos específicos:

- Identificar perfis dos discentes nos cursos;
- Vislumbrar o ambiente acadêmico, de forma dinâmica;
- Permitir aos usuários a geração de relatórios e gráficos gerenciais;
- Realizar pesquisas complexas com a base de dados; e
- Permitir o acesso a informações do banco de dados.

CAPÍTULO 3– PROPOSTA DA PESQUISA

3.1- Descrição da Pesquisa

Com o avanço da tecnologia e o surgimento de novas necessidades, antes consideradas irrelevantes no processo de tomada de decisão, houve a necessidade de transferir para o usuário final um papel mais ativo na manipulação de dados em busca de resultados que agregassem valores dentro do domínio estratégico de uma Instituição.

Conforme a evolução no processamento dos dados, percebeu-se que a análise minuciosa e focada nos dados poderia aperfeiçoar resultados e proporcionar conhecimento em novas tendências de mercado. Baseando-se em tal fato, a busca do conhecimento centrado em informações consistentes para tomada de decisões passou a acarretar forte impacto dentro de Instituições, que vêm nesta necessidade um investimento em ações futuras.

Como os sistemas transacionais, inseridos nas Instituições com o objetivo de garantir as operações do dia a dia, não registravam históricos organizados e nem geravam informações estratégicas, houve a necessidade de agrupar tais informações e centralizá-las em um único sistema, disponibilizando um histórico multidimensional, ou seja, a mesma informação poderia ser vista de dimensões diferentes. Para este tipo de Sistema de Apoio à Decisão foi dado o nome de *Data Warehouse*.

Baseado no conceito de *Business Intelligence*, o *Data Warehouse* surgiu, assim como outros conceitos diretamente ligados a ele, para facilitar o acesso até as informações de mais minuciosa análise possível, muitas vezes solicitadas pelos dirigentes organizacionais.

Não se pode, hoje, considerar que uma Instituição que almeja alcançar ou manter a posição de líder em seu segmento, que esta o faça sem utilizar os princípios estabelecidos em um processo de BI, pois este tem como foco a informação, vital para a condução dos negócios e para a tomada de decisões estratégicas. Entender como este processo e as tecnologias aliadas a ele

funcionam, auxiliará os tomadores de decisão a descobrir novas tendências e caminhos para competir numa economia globalizada, onde a concorrência é acirrada, prestando melhores serviços, com maior rapidez e sem aumento dos custos.

Focando o Censo de Educação Superior, feito todos os anos para o Ministério da Educação e Cultura, foi identificada uma excelente oportunidade de aplicar os conceitos acima mencionados, pois, diversas informações são coletadas e, por muitas vezes, não aproveitadas de forma a identificar oportunidades de melhoria ou focos em ações determinantes no processo de preservação e manutenção dos discentes. Com isso, a proposta de utilizar ferramentas de BI na Universidade Católica de Brasília, surgiu com o objetivo de gerar conhecimentos por meio da coleta de dados, para obtenção de informações coesas e de grande valia no processo decisório da alta gerência da IES.

Aliado aos conceitos computacionais e de gestão da informação, um dos fatores mais importantes na manipulação de um *Data Warehouse* e que representa o sucesso do resultado apresentado por uma decisão tomada corretamente é a credibilidade da informação. Tal credibilidade influencia diretamente o resultado final do processo decisório.

Além das camadas básicas de um DW, os dados passam por diversos processos de transformação para fins de integração, consistência e certificação de sua fidedignidade. Toda a arquitetura e abrangência de suas ferramentas serão retratadas, com a finalidade de orientar e apoiar o processo decisório.

3.2 - Resultados esperados

Dentre os resultados esperados com a proposta do projeto, a utilização de ferramentas de *Business Intelligence* proporcionará benefícios com impacto produtivo na Instituição, tais como:

- Combinações de dados;

- Compartilhamento de informações entre diferentes grupos de usuários;
- Respostas pontuais sobre os negócios;
- Análises de diferentes informações para se chegar a uma unificação; e
- Vantagem competitiva pela melhor exploração dos benefícios oferecidos pela ferramenta, transformando dados em conhecimento do negócio.

3.3 - Relação Custo x Benefício

Com a identificação de novos nichos de mercados e oportunidades de melhorias disponibilizadas com o auxílio das ferramentas de *Business Intelligence*, o aumento do rendimento e credibilidade da Instituição é fortemente esperado, através da conquista e fidelização dos discentes da Universidade Católica de Brasília.

Cabe a alta gerência, com base nas informações geridas pelas ferramentas apresentadas, administrar a estratégia de negócio que garanta um ensino com qualidade, conquistando mercados considerados intangíveis.

3.4– Áreas afetadas pela pesquisa

O benefício da pesquisa se dará à alta gerência da Universidade Católica de Brasília, que envolve reitoria, pró-reitoria e mantenedora, ficando a manipulação da ferramenta disponível para o Pesquisador Institucional.

Cabe a alta gerência definir os possíveis setores que poderão usufruir das informações geridas pela ferramenta, no intuito de aplicar as decisões de negócios necessárias para o trabalho de fidelização e conquista de novos discentes.

CAPÍTULO 4 – METODOLOGIA DO TRABALHO

No desenvolvimento do projeto será realizada uma pesquisa científica, onde serão analisadas referências bibliográficas em todos os assuntos pertinentes ao tema abordado, respeitando as normas estabelecidas pela Associação Brasileira de Normas Técnicas (ABNT).

Do ponto de vista de sua natureza, será uma pesquisa aplicada, envolvendo verdades e interesses do Censo da Educação Superior da Universidade Católica de Brasília.

Por considerar a análise dos dados informados no Censo, a pesquisa abordará um caráter quantitativo, enumerando os principais indicadores pertinentes à tomada de decisões.

Quanto aos fins, possuirá características metodológicas, pois irá capturar informações dos discentes e as manipulará em busca da construção de instrumentos que identifiquem perfis.

Em relação aos meios de investigação, será utilizada a técnica de pesquisa de campo, onde os participantes do projeto fazem parte do estudo realizado, complementada pela investigação documental, observando documentos publicados pelo Ministério da Educação.

CAPÍTULO 5 – DETALHAMENTO DO TRABALHO – ESTUDO TEÓRICO

Na Educação Superior, tão importante quanto à oferta de um ensino de qualidade, de uma infra-estrutura adequada, de um corpo docente capacitado e de uma coordenação competente, é a qualidade de seus controles e registros, que permitirão às Instituições alcançarem um alto nível de segurança e eficiência nas atividades desempenhadas e nas informações prestadas à comunidade acadêmica e aos órgãos oficiais de avaliação e controle.

Para vislumbrar os dados sobre a educação superior, a tecnologia da informação tem se mostrado muito necessária e presente, oferecendo aos gestores, subsídios para o planejamento das ações acadêmicas e administrativas.

Atualmente, o maior diferencial entre Instituições de qualidade que buscam aumentar sua participação no mercado é a informação, que deve ser coesa, segura e disponibilizada em tempo hábil para a tomada de decisão. As Instituições têm o mesmo público alvo e visam aumentar suas margens de lucro.

Para que este processo ocorra da forma mais correta possível, é preciso cuidar do equilíbrio entre os investimentos empregados e os resultados obtidos. Além disso, a estratégia de implementação deve estar alinhada com os objetivos institucionais.

Com essa finalidade, as ferramentas disponibilizadas pelo *Business Intelligence* se apresentam com grande importância para as Instituições que almejam alta competitividade no mercado, uma vez que, com elas, pode-se obter antecipação aos concorrentes e entendimento das variáveis do mercado.

5.1 – Censo

A palavra Censo se origina do latim *Censu*, que segundo o dicionário Aurélio é o “Conjunto de dados estatísticos dos habitantes de uma cidade, província, estado, nação, entre outras, com todas as características.” O levantamento censitário é enquadrado como pesquisa quantitativa, que descreve as características de uma determinada situação.

5.1.1 – Histórico

O primeiro ato regulatório do censo demográfico no Brasil é datado em 1846, que fixava um intervalo de oito anos para coleta de dados. Para viabilizar a execução do censo foram necessários cinco anos de preparação, sendo programado para ocorrer em 1852. Em pleno início de execução do censo, a população se revoltou e impediu a realização do censo, por acreditar que o Decreto 797 de Junho de 1851, visava à escravidão dos homens de cor. A realização do censo então foi postergada por mais vinte anos.

Em 1872 foi realizado o primeiro recenseamento nacional do País, denominado Recenseamento da População do Império do Brasil. Regulamentado em 1970, o novo regulamento surgiu com a necessidade de ampliar as estatísticas no país, principalmente após a implantação do registro civil de nascimento, casamentos e óbitos. O censo cobriria todo o território nacional a cada dez anos.

Fundado em 1936 o Instituto Brasileiro de Geografia e Estatística – IBGE modernizou o levantamento censitário no País, ampliando a abrangência temática do questionário para itens de relevância social e econômica do País.

O IBGE é uma instituição da administração pública federal, subordinado ao Ministério do Planejamento, Orçamento e Gestão. Tem como missão “Retratar o Brasil com informações necessárias ao conhecimento da sua realidade e ao exercício da cidadania.”.

5.1.2 – Censo da Educação Superior

O levantamento estatístico da Educação de Ensino Superior surgiu em 1956, por meio do decreto 38.661, em um trabalho conjunto com o Instituto Brasileiro de Geografia e Estatística – IBGE e Ministério da Educação, que buscavam sistematizar e otimizar as estatísticas educacionais no País. Nasce neste momento no Ministério da Educação o Serviço de Estatísticas da Educação e Cultura (SEEC).

Coube a Coordenação do Sistema de Estatística da Educação Superior (SEEC), definir os instrumentos de coletas a serem realizados pelo IBGE, em todas as instituições de ensino superior do País. A coleta era realizada anualmente, utilizando-se de questionários que deveriam ser respondidos a máquina ou mão e devolvidos via correio ao Ministério da Educação e Cultura. Os questionários coletavam dados Cadastrais, Recursos Humanos, Alunado e de Educação Física e Desportos da Instituição de Ensino Superior. Ao recebê-los das Instituições de Ensino Superior do País, a SEEC realizava o processamento, que era divulgado em publicações do MEC.

Com a criação da Lei de Diretrizes e Bases da Educação Nacional, intitulada LDB (Lei 9.394 de 20 de dezembro de 1996), fica estabelecido no art. 9º, inciso V que *“a união incubiar-se-á de coletar, analisar e disseminar informações sobre educação”*.

Em 1997, o Instituto Nacional de Estudos e Pesquisas Educacionais - INEP é transformado em autarquia federal, por meio da lei 9.448, de 14 de março de 1997. Fica estabelecido em seu art. 1º, inciso I, que o INEP será o órgão responsável por *“organizar, e manter o sistema de informações e estatísticas educacionais”*.

Com o advento da internet, o Ministério da Educação através da Portaria Ministerial nº 971, de 22 de agosto de 1997, no seu art. 3º, tornou obrigatório o encaminhamento ao INEP, anualmente uma relação de dados sobre a IES e seu corpo docente e discente, através da internet ou por meio de disquete para as instituições que não possuíam acesso a internet.

No ano de 2000 é criada a base de dados corporativa do INEP, intitulado Sistema Integrado de Informações Educacionais (SIEDSUP), que contém em um de

seus subsistemas o Censo da Educação Superior. O SIEDSUP aprimorou a coleta de dados através de formulários eletrônicos disponibilizados no site www.ensinosuperior.inep.gov.br, que deveria ser acessado pela IES, com a utilização de senha individual.

Em 2001, através da Portaria Ministerial nº 2517 de 22 de novembro de 2001, fica estabelecido que o dirigente da IES deverá nomear um Pesquisador Institucional, que será o elo entre a IES e o Ministério da Educação para atender as demandas relativas ao INEP, inclusive o censo da educação superior. O preenchimento e validação do censo é pré-requisito para as diversas demandas do MEC, seja um pedido de reconhecimento, inscrição de alunos no exame nacional de cursos, avaliação institucional, adesão ao FIES e Prouni, entre outros.

A concepção do Censo da Educação Superior transcende ao Ministério da Educação e Cultura e as Instituições de Ensino Superior do País, atende em potencial os candidatos a Educação Superior, que através do Censo são capazes de vislumbrar os perfis das instituições no país, obtendo, por exemplo, informações sobre biblioteca, titulação de docentes e quantidade de concluintes, democratizando a informação do ensino superior, o que demonstra um perfil mais exigente dos candidatos, que prezam pela melhoria contínua da qualidade. [SINAES, 2007.]

5.2 – Sistemas de Apoio à Decisão

Segundo o dicionário Aurélio “Decisão é o ato ou efeito de decidir; é a capacidade de tomar decisões.”. Além desta definição, pode-se conceituar, segundo o mesmo autor, “*Apoio é tudo que serve de sustentáculo, de suporte; auxílio, socorro, amparo*”. Utilizando as palavras no mesmo contexto e vislumbrando o ambiente computacional, pode-se inserir o conceito de Sistemas de Apoio à Decisão (SAD) que, de forma resumida, entende-se que são ferramentas desenvolvidas com foco no auxílio à tomada de decisões, geralmente, por parte dos líderes das empresas onde são aplicados.

Da mesma forma, Binder (1994) descreve Sistemas de Apoio à Decisão como “sistemas mais complexos que permitem total acesso à base de dados corporativos, modelagem de problemas, simulações e possuem uma interface amigável. Além disso, auxiliam o executivo em todas as fases de tomada de decisão, principalmente, nas etapas de desenvolvimento, comparação e classificação dos riscos, além de fornecer subsídios para a escolha de uma boa alternativa”.

Em alguns casos, pela dificuldade na adoção de caminhos a serem tomados para melhoria da execução e pela alta diversidade e elevado número de informações, podem ser usados como ferramentas chaves, de *software*, no processo decisório estratégico de uma instituição.

Muitos são os fatores que influenciam, apóiam e guiam um líder a tomar uma decisão frente a problemas que se refletem tanto internamente quanto externamente ao ambiente institucional, tais como: competitividade, necessidade de integração de dados, entre outros. Isso porque toda decisão envolve riscos e conseqüências que devem ser analisados minuciosamente para que tenham resultados positivos e atuem como o esperado. Independente dos fatores que serão tidos como base para a tomada de decisão, o que se espera, em todos os casos, é o êxito na execução e o sucesso no resultado.

Bispo (1998), em seus estudos, menciona que desde o início do processo decisório na civilização “*o homem sempre procurou algo que o auxiliasse*”. Bispo afirma que “*ele buscava nas divindades esse auxílio para que pudesse tomar decisões consideradas sábias.*”.

Atualmente, tendo o conceito fortalecido por profissionais da área de Tecnologia da Informação (TI) e pessoas que entendem a real e verdadeira importância que uma informação pode agregar ao negócio, os Sistemas de Apoio à Decisão vêm se fortalecendo e tornando-se “*extremamente úteis no processo de tomada de decisão, pois possibilitam a obtenção de dados com maior qualidade e maior velocidade e, em certos casos, poderão até sugerir novos caminhos decisórios.*” [Binder, 1994].

Dentre os fatores considerados de fundamental relevância no processo decisório de uma instituição citam-se:

- **Tempo:** em uma era onde a tecnologia atua fortemente em todos os ramos das empresas, não se pode desconsiderar que quem tem uma boa execução baseada em informações estratégicas pode sair na frente com lançamentos de novos conceitos e idéias;
- **Novas ferramentas:** com o melhoramento nos processos de coleta e transformação de dados e informações, gerando conhecimento, pode-se garantir um atendimento rápido e resultado em curto prazo;
- **Preço x Qualidade:** ambas as características aliadas geram total retorno à instituição; e
- **Concorrência:** sempre respeitada, serve como referência e base para o incremento nas decisões.

Além dos fatores mencionados acima, pode-se mencionar outros de igual relevância, como por exemplo: meio-ambiente, legislações, cultura organizacional, entre outros.

Segundo Binder (1994), “o desenvolvimento e o uso de um Sistema de Apoio à Decisão requerem certo ambiente bem planejado para que possam prosperar.” Ele menciona ainda que “um SAD não é um sistema simples e também não são simples as tarefas as quais ele se propõe a resolver”.

Antes de qualquer coisa, para que um SAD atenda sua proposta e, conseqüentemente, obtenha sucesso, é necessário que seja aceito pela instituição. Quando se menciona a expressão verbal “ser aceito” a referência não se faz somente a quem, de fato, desenvolve e põe em prática, por meio de seus processos, o conceito do sistema. A expressão aponta aceitação por todos os membros envolvidos no amplo contexto de sistema de apoio à decisão, partindo dos desenvolvedores e profissionais selecionados para manipularem a ferramenta, passando pelos usuários do sistema, até a alta gerência, que receberá as informações necessárias à tomada de decisões em estágio avançado.

Dentre as características que um ambiente tecnológico deve possuir para suportar um SAD, de acordo com Binder (1994), destacam-se: o suporte dos profissionais de TI, os sistemas de informação eficientes e eficazes e a integração da rede de sistemas de informação da empresa.

Partindo do pressuposto que o SAD adotado por uma instituição traz um conceito não novo, mas diferente do que comumente é usado, pode-se levar em consideração que seus usuários, por não estarem habituados a utilizar esse tipo de ferramenta, necessitarão de apoio para manipularem a ferramenta. Esse apoio deve partir da área responsável pela execução correta dos processos requeridos pela alta gerência, ou seja, as pessoas responsáveis por “guiar” o conceito daquilo que realmente deve-se executar com a ferramenta. Além disso, como principal apoio a um SAD, deve-se ter o suporte dos profissionais tecnicamente capazes de auxiliar os usuários, quer seja em atividades mais simples, como manuseio das opções oferecidas pela ferramenta, quer seja em atividades mais complexas, como manuseio de banco de dados, requeridos em qualquer sistema de apoio à decisão.

Sabe-se que, assim como nos seres humanos, a eficiência de um sistema está ligada as ações realizadas e a eficácia aos resultados produzidos. Assim, sendo o SAD um sistema que serve como uma ponte de informação entre aquilo que a empresa possui e o que ela deseja (utilizando para isso uma informação clara e precisa), é necessário que todo dado requerido e utilizado pelo sistema esteja armazenado (provavelmente no banco de dados da instituição) de forma correta e atualizada, pois é desta base que o SAD retirará informações necessárias às análises futuras. Sendo assim, com a integridade dos dados, o SAD poderá garantir informações precisas, confiáveis e em tempo adequado.

É muito importante que, quem trabalha com um sistema de apoio à decisão, tenha em mente que toda e qualquer informação gerada por meio da ferramenta utilizada pode e deve ser repassada às diversas áreas da instituição. Isso porque, o ideal de uma ferramenta desse tipo é a integração com outros sistemas utilizados, proporcionando um maior conhecimento do negócio e decisões mais apropriadas de acordo com cada caso específico. Ao surgirem dúvidas de quais informações devem ser geradas, como gerá-las, para quem gerá-las e quando gerá-las, Binder (1994) menciona: “estas questões são fatores críticos para o sucesso de qualquer tomada de decisão e provam que existe a real necessidade de integração dos sistemas da empresa. Esta integração permite o tratamento das mais diversas informações, a seleção de dados relevantes e o direcionamento da informação para a pessoa certa, na hora certa, possibilitando maior competitividade e melhor aproveitamento das oportunidades de negócio. O uso de Sistema de Apoio à

Decisão em um ambiente deste tipo vai fornecer maior velocidade e melhor qualidade às informações, fatores indispensáveis no mundo extremamente competitivo em que vivemos”.

De acordo com Binder (1994), pode-se estruturar um SAD a partir de duas partes: a que trata dos níveis de tecnologia, do pessoal envolvido e da abordagem de desenvolvimento e a que trata das concepções das diferentes pessoas envolvidas no processo.

Na primeira parte, pode-se subdividir os níveis de tecnologia em três, sendo eles o SAD específico, os geradores de SAD e as ferramentas para SAD. "Os SAD específicos são os sistemas de apoio à decisão propriamente ditos, utilizados pelo usuário final. Com os geradores de SAD constroem-se SAD específicos e com as ferramentas podemos desenvolver tanto os geradores quanto os aplicativos em si" [Binder, 1994].

Uma observação muito importante citada por Binder (1994) é a diferença entre SAD e sistemas tradicionais. Para ele, tal diferença se deve ao fato de que os sistemas de apoio à decisão são caracterizados pela flexibilidade e adaptabilidade às mudanças. Tais mudanças ocorrem, não só no problema, mas também no contexto em que ele está inserido e na forma de encará-lo. Portanto, os usuários devem ter um envolvimento ainda maior no processo de um desenvolvimento de um SAD, podendo chegar ao papel de projetista.

Baseando-se nos três níveis de tecnologia mencionados, pode-se distribuir o pessoal envolvido no desenvolvimento de um SAD de acordo com a tabela abaixo:

Tabela 1 - Sistemas de Apoio à decisão

Sistemas de Apoio à Decisão		
Específico	Gerador	Ferramenta
Gerente: responsável pela tomada de decisão.	Projetista: conhecedor da área onde o problema está inserido.	Criador: desenvolvedor da ferramenta.
Intermediário: auxilia o usuário.	Suporte: desenvolvedor de recursos.	

Com o trabalho desenvolvido diariamente com a ferramenta, observa-se que um mesmo profissional pode assumir diversos papéis no desenvolvimento de um SAD. Pelo fato da ferramenta exigir profundo conhecimento das informações que estão sendo tratadas, o projetista passa a ser um profissional da área de atuação

foco da ferramenta (como por exemplo, profissionais da área de marketing de uma instituição), contando este com ajuda específica de profissionais da TI.

Pelo fato da definição do problema inicial estar nas mãos do projetista, juntamente com o usuário, para, a partir daí, iniciar-se o desenvolvimento do sistema que apoiará a idéia conceitual da ferramenta final, a abordagem de desenvolvimento ideal é a iterativa, onde um sistema inicial utilizável é desenvolvido e vai crescendo de acordo com as necessidades apresentadas. O mais adequado é que todas as possíveis exigências sejam centralizadas e formalizadas, para que sirvam de documento para possíveis alterações de versões de sistemas.

Ao se desenvolver um SAD deve-se levar em consideração o que se definiu como sendo a segunda parte estrutural de um sistema desta natureza: as concepções das diferentes pessoas envolvidas no processo.

Do ponto de vista do usuário, o que se percebe, na maioria dos casos, são exigências de funcionamento, sendo elas, segundo Binder (1994):

- Um SAD deve servir de apoio ao processo decisório, enfatizando os processos não estruturados e/ ou semi-estruturados;
- Um SAD deve englobar todos os níveis do processo decisório de um gerente, realizando integração destes níveis, quando for necessário;
- Um SAD deve apoiar tanto decisões tomadas individualmente, quanto em grupo;
- Todas as fases de um processo decisório devem ser auxiliadas por um SAD;
- Um SAD deve dar apoio a diversos processos de tomada de decisão, para que o escopo de utilização seja bastante amplo; e
- Um SAD deve ser fácil de usar.

O ponto de vista do projetista aponta para os recursos de desenvolvimento. Seu principal interesse é o de utilizar as ferramentas e os geradores da melhor maneira possível, para dar apoio a processos decisórios [Binder, 1994]. A partir de tal ponto de vista, foi elaborado, com o decorrer dos anos, um modelo conceitual dos componentes de um SAD, com o objetivo de facilitar o trabalho de desenvolvimento e de manutenção do mesmo [Binder, 1994]:

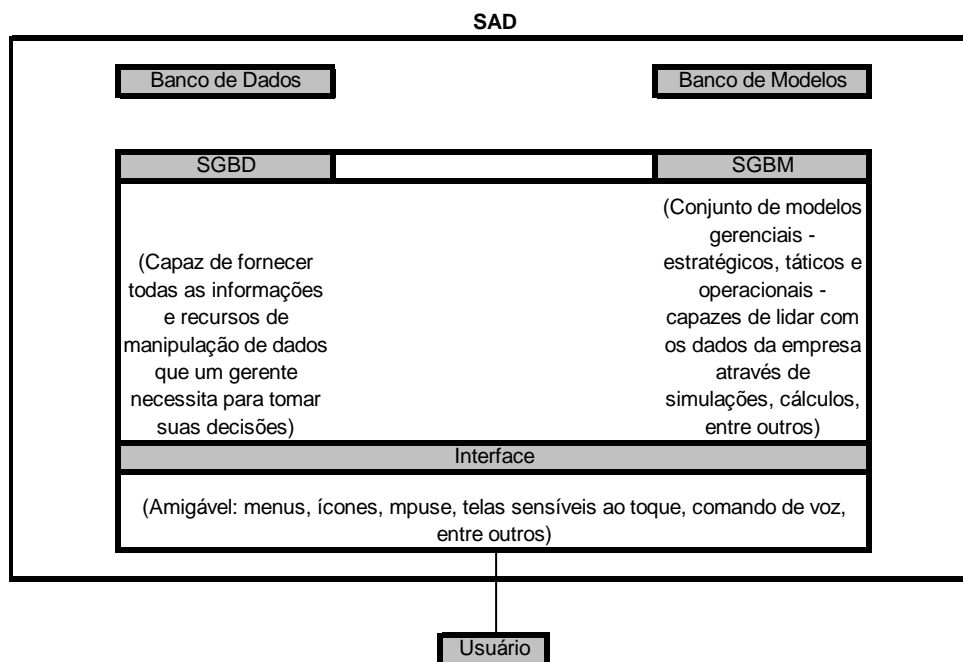


Figura 3 - Modelo Conceitual – SAD [Binder, 1994]

Já do ponto de vista do criador, o fator primordial é a tecnologia de desenvolvimento, ou seja, as ferramentas que poderão ser utilizadas quando forem criados os SAD's específicos e os geradores de SAD.

Os Sistemas de Apoio à Decisão não só fornecem informações para tomada de decisões, mas também contribuem e influenciam o processo. Um SAD deve fornecer e analisar alternativas, pesquisar históricos de decisões tomadas e auxiliar a resolução de problemas estruturados [Colaço, 2004]. Por esses fatores, um SAD permeia nos três níveis existentes dentro de uma instituição: estratégico, tático e operacional.

Como símbolo de Sistemas de Apoio à Decisão dentro das organizações, surge atualmente forte, o conceito de *Business Intelligence* (BI), que, de forma mais ampla, pode ser entendido como a utilização de variadas fontes de informação para se definir estratégias de competitividade nos negócios da empresa [Barbieri, 2001].

5.2.1 - *Business Intelligence*

O termo *Business Intelligence* surgiu em 1989 e representa a habilidade de se estruturar, acessar e explorar informações, normalmente guardadas em *Data Warehouses* ou *Data Marts*, com o objetivo de desenvolver percepções, entendimentos, conhecimentos, os quais podem produzir um melhor processo de tomada de decisão [Barbieri, 2001].

Com o avanço da tecnologia e com as reais necessidades de apresentar um diferencial competitivo no mercado, muitas empresas passaram a evoluir em suas formas de armazenamento de dados, proporcionando um melhoramento em seus processos de negócio, com base nos sistemas de gestão de bases de dados. Tais dados tornaram-se ativos extremamente importantes em todos os processos internos e externos, pois passaram a influenciar nas decisões tomadas. O BI surgiu neste contexto, permitindo que os dados fossem trabalhados e transformados em informações valiosas.

De acordo com informações disponibilizadas pela DW Brasil, no artigo Histórico do *Data Warehouse*, a descoberta do conceito de BI permitiu dividir os sistemas informatizados em dois grupos, sendo eles:

- Sistemas transacionais: sistemas que tratam o negócio. Esse tipo de sistema dá suporte e garante a operação da Instituição; e
- Sistemas de suporte à decisão: sistemas que analisam o negócio. Esse tipo de sistema ajuda a interpretar os fatos ocorridos e a decidir sobre estratégias futuras.

Enquanto BI é mais compartimentada, objetiva, focada em estruturas definidas e transforma dados em informação, produzindo cubos, OLAP's e relatórios, surge, no mesmo contexto, o conceito de Gerência do Conhecimento, que trabalha o ativo de informações, independentemente da sua forma, estrutura e domínio, incumbindo-se de realizar as devidas combinações, compilações, subscrições e distribuição, transformando informações em conhecimento. Nesse cenário, entram informações estruturais, documentos de fontes díspares, fatos,

opiniões e conhecimentos empíricos não formalmente documentados [Barbieri, 2001].

Sun Tzu, em seu livro *A Arte da Guerra*, menciona que se deve “deter todo o conhecimento de suas fraquezas e virtudes, além de todo o conhecimento das fraquezas e virtudes do inimigo. A falta deste conhecimento pode resultar na derrota”. Com essa informação, alinhada ao conceito foco trabalhado, percebe-se a real importância da junção do trabalho realizado pelos ativos tangíveis e intangíveis de uma organização, pois, com eles, pode-se gerar o conhecimento total necessário para enfraquecer o “inimigo” ao mesmo tempo em que se fortalece a “equipe aliada”. Quando mencionado o termo “ativo tangível” deve-se compreender tudo aquilo que está a favor da organização, mas que pode ser alterado ou destruído a qualquer momento, como códigos de sistemas, por exemplo. Já o termo “ativo intangível” refere-se ao capital intelectual existente na organização, o conhecimento das pessoas, aquilo que não pode ser destruído e que, sem ele, não é possível o desenvolvimento lógico estratégico organizacional.

A figura 4 ilustra, de acordo com Barbieri (2001), de forma resumida, uma visão geral das relações entre BI e Gerência do Conhecimento, demonstrando a relação existente entre os ativos organizacionais:

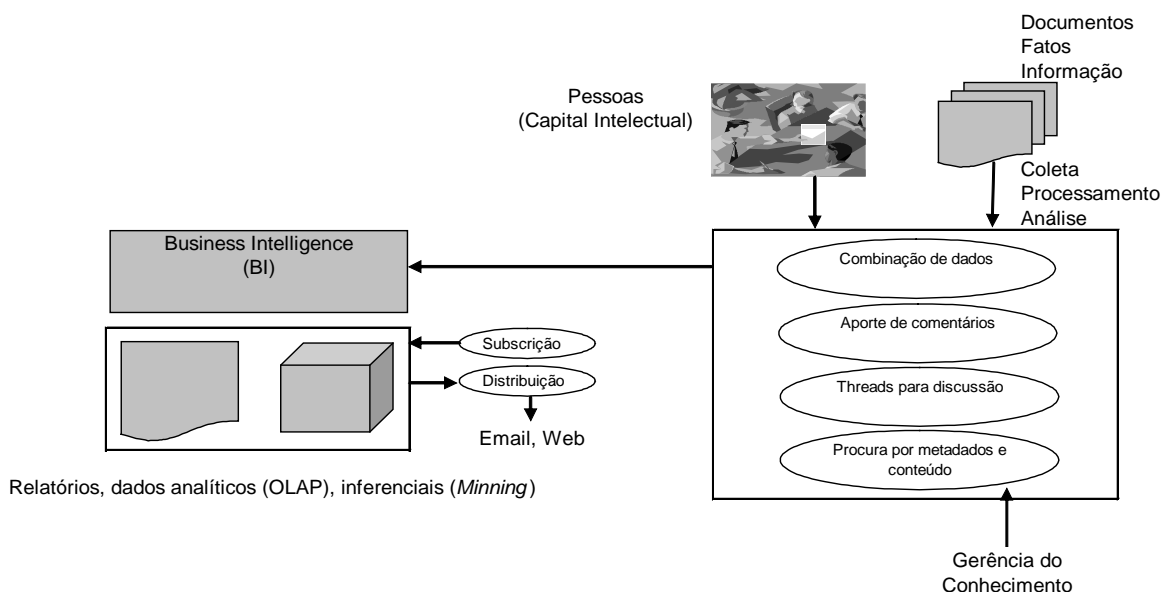


Figura 4 - Visão geral das relações entre BI e Gerência do Conhecimento [Barbieri, 2001]

O BI apóia-se em modelos de dados multidimensionais para modelar a informação em fatos de negócio e transportá-los para *data warehouses* ou *data marts*. Com as informações moldadas de acordo com as necessidades apresentadas, são fornecidas aos responsáveis pela tomada de decisão, ferramentas para análises das várias dimensões presentes nos fatos de negócio que proporcionam um novo guia de segmento para definição de estratégias. Além disso, os conhecimentos de BI levam a extração de conhecimento detalhada e aprofundada (*Data Mining*), possibilitando previsões de acontecimentos baseados em possibilidades previamente definidas.

Em entrevista a revista *Computer World*, no ano de 2006, Howard Dresner, pai da expressão *Business Intelligence*, menciona que BI está no meio de dois extremos, sendo a informação estruturada de um lado e o usuário do outro. Para que a disseminação da informação ocorra de forma correta, pontual e focada, é necessário o comprometimento daqueles que estão envolvidos com o negócio. Howard complementa em tal entrevista que não é a tecnologia que prende a adoção do BI e, sim, a cultura da organização e mostra, ainda, alguns números relevantes sobre o tema, conforme figuras 5 e 6:

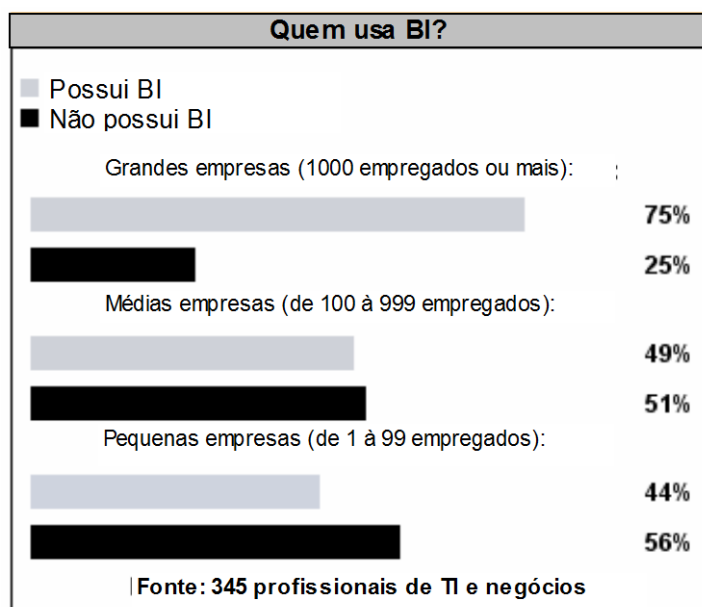


Figura 5 - Quem usa BI? [Revista Computer World, 2006]

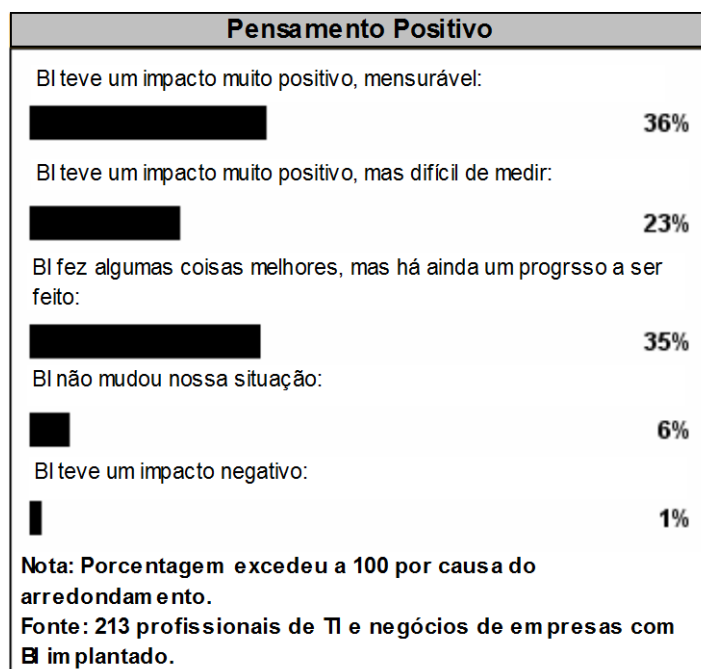


Figura 6 - Pensamento Positivo - BI [Revista Computer World, 2006]

Com os números apresentados acima, pode-se identificar que um percentual elevado de empresas, já no ano de 2006, passou a dar maior atenção às soluções de BI, onde, de acordo com a proporção, um número muito pequeno de insatisfação ou indiferença foi apresentado.

De acordo com a empresa SOL7 – BI, especialista em soluções de *Business Intelligence* baseadas em *software* livre, as características que merecem destaque dentro do tema podem ser enumeradas da seguinte forma:

- Extrai e integra dados de múltiplas fontes;
- Faz uso da experiência (dados históricos);
- Analisa dados contextualizados;
- Trabalha com hipóteses;
- Procura relações de causa e efeito;
- Transforma os registros obtidos em informação útil para o conhecimento empresarial; e
- Seus conceitos são aplicáveis a todo tipo de empresa, independente de seu porte, faturamento ou segmento.

Seguindo os princípios de outra empresa, a Microsoft, que aponta que as soluções de *Business Intelligence* se baseiam no desempenho aprimorado nos negócios através da tomada de decisões aprimorada em toda a organização e que quando se sabe que a percepção dos dados corporativos é sólida, bem-informada e completa, pode-se confiar em cada decisão que toma, com um nível de confiança ao alcance, aprimorando o desempenho dos negócios, criando vantagens competitivas e alcançando objetivos corporativos, consegue-se entender os três passos fundamentais considerados por esta empresa nas soluções para tomada de decisões: qualidade de dados, ferramentas adequadas e decisões alinhadas com os objetivos corporativos. A qualidade dos dados retrata a integridade e veracidade de tudo aquilo que será repassado como informação aos responsáveis pela tomada de decisão. As ferramentas a serem utilizadas deverão estar de acordo com a necessidade apresentada pela instituição, proporcionando facilidade em sua manipulação e geração de resultados. Já as decisões, por estarem alinhadas aos objetivos estratégicos, deverão ser embasadas por dados históricos contidos e repassados pela ferramenta de BI mais adequada.

Pode-se entender o funcionamento de um sistema de BI de acordo com a seguinte figura:



Figura 7 - Funcionamento BI [SOL7 - BI]

De acordo com a figura 7, devem-se entender as seguintes representações:

- Base de Dados: base de dados transacional;
- ETL: extração, transformação, carga e sincronização;
- *Data Warehouse*: os dados são limpos, consistidos, unificados, sumarizados e organizados em uma base multidimensional;
- OLAP: facilita o acesso e a extração das informações, que poderão ser apresentadas em planilhas, gráficos ou em relatórios.

As Bases de Dados do BI são de extrema utilidade na formatação de estruturas capazes de serem implantadas e entendidas pelos gerenciadores de bancos de dados. A abordagem tradicional de dados sempre primou pela representação de estruturas que melhor se ajustassem às características transacionais dos processamentos. Com o desenvolvimento de outras necessidades, alavancadas por aspectos de competitividade e busca de diferenciais de negócios e conseqüente tomada de decisão, esses modelos se mostraram inadequados, pois suas características de pulverização de informações por estruturas diferentes, motivadas pelos rigores das regras de normalização de dados se mostraram imperfeitas para os processamentos demandados pela ótica dimensional. Com isso, surgiu a modelagem dimensional de dados, que permitiu a formatação estrutural mais voltada para os muitos pontos de entradas específicos (dimensões) e menos para os dados granulados em si (fatos). As estruturas modeladas dimensionalmente, armazenadas em *Data Warehouse* ou *Data Marts* e interpretadas pela ótica analítica das ferramentas de OLAP ou pelo prisma inferencial das ferramentas de *Data Mining*, atinge o objetivo proposto pelas premissas de BI. [Barbieri, 2001]

Importantes conceitos surgiram com o desenvolvimento do processo de sistemas de apoio à decisão, dentre eles o de *Data Warehouse*, cuja tradução literal significa Armazém de Dados, mas pode ser definido como um banco de dados, destinado a sistemas de apoio à decisão e cujos dados foram armazenados em estruturas lógicas dimensionais, possibilitando o seu processamento analítico por ferramentas especiais, sendo estas OLAP e *Data Mining*. [Barbieri, 2001]

5.2.2 - *Data Warehouse*

Atualmente, devido à forte concorrência no mercado, um dos fatores primordiais, já mencionado neste projeto, a ser levado em consideração para a tomada de decisão, é o tempo levado para identificar, analisar e colocar em prática um plano de ação baseado em informações estratégicas. O *Data Warehouse* (DW) surge neste contexto, visando a agilidade do conhecimento de informações

confiáveis, com foco na análise estratégica e agilidade nas ações a serem desenvolvidas em benefício da Instituição. Pode-se dizer, ainda, que o DW surgiu “objetivando integrar dados de múltiplas fontes, um processo de análise com informação de qualidade sem impacto para o ambiente operacional e um atendimento a diferentes tipos de usuários com agilidade e flexibilidade.” [Colaço, 2004]

Possuir um *Data Warehouse* significa ter um enorme banco de dados com informações organizadas. Tais informações representam o histórico realizado de dados, podendo este ser consultado para ajuda no processo de tomada de decisão.

De acordo com Barbieri (2001), todo e qualquer DW apresenta três objetivos específicos em comum, sendo eles:

- Uma base de dados preparada em vários níveis de granularidade (nível de detalhe existente em um DW) e obtida a partir dos sistemas estruturados;
- Armazenamento de dados em vários graus de relacionamento e sumariação, de forma a facilitar e agilizar os processos de tomada de decisão por diferentes níveis gerenciais; e
- Os dados, oriundos dos sistemas de informação de produção, deverão estar “mastigados”, integrados e disponíveis, permitindo diversas formas de consultas, através dos mecanismos amistosos das ferramentas de usuários.

Segundo Bill Inmon, considerado o pai do conceito de *Data Warehouse* e autor de vários livros a respeito do assunto, um DW é “um lugar onde a informação é o sujeito da ação (não o predicado), é integrada, é dependente do momento em que foi criada (histórico) e não é volátil, isto é, mais dados podem ser adicionados, mas os dados antigos não mudam.” [Revista Fator Brasil, 2007]

Analisando as características enunciadas por Inmon e, baseando-se nas idéias apresentadas por Colaço (2004), cada item pode ser especificado da seguinte forma:

- Orientado por temas: o DW armazena informações necessárias para o processo de suporte à decisão. Essas informações são organizadas pelos temas importantes para o negócio da Instituição;
- Informação Integrada: após a arquitetura ter sido criada e a infraestrutura ter sido construída adequadamente, o DW deve consolidar dados de diversas origens e, tais dados, devem ser perfeitamente integrados para que, ao serem armazenados, assumam uma única convenção, ou seja, uma padronização de representação única deve ser criada para os dados de todos os sistemas que formarão a base de dados do DW;
- Histórico: projetado para recordar informações, em um DW os dados são carregados como fotos da base de dados operacional do momento, ou seja, cada ocorrência e cada mudança são consideradas como um novo registro. Os dados não são atualizados e podem ser comparados ao longo do tempo;
- Não volátil: depois que os dados estão no DW não poderão ser atualizados ou alterados, apenas acessados. Os novos dados serão absorvidos, integrando-se aos dados existentes. O DW permite apenas a carga inicial dos dados e a consulta aos mesmos.

De acordo com Barbieri (2001), existem alguns fatores que devem ser considerados críticos em projetos de Data Warehouse, sendo eles: foco bem definido, patrocinador forte, dados necessários, envolvimento dos usuários, boa equipe de projeto, arquitetura tecnológica, marketing e acompanhamento. Todos os fatores mencionados são merecedores de total atenção dentro das Instituições, pois podem influenciar diretamente o resultado final daquilo que se almeja com a utilização do DW.

É fundamental em todas as atividades envolvidas com DW que, na sua implantação, seja estabelecido e bem definido um foco, pois, caso contrário, informações desnecessárias serão geradas e repassadas, causando desperdício de tempo e acúmulo de dados. Como Barbieri (2001) menciona em seu livro, “a falta de objetivo ou de um foco melhor definido sobre os produtos a serem entregues são

fatores primordiais para o insucesso de qualquer projeto, e não seria diferente nos de *Data Warehouse*.”

Para que a implantação do DW ocorra de forma completa, é extremamente necessário que um dos usuários das informações fornecidas por ele seja alguém que participe do nível estratégico da Instituição. Participando como um patrocinador da idéia, este usuário, presente em todos os eventos que envolvam discussões a respeito do assunto, influenciará e contará com o apoio dos demais níveis (tático e operacional), fazendo com que o projeto se fortaleça e aumente sua credibilidade.

Todos os dados necessários às análises e manipulação para geração de informações devem, previamente, existir na base de dados da Instituição. “O DW não produz dados, a menos daqueles obtidos *in-flight* por cálculos. É fundamental que o mapeamento das fontes de dados seja feito, com rigoroso critério de qualidade, certificando-se da natureza dos dados, sua periodicidade de atualização, sua qualidade atual, seus sistemas mantenedores e a perspectiva de duração daquela fonte de dados.” [Barbieri, 2001]. Os riscos que serão assumidos com o projeto só deverão ser compromissados após tais verificações.

O comprometimento de todos os usuários envolvidos também é apontado como um dos fatores críticos de sucesso, por estar diretamente ligado ao seu bom desempenho. De nada adianta possuir excelentes ferramentas se seus usuários não compram a idéia e não acreditam nos benefícios oferecidos por ela. A motivação deve vislumbrar os alcances do projeto.

A equipe desenvolvedora do projeto deve ser “coesa, motivada, fortemente associada ao projeto, com condições para realização de suas várias etapas.” [Barbieri, 2001]. Além disso, como menciona Barbieri (2001), a equipe tecnológica deve possuir profissionais devidamente capacitados para solucionar problemas relacionados a Banco de Dados, comandos SQL agigantados, conexões entre máquinas de protocolos diferentes, problemas de tempo de resposta, entre outros. Pelo lado funcional, a equipe deve possuir analistas com grande conhecimento dos aplicativos que darão origem aos dados do DW e conhecimento em relação ao negócio abordado.

Uma boa arquitetura tecnológica deve ser definida. “De nada adiantará um belo modelo dimensional, cheio de tabelas Fato e tabelas Dimensão, bem

balanceadas, se o ferramental de suporte físico não condisser com as demandas do projeto.” [Barbieri, 2001]. Uma plataforma adequada deverá ser cuidadosamente analisada.

Feitas todas as análises e gerados os primeiros resultados, a divulgação do que foi obtido com o projeto (através de disponibilização na rede interna, fixação de gráficos em murais, apresentações às equipes separadamente, dentre outros métodos) deve ser realizada imediatamente. O marketing em relação aos pontos favoráveis apresentados pelo projeto deve ser realizado para ressaltar sua importância.

Como todo e qualquer projeto desenvolvido, o acompanhamento da utilização do DW deve ser feito e os possíveis motivos de declínio na sua utilização analisados. Caso necessário, alterações devem ser feitas.

O funcionamento de um *Data Warehouse* pode ser melhor analisado quando especifica-se as camadas que o alimentarão e o que pode ser produzido por ele. A arquitetura básica pode ser representada de acordo com a figura 8:

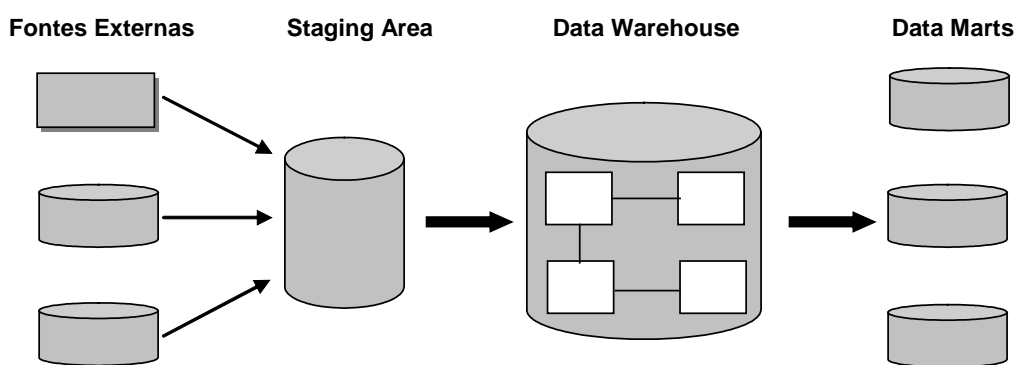


Figura 8 - Camadas da arquitetura de um ambiente de DW (www.sqlmagazine.com.br)

De acordo com a figura acima, pode-se extrair as seguintes informações a respeito de cada camada:

- Fontes Externas: são todos os dados armazenados em sistemas transacionais e fornecidos pela Instituição que deverão ser analisados em busca de um resultado final;

- *Staging Area* (Área de Estágio): é o local intermediário entre as fontes externas e o DW. Nesta área os dados sofrem integração, limpeza e depois são exportados para o DW;
- *Data Warehouse*: é onde ocorre a reunião dos dados que foram extraídos, transformados e carregados das fontes externas;
- *Data Marts*: são camadas que permitem o acesso a um tipo específico de dado.

Baseando-se em informações contidas no site da DW Brasil, os bancos de dados transacionais armazenam as informações das transações diárias da Instituição e são utilizados por todos os funcionários para registrar e executar operações pré-definidas, por isso seus dados podem sofrer constantes mudanças. Por não ocorrer redundância nos dados e as informações históricas não ficarem armazenadas por muito tempo, este tipo de banco de dados não exige grande capacidade de armazenamento. Já um DW armazena dados analíticos, destinados às necessidades da gerência no processo de tomada de decisões. Isto pode envolver consultas complexas que necessitam acessar um grande número de registros, por isso é importante a existência de muitos índices criados para acessar as informações da maneira mais rápida possível. Como um DW armazena informações históricas de muitos anos, deve ter uma grande capacidade de processamento e armazenamento dos dados que se encontram de duas maneiras: detalhados e resumidos.

Com a visualização da arquitetura de um DW, observa-se que um novo conceito surge, no intuito de proporcionar ao usuário uma abordagem centralizada de dados, que poderão ser consultados através de uma Ferramenta de Apoio à Decisão. Tal conceito reflete o objetivo de um *Data Mart*.

5.2.3 - Data Mart

Um *Data Mart* é um subconjunto lógico de um DW, um DW setorial [Colaço, 2004]. Ainda segundo Colaço, os *Data Marts* são muito úteis nas seguintes condições:

- Os dados devem estar segregados para melhorar o desempenho do sistema do ponto de vista do usuário;
- Deve existir uma cópia dos dados onde só pessoas com autorização devem ter o privilégio de acessá-las; e
- Em um ambiente corporativo, é importante fortalecer o conceito de propriedade dentro do banco de dados. Diferentes setores serão responsáveis por diferentes *Data Marts*.

De acordo com Barbieri (2001), tanto o *Data Warehouse* quanto o *Data Mart* se referem às estruturas dimensionais de dados com o objetivo de prover análises diferenciais. Ambos podem ser definidos como espécies do mesmo tipo, ficando a diferença entre os dois centrada no escopo do projeto e nos limites de suas abrangências.

Bispo (1998) apresenta uma tabela comparativa entre *Data Mart* e *Data Warehouse* - sugerida por Strange (1998) - onde o principal objetivo é a justificativa da implantação de um *Data Mart* no lugar de um *Data Warehouse*, pois o primeiro apresenta resultados mais rápidos, uma vez que demora entre quatro e doze meses para ser implementado, enquanto o segundo tem prazo de implementação completa variante entre um à cinco anos:

Tabela 2 - Tabela comparativa entre Data Mart e Data Warehouse

Propriedades	Data Warehouse	Data Mart
Âmbito	aplicações neutras	aplicações específicas
	centralizado e compartilhado	departamento ou área de usuário
Assuntos	múltiplas áreas	área única
Fontes de dados	muitas	poucas
Tempo de implementação	<ul style="list-style-type: none"> • 9 a 18 meses para 1º estágio (duas ou três áreas de assuntos); • múltiplos estágios de implementação; 	4 a 12 meses
Características	flexível	restritivo
	nível estratégico	nível tático

Em um ambiente onde, a cada dia, a busca por facilidades na forma de encontrar informações para auxílio na tomada de decisões é fator primordial, a organização no armazenamento dos dados específicos passa a ser de extrema importância e, para isso, “o termo *Data Mart* (Mercado de Dados) significa, nesse cenário, depósito de dados que atende a certas áreas específicas da empresa e voltados (também) para o processo decisório gerencial” [Barbieri, 2001].

Atualmente, é necessário que se perca menos tempo com o acesso e a análise dos dados e o foco seja maior na elaboração de estratégias de negócios, baseando-se em informações e fatos analisados e comprovados. Para isso, segundo Bispo (1998), o *Data Warehouse* sozinho não pode realizar todas as consultas e análises que seus usuários necessitam. Neste contexto, surgem os conceitos OLAP e OLTP, que são ferramentas que proporcionam análises mais sofisticadas nos dados.

5.2.4 – OLTP

OLTP é caracterizado por registrar todas as transações contidas em uma determinada operação Institucional, agindo em execuções e tarefas do dia a dia, otimizando os bancos de dados. Essa ferramenta opera com dados que movimentam o negócio em tempo real.

O OLTP é uma ferramenta relacional, orientada para o processo, trabalhando com dados do presente e processando um registro de cada vez. Tem como principal função alimentar a base de dados que compõe o OLAP. [SISNEMA 2004].

De acordo com Alcântara (2006), essas ferramentas são excelentes para administrar o dia a dia das empresas, mas pecam quando o objetivo é o planejamento estratégico. Os relatórios que os sistemas OLTP fornecem são restritos a uma visão bidimensional do negócio, que não possibilita aos tomadores de decisão a flexibilidade que necessitam na análise da Instituição como um todo. A figura 9 esboça um exemplo de relatório bidimensional fornecido por uma ferramenta OLTP:

Produto	Região	Data	Valor
P1	AM	02/03/2003	250,00
P2	AM	16/12/2005	1250,00
P3	BA	20/07/2004	130,00
P4	BA	10/01/2003	960,40
P5	DF	07/05/2005	7950,00
P2	ES	06/06/2006	24,00
P3	SP	23/07/2004	1204,00
P2	SP	16/12/2004	532,00
P4	DF	02/03/2006	4500,50

Figura 9 - Relatório Bidimensional (www.microsoft.com)

Segundo Colaço (2004), os bancos de dados desenvolvidos para os Processamentos de Transação *On-Line* geralmente são considerados inapropriados para *Data Warehouses*. Eles não podem ser repositórios de fatos e dados históricos, não atendem satisfatoriamente a consultas e a recuperação rápida dos dados é praticamente impossível. Tal fato se justifica pelos usuários de OLTP terem de lidar com um registro de cada vez, o que faz com que a mesma tarefa seja executada inúmeras vezes.

Ao se utilizar ferramentas OLTP é necessário atentar-se a dois fatores de primordial importância na implantação de tecnologias diferenciadas: segurança e custo. A disponibilidade plena das informações que essa ferramenta propicia deixa os dados disponíveis a serem acessados por pessoas mal intencionadas. Além disso, a menor falha na ferramenta pode causar uma série de problemas, causando perda de tempo e dinheiro.

5.2.5 – OLAP

OLAP teve seu conceito criado em 1993 e é caracterizado por um conjunto de técnicas utilizadas para tratar informações contidas em um *Data Warehouse* [Colaço, 2004]. O Processamento Analítico *On-Line* apresenta, ainda, segundo

Barbieri (2001), a característica de trabalhar os dados, com operações dimensionais, possibilitando uma forma múltipla e combinada de análise.

O termo OLAP surgiu como antítese que complementa o termo OLTP. Essas ferramentas são implantadas sobre modelos multidimensionais, o que permite análises complexas e consultas com rápido tempo de resposta. [DW Brasil, 2003].

O processamento OLAP provê acesso aos dados corporativos de um *Data Warehouse* com total segurança e controle e provê aos usuários todas as flexibilidades existentes em programas dedicados à análise de dados. [Colaço, 2004].

Por ter uma função de análise, os dados de um sistema OLAP são provenientes de um ou mais sistemas OLTP e arquivos (planilhas, e-mails, XML, dentre outras fontes) em um processo conhecido como Carga Incremental. Esse processo normalmente é agendado para ocorrer de tempos em tempos e esse tempo é determinado pela quantidade de dados já existentes. [Alcântara, 2006].

A ferramenta OLAP permite aos usuários o cruzamento de informações diversas na busca resultados capazes de alterar o curso das estratégias definidas. Por isso, a utilização de visões multidimensionais, conhecidas como Cubos, é de fundamental importância no processo de análise de dados. A figura 10 ilustra claramente a enorme quantidade de informações retiradas de um relatório multidimensional:

		Year ▾ Quarter Month										
		1996		1997		1998						
		Quarter 1		Quarter 2		Quarter 1		Quarter 2				
Category Name ▾	Product Name	Ship Country ▾	Ship Region	Ship City	Unit Price	Quantity	Unit Price	Quantity	Unit Price	Quantity	Unit Price	Quantity
☑ Beverages					2033,8	1842	5058,35	3996	3622	2381	1097,7	1097
☑ Condiments					705,6	962	2211,85	2895	1037,4	936	650,4	650,4
☑ Confections					1333,4	1357	3547,63	4137	2059,63	1795	608,6	608,6
☑ Dairy Products					1857,6	2086	5070	4374	1632,1	1649	1316,1	1316,1
☑ Grains/Cereals	☑ Filo Mix				16,8	48	102,2	313	49	105	2,2	2,2
	☑ Gnocchi di nonna Alice				212,8	96	1178	971	304	171	7,7	7,7
	☑ Gustafs Knäckebröd				16,8	6	184,8	209	63	108	2,2	2,2
	☑ Ravioli Angelo				93,6	133	148,2	124	136,5	129	3,3	3,3
	☑ Singaporean Hokkien Fried Mee				54,6	37	215,6	451	56	97	7,7	7,7
	☑ Tunnbröd				21,6	105	82,8	287	45	176	1,1	1,1
	☑ Wimmers gute Semmelknödel				106,4	124	425,6	281	266	158	13,3	13,3
	Total				522,6	549	2337,2	2636	919,5	944	38,3	38,3
☑ Meat/Poultry					1294,3	950	3795,48	2189	1473,09	527	854,4	854,4
☑ Produce	☑ Longlife Tofu				40	141	54	116	10	20	1,1	1,1
	☑ Manjimup Dried Apples	☑ Argentina					53	7				
		☑ Austria					53	120				
		☑ Brazil			84,8	37						
			☑ RJ									
			☑ SP	Campinas			53	28				
				Total			53	28				
			Total		84,8	37	53	28				
		☑ Canada										5
		☑ Denmark					53	30				
		☑ France			42,4	20	106	35	106	7		
		☑ Germany			49,4	40	59	48	59	9		
		☑ Ireland										

Figura 10 - Relatório Multidimensional (www.microsoft.com)

Além de oferecer uma visão multidimensional dos dados, permitindo o cruzamento de todos os tipos de informações, as ferramentas OLAP apresentam outras características, definidas pelo criador do conceito OLAP, Dr. Codd (1993), e referenciadas por Colaço (2004), sendo elas:

- **Transparência:** OLAP deve atender a todas as solicitações do analista, não importando de onde os dados virão. Todas as implicações devem ser transparentes para os usuários finais;
- **Acessibilidade:** as ferramentas OLAP devem permitir conexão com todas as bases de dados legadas. A distribuição de informações deve ser mapeada para permitir o acesso a qualquer base;
- **Desempenho de informações consistentes:** as ferramentas OLAP devem possuir conhecimento sobre todas as informações armazenadas para que possa disponibilizar sem complexidade para o usuário final, qualquer tipo de consulta;
- **Arquitetura cliente-servidor:** OLAP deve ser construída em arquitetura cliente-servidor para que possa atender a qualquer usuário em qualquer ambiente operacional;
- **Dimensionalidade genérica:** deve ser capaz de tratar informações em qualquer quantidade de dimensões;
- **Manipulação de dados dinâmicos:** devido ao grande volume de informações armazenadas nas diversas dimensões de um modelo multidimensional, é comum a anulação dos dados, e então essas células nulas devem ser tratadas para evitar custos com memória;
- **Suporte a multiusuários:** nas grandes organizações é comum vários analistas trabalharem com a mesma massa de dados;
- **Operações ilimitadas em dimensões cruzadas:** as ferramentas OLAP devem ser capazes de navegar nas diversas dimensões existentes;
- **Manipulação intuitiva dos dados:** o usuário deverá ser capaz de manipular os dados livremente, sem necessitar de qualquer tipo de ajuda;

- Flexibilidade nas consultas: o usuário deverá ter a flexibilidade para efetuar qualquer tipo de consulta; e
- Níveis de dimensão e agregação ilimitados: devido às várias dimensões existentes, deve haver vários níveis de agregação dos dados.

Os dados de uma ferramenta OLAP se baseiam nas informações contidas nos diversos *softwares* utilizados pela Instituição. Por estes *softwares* poderem apresentar a mesma informação apresentada de maneiras diferentes, a OLAP possui uma camada onde estão os dados originais e outra camada onde estão os dados tratados, sendo esta o *Data Warehouse*. Para extrair os dados das fontes de dados e armazená-los no *Data Warehouse*, é necessária uma etapa de extração e transformação dos dados. Normalmente esta tarefa é realizada por um *script* ou programa feito especificamente para o cliente e que é agendado para executar de tempos em tempos. Para que o processo de migração dos dados da fonte de dados para o *Data Warehouse* seja simplificado, é recomendado que se utilize *Data Marts*, agrupando informações em grupos específicos. As ferramentas OLAP consultam os *Data Marts* e exibem os dados para os tomadores de decisão da Instituição. A figura a seguir demonstra esse processo:



Figura 11 - Arquitetura de um sistema OLAP (www.microsoft.com)

Existem diversos tipos de ferramentas disponíveis no mercado e, por incrível que pareça, a mais popular dentre elas é o Microsoft Excel, que possui um componente feito especialmente para este fim, a *Pivot Table*. [Alcântara, 2006]. A *Pivot Table*, ou Tabela Dinâmica, permite a manipulação e cruzamento de informações, gerando resultados de acordo com a necessidade e solicitação dos tomadores de decisão.

Dentro da estrutura funcional de uma ferramenta OLAP, é oferecido um suporte para funções de derivação de dados complexos que recebe o nome de *Slice* (significa analisar determinada fatia do cubo OLAP) *and Dice* (significa alterar a visão de um cubo OLAP). Segundo Colaço (2004) “o suporte *Slice and Dice* é uma das principais características de ma ferramenta OLAP. Ela serve para modificar a ordem das dimensões, alterar linhas por colunas de maneira a facilitar a compreensão dos usuários”. O *Slice and Dice* compreende as seguintes operações:

- *Ranging*: é a operação responsável por alterar o resultado das consultas, inserindo novas posições ou removendo as que estão em foco;
- *Drill Down*: significa descer um nível hierárquico em uma dimensão;
- *Drill Up*: significa subir um nível hierárquico em uma dimensão;
- *Drill Across*: significa analisar um nível intermediário dentro de uma mesma dimensão;
- *Drill Throught*: significa alternar a análise de uma dimensão para outra;
- *Rotation*: altera a forma de visualização das informações; e
- *Ranking*: filtra as informações que se deseja ver.

Existem, ainda, quatro categorias de ferramentas OLAP, sendo elas, de acordo com informações retiradas do artigo de Ivã Cielo, publicado no site www.datawarehouse.inf.br:

- *ROLAP (Relational On-Line Analytical Processing)*: possuem uma engenharia de acesso aos dados e análise OLAP com uma arquitetura um pouco diferente. A consulta é enviada ao servidor de banco de dados relacional e processada no mesmo, mantendo o cubo no servidor. O que nota-se nesse caso é que o processamento OLAP se

dará somente no servidor. A principal vantagem dessa arquitetura é que ela permite analisar enormes volumes de dados, em contrapartida uma grande quantidade de usuários acessando simultaneamente poderá causar sérios problemas de desempenho no servidor;

- **MOLAP** (*Multidimensional On-Line Analytical Processing*): com um servidor multidimensional o acesso aos dados ocorre diretamente no banco, ou seja, o usuário trabalha, monta e manipula os dados do cubo diretamente no servidor. Isso traz grandes benefícios aos usuários no que diz respeito a desempenho, mas tem problemas com escalabilidade, além de ter um custo alto para aquisição.
- **HOLAP** (*Hybrid On-Line Analytical Processing*): nada mais é do que uma mistura de tecnologias onde há uma combinação entre ROLAP e MOLAP. A vantagem é que com a mistura de tecnologias pode-se extrair o que há de melhor de cada uma, ou seja, a alta performance do MOLAP com a escalabilidade melhor do ROLAP.
- **DOLAP** (*Desktop On-Line Analytical Processing*): o ganho com essa arquitetura é o pouco tráfego que se dá na rede, visto que todo o processamento OLAP acontece na máquina cliente, e a maior agilidade de análise, além do servidor de banco de dados não ficar sobrecarregado, sem incorrer em problemas de escalabilidade. A desvantagem é que o tamanho do micro-cubo não pode ser muito grande, caso contrário a análise passa a ser demorada e a máquina do cliente pode não suportar em função de sua configuração.

As opções de estratégias de armazenamento de informações e implementação de estruturas dimensionais no DW, definidas acima, podem ser representadas da seguinte forma:

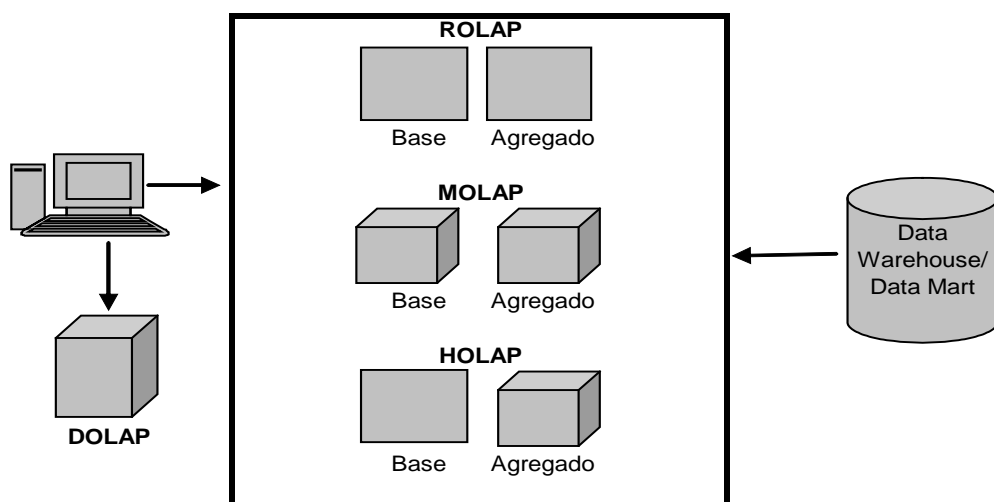


Figura 12 - Opções de armazenamento/ implementação de estruturas dimensionais (Barbieri, 2001)

5.2.6 – OLAP x OLTP

As principais diferenças entre os processamentos OLAP e OLTP podem ser visualizadas na tabela a seguir:

Tabela 3 - Diferenças entre OLAP e OLTP

OLAP	OLTP
Relevância para dados históricos	Mantém usualmente a situação corrente
Necessidade de ver o dado sob diferentes perspectivas: aplicações dinâmicas	Voltado para velocidade e automação de funções repetitivas
Atualizações quase inexistentes, apenas novas inserções	Atualizações em grande número
Baseado em dados históricos, consolidados e freqüentemente totalizados	Baseado em transações
Operações de agregação e cruzamentos	Alto nível de detalhe

5.2.7 – *Data Mining*

O processo de *Data Mining*, ou Mineração de Dados, é referenciado por Colaço (2004) como responsável por automatizar a detecção de padrões relevantes em um banco de dados. Segundo ele, ainda, o diferencial da utilização das técnicas de *Data Mining* está na extração de informações úteis, valiosas e previamente definidas. Na maioria das técnicas de exploração dos dados, os resultados apresentados são dados já conhecidos.

O processo de *Data Mining* tem certas diferenças em relação a outras ferramentas de BI. Segundo Barbieri (2001), “enquanto as técnicas OLAP objetivam trabalhar os dados existentes, buscando consolidações em vários níveis, trabalhando fatos em dimensões variadas, a técnica de *Mining* busca algo mais que a interpretação dos dados existentes. Visa fundamentalmente realizar inferências, tentando como que “adivinhar” possíveis fatos e correlações não explicitadas nas montanhas de dados de um DW/ DM.”.

O *Data Mining* utiliza técnicas de estatística e inteligência artificial para construir modelos que predizem o comportamento dos envolvidos em uma determinada análise. Podem-se agrupar suas principais técnicas, de acordo com informações disponibilizadas por Coutinho (2003), através da DW Brasil, nos seguintes grupos:

- Redes Neurais: oferece o mais profundo poder de mineração. Tentam construir representações internas de modelos ou padrões achados nos dados, mas essas representações não são apresentadas para o usuário. Estruturalmente, uma rede neural consiste em um número de elementos interconectados e organizados em camadas que aprendem pela modificação da conexão firmemente conectando as camadas. Embora apresentem o mais avançado poder de mineração, muitos não podem fazer uso delas porque os resultados finais não podem ser explicados;
- Indução de regras: é à detecção de tendências dentro de grupos ou regras de dados. As regras são apresentadas ao usuário como uma

lista “não encomendada”. Vários algoritmos e índices são colocados para executar esse processo;

- Árvores de decisão: são meios de representar resultados de *Data Mining* na forma de árvore, e que lembram um gráfico organizacional horizontal;
- Análises se séries temporais: incorpora um envolvimento muito forte do usuário, exigindo engenheiros experientes, para construir modelos que descrevem o comportamento do dado através dos métodos clássicos de matemática; e
- Visualização: mapeia o dado sendo minerado de acordo com dimensões especificadas. Nenhuma análise é executada pelo programa de *Data Mining* além de manipulação estatística básica.

Os passos básicos a serem seguidos para realização de um processo *Data Mining*, são descritos por Navega (2002) em seu artigo, sendo representado de acordo com a figura 13, onde a partir de fontes de dados (*DataBases*) efetua-se uma limpeza e os dados são armazenados em *Data Warehouses* ou *Data Marts*. A partir desses repositórios é possível selecionar informações para passarem pelo processo de mineração. Um analista responsável pelas análises do negócio deve, então, refinar e conduzir o processo *Mining*, gerando visualizações de resultados relevantes a condução de ações.

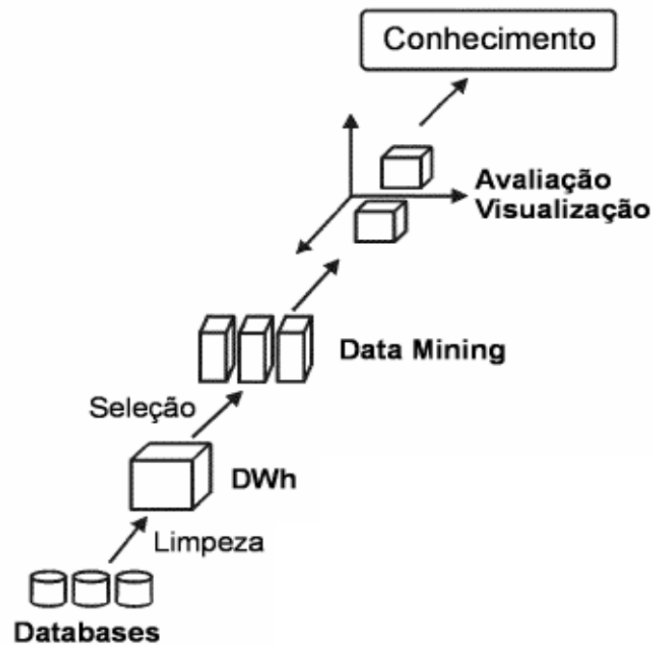


Figura 13 - Passos do Data Mining (Navega, 2002)

Para revelar estruturas de conhecimento coesas que geram conhecimento e que possam guiar decisões em condições de certeza, baseada em resultados obtidos através do cruzamento de dados armazenados, o conceito de *Data Mining* está se tornando cada vez mais popular como uma ferramenta de gerenciamento de informação, sendo um grande auxílio no processo de tomada de decisão.

CAPÍTULO 6 – ESTUDO DE CASO

Este trabalho foi idealizado e construído no âmbito da Secretaria Acadêmica da Universidade Católica de Brasília, utilizando-se do Sistema Integrado de Informações da Educação Superior – SIEdSup, que é uma ferramenta *on-line* disponibilizadas pelo Ministério da Educação e Cultura - MEC, bem como a consultas ao Sistema Acadêmico da Universidade Católica de Brasília, denominado *Stratus*, módulo Graduação.

Primeiramente foi feita uma pesquisa bibliográfica sobre os temas chaves para a evolução do trabalho, onde se realizou um encadeamento das informações relativas ao censo do ensino superior, com ênfase no corpo discente.

O Ensino Privado Superior Brasileiro está em momento de expansão acelerada. A Universidade Católica de Brasília, observando este crescimento, tem priorizado sua qualidade de ensino. Acredita-se que, com a validação deste estudo de caso, será possível galgar uma nova visão de mercado para IES, que, utilizando a manipulação de dados enviados ao Ministério da Educação e Cultura, poderá gerar análises, relatórios e monitoramentos em busca de soluções estratégicas para a alta gerência da Universidade Católica de Brasília, com o objetivo de fidelizar os discentes da Instituição.

6.1 – Sistema Existente

A Universidade Católica de Brasília, focada em seu sistema de controle acadêmico, não dispõe de uma ferramenta de apoio à decisão capaz de abstrair e tratar as informações enviadas anualmente ao Ministério da Educação e Cultura. No módulo Graduação, do sistema *Stratus*, são extraídos os dados acadêmicos, para os lançamentos necessários ao censo da educação superior, requeridos anualmente.

O Ministério de Educação e Cultura habilita para as Instituições de Ensino Superior do país o Sistema Integrado de Informações da Educação Superior –

SIEDSUP, que é uma ferramenta *on-line*, receptora dos dados censitários das IES do país. Como gestor e proprietário da ferramenta, o Ministério da Educação e Cultura manipula esta base de dados, obtendo a visão macro da situação acadêmica no país, para aplicação de políticas públicas no ensino superior.

A seguir são apresentadas as principais informações requeridas para a base de dados do censo do ensino superior, relativo aos cursos de graduação do país:

- Número de vagas por processos seletivos e turno;
- Número de candidatos por processo seletivo, turno e sexo;
- Número de ingressantes por curso, turno e sexo;
- Número de ingressantes por sexo e faixa etária;
- Total de alunos matriculados por turno e sexo;
- Total de alunos matriculados portadores de necessidades especiais,
- Total de alunos com matrículas trancadas;
- Total de evasão de alunos; e
- Total de concluintes.

6.2 – Detalhamento dos Problemas Diagnosticados

Em conformidade com a descrição apresentada no capítulo 1, referente aos problemas diagnosticados, será apresentada a especificação de cada item mencionado:

- **Os dados informados no Censo não são utilizados nas estratégias para a tomada de decisão:** a manipulação é feita apenas pelo Pesquisador Institucional, que envia ao Ministério da Educação e Cultura as informações requeridas, sem que haja o conhecimento da alta gerência no que tange as informações censitárias inseridas no sistema SiedSup do MEC;

- **Inexistência de acompanhamento dos dados informados no decorrer do ano que antecede ao censo:** a preocupação em levantar as informações é concentrada apenas no período de divulgação de abertura do censo, não havendo um trabalho posterior que acompanhe a confiabilidade das informações requeridas pelo MEC;
- **Impossibilidade de gerar relatórios ou gráficos com as informações censitárias:** pelo tratamento diferenciado das informações requeridas pelo Ministério da Educação e Cultura e as informações disponibilizadas pelo sistema acadêmico da UCB, não há possibilidade de gerar relatórios que caracterizem as informações censitárias; e
- **Macro visão do MEC x Micro visão da IES:** por meio do censo, o Ministério da Educação e Cultura visualiza as Instituições de Ensino Superior do país, aplicando, assim, suas políticas públicas educacionais, geradas por meio de sua macro visão. Já a UCB, por não manipular as informações que são transmitidas ao MEC, não é capaz de obter e usufruir de uma micro visão que poderia lhe ser de extrema utilidade no âmbito acadêmico e institucional, a ser utilizada pela alta gerência, na viabilização de estruturas para o apoio à decisão em relação ao gerenciamento de estratégias para fidelização de discentes.

6.3 – Proposta de solução

Baseado nos problemas diagnosticados, propõe-se uma solução que viabilize a visão interna da Instituição de forma a agregar informações que gerem conhecimento. Tal solução é baseada na garantia de um suporte interno à Instituição e, por meio dele, busca-se possibilitar a mesma visão do MEC, porém de forma minuciosa. O objetivo é gerar subsídios para a tomada de decisão, a serem disponibilizadas para a alta gerência.

A seguir serão descritos alguns exemplos de informações chaves que poderão ser trabalhadas e focadas no intuito de identificar reais oportunidades no negócio:

- Faixa etária: possibilita a identificação de público alvo;
- Evasão x Ingresso: quantidade de alunos matriculados e a evasão a cada semestre;
- Sexo: identifica perfil feminino ou masculino em cada curso oferecido;
- Ingresso x Concluintes: identifica a quantidade de alunos que ingressaram na instituição e, destes, quantos concluíram;
- Identificar métricas que obtenham, por exemplo, a evolução das notas de candidatos ao processo seletivo em períodos estabelecidos, bem como a identificação da média de notas dos alunos em determinado curso; e
- Necessidades especiais: identifica as necessidades mais recorrentes dos alunos e, com isso, possibilita a estrutura mais adequada à necessidade.

6.4 – Modelo Diagrama de Classes

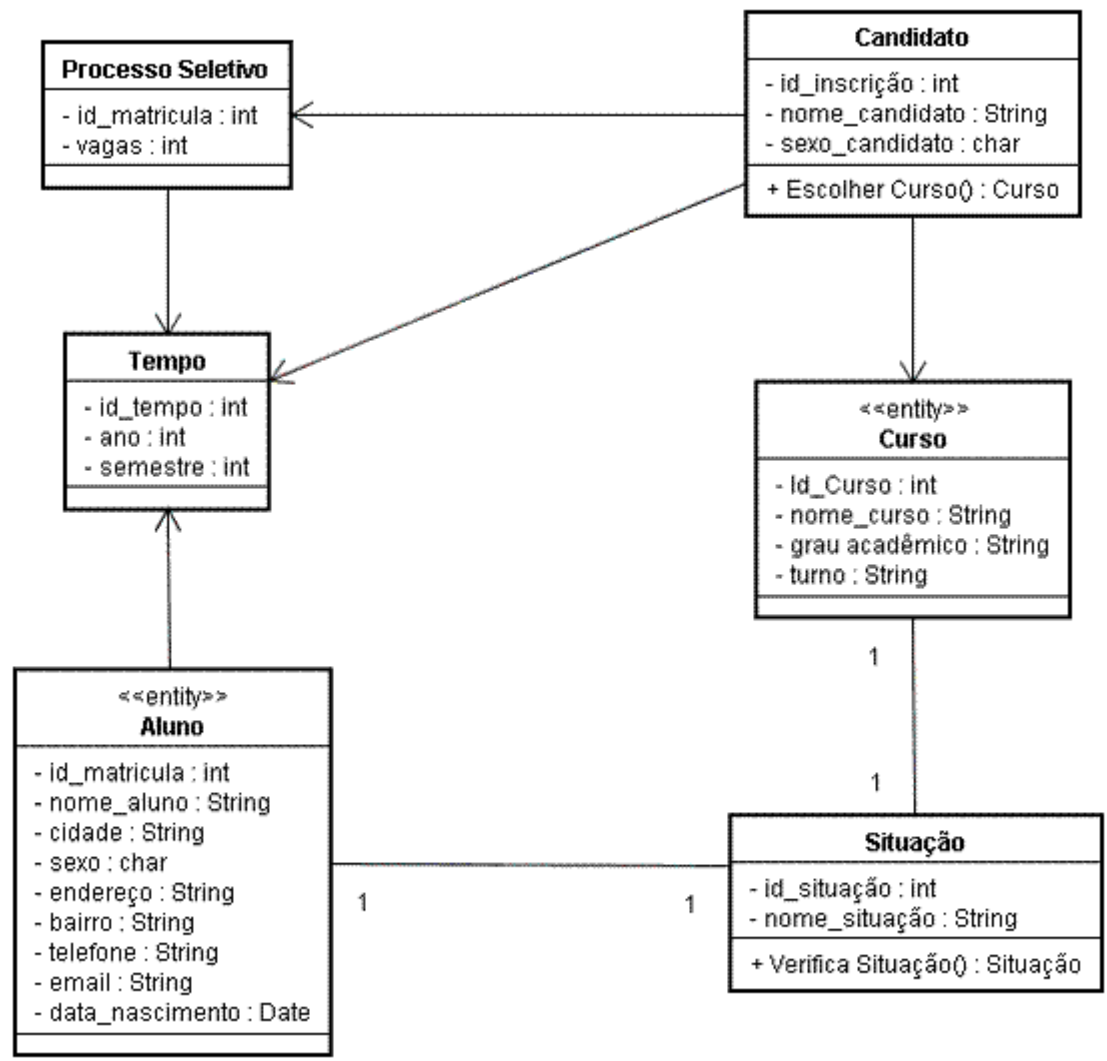


Figura 14 - Modelo Diagrama de Classes

6.5 – Modelo Entidade Relacionamento

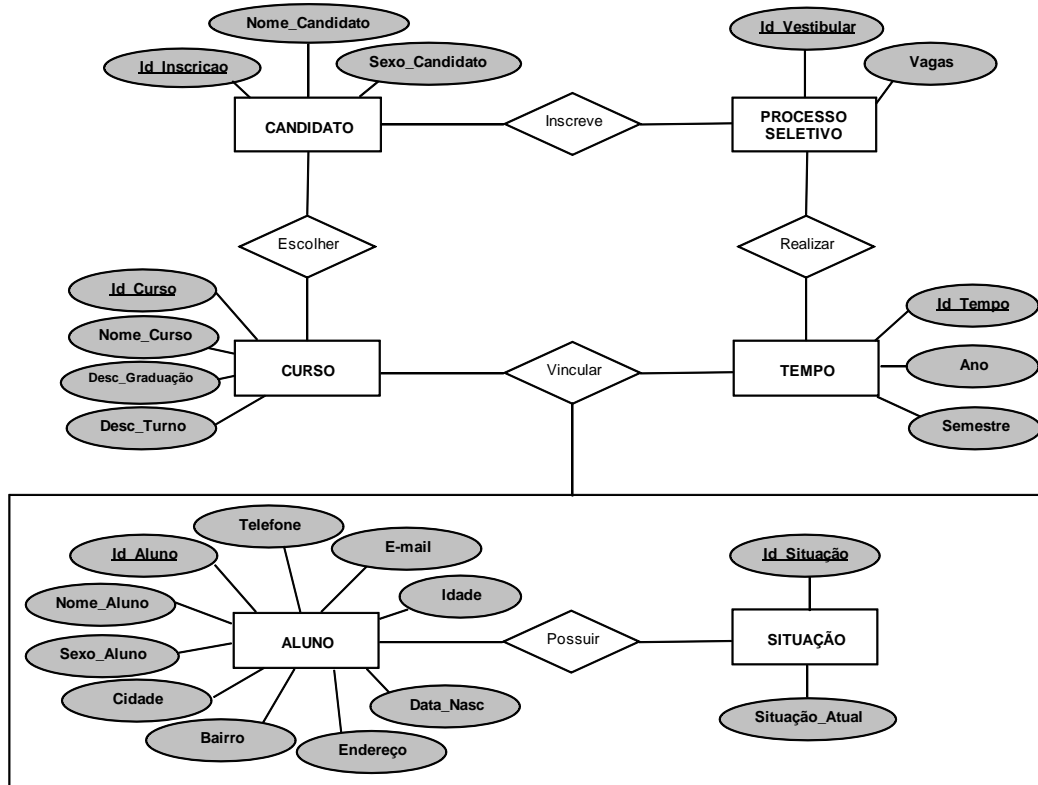


Figura 15 - Modelo Entidade Relacionamento

6.6 – Modelagem do *Data Warehouse*

Para ser útil, o *Data Warehouse* deve ser capaz de responder a consultas avançadas de maneira rápida, sem deixar de evidenciar detalhes relevantes à resposta. Considerando que dados multidimensionais podem ser armazenados e representados em estruturas relacionais, será utilizado o modelo Estrela (*Star Schema*), onde todas as tabelas se relacionam com a tabela Fato, o que fornece um acesso mais rápido aos dados, tornando mais fácil a navegação.

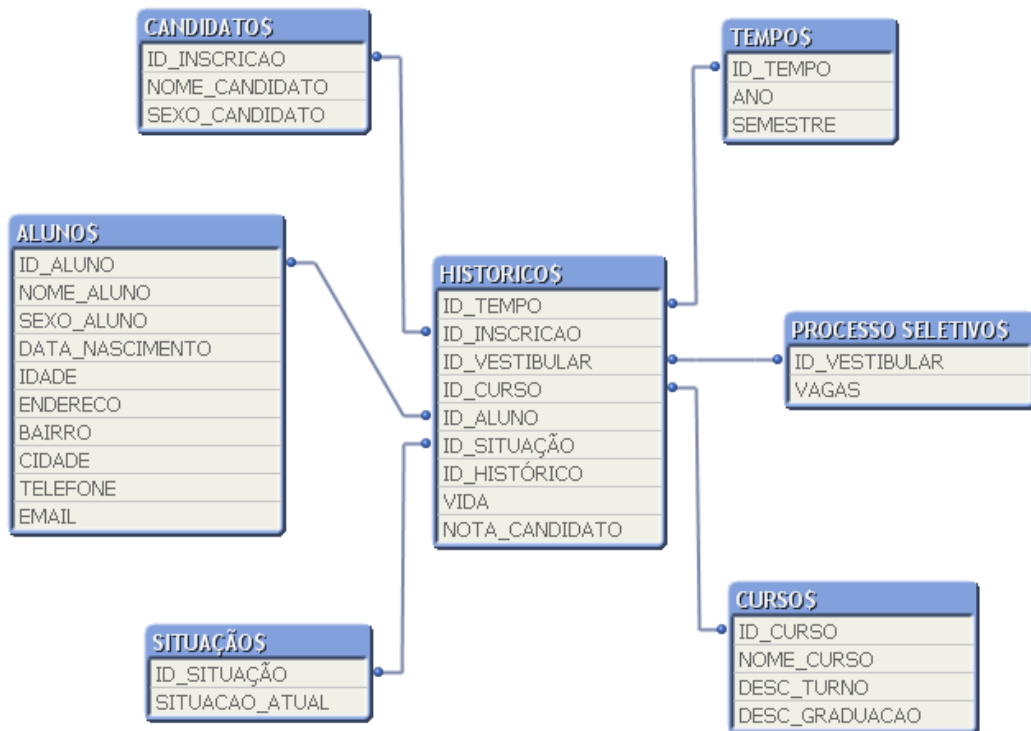


Figura 16 - Modelagem do *Data Warehouse*

6.6.1 – Descrição das tabelas do Modelo *Data Warehouse*

As tabelas fato e dimensão foram acrescentadas focando os pontos-chaves a serem analisados em busca de informações relevantes ao processo decisório. Desta forma, seguem tabelas com suas respectivas descrições:

Tabela 4- Descrição da Tabela Fato: Histórico

Tabela	Atributo	Tipo
Historico	id_historico (PK)	integer
	id_vestibular	integer
	id_inscricao	integer
	id_aluno	integer
	id_curso	integer
	id_situacao	integer
	id_tempo	integer
	nota_candidato	float
	vida	float

Tabela 5 - Descrição da Tabela Dimensão: Processo Seletivo

Tabela	Atributo	Tipo
Processo Seletivo	id_vestibular (PK)	integer
	vagas	integer

Tabela 6 - Descrição da Tabela Dimensão: Candidato

Tabela	Atributo	Tipo
Candidato	id_inscricao (PK)	integer
	nome_candidato	Varchar(50)
	sexo_candidato	Char (1)

Tabela 7 - Descrição da Tabela Dimensão: Aluno

Tabela	Atributo	Tipo
Aluno	id_aluno (PK)	integer
	nome_aluno	Varchar(50)
	sexo_aluno	Char (1)
	endereço	Varchar(50)
	bairro	Varchar(20)
	cidade	Varchar(20)
	telefone	Varchar(20)
	email	Varchar(50)
	idade	integer
	data_nascimento	data

Tabela 8 - Descrição da Tabela Dimensão: Curso

Tabela	Atributo	Tipo
Curso	id_curso (PK)	integer
	nome_curso	varchar (40)
	desc_graduacao	varchar (20)
	desc_turno	varchar (10)

Tabela 9 - Descrição da Tabela Dimensão: Tempo

Tabela	Atributo	Tipo
Tempo	id_tempo (PK)	integer
	ano	integer
	semestre	integer

Tabela 10 - Descrição da Tabela Dimensão: Situação

Tabela	Atributo	Tipo
Situacao	id_situacao (PK)	integer
	situacao_atual	varchar (20)

De acordo com as tabelas relacionadas podem-se identificar os atributos especificados dentro do seguinte contexto:

- Id_Histórico: identificação única que um registro terá dentro do Histórico;

- Nota_Candidato: métrica para identificar a média do candidato em períodos pré-estabelecidos;
- Vida: métrica para identificar a média do aluno, armazenada no índice VIDA, em determinado curso e período;
- Id_Vestibular: identificação única do Processo Seletivo;
- Vagas: quantidade de vagas disponibilizadas para cada curso no processo seletivo;
- Id_Inscrição: identificação única do candidato ao se inscrever no processo seletivo;
- Nome_Candidato: nome completo fornecido pelo candidato no ato da inscrição;
- Sexo_Candidato: identificação do sexo do candidato;
- Id_Aluno: identificação única que o estudante terá dentro da Instituição;
- Nome_Aluno: nome completo fornecido pelo aluno no ato da matrícula;
- Cidade: cidade onde o aluno reside;
- Sexo_Aluno: identificação do sexo do aluno;
- Endereço: local onde reside o aluno;
- Bairro: bairro onde reside o aluno;
- Telefone: telefone de contato fornecido pelo aluno;
- E-mail: endereço eletrônico fornecido pelo aluno;
- Idade: identificação da idade do aluno;
- Data_Nascimento: armazenará o dia, mês e ano de nascimento do aluno;
- Id_Curso: identificação única que o curso possui dentro da Instituição;
- Nome_Curso: nomenclatura do curso na Instituição;
- Desc_Graduação: tipo da formação de determinado curso;

- Desc_Turno: turno em que o curso é oferecido;
- Id_Tempo: identificação única para a dimensão tempo;
- Ano: registrará o ano em que ocorreu determinado registro;
- Semestre: registrará o semestre em que ocorreu determinado registro;
- Id_Situação: identificação única da situação do aluno na Instituição; e
- Situação_Atual: nomenclatura dos tipos de situações existentes na Instituição.

6.6.2 – Exemplos das tabelas do Modelo *Data Warehouse*

Seguem abaixo alguns exemplos dos conteúdos das tabelas dimensões apresentadas:

Tabela 11 - Exemplo da Tabela Processo Seletivo

id_vestibular	vagas
12008	50
22006	50
22001	50
12004	50
12005	50

Tabela 12 - Exemplo da Tabela Candidato

id_inscricao	nome_candidato	sexo_candidato
12345678999	JOSE ALUIZIO FERREIRA LIMA	M
98765432144	CINARA VIRGINIA TINOCO	F
56978412433	EGILDO DE OLIVEIRA LIMA	M
25896314788	CATIA MARIA DA SILVA	F
97968811104	MARCOS SILVA MACEDO	M

Tabela 13 - Exemplo da Tabela Aluno

Id_Aluno	Nome_Aluno	Sexo_Aluno	Cidade	Endereço	Bairro	Telefone	E-mail	Data_Nasc	Idade
111111111	JOSE ALUIZIO FERREIRA	M	CEILÂNDIA	2 AV BL 1500 CS 09	CEILANDIA SUL	376-0307	aluno@ucb.br	23-mar-78	30
222222222	CINARA VIRGINIA TINOCO	F	TAGUATINGA	3ª AV BL 990 LT 1000B APTO 302	PARK WAY	563-8299	aluno@ucb.br	24-ago-80	28
333333333	EGILDO DE OLIVEIRA LIMA	M	LAGO SUL	3ª AVENIDA 02 LT K ED. MULT	PARK WAY	3388897	aluno@ucb.br	15-jul-84	24
444444444	CATIA MARIA DA SILVA	F	GAMA	ADE CJ 07 LT 10 APT 102	SETOR LESTE	3751865	aluno@ucb.br	14-set-81	27
555555555	MARCOS SILVA MACEDO	M	BRASILIA	AE 2A CONJ B LJ 03	ASA NORTE	81424957	aluno@ucb.br	24-dez-80	28

Tabela 14 - Exemplo da Tabela Curso

id_curso	nome_curso	desc_graduacao	desc_turno
327	Sistemas de Informacao	Bacharel	Noturno
328	Medicina	Bacharel	Integral
302	Pedagogia	Licenciatura	Noturno
330	Biomedicina	Bacharel	Matutino
331	Farmácia	Bacharel	Vespertino

Tabela 15 - Exemplo da Tabela Tempo

Id_Tempo	Ano	Semestre
1	1999	1
2	1999	2
3	2000	1
4	2000	2
5	2001	1

Tabela 16 - Exemplo da Tabela Situação

id_situacao	situacao_atual
1	matriculado
2	desistente
3	trancado
4	cancelado
5	transferido
6	formado

6.7 – Ferramentas utilizadas

As ferramentas que foram efetivamente utilizadas para atender às demandas deste estudo de caso e que mostraram um desempenho de acordo com os resultados esperados foram as seguintes:

- *JUDE*: ferramenta case utilizada para geração dos diagramas necessários para melhor entendimento das necessidades apresentadas;
- *Microsoft Access*: utilizado para criação do banco de dados;
- *Microsot Excel*: utilizado para administração do banco de dados; e
- *Qlik View*: utilizado para o processamento das informações do banco de dados, gerando as visões necessárias ao andamento do negócio.

É importante ressaltar que estas ferramentas foram pré-selecionadas e, posteriormente, o estudo foi aprofundado, por meio de referências teóricas e análises práticas.

Primeiramente, foi utilizada a ferramenta UML *JUDE* para criação do diagrama de classes, tornando possível uma melhor visualização do escopo do projeto. A ferramenta é *free* e seu desempenho não deixa nada a desejar em vista de outras ferramentas que possuem o mesmo objetivo.

O *Microsoft Access* possui fácil acesso por parte de qualquer usuário, pois vem incluso no *Pacote Office Professional*. Sua utilização garante o desenvolvimento rápido de aplicações, trabalhando com modelagem de dados e lógica de programação. Por apresentar tais características, optou-se pela escolha desta ferramenta, uma vez que foi necessária a migração dos dados referenciais para criação do banco de dados a ser utilizado.

A ferramenta *Microsoft Excel* apresenta as mesmas características de acesso do *Microsoft Access* e, por isso, também foi utilizada, no intuito de facilitar a administração do banco de dados, manipulando inserções e exclusões de registros e dados considerados desnecessários.

Para tornar possível a visualização e melhor entendimento das oportunidades oferecidas pelo trabalho focado e aprofundado em informações relevantes ao processo de tomada de decisão, foi identificada uma ferramenta que tornou capaz o cruzamento de informações e posteriores análises, baseadas em tabelas e gráficos. Tal ferramenta, chamada *Qlik View*, é voltada para o desenvolvimento de aplicativos de posicionamento estratégico, ou seja, *Business Intelligence*. Ela não é uma ferramenta *free*, porém é disponibilizada na internet uma versão *Trial*, que pode ser utilizada por 15 dias não consecutivos.

A ferramenta reúne, em um rico ambiente de desenvolvimento, as ferramentas de extração, transformação e entrada de dados (ETL) totalmente integrados. O *QlikView* é capaz de carregar, pesquisar e analisar grandes quantidades de dados em segundos. A velocidade do carregamento e atualização dos dados também é rápida, pois é incrementada pela utilização de cargas incrementais (o programa faz o carregamento de todos os dados na memória e salva em um arquivo próprio).

A tecnologia AQL (*Associative Query Logic*), patenteada pela *Qlik Tech* – empresa criadora do *Qlik View* - constrói e mantém um banco de dados não relacional, associativo e com uma base de dados altamente eficiente alojada na memória RAM. A vantagem da arquitetura AQL é que a fonte de dados é retida e imediatamente disponibilizada *off line* para análise de todo caminho da fonte transacional. O resultado é uma poderosa capacidade analítica provida por uma interface para o usuário altamente intuitiva que encoraja a exploração e a criatividade.

Neste estudo de caso, após a criação do banco de dados no *Microsoft Access*, todas as tabelas foram salvas com extensão *.xls*. Em seguida à conclusão do escopo de dados a serem analisados, a utilização do *Qlik View* pôde ser iniciada por meio do desenvolvimento de um script, gerado na sintaxe SQL, que proporcionou a carga dos dados.

Logo em seguida, a representação relacional foi gerada pela ferramenta, de acordo com o Modelo do *Data Warehouse*, especificado na figura 16.

Após certificação da correta representação das tabelas carregadas, foi possível desenvolver as interfaces gráficas, que forneceram informações ricas, flexíveis e funcionais.

A seguir serão mostradas algumas visões fornecidas pela ferramenta *Qlik View*, onde foram trabalhados indicadores chaves do negócio, no intuito de identificar diferentes cenários dos discentes e respectivos cursos da Universidade Católica de Brasília, que proporcionaram aplicações interativas de análises de dados:

- **Informações de cursos oferecidos:**

ID_CURSO	NOME_CURSO	DESC_GRADUACAO	DESC_TURNO
300	Formação Geral	Disciplinas Isoladas	Matutino
301	Pedagogia	Licenciatura Curta	Noturno
302	Pedagogia	Licenciatura Plena	Noturno
303	Administração	Bacharel	Matutino
304	Ciências Econômicas	Bacharel	Noturno
305	Sup de Tec em Processamento de Dados	Tecnólogo	Noturno
310	Ciência da Computação	Bacharel	Matutino
311	Filosofia	Licenciatura Plena	Noturno
312	Letras	Licenciatura Plena	Noturno
313	Educação Física	Licenciatura Plena	Noturno
314	Relações Internacionais	Bacharel	Matutino
315	Comunicação Social	Bacharel	Matutino
316	Direito	Bacharel	Noturno
317	Ciências Contábeis	Bacharel	Noturno
318	Fisioterapia	Fisioterapeuta	Integral
319	Nutrição	Bacharel	Vespertino
320	Psicologia	Psicólogo	Vespertino
321	Odontologia	CirurgiãoDentista	Integral
322	Matemática	Licenciatura Plena	Noturno
323	Física	Licenciatura Plena	Vespertino
324	Química	Licenciatura Plena	Noturno
325	Ciências Biológicas	Bacharel	Matutino
326	Engenharia Ambiental	Engenheiro Ambiental	Matutino
327	Sistemas de Informação	Bacharel	Noturno
328	Medicina	Médico	Integral

Figura 17 - Cursos oferecidos (Visão na Ferramenta Qlik View)

▪ **Informações dos discentes:**

ID_ALUNO	NOME_ALUNO	SEXO...	DATA_NASCIM...	IDADE	ENDERECO	BAIRRO	CIDADE	TELEFONE	EMAIL
UC03000135	ALINE VASCONCELOS LOPES MADUREI...	F	14/8/1975	33	QNN 37 CONJ. G CASA 01	CEILANDIA NO...	AGUAS CLARAS	84463161	mariana@...
UC07000733	MARCELA MAGALHAES SCAFUTO	F	17/2/1989	19	QUADRA 17 CASA 60	SETOR LESTE	AGUAS CLARAS	3384472181492480	walderse
UC08069435	WOLYSSON HIYANE CARVALHO	M	16/12/1985	23	SGN 309 BL O APT 607	ASA NORTE	AGUAS CLARAS	6132730416	flaviafagg
UC03000291	BRUNO ALVES DE ARAUJO	M	5/9/1980	28	QNR 12 CONJ V CASA 31	ST. P SUL	AGUAS CLARAS	6133779993	lucienebr
2002004927	PRISCILA DIAS NASCIMENTO	F	18/4/1956	52	QE 13 CONJUNTO D CASA 15	GUARÁ II	AGUAS CLARAS	613822673	fabiodata
UC08068226	DANIANNA SIQUEIRA BATISTA	F	9/2/1962	46	QNL 10 CONJ D CASA 17	LNORTE	TAGUATINGA	3352293781267611	luciano_c
UC03000402	TATIANNA CARDIM MORAIS FINO	F	12/11/1990	18	SGN 209 BL G APT 307	ASA NORTE	TAGUATINGA	2733721	andrealv
UC08068821	RENATO SANTIAGO AMANCIO	M	21/3/1986	22	QUADRA 104 LOTE 08/10 BLOCO A APTO 202	AGUAS CLARAS	TAGUATINGA	614352831	paulosou
UC08068528	ERICA FERREIRA ROCHA	F	1/2/1984	24	RUA 21 SUL LOTE 08 BLOCO D APT 701 AVENIDA ARAUCA...	AGUAS CLARAS	TAGUATINGA	4356312	ninalasn
UC06000713	DAISE DINIZ DE PAULA RIBEIRO	F	12/8/1980	28	QR 506 CONJ 13 CS 20	SAMAMBAIA SUL	TAGUATINGA	3357192196624547	ana.filgus
2002004731	IZABELA GUIMARAES CUNHA E SILVA	F	18/4/1981	27	QNL 07 CONJUNTO F CASA 14	TAGUATINGA N...	TAGUATINGA	3365167 / 91027938	clarissa_
UC03000364	FERNANDO AGUIAR MEIRELES	M	9/2/1983	25	QR 412 CONJ 05 CASA 05	SAMAMBAIA NO...	TAGUATINGA	4581480	aluno@u
UC04001267	ANNE CURVELLO DE MENDONÇA FONS...	F	20/3/1970	38	AOS 05 BLOCO E APTO 202	OCTOGONAL	TAGUATINGA	3610927	aluno@u
UC08069354	CAROLINA ALBUGUERQUE SILVA	F	21/11/1975	33	SMPW QUADRA 08 CONJUNTO 03 LOTE 02 CASA 01	PARK WAY	TAGUATINGA	3802327	aluno@u
2002005192	GRAZIELE DOS ANJOS SILVA	F	29/12/1976	32	SHRF 02 COND. VIA ARAGUAIA 523 A	RIACHO FUND...	TAGUATINGA	33994897	resozinhc
UC07000491	FABIO BRUNO ALVES	M	3/11/1985	23	QS 11 CONJ J CS 39	AREAL	TAGUATINGA	613567715	amandinl
UC08068846	CATIA DOS SANTOS VIEIRA	F	15/7/1987	21	QN 01 CJ 03 CS 07	RIACHO FUND...	TAGUATINGA	84257078	claudia.1
2000003915	MARISTELA FERREIRA DO NASCIMENTO	F	25/5/1985	23	QS 12 CONJUNTO 4B CASA 12	RIACHO FUND...	TAGUATINGA	613992164	izaleinh:
UC04000562	ANA PAULA GOMES CORREIA CARNEIRO	F	13/12/1980	28	Q 70 CS 10 RUAS CENTRO	CENTRO	TAGUATINGA	84759257	ybytyra@
2002005711	CARLA ALVES GOMES	F	14/3/1984	24	RUA 22 NORTE, LOTE 04, RESIDENCIAL VILLA LORENZA	AGUAS CLARAS	TAGUATINGA	3436101281346374	ipivello@

Figura 18 - Informações dos discentes (Visão na Ferramenta Qlik View)

▪ **Identificação de público alvo:**

- Gráfico da representação:

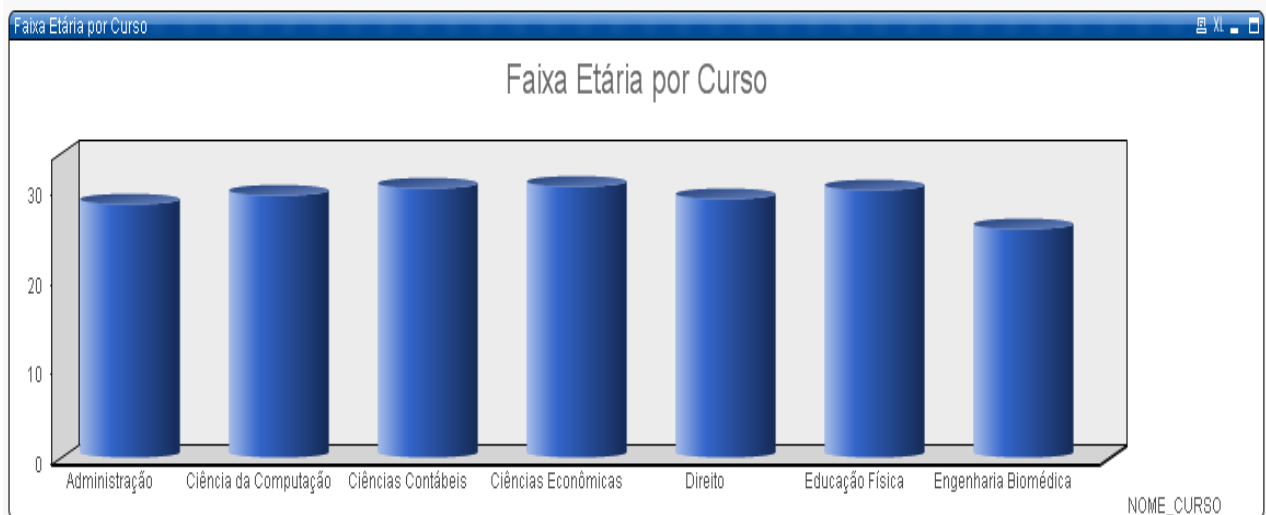


Figura 19 - Identificação de público alvo - Gráfico: Faixa Etária (Visão na Ferramenta Qlik View)

- Tabela e filtro da representação:

ID_ALUNO	NOME_ALUNO	SEXO_ALUNO	DATA_NASCI...	IDADE	DESC_TURNO	SITUACAO_ATUAL
UC03000291	BRUNO ALVES DE ARAUJO	M	5/9/1980	28	Noturno	A
2002004927	PRISCILA DIAS NASCIMENTO	F	18/4/1956	52	Noturno	D
UC08068226	DANIANA SIQUEIRA BATISTA	F	9/2/1962	46	Noturno	A
UC08068821	RENATO SANTIAGO AMANCIO	M	21/3/1986	22	Matutino	A
UC06000713	DAISE DINIZ DE PAULA RIBEIRO	F	12/8/1980	28	Matutino	A
2002004731	IZABELA GUIMARAES CUNHA E SILVA	F	18/4/1981	27	Matutino	D
UC03000364	FERNANDO AGUIAR MEIRELES	M	9/2/1983	25	Noturno	D
UC08069354	CAROLINA ALBUQUERQUE SILVA	F	21/11/1975	33	Matutino	D
UC07000491	FABIO BRUNO ALVES	M	3/11/1985	23	Matutino	D
UC08068846	CATIA DOS SANTOS VIEIRA	F	15/7/1987	21	Noturno	A
UC07000482	LEANDRO MAGALHAES TRINDADE	M	5/6/1980	28	Noturno	A
UC08069087	AGENEY GOMES BREDOFF	M	16/3/1973	35	Noturno	A
UC06000144	JÓÃO BENEDITO DA CUNHA ALVES	M	24/7/1988	20	Noturno	A
UC07001349	JOSÉ REGINALDO DE OLIVEIRA	M	19/6/1978	30	Noturno	A
UC08069184	ADEMÁRIO DE JESUS CHAGAS	M	22/9/1987	21	Noturno	A

Figura 20 - Identificação de público alvo – Tabela e Filtro: Faixa Etária (Visão na Ferramenta Qlik View)

- Identificação de perfil:

- Gráfico e tabela com seleção múltipla da representação:

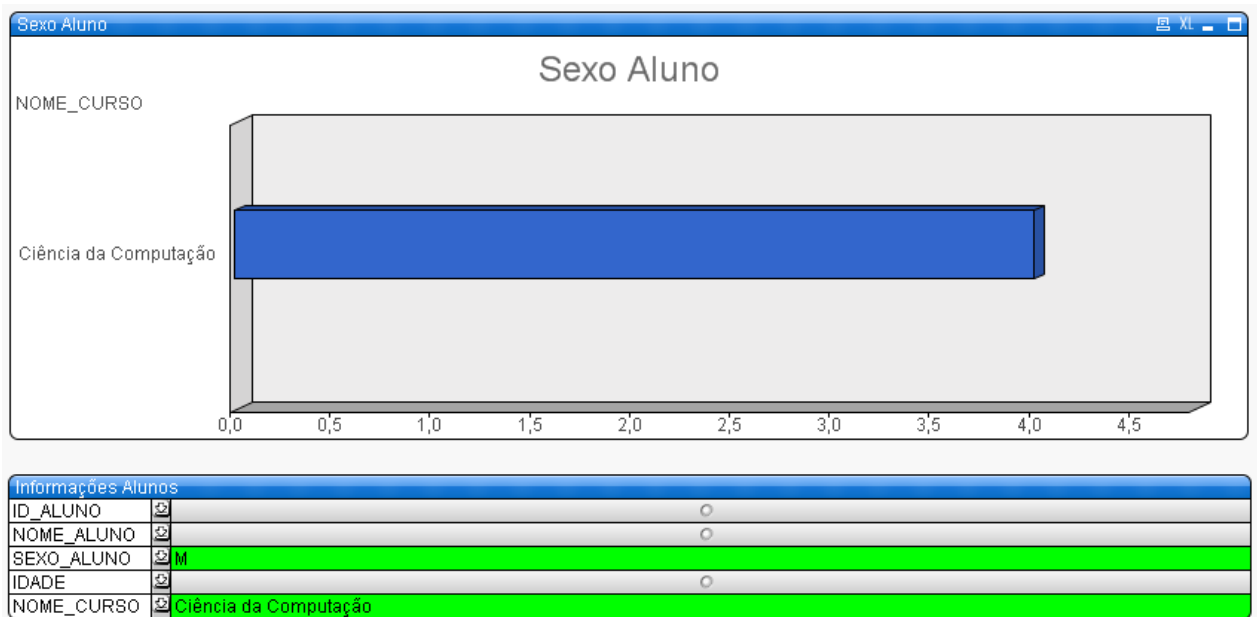


Figura 21 - Identificação de Perfil – Gráfico e Tabela com Seleção Múltipla: Sexo (Visão na Ferramenta Qlik View)

▪ **Aproveitamento do índice VIDA:**

- Gráficos da representação:

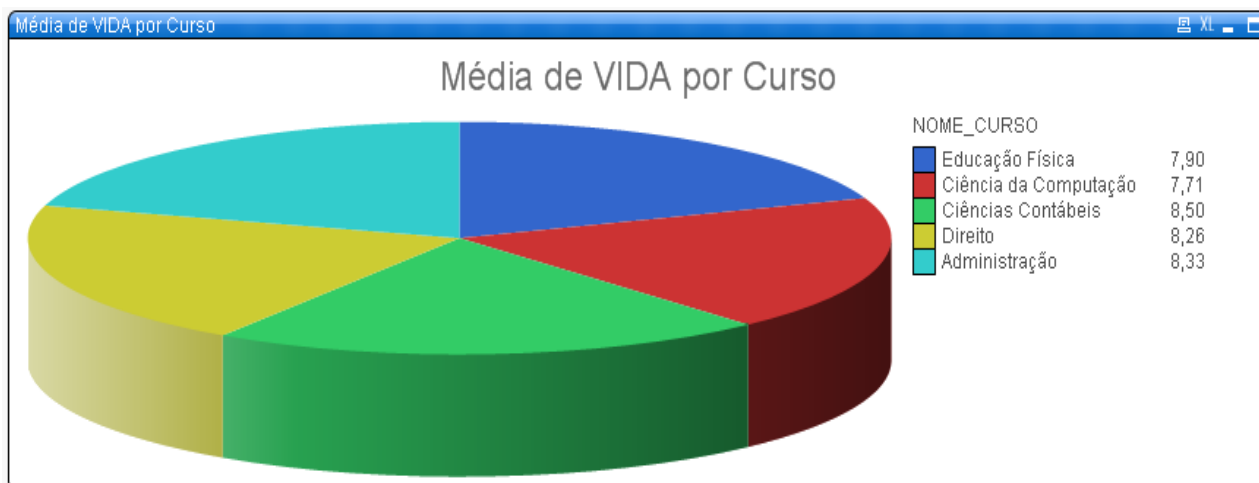


Figura 22 - Aproveitamento do índice VIDA – Gráfico: Curso (Visão na Ferramenta Qlik View)

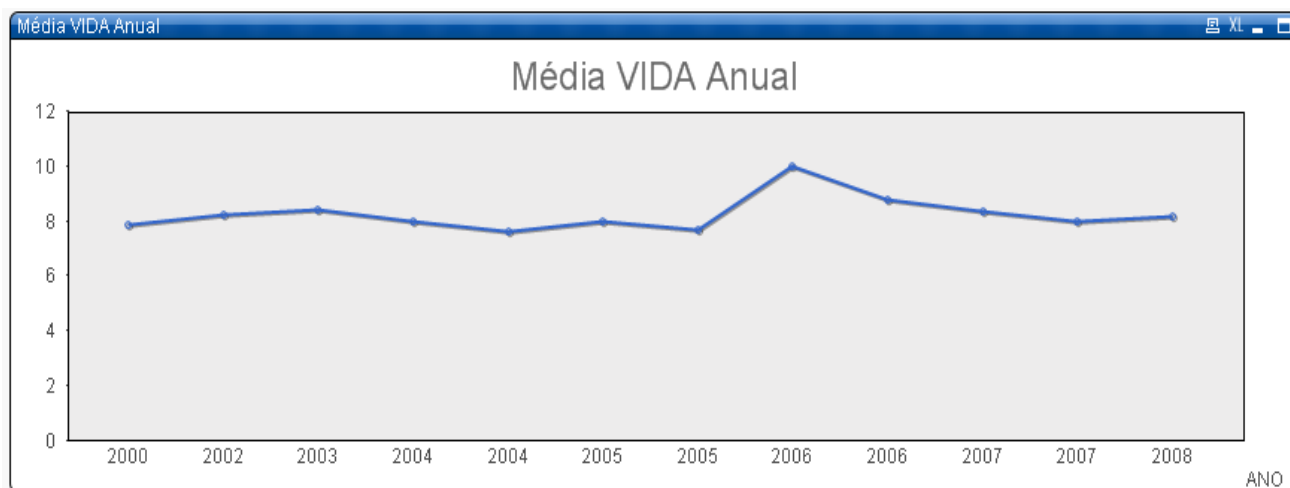
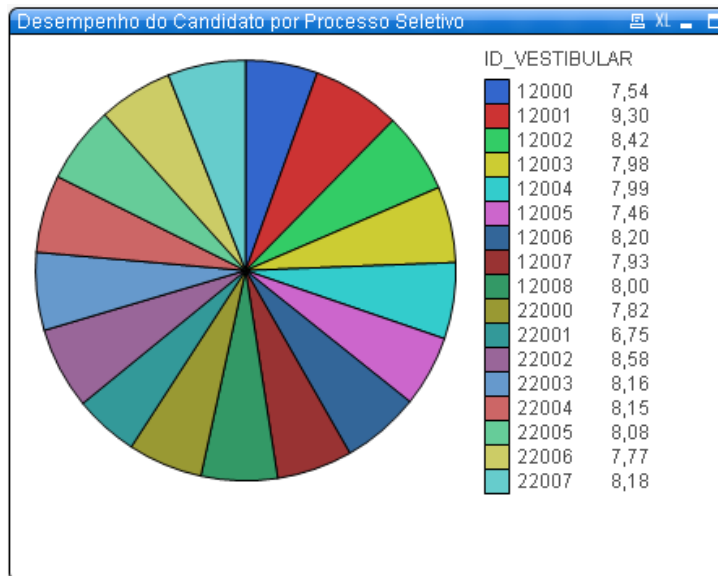


Figura 23 - Aproveitamento do índice VIDA – Gráfico: Ano (Visão na Ferramenta Qlik View)

▪ **Desempenho de candidatos em processos seletivos:**



○ Gráfico da representação:

Figura 24 - Desempenho de candidatos por Processo Seletivo – Gráfico (Visão na Ferramenta Qlik View)

○ Tabela e filtros da representação:

INFORMAÇÕES CANDIDATOS			
ID_INSCRICAO	NOME_CANDIDATO	SEXO_CANDIDATO	NOTA_CANDIDATO
4	BRUNO ALVES DE ARAUJO	M	7
5	PRISCILA DIAS NASCIMENTO	F	8,4
6	DANIANA SIQUEIRA BATISTA	F	8,8
8	RENATO SANTIAGO AMANCIO	M	8,9
10	DAISE DINIZ DE PAULA RIBEIRO	F	8,2
11	IZABELA GUIMARAES CUNHA E SILVA	F	8,3
12	FERNANDO AGUIAR MEIRELES	M	7,5
14	CAROLINA ALBUQUERQUE SILVA	F	8,2
16	FABIO BRUNO ALVES	M	10
17	CATIA DOS SANTOS VIEIRA	F	8,7
22	LEANDRO MAGALHAES TRINDADE	M	8
23	AGENEY GOMES BREDOFF	M	8,3
24	JOÃO BENEDITO DA CUNHA ALVES	M	8,8
25	JOSÉ REGINALDO DE OLIVEIRA	M	5,9
26	ADEMÁRIO DE JESUS CHAGAS	M	9,9
28	ANTONIO JOSÉ DA SILVA	M	7,9

ID_VESTIBULAR	NOME_CURSO
12000	Administração
12001	Ciência da Computação
12002	Ciências Contábeis
12003	Ciências Econômicas
12004	Direito
12005	Educação Física
12006	Engenharia Biomédica
12007	Administração Instituições Financeiras
12008	Agronegócios
22000	Biomedicina
22001	Ciências Biológicas
22002	Comunicação Social
22003	Curso Sup. de Tec. em Gestão Comercial
22004	Disciplinas Isoladas
22005	Enfermagem
22006	Engenharia Ambiental
22007	Farmácia

NOTA_CANDIDATO	
Média	7,93
Mínimo	5,9
Máximo	10

Figura 25 - Desempenho de candidatos por Processo Seletivo – Tabelas e filtros (Visão na Ferramenta Qlik View)

CAPÍTULO 7– CONCLUSÃO

Na oportunidade de participar da análise de requisitos para elaboração do Censo de Educação Superior do Ministério da Educação e Cultura, idealizou-se a realização deste projeto, no intuito de prover soluções de apoio à decisão para alta gerência, por meio da utilização de ferramentas que manipulem a base de dados do censo.

Baseado no *Business Intelligence*, o mercado de tecnologia oferece diversas ferramentas de apoio à decisão, que proporcionam às organizações análises avançadas aos tomadores de decisão, por meio da exploração de massa de dados complexas, de modo a proporcionar conhecimento rápido e objetivo.

A base de dados acadêmica do discente, disponível do sistema *Stratus* da Universidade Católica de Brasília, grava apenas o último registro do aluno, perdendo todo o seu histórico no decorrer do período, o que ocasiona perda de informações importantes para análise da tomada de decisão. A preocupação em armazenar estes dados em banco iniciou-se no ano de 2002, no entanto com a mesma característica. Todas as informações anteriores a este período estão em arquivos de papel, dispostas às interferências naturais do tempo, passíveis a desastres, tais como: incêndios, inundações, entre outros.

Utilizar sistemas de apoio à decisão garante a instituição o tratamento da grande massa de dados, que, por meio de filtros, gerem a compreensão do cenário atual, com o objetivo de manter a fidelização dos discentes.

Diante disso, pode-se afirmar que a aplicação de uma solução *Business Intelligence* capacita a instituição uma representação de bancos de dados organizados, com histórico de ocorrências, que auxiliem consultas focadas em estratégias no processo de tomada de decisão para alta gerência.

Com a utilização da ferramenta proposta, os dados informados no Censo poderão ser utilizados estrategicamente no processo de tomada de decisão, proporcionando visões que capacitem os níveis estratégico, tático e operacional da Instituição nas suas ações, focadas no acompanhamento do histórico das informações fornecidas, gerando relatórios, gráficos e análises baseadas nas

informações censitárias, proporcionando uma visão minuciosa das oportunidades oferecidas.

O resultado desta aplicação garante a UCB um diferencial frente às outras instituições de ensino superior da região, capacitando à mesma uma visão ampla do negócio, com a observação de nichos de mercado, que potencialize os recursos por meio de soluções estratégicas que aprimorem um ensino de qualidade.

7.1. Trabalhos Futuros

Projetar um sistema acadêmico, que viabilize o armazenamento de todos os registros ocorridos em um determinado período, mantendo a visão precisa e fidedigna dos dados, para a construção de um *Data Warehouse*.

A ferramenta utilizada deve ser capaz de capturar, analisar e entender as tendências dos discentes em tempo real, por meio do cruzamento das dimensões envolvidas no processo, bem como na criação de novas dimensões que poderão surgir no decorrer da utilização. Além disso, deve proporcionar uma interface rica, flexível e poderosa, tornando a manipulação de fácil acesso a qualquer usuário.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ALCANTARA, A. **Business Intelligence – Produzindo Resultados**. Brasília: 2003.
- [2] ALCANTARA, I. **Business Intelligence: Elevando a gestão dos negócios a um novo patamar**. 2006.
- [3] BARBIERI, C. **BI – Business Intelligence: Modelagem e Tecnologia**. Rio de Janeiro: Editora Axcel Books, 2001.
- [4] BINDER, F. V. **Sistemas de Apoio à Decisão**. São Paulo: Editora Érica, 1994.
- [5] BISPO, C. A. F. **Uma Análise da Nova Geração de Sistemas de Apoio à Decisão**. São Paulo: 1998.
- [6] Catalogo 2004/2005 - Universidade Católica de Brasília – UCB
- [7] COLAÇO, M. J. **Projetando Sistemas de Apoio à Decisão Baseados em Data Warehouse**. Rio de Janeiro: Editora Axcel Books, 2004.
- [8] FERREIRA, A. B. H. **Dicionário Aurélio**. Rio de Janeiro: Editora Nova Fronteira.
- [9] MOREIRA, E. **Modelo Dimensional para Data Warehouse**. Minas Gerais: 2006.
- [10] NAVEGA, S. **Princípios Essenciais do Data Mining**. São Paulo: 2002.
- [11] [SINAES, 2007] SINAES - Sistema Nacional de Avaliação da Educação Superior “Da Concepção à Regulação”. 4^o edição, Ministério da Educação e Cultura, 2007.

Internet

<http://www.ensinosuperior.inep.gov.br/>, acessado em 28 de agosto de 2007.

<http://sinaes.inep.gov.br/sinaes/>, acessado em 29 de agosto de 2007.

<http://portal.mec.gov.br/> , acessado em 27 de agosto de 2007.

<http://www.phdbrasil.com.br/>, acessado em 04 de setembro de 2007.

http://brazil.emc.com/news/artigos_archive/2001/122001_datawarehouse.jsp,
acessado em 11 de setembro de 2007.

<http://www.revistafatorbrasil.com.br/>, acessado em 20 de setembro de 2007.

<http://www.dwbrasil.com.br/>, acessado em 20 de setembro de 2007.

<http://www.sqlmagazine.com.br>, acessado em 25 de setembro de 2007.

<http://www.teses.usp.br/teses/disponiveis/18/18140/tde-04042004-152849>,
acessado em 11 de outubro de 2007.

[<http://www.computerworld.uol.com.br/>, acessado em 11 de outubro de 2007.

<http://www.computerworld.com/>, acessado em 11 de outubro de 2007.

<http://www.sol7.com.br/>, acessado em 11 de outubro de 2007.

<http://www.microsoft.com/brasil/servidores/bi/default.aspx>, acessado em 25 de
outubro de 2007.

<http://www.lc.leidenuniv.nl/awcourse/oracle/server.920/a96520/concept.htm>,
acessado em 25 de outubro de 2007.

<http://www.devmedia.com.br/> , acessado em 23 de outubro de 2007.

<http://www.suntzu.com.br/>, acessado em 19 de outubro de 2007.

<http://www.datamodelling.com.br/>, acessado em 01 de outubro de 2007.

<http://www.sisnema.com.br/>, acessado em 08 de novembro de 2007.

<http://www.wikipedia.org/>, acessado em 06 de novembro de 2007.

<http://www.microsoft.com/brasil/technet/colunas/default.mspx>, acessado em 16 de
novembro de 2007.

<http://www.datawarehouse.inf.br/artigos/olap2.asp>, acessado em 14 de novembro de 2007.

<http://www.edudatabrasil.inep.gov.br/>, acessado em 20 de novembro de 2007.

www.microsoft.com , acessado em 20 de novembro de 2007.

<http://www.microstrategy.com.br/>, acessado em 19 de maio de 2008.

<http://www.nordicasoftware.com.br>, acessado em 25 de maio de 2008.

<http://www.itg.com.br>, acessado em 14 de junho de 2008.