



Instituto Politécnico de Castelo Branco  
Escola Superior de Tecnologia



# Protótipo de DataWarehouse a aplicar numa Escola Secundária/3 de Ensino

Cristiana Isabel Morgado e Silva

Trabalho de Mestrado  
Desenvolvimento de Software e Sistemas Interactivos

Trabalho Efectuado sob a orientação do  
Professor Doutor Eurico Lopes

Fevereiro 2011

Instituto Politécnico de Castelo Branco  
Escola Superior de Tecnologia

# PROTÓTIPO DE *DATAWAREHOUSE* A APLICAR NUMA ESCOLA SECUNDÁRIA/3 DE ENSINO

Cristiana Isabel Morgado e Silva

Dissertação apresentada ao Instituto Politécnico de Castelo Branco para cumprimento dos requisitos necessários à obtenção do grau de Mestre em **Desenvolvimento de *Software* e Sistemas Interactivos**, realizada sob a orientação científica do Doutor Eurico Lopes, Professor Coordenador da Unidade Técnico-Científica de Informática da Escola Superior de Tecnologia do Instituto Politécnico de Castelo Branco.

Fevereiro de 2011



*A ti, Ricardo.*



## AGRADECIMENTOS

Ao Professor Doutor Eurico Lopes, meu orientador, pela competência científica e acompanhamento do trabalho, pela disponibilidade revelada, assim como pelas críticas, correcções e sugestões relevantes feitas durante a orientação.

Um agradecimento muito especial ao Director da Escola Secundária/3 de Amato Lusitano, Professor João Belém, por ter autorizado o acesso aos dados internos da escola, que foram objecto de estudo.

Ao Pedro Vaz, pela ajuda na recolha bibliográfica e sua disponibilização.

À Cristina Canavarro, pela amizade e exemplo, que incentiva esta minha caminhada pelo mundo do saber.

A todos os que me querem bem, familiares e amigos, que me apoiaram incondicionalmente.

À minha mãe Emília e ao meu filho Ricardo, pela paciência, pelo constante apoio, motivação e incentivo nesta fase da minha vida.



## RESUMO

Nos tempos que correm é importante que se reflecta sobre a importância de um complemento da visão pedagógica à visão estratégica, a fim de uma maior eficácia aliada ao conhecimento para enfrentar o mercado inovador e competitivo de hoje.

Com o intuito de fornecer ferramentas que auxiliem um planeamento estratégico nas escolas, pressupôs este trabalho o desenvolvimento inovador de um depósito de dados a larga escala que permita de uma forma imediata a resposta a questões subjacentes às decisões a tomar por parte dos agentes pedagógicos.

O estudo proposto por esta dissertação, pretende sensibilizar no interesse e vantagem da construção, implementação, gestão e utilização de um sistema de apoio à decisão em escolas.

Desta forma foi projectado e desenvolvido um protótipo de *Data Warehouse* (DW), através da plataforma de *Business Intelligence* (BI) da *Microsoft*, aplicando as tecnologias de modelação multidimensional, onde após selecção, preparação e limpeza dos dados extraídos dos sistemas transaccionais da Escola Secundária/3 de Amato Lusitano, se carregou o *Data Warehouse* e se procederam às respectivas análises através de tecnologias *Online Analytical Processing* (OLAP). Para disponibilização/construção de relatórios implementou-se um *front-end*, através da plataforma colaborativa *SharePoint*.

PALAVRAS-CHAVE: *Data Warehouse*. *Business Intelligence*. Sistemas de Informação. Plataforma *BI* da *Microsoft*.



## ABSTRACT

Nowadays it is important to reflect on the importance of a pedagogic vision complement to the strategic vision in order to be more efficient as well as to know how to face today's innovative and competitive market.

It was my intention to supply tools to support strategic planning in schools so, this work aims for an innovative developing for a database that allows immediate response to underlying questions to the decisions to be taken by the pedagogic agents.

The study proposed in this dissertation intends to sensitize to the interest and advantage of the creation, implementation, management and use of decision support systems at schools.

This way a prototype of Data Warehouse (DW) was projected and developed, through the Microsoft platform of Business Intelligence (BI), using the multidimensional modelling technologies where, after having selected, prepared and cleaned the information obtained from the transactional systems of Escola Secundária/3 de Amato Lusitano, Data Warehouse was loaded and the technologies Online Analytical Processing (OLAP) were used to make the respective analyses. To build/make available reports a frontend was implemented through the collaborative platform SharePoint.

**KEYWORDS:** *Data Warehouse. Business Intelligence. Information Systems. Microsoft BI.*



## ÍNDICE

Dedicatória .....	iii
Agradecimentos.....	v
Resumo .....	vii
<i>Abstract</i> .....	ix
Índice .....	xi
Lista de figuras ou ilustrações .....	xiii
Lista de tabelas .....	xvii
Siglas ou abreviaturas .....	xix
<b>1 Introdução e objectivos .....</b>	<b>1</b>
1.1 Introdução .....	1
1.2 Definição do tema .....	3
1.3 Organização da dissertação .....	3
1.4 Enquadramento do trabalho .....	3
1.5 Objectivos do trabalho .....	4
1.6 Cronograma.....	5
<b>2 Revisão bibliográfica .....</b>	<b>7</b>
2.1 Introdução .....	7
2.2 Enquadramento teórico.....	7
2.3 Caracterização de um <i>Data Warehouse</i> .....	10
2.3.1 Orientação por assunto .....	11
2.3.2 Integração .....	12
2.3.3 Não volátil .....	13
2.3.4 Variação no tempo.....	14
2.4 Modelação Multidimensional .....	16
2.5 ETL - <i>Extract, Transform and Load</i> .....	20
2.6 Cubos .....	21
2.7 Técnicas de análise multidimensional.....	22
2.8 Metadados .....	23
2.9 Architecturas, abordagens e metodologias de implementação de DW .....	24

<b>3</b>	<b>Metodologias e ferramentas utilizadas.....</b>	<b>27</b>
3.1	Introdução .....	27
3.2	Apresentação do caso de estudo .....	27
3.3	Gestão estratégica.....	30
3.4	Infra-estrutura tecnológica.....	36
3.5	Outras ferramentas de <i>Business Intelligence</i> .....	38
<b>4</b>	<b>Modelação.....</b>	<b>41</b>
4.1	Introdução .....	41
4.2	Nível mais elevado de descrição do projecto .....	41
4.3	Componentes básicos .....	43
4.4	Modelo conceptual de dados.....	44
4.5	Modelo lógico de dados .....	46
4.6	Descrição das tabelas .....	49
4.6.1	Tabela de dimensão <i>DimData</i> .....	49
4.6.2	Tabela de dimensão <i>DimAluno</i> .....	50
4.6.3	Tabela de dimensão <i>DimMatricula</i> .....	51
4.6.4	Tabela de dimensão <i>DimDisciplina</i> .....	53
4.6.5	Tabela de dimensão <i>DimProfessor</i> .....	54
4.6.6	Tabela de facto <i>FactoAvaliacao</i> .....	55
4.6.7	Tabela de facto <i>FactoProfissional</i> .....	56
<b>5</b>	<b>Implementação do Protótipo de DW.....</b>	<b>59</b>
5.1	Introdução .....	59
5.2	Ciclo de vida do projecto.....	59
5.3	A plataforma <i>Business Intelligence</i> da <i>Microsoft</i> .....	60
5.4	Integração de dados - ETL.....	63
5.5	Gestão de metadados .....	65
5.6	Análise multidimensional e criação dos cubos de dados.....	66
5.7	Elaboração de relatórios.....	67
5.8	Disponibilização dos relatórios.....	68

<b>6</b>	<b>Conclusões .....</b>	<b>75</b>
6.1	Reflexão crítica .....	75
6.2	Avaliação do trabalho realizado.....	77
6.3	Trabalho futuro .....	78
6.4	Considerações finais.....	79
	<b>Referências bibliográficas.....</b>	<b>83</b>



## LISTA DE FIGURAS OU ILUSTRAÇÕES

Figura 1.1 - Objectivo do Projecto. Fonte: Elaboração própria. ....	4
Figura 1.2 - Cronograma do Projecto. Fonte: Elaboração própria.....	5
Figura 2.1 - Diagrama do ciclo de vida de um DW. Fonte: Adaptado de (Kimball et al., 2002).....	11
Figura 2.2 - Orientação por assunto. Fonte: Adaptado de (Inmon, 2005). ....	12
Figura 2.3 - Integração de dados. Fonte: Elaboração própria.....	13
Figura 2.4 - Não volatilidade. Fonte: Adaptado de (Inmon, 2005). ....	14
Figura 2.5 - Variação no tempo. Fonte: Elaboração própria. ....	14
Figura 2.6 - Nível de detalhe da granularidade. Fonte: Adaptado de (Inmon, 2005). ....	17
Figura 2.7 - Esquema em Estrela. Fonte: Elaboração própria.....	18
Figura 2.8 - Esquema em Floco de Neve. Fonte: Elaboração própria. ....	19
Figura 2.9 - Esquema em Constelação. Fonte: Elaboração própria. ....	19
Figura 2.10 - Representação visual de um cubo. Fonte: Elaboração própria. ....	22
Figura 3.1 - Vista aérea da ESAL. Fonte: ESAL. ....	28
Figura 3.2 - Perspectivas de desempenho do BSC. Fonte: Adaptado de (Kaplan & Norton, 1992).....	32
Figura 3.3 - Modelo de BSC para organizações sem fins lucrativos. Fonte: (Kaplan & Norton, 2000). ....	33
Figura 3.4 - Perspectivas BSC da ESAL. Fonte: Elaboração própria. ....	34
Figura 3.5 - Infra-estrutura tecnológica do DW. Fonte: Elaboração própria. ....	37
Figura 3.6 - Plataforma <i>Business Intelligence</i> da <i>Microsoft</i> . Fonte: Elaboração própria.....	38
Figura 4.1 - Diagrama de caso de uso do protótipo DW. Fonte: Elaboração própria. ....	42
Figura 4.2 - Diagrama de pacotes. Fonte: Elaboração própria.....	42
Figura 4.3 - Diagrama de componentes. Fonte: Elaboração própria. ....	44
Figura 4.4 - Modelo conceptual inicial do <i>Data Warehouse</i> . Fonte: Elaboração própria. ....	45
Figura 4.5 - Modelo conceptual final do <i>Data Warehouse</i> . Fonte: Elaboração própria.....	46
Figura 4.6 - Esquema em estrela dos dados dos Cursos Profissionais. Fonte: Elaboração própria. ....	46
Figura 4.7 - Modelo lógico do <i>Data Warehouse</i> . Fonte: Elaboração própria. ....	47
Figura 4.8 - Diagrama de classes do modelo dimensional. Fonte: Elaboração própria. ....	48

Figura 4.9 - Dimensão <i>DimData</i> . Fonte: Elaboração própria. ....	49
Figura 4.10 - Representação do <i>DataPeriodo</i> no modelo multidimensional. Fonte: Elaboração própria. ....	49
Figura 4.11 - Parte dos atributos da dimensão <i>DimAluno</i> . Fonte: Elaboração própria. ....	50
Figura 4.12 - Parte dos atributos da dimensão <i>DimMatricula</i> . Fonte: Elaboração própria. ....	51
Figura 4.13 - Representação gráfica de uma hierarquia. Fonte: Elaboração própria. ....	52
Figura 4.14 - Representação da informação de uma hierarquia. Fonte: Elaboração própria. ....	53
Figura 4.15 - Dimensão <i>DimDisciplina</i> . Fonte: Elaboração própria. ....	53
Figura 4.16 - Dimensão <i>DimProfessor</i> . Fonte: Elaboração própria. ....	54
Figura 4.17 - Atributos da tabela factó <i>FactoAvaliacao</i> . Fonte: Elaboração própria. ....	55
Figura 4.18 - Atributos da tabela factó <i>FactoProfissional</i> . Fonte: Elaboração própria. ....	57
Figura 5.1 - Ciclo de vida do protótipo de DW. Fonte: Elaboração própria. ....	60
Figura 5.2 - Infra-estrutura tecnológica do protótipo de DW. Fonte: Elaboração própria. ....	61
Figura 5.3 - Criação da tabela <i>FactoAvaliacao</i> do DW. Fonte: Elaboração própria. ....	62
Figura 5.4 - Estrutura de tabelas do DW. Fonte: Elaboração própria. ....	63
Figura 5.5 - Processo de ETL. Fonte: Elaboração própria. ....	64
Figura 5.6 - Criação de pacotes de dados no <i>Business Intelligence Development Studio</i> . Fonte: Elaboração própria. ....	65
Figura 5.7 - Cubo de dados gerado no <i>Business Intelligence Development Studio</i> . Fonte: Elaboração própria. ....	66
Figura 5.8 - Teste do Cubo de dados no ambiente de trabalho. Fonte: Elaboração própria. ....	67
Figura 5.9 - Construção de relatórios no ambiente de trabalho. Fonte: Elaboração própria. ....	68
Figura 5.10 - Interface do <i>front-end</i> de disponibilização de relatórios. Fonte: Elaboração própria. .	69
Figura 5.11 - Relatório pré-definido da secção Relatórios Estáticos. Fonte: Elaboração própria. ....	70
Figura 5.12 - Relatório <i>Ad-Hoc</i> elaborado no <i>Report Builder 2.0</i> . Fonte: Elaboração própria. ....	71
Figura 5.13 - Exportação de relatórios para outros formatos. Fonte: Elaboração própria. ....	72
Figura 5.14 - Relatório produzido através do <i>Excel Services</i> . Fonte: Elaboração própria. ....	73

## LISTA DE TABELAS

Tabela 2.1 - Base dados transaccional vs multidimensional. Fonte: Elaboração própria. ....	15
Tabela 2.2 - Tipos de metadados e sua proveniência. Fonte: Elaboração própria.....	24
Tabela 3.1 - Análise SWOT. Fonte: Elaboração própria.....	31
Tabela 3.2 - Objectivos Estratégicos e Indicadores. Fonte: Elaboração Própria. ....	35
Tabela 4.1 - Descrição das métricas. Fonte: Elaboração própria.....	56



## ABREVIATURAS OU SIGLAS

BI	<i>Business Intelligence</i>
BSC	<i>Balanced Scorecard</i>
DBMS	<i>Data Base Management System</i>
DM	<i>Data Mart</i>
DW	<i>Data Warehouse</i>
E-R	Entidade Relacionamento
ESAL	Escola Secundária/3 de Amato Lusitano
ETL	<i>Extract, Transform and Load</i>
MDX	<i>MultiDimensional eXpressions</i>
OLAP	<i>Online Analytical Processing</i>
OLTP	<i>Online Transaction Processing</i>
SQL	<i>Structured Query Language</i>
SWOT	<i>Strengths, Weaknesses, Opportunities, and Threats</i>



# 1 Introdução e objectivos

---

## 1.1 Introdução

Nas últimas décadas, tem-se assistido a grandes mudanças nos mais variados campos do conhecimento. A crescente globalização, que tem atingido todos os segmentos da sociedade, leva a um relacionamento entre “cliente” e “fornecedor” mais complexo, pois quem ousa definir regras é o cliente. Este, tem consciência do seu poder e tem à sua disposição todas as informações necessárias, podendo assim comparar com os concorrentes locais, regionais e globais.

Esta descrição, também ela se adapta à realidade das escolas, privadas ou públicas, que até então viviam num ambiente protegido e de poucas mudanças.

As organizações que não introduziram na sua cultura conceitos de excelência nos seus produtos e serviços, incluindo o atendimento ao cliente, podem estar a comprometer a sua sobrevivência (Santos, 2008). As influências destas mudanças também estão a ser vividas pela comunidade escolar.

As escolas, que até então tinham a sua preocupação voltada para a área pedagógica, encontraram dificuldades nesta nova maneira de abordar o seu “cliente” (pais e filhos/alunos), que agora, por norma, está mais criterioso: na escolha do serviço que será prestado; na qualidade do ensino ministrado e de como ele chega; quais as tecnologias adoptadas; qual o método de ensino entre outros.

Este novo perfil do “cliente” leva-nos a uma forte reflexão sobre algumas velhas práticas que já não são eficazes.

A palavra de ordem passa a ser qualidade. Há que compreender que não basta reajustar o contexto interno no que diz respeito à visão pedagógica. É preciso um planeamento estratégico que dê suporte a todas estas mudanças, podendo assim ir além das fronteiras da própria escola e integrar-se no cenário actual.

As escolas começam já a prestar mais atenção ao que se passa na sociedade e algumas delas já se movimentam para planear a estratégia que possa anexar valor para o “cliente”, o aluno (Estevão, 1999).

Uma das principais mudanças a incutir no cenário da educação é que, além da visão pedagógica, a escola precisa de conquistar também uma forte visão empresarial (Estevão, 1999). Isto significa que a visão pedagógica deve ser complementada com a visão estratégica para que possa dar sustento à organização, assegurando-lhe condições de competir num mercado altamente disputado, como tem sido o da educação.

Não se pretende com esta visão estratégica descaracterizar os aspectos culturais da visão de sempre atribuída à organização escola, mas, o que se propõe é, complementá-la visando o crescimento empresarial, sem nunca esquecer que a sua missão é prestar serviços educacionais de modo a desenvolver as potencialidades do educando, assegurando-lhe a formação para o exercício da cidadania, progresso no trabalho e na vida académica, e a consciência do seu papel como agente da história.

É importante que se reflita sobre a importância de um complemento da visão pedagógica à visão estratégica, a fim de maior eficácia aliada ao conhecimento para enfrentar o mercado inovador e competitivo de hoje.

Mas, a aplicabilidade directa de abordagens do planeamento e gestão estratégica tal como as conhecemos do sector empresarial não são de todo viáveis no campo da educação, pois estas têm por objectivo o aumento da produtividade e competitividade, cujas variáveis de manipulação se prendem com os mercados e produtos visando margens de lucro (Estevão, 1999).

Acresce que há toda uma tradição no planeamento e definição de políticas e orientações educativas centralizadas nos serviços centrais do Ministério da Educação, ficando apenas a operacionalização das mesmas à responsabilidade de cada escola adaptando-as à sua realidade, sem contudo existir um departamento de planeamento onde se desenvolvam estratégias próprias.

Atendendo a esta cultura de dependência de decisões estratégicas impostas do exterior, tão enraizada, tem sido difícil a implementação de uma gestão estratégica diferenciadora nas escolas, que poderia viabilizar-se expressando-se em projectos de escola eficazes se estas tivessem verdadeira autonomia.

## 1.2 Definição do tema

O *Data Warehouse* como instrumento de apoio ao planeamento, auxiliando a criação de cenários de decisão em instituições de ensino, em particular na Escola Secundária/3 de Amato Lusitano de Castelo Branco. Pretende-se, assim, integrar as diferentes fontes de informação dispersas na escola, num único local, proporcionando o fácil acesso e uma rápida resposta a questões que se lhe coloquem.

## 1.3 Organização da dissertação

A presente dissertação encontra-se estruturada em sete capítulos.

Além do presente capítulo que pretende enquadrar o projecto contextualizando o tema da dissertação e os objectivos propostos, o segundo capítulo apresenta uma revisão da literatura onde são abordados os principais conceitos que irão dar sustento ao suporte teórico da pesquisa efectuada para a concretização do protótipo de *Data Warehouse* com a apresentação de metodologias de autores conceituados na área em causa.

No terceiro capítulo é feita uma caracterização da ESAL, caso de estudo real que sustenta o protótipo e ainda é apresentada a infra-estrutura tecnológica que irá ser implementada.

O quarto capítulo aborda a modelação do DW segundo o paradigma da modelação utilizado na linguagem *Unified Modeling Language* para o modelo multidimensional, abordando ainda alguns conceitos base.

Segue-se no capítulo quinto, com um conteúdo ainda mais específico, o desenvolvimento detalhado do protótipo de DW, incluindo a sua especificação, implementação e metodologia utilizada.

Por último, o sexto capítulo aborda conclusivamente os passos efectuados durante a execução do protótipo, uma reflexão crítica onde são apresentados os principais contributos deste trabalho de projecto, reconhecendo-lhe algumas limitações e apontando sugestões para trabalhos futuros.

## 1.4 Enquadramento do trabalho

O projecto descrito neste relatório visou a construção de um protótipo de *Data Warehouse*, ou seja, uma plataforma de análise multidimensional com um ambiente preciso de dados, para dar suporte aos processos/questões quer internos quer de avaliação exigidos pelo Ministério da Educação e facilitar o processo de análise de correlação, multivariado, estatístico ou outro, aos dados internos da ESAL, quer ao Director ou a qualquer outro elemento pedagógico da comunidade escolar.



Figura 1.1 - Objectivo do Projecto. Fonte: Elaboração própria.

O principal objectivo deste projecto foi a construção de uma infra-estrutura multidimensional como protótipo base, que permitisse o suporte supracitado (Figura 1.1).

## 1.5 Objectivos do trabalho

O estudo proposto por esta dissertação, pretende sensibilizar no interesse e vantagem da construção, implementação, gestão e utilização de um sistema de apoio à decisão em escolas, em particular na ESAL.

De uma forma geral, os objectivos a considerar neste trabalho foram:

- Projectar e desenvolver um protótipo de *Data Warehouse*;
- Aplicar as tecnologias de modelação multidimensional;
- Seleccionar, preparar, limpar e carregar o *Data Warehouse* com os dados extraídos dos sistemas transaccionais da Escola Secundária/3 de Amato Lusitano, para análise através de tecnologias OLAP;
- Produzir e disponibilizar relatórios que possibilitem extrair informação coerente.

## 1.6 Cronograma

O plano de trabalho proposto para o desenvolvimento do projecto pode ler-se na Figura 1.2.

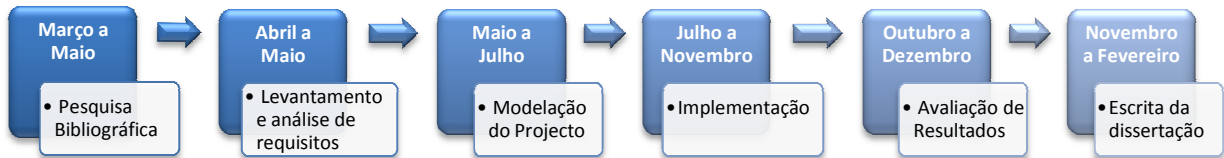


Figura 1.2 - Cronograma do Projecto. Fonte: Elaboração própria.

Fases do Projecto:

Fase 1 - Pesquisa bibliográfica sobre o tema, bem como análise pormenorizada do estado da arte no que diz respeito à implementação e operacionalização do *Data Warehouse*.

Fase 2 - Tem por objectivo a análise e especificação dos requisitos funcionais do *Data Warehouse*. É uma fase, antes de mais, de conhecimento da ESAL e das necessidades de análise e de informação dos utilizadores finais. Será em função destes requisitos, e da análise mais profunda dos sistemas operacionais existentes, que deverão ser então, esboçados o modelo conceptual e lógico do *Data Warehouse*.

Fase 3 - Compreensão das fontes e bases de dados operacionais e escolha da arquitectura e infra-estrutura tecnológica a implementar. Nesta terceira fase pretende-se efectuar a modelação multidimensional que constituirá o *Data Warehouse*. Esta modelação, conjuntamente com uma análise profunda das fontes de dados, tanto ao nível de sistemas que os originam como dos próprios dados em si, deverá ser feita de modo a poder elaborar uma estratégia para a selecção, extracção, transformação e carregamento dos dados.

Fase 4 - Montagem de toda a infra-estrutura do *Data Warehouse*. É feito o desenvolvimento dos componentes de suporte à operacionalização do *Data Warehouse* destacando-se os processos de extracção, transformação e integração de dados. É ainda nesta fase que se fará a análise multidimensional, as interrogações em *Structured Query Language* (SQL) e *MultiDimensional eXpressions* (MDX) aos cubos de dados e a transposição dos resultados para os relatórios.

Fase 5 - Disponibilização dos relatórios produzidos através de uma ferramenta colaborativa, avaliando os resultados obtidos.

Fase 6 - Consolidação da escrita da dissertação.

## 2 Revisão bibliográfica

---

### 2.1 Introdução

Neste capítulo são abordados os principais conceitos que irão dar sustento ao suporte teórico da pesquisa efectuada para a concretização do protótipo de *Data Warehouse* (DW).

Assim, é feito o enquadramento teórico e histórico dos DW, passando à sua caracterização segundo alguns autores, em especial Inmon e Kimball que norteiam o estudo de DW, dando alguma ênfase às características base apontadas por Inmon, como a orientação por assunto, a integração, a não volatilidade e a variação no tempo dos dados. É feita em seguida uma descrição do conceito de Modelação Multidimensional (MM), explicando em que consiste esta técnica aplicada aos DW, com os vários esquemas possíveis, estrela, floco de neve e constelação.

Faz-se uma abordagem ao processo de *Extraction, Transformation and Loading* (ETL) dos dados dos sistemas transaccionais para o ambiente de DW, onde após esta caracterização é introduzido o conceito de cubo de dados bem como das técnicas de análise multidimensionais associadas à navegação entre os diferentes níveis de granularidade de um cubo de dados.

Por fim é explicado o conceito de metadados, os diversos tipos e a sua proveniência no processo de desenvolvimento de um DW.

Este capítulo termina falando sobre as arquitecturas, abordagens e metodologias de implementação de um DW segundo os autores Inmon e Kimball.

### 2.2 Enquadramento teórico

Para a compreensão do conceito de *Data Warehouse* no âmbito das Tecnologias da Informação, é necessário retroceder no tempo e lembrar a evolução dos processos tecnológicos nas últimas décadas. Terminada a década de 60, o uso de computadores tornava-se indispensável nas grandes organizações. Na época, apenas era possível correr uma aplicação de cada vez, onde cada

aplicação era executada sobre ficheiros mestres. As aplicações resumiam-se na sua maioria, a relatórios e programas concebidos em COBOL. O uso de cartões perfurados era comum. Os ficheiros mestres eram armazenados em fitas magnéticas, adequadas para o armazenamento de um grande volume de dados a baixo custo, apresentando o inconveniente de apenas permitirem um acesso sequencial (Machado, 2004).

No início da década de 70, dá-se uma revolução no campo do armazenamento e acesso a dados com a chegada da tecnologia de armazenamento em disco. Desenvolve-se igualmente um novo *software*, os sistemas de gestão de bases de dados que aliados à nova tecnologia de armazenamento, geram uma mudança de visão e de paradigma das organizações que passam para um processamento baseado em dados, centralizando todas as actividades em computadores (Machado, 2004).

Os computadores passam a conferir vantagem competitiva e maior eficiência às organizações, iniciando-se uma nova era no sector dos negócios, sustentados com este desenvolvimento dos sistemas de informação.

Profundos aperfeiçoamentos tecnológicos, nas décadas de 70 e 80, promoveram novos sistemas de informação a menor custo e bem mais poderosos. As pessoas nas organizações passaram a ter acesso aos microcomputadores e a rigidez no seu acesso diminuiu consideravelmente. A implementação de novas tecnologias, a banalização na utilização de computadores e o desenvolvimento de linguagens de quarta geração, conduziram a um papel mais activo da parte do utilizador final, controlando este, directamente os sistemas e os dados fora do domínio do processamento de dados (Machado, 2004). As organizações perceberam no seguimento deste evoluir tecnológico que, os seus dados poderiam ser analisados de uma forma optimizada, ou seja, descobriram que podiam incrementar os seus recursos com processos de análise de dados.

Com a alteração da visão que até então era atribuída ao conjunto de informações - sistemas, surge a distinção entre sistemas que tratam o negócio, que são os que sustentam o dia-a-dia do negócio da organização e que garantem a sua operacionalidade, os chamados Sistemas Transaccionais e os sistemas que analisam o negócio, que são os sistemas que ajudam a interpretar o que ocorreu e a decidir sobre estratégias futuras para a organização, designados por Sistemas de Apoio à Decisão (Turban et al., 2005).

Várias são as solicitações que são colocadas aos sistemas transaccionais com o desenvolvimento de novas ferramentas de análise da informação, mas com alguma ineficácia, dado que os sistemas foram originalmente desenvolvidos para garantir a operacionalidade da organização, não estando aptos para gerar e armazenar as informações estratégicas necessárias a processos analíticos de dados, também designados por BI - *Business Intelligence*. Dando resposta a este novo paradigma, são desenvolvidos programas extractores, que permitem copiar dados dos sistemas

transaccionais, com o intuito de os trabalhar noutros ambientes. Na grande maioria essas extracções ocorriam para ficheiros intermediários, onde as informações eram trabalhadas, acabando por originar uma falha na integridade das informações, provocando muitas vezes, uma falta de credibilidade dos dados, repercutindo-se na produtividade e na informação publicada com valores diferentes. Não esquecendo ainda o imenso volume de dados gerados pelos sistemas transaccionais e a diversidade dos sistemas implantados nas organizações, que levavam a que as pesquisas realizadas fossem produzidas muito lentamente (Machado, 2004).

Os relatórios produzidos como resultado dessas pesquisas, na época, eram sinónimos de muitas horas no computador, pois era necessário que fossem extraídos os dados de vários sistemas, muitas vezes incompatíveis entre si. Apesar dos motivos indicados, é de realçar que é possível a prática de BI com os sistemas operacionais da organização, bem como com outras fontes de dados como folhas de cálculo, embora se trate de procedimentos que implicam a possibilidade de equívocos, já que esses dados são provenientes de fontes independentes, e sem relação de integridade entre eles. No âmbito das decisões estratégicas, falhava também o registo dos factos históricos nos sistemas transaccionais, pois estes trabalhavam com uma situação instantânea dos negócios. Para solucionar esta situação, começou por se analisar a forma como centralizar numa base de dados a informação dos sistemas transaccionais, de modo a integrar todos os dados da organização. Era ainda necessário manter inalterado o histórico das informações e permitir que a base de dados fosse disposta dimensionalmente, ou seja, que o analista pudesse visualizar um mesmo facto através de diferentes dimensões. A essa modalidade de Sistema de Apoio à Decisão foi dado o nome de DW que não é um *software* ou produto de *hardware*, mas sim um ambiente computacional onde os utilizadores controlam directamente os dados que necessitam para melhor tomar as suas decisões. De realçar que as perspectivas e técnicas necessárias para projectar um DW são profundamente diferente dos sistemas transaccionais. Os utilizadores, o conteúdo dos dados, a estrutura dos dados, o *hardware* e o *software*, a administração, a gestão dos sistemas, o ritmo diário, as solicitações, as respostas e o volume de informações são diferentes. Numa economia globalizada, onde cada vez mais a concorrência se estreita, o entendimento dessa tecnologia poderá ajudar as organizações a descobrir novos caminhos e tendências para melhor poder competir, resultando em melhores serviços, em menor tempo e sem grande aumento de custos (Inmon, 2005).

Foi, pela década de 1990, que surgiram os primeiros conceitos documentados para os DW, propostos pelos seus pioneiros e visionários Ralph Kimball e William Inmon. Estes princípios norteadores são ainda hoje a base de todos os Sistemas de *Data Warehouse*, facultando as intenções e motivos subjacentes aos projectos de criação de DW (Silvers, 2008).

## 2.3 Caracterização de um *Data Warehouse*

O DW é definido, clara e objectivamente em (Inmon, 2005), como um sistema de dados orientado por assuntos, integrados, não voláteis e variantes no tempo. Um DW tem por objectivo integrar dados de diferentes fontes e formatos, não sendo construído para suportar o processo funcional ou operacional da empresa, ou seja, não é o fim, mas é o meio para facilitar o uso da informação (Kimball et al., 1998). Estas características quando devidamente exploradas, afirmam o DW como um meio tecnológico capaz de conceder vantagens estratégicas às organizações no que respeita à tomada de decisão, por isso, interessa que a sua implementação seja um sucesso e a sua manutenção uma realidade. O sucesso destes sistemas de dados prevê a materialização de um conjunto de pressupostos, como sejam entre outros: o alinhamento com a estratégia da organização; a engenharia de requisitos adequada; a solidez e disponibilidade tecnológica que alicerça o DW; as técnicas de modelação utilizadas e a sustentação em processos de extracção e refrescamento dos dados de modo eficaz e eficiente.

Os sistemas de DW situam-se no mais alto plano da acção estratégica das organizações, pelo que apresentam características muito específicas no domínio dos sistemas de informação, afastando-se do plano estritamente operacional e logístico. A forma como este tipo de sistema interage com os seus utilizadores imediatos, baseia-se essencialmente nos meios humanos condutores do negócio, os agentes de decisão, e é realizada apenas com o intuito de lhes proporcionar informações. Os utilizadores apresentam-se assim, como meros consumidores de informação e conhecimento. A interacção com o sistema, é feita geralmente através de ferramentas que potenciam o processamento analítico dos dados *Online Analytic Processing* (OLAP). Os sistemas de DW assumem o papel de concentradores dos vários dados dispersos pela organização, facultando-os de um modo coeso e adequado, posto que são armazenados de uma forma integrada e homogeneizada.

Desta forma, a tomada de decisões, assenta primordialmente nos resultados oferecidos pelos sistemas de DW. Atinge-se um patamar onde são exigidos resultados incontestáveis, no que respeita à correcção, relevância, acessibilidade, completude, segurança, frescura e consistência das informações, balizando uma área de conhecimento específica.

O mercado actual acarreta o dinamismo das acções do negócio e implica, entre outros, o manuseamento de dados de origem interna e externa aos sistemas informáticos; o cruzamento destes pelos diversos actores da organização; a prontidão da resposta; a sintonia entre os indicadores organizacionais gerais e a pormenorização dos dados que suportam esses indicadores. O Sistema Operacional, com todas as suas aplicações de trabalho diário nas organizações, não responde adequadamente perante estas solicitações e como tal não satisfaz os anseios dos actores da organização, atendendo a que a natureza primária destes sistemas se encontra orientada para a organização e armazenamento dos dados (Kimball et al., 2002). A sua estrutura interna apresenta

dificuldades impeditivas de manipulação eficaz e eficiente dos enormes volumes de dados armazenados nas organizações e, normalmente, necessários para a tomada de decisão. Atendendo à estrutura não volátil e integrada dos dados que caracterizam os sistemas de DW, são assim superadas estas dificuldades. Deste modo os sistemas de DW, permitem interrogações complexas orientadas aos assuntos da organização, focalizadas em espaços temporais mais ou menos alargados e agindo de modo integrado (Inmon, 2005).

A construção de um DW é o processo que combina as necessidades de informação de uma comunidade de utilizadores com os dados que realmente estão disponíveis (Kimball et al., 2002). O desenvolvimento de um DW é uma actividade complexa e deve ser cuidadosamente planeado. Construí-lo não é apenas um exercício de integração de sistemas (Singh, 2001). Essa construção está apoiada num processo sistemático, integrando algoritmos, ferramentas e técnicas, até uma arquitectura especialmente concebida, para facilitar o armazenamento de grandes volumes de dados e viabilizar a obtenção de informação direccionada à tomada de decisão a partir da execução de consultas complexas, num restrito tempo de resposta. Este processo sistemático, no qual é desenvolvido um DW, é denominado de *Data Warehousing*.

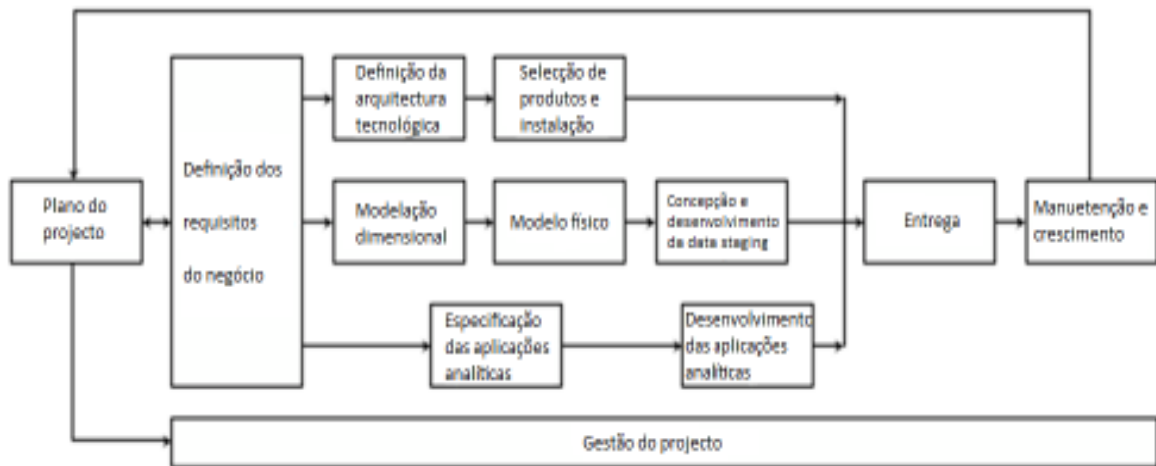


Figura 2.1 - Diagrama do ciclo de vida de um DW. Fonte: Adaptado de (Kimball et al., 2002).

Um DW tem algumas características importantes, como a orientação por assunto, a integração, o facto de não ser volátil e a variação no tempo. Essas são as principais características do ambiente DW (Inmon, 2005).

### 2.3.1 Orientação por assunto

Enquanto as bases de dados operacionais têm os seus dados centrados nas aplicações das organizações, ou seja nas operações do dia-a-dia da organização, o DW tem os seus dados centrados

nos assuntos da empresa que são relevantes ao processo de tomada de decisão. Antes de iniciar a construção de um DW, na fase de análise, é muito importante que se discuta com os utilizadores quais são os seus objectivos, para conseguir fazer uma classificação por assunto.

Outra diferença entre os dados orientados a aplicações e o DW, é que os orientados a aplicações possuem detalhes que satisfazem os requisitos imediatistas do processamento funcional, enquanto que os DW têm por objectivo tratar assuntos estratégicos para a organização. Assunto é, portanto, o conjunto de informações relativas a determinada área estratégica de uma organização (Machado, 2004). Assim, para uma melhor tomada de decisão, as informações devem estar organizadas por assuntos.

Durante o desenvolvimento de um DW deve procurar-se a identificação desses assuntos como forma de facilitar o uso do sistema no futuro, como se ilustra na Figura 2.2. A orientação por assunto dos dados permite ainda a um DW manter intacta a sua arquitectura no decorrer do seu ciclo de vida, permitindo-lhe absorver mudanças inevitáveis sem contudo sofrer mudanças drásticas na sua arquitectura (Silvers, 2008).



Figura 2.2 - Orientação por assunto. Fonte: Adaptado de (Inmon, 2005).

### 2.3.2 Integração

A integração de dados é uma das características mais importantes de um DW. Todo o processo de alimentação dos dados é efectuado de forma que as muitas inconsistências das aplicações sejam desfeitas. Como o DW em geral é alimentado por múltiplas fontes diferentes de dados, para que estas possam interagir é necessária uma representação única para estes dados. É assim necessário fazer a integração dos dados das diversas fontes, para que se obtenha a consistência de nomes, a

consistência de variáveis de medida, a consistência da codificação das estruturas, a consistência dos atributos físicos dos dados, entre outros (Figura 2.3). Para tal, utilizam-se técnicas de limpeza e integração dos dados para assegurar essa consistência.



Figura 2.3 - Integração de dados. Fonte: Elaboração própria.

Para que os dados sejam inseridos num DW é preciso que estes sejam integrados, onde essa integração é tida como o processo pelo qual os dados são modificados permitindo a sua inserção no repositório.

### 2.3.3 Não volátil

Os dados, uma vez escritos num DW, nunca são excluídos ou actualizados (Silvers, 2008). Num DW os dados não sofrem quaisquer alterações, armazenando o histórico imutável do ambiente organizacional.

No ambiente operacional, os dados estão em constante actualização. Contudo, os dados existentes no DW apresentam um conjunto de características muito diferentes, estes são carregados, normalmente em grandes quantidades, e apenas acedidos (Figura 2.4). A actualização dos dados, geralmente, não é comum num ambiente de *Data Warehouse* (Inmon, 2005).

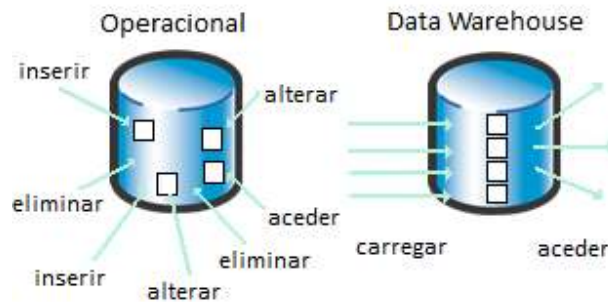


Figura 2.4 - Não volatilidade. Fonte: Adaptado de (Inmon, 2005).

Nos sistemas operacionais, são realizadas diversas operações sob uma base de dados, num DW apenas duas operações são lícitas: inserção e consulta de dados. A operação de inserção consiste nos processos de integração dos dados provenientes das diversas fontes no DW, após as transformações necessárias para adaptação da informação. Já as operações de consulta desencadeiam a apresentação desses dados.

#### 2.3.4 Variação no tempo

Um *Data Warehouse* é variável com o tempo atendendo a que os dados nele contidos se referem a um momento específico. Isto significa, que quando o dado é carregado, recebe na sua chave, uma unidade temporal. Os analistas de negócios em sistemas com este tipo de armazenamento conseguem fazer análises de tendências, podendo visualizar num determinado horizonte temporal as variações de uma informação específica (Inmon, 2005). O horizonte de tempo dos sistemas transaccionais é relativamente pequeno em relação ao DW (Figura 2.5).



Figura 2.5 - Variação no tempo. Fonte: Elaboração própria.

A variação no tempo é uma característica que permite a análise temporal das informações. Esta característica está associada à forma como os dados são inseridos num DW. A carga dos dados pode acontecer anualmente, mensalmente, semanalmente, variando de acordo com a necessidade de cada organização. Todos os DW's irão possuir uma análise relacionada a informações temporais.

Um DW pode ser pensado como uma série sofisticada de instantâneos, onde cada instantâneo é recolhido num dado momento no tempo. O efeito criado pela série de instantâneos é que o DW tem uma sequência histórica de actividades e eventos, algo não de todo evidente num ambiente operacional, em que apenas os mais actuais são encontrados (Inmon, 2005). Isso significa que é possível uma análise do desempenho da organização no tempo.

Os sistemas transaccionais de uma organização estão preparados para disponibilizar informação no exacto momento do acesso. Alguns desses sistemas também realizam consultas com variação temporal, porém a principal diferença é a volatilidade da informação. A informação é actualizada no momento do acontecimento e o registo temporal é modificado. Porém, num DW, os dados nele armazenados, não sofrem actualizações, permanecendo inalterada a informação para sempre, característica entre outras realçadas na Tabela 2.1. Desta forma, uma análise sobre determinado assunto num período temporal mais ou menos alargado, apresentará sempre o mesmo resultado.

CARACTERÍSTICAS	Base Dados Transaccional	Base Dados Multidimensional
Finalidade	Registo de operações diárias	Análise do negócio
Uso	Operacional	Informativo
Tipo de processamento	OLTP	OLAP
Unidade de trabalho	Inserir, alterar, eliminar, consultar	Carregar, consultar
Número de utilizadores	Milhares	Centenas
Tipo de utilizador	Operadores	Decisores
Condições dos dados	Dados operacionais	Dados Analíticos
Volume	Megabytes - gigabytes	Gigabytes - terabytes
Histórico	60 a 90 dias	5 a 10 anos
Granularidade	Detalhados	Detalhados e resumidos
Redundância	Não ocorre	Ocorre
Estrutura	Estática	Variável
Manutenção desejada	Mínima	Constante
Acesso a registos	Dezenas	Milhares
Actualização	Contínua (tempo real)	Periódica (em <i>batch</i> )
Integridade	Transacção	A cada actualização

Tabela 2.1 - Base dados transaccional vs multidimensional. Fonte: Elaboração própria.

Com o evoluir dos tempos, as regras do negócio da organização podem sofrer alterações devido a mudanças impostas pela competitividade ou por outras circunstâncias, mas o histórico da informação permanecerá inalterado.

Os metadados, definição de dados sobre dados ou de regras sobre os dados, devem ser actualizados de forma a acompanharem o evoluir das regras do negócio, sendo imperativo que as regras antigas se mantenham para que os órgãos de decisão consigam realizar uma análise coerente e correcta para todo o histórico organizacional.

## 2.4 Modelação Multidimensional

Tradicionalmente, os sistemas de dados são representados através da modelação Entidade-Relacionamento (E-R). Este tipo de modelação remonta aos anos 70 tendo sido proposto na época como um caminho para unificar as relações entre bases de dados. O modelo evoluiu e actualmente é utilizado sem restrição para o desenho de bases de dados. O Modelo Entidade-Relacionamento é um modelo em rede que descreve através de diagramas e conceitos os dados armazenados num sistema com um alto nível de abstracção (Yourdon, 1992). Esta técnica baseada em diagramas que tem sido amplamente usada na documentação de sistemas transaccionais, não pode ser aplicada aos sistemas analíticos dadas as características intrínsecas do DW. O modelo E-R não facilita o processo de compreensão do modelo pelos utilizadores (Kimball et al., 2002). Esse modelo é orientado à não redundância de dados, através da normalização das tabelas, enquanto que por definição um *Data Warehouse*, por ser um armazém de dados não volátil, tem redundância de dados. Em sistemas de DW, por não haver actualização de dados, a redundância demonstra ser um artifício para maximizar a eficiência das consultas à base de dados (Kimball et al., 2002). Atendendo a esses problemas, surgiu a necessidade de aplicar um novo processo de modelação, onde os dados do negócio da organização tivessem fácil compreensão pelos utilizadores. O Modelo Multidimensional é definido por uma técnica de concepção e visualização de um modelo de dados, onde através de um conjunto de medidas se descrevem aspectos comuns de negócios. Esta técnica é utilizada em particular para resumir e reestruturar dados que facilitem o cálculo e a selecção das informações e a sua apresentação segundo visões que suportem uma análise desses mesmos dados. Um modelo multidimensional é formado por três elementos básicos: factos, dimensões e medidas (Machado, 2004).

Em vários aspectos a Modelação Multidimensional é mais simples, mais expressiva e mais fácil de entender comparativamente à modelação ER.

A tabela de factos é constituída por atributos numéricos e pelas chaves estrangeiras que a ligam às tabelas de dimensões. Esta tabela está bastante normalizada e contém geralmente uma enorme quantidade de registos. As tabelas de dimensões contêm os atributos descritores dos

atributos da tabela de factos, e existem tantas, quantas vertentes sob as quais se pretende analisar os factos, pois são as diferentes perspectivas envolvidas. Estas tabelas são fortemente desnormalizadas, possuindo geralmente muitos atributos, contendo no entanto poucos registos quando comparadas com a tabela de factos. Ao contrário dos factos, as dimensões normalmente não possuem atributos numéricos pois são basicamente descrições ou classificações referentes aos elementos que participam de um facto. Uma característica importante de destaque é que as dimensões podem ser divididas em hierarquias.

As medidas, ou métricas, são os atributos numéricos que representam um facto. Através delas é possível analisar o desempenho de um indicador de negócios relativo às dimensões que participam do facto. Uma medida é determinada pela combinação das dimensões que participam de um facto. Desenvolver um DW é uma questão de unir as necessidades dos seus utilizadores com a realidade dos dados disponíveis (Kimball et al., 1998).

O nível de detalhe ou de resumo dos dados armazenados no DW, isto é, a granularidade, é uma preocupação que tem de ser definida de acordo com o nível de detalhe da informação a ser obtida (Inmon, 2005), estando dependente dos objectivos da organização. A definição da granularidade influencia o volume de dados armazenados no DW, bem como o desempenho do tipo de consulta a executar, como o ilustra a Figura 2.6. Quanto mais detalhado for o dado, mais baixa será a granularidade, quanto menos detalhado for, mais alto será o nível de granularidade. Quando é definido um nível de detalhe muito baixo, é preciso ter em conta que estes dados já não poderão ser decompostos para uma análise com um maior nível de detalhe. No entanto, será sempre possível agrupar os dados se a granularidade for baixa para que sejam visualizados com um menor nível de detalhe, mais sumarizados.



Figura 2.6 - Nível de detalhe da granularidade. Fonte: Adaptado de (Inmon, 2005).

O Modelo Multidimensional representa o registo de dados em tabelas, relacionando factos e dimensões. Esta forma de modelar a informação pode apresentar diferentes arranjos, denominados esquemas. De acordo com a complexidade das relações existentes entre tabelas de factos e tabelas de dimensões, os esquemas apresentam diferentes configurações, referimo-nos ao esquema em estrela, ao esquema em floco de neve e ao esquema em constelação.

O esquema em estrela (Figura 2.7) é o que apresenta o modelo mais simples. Possui uma tabela central com os factos à qual estão ligadas as tabelas das dimensões, não existindo quaisquer relações entre as tabelas de dimensões.

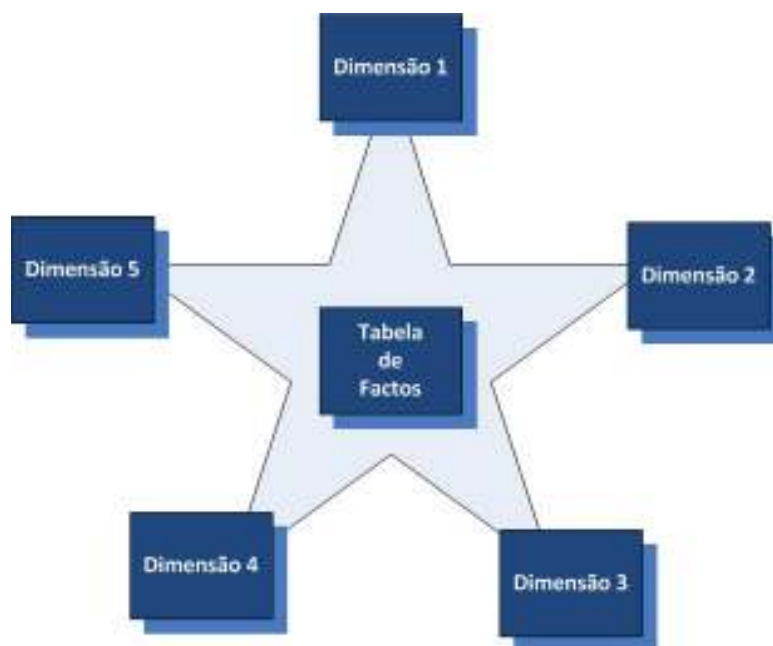


Figura 2.7 - Esquema em Estrela. Fonte: Elaboração própria.

O esquema em floco de neve (Figura 2.8) é um refinamento do esquema em estrela, onde parte da hierarquia dimensional é normalizada num conjunto de tabelas de dimensão mais pequenas, numa disposição alusiva a um floco de neve. Neste esquema as tabelas de dimensão são normalizadas em múltiplas tabelas de pesquisa, cada uma representando um nível na hierarquia da dimensão. Devido ao facto de estar normalizado, o modelo evita redundância de dados nas tabelas de dimensão. Este modelo não é muito utilizado dado que introduz complexidade na exploração dos dados (Machado, 2004).

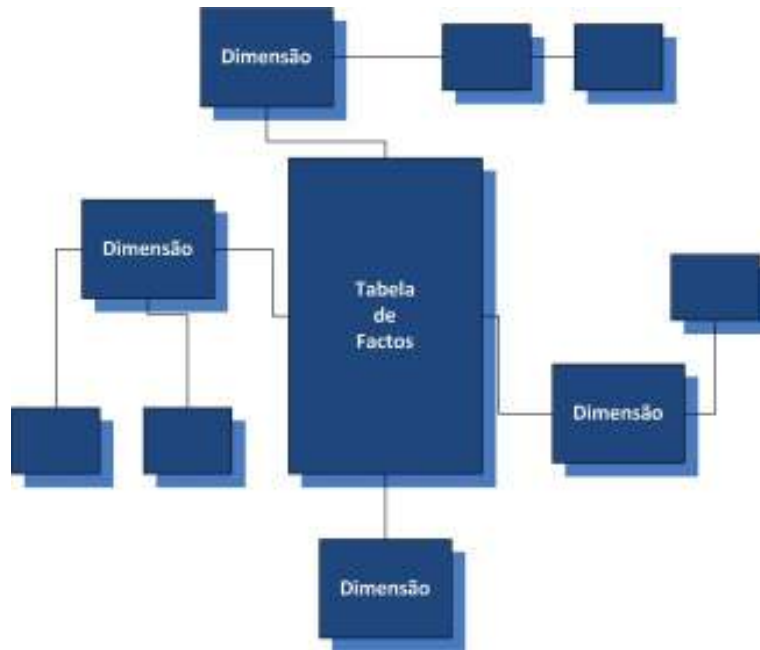


Figura 2.8 - Esquema em Floco de Neve. Fonte: Elaboração própria.

O esquema em constelação (Figura 2.9) apresenta uma série de esquemas em estrela ligados entre si, em função de tabelas de dimensões que cada esquema possa ter em comum com outro. Na realidade, os esquemas em estrela podem estar ligados entre si por mais de uma tabela de dimensão que tenham em comum.

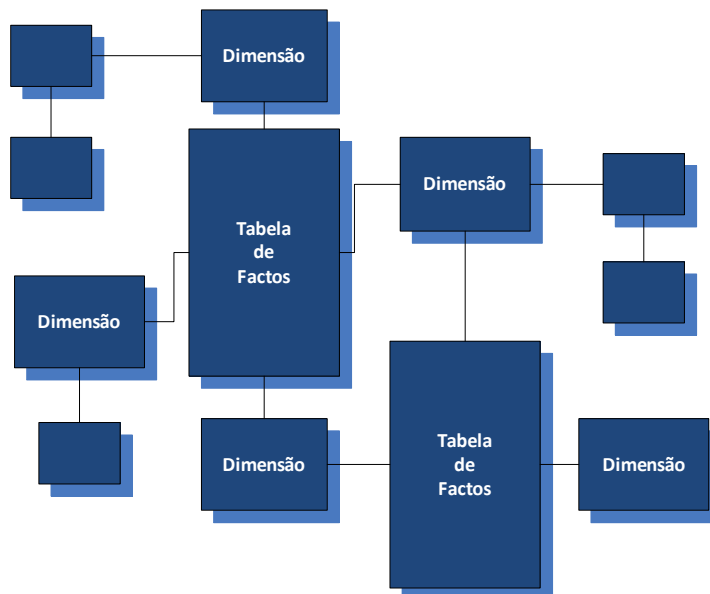


Figura 2.9 - Esquema em Constelação. Fonte: Elaboração própria.

Kimball e Inmon chegaram ao mesmo conjunto de princípios de construção de DW, mas cada um usando técnicas de modelação distintas. Kimball criou o Modelo de Dados Dimensional, também conhecido como o esquema em estrela (Kimball et al., 2002) enquanto Inmon dá preferência à modelação onde é garantida a Terceira Forma Normal no modelo de dados (Inmon, 2005).

## **2.5 ETL - *Extract, Transform and Load***

O processo denominado *Extraction, Transformation and Loading* (ETL) consiste na obtenção de dados dos sistemas de processamento transaccionais (OLTP) para o ambiente de *Data Warehouse*. O processo de extracção envolve a leitura e compreensão dos dados de origem além de uma cópia dos dados necessários ao DW na *Data Staging Area* (Kimball et al., 2002). A *Data Staging Area* permite a ligação entre os Sistemas OLTP e o DW.

Os dados operacionais brutos são transformados num formato específico, pré-definido no DW, prontos para serem consultados e utilizados pelo utilizador. Os dados na *Data Staging Area* sofrem inúmeras transformações, como a limpeza dos dados (correções de erros ortográficos, resolução de conflitos de domínio, tratamento de elementos em falta ou a análise em formatos padrão), combinação de dados de várias origens, eliminação de dados duplicados, entre outras. A *Data Staging Area* é dominada pelas simples actividades de triagem, classificação e processamento sequencial, não se baseando na tecnologia relacional, podendo ser formada por um simples sistema de ficheiros (Kimball et al., 2002).

O processo de integração pode assim ser dividido em quatro fases: análise e compreensão dos dados a partir das diferentes fontes de dados; preparação e recolha dos dados para a área de teste; combinação, limpeza e transformação dos dados, e finalmente, personalização dos dados de acordo com a estrutura do DW.

Os dados no DW são organizados em modelos dimensionais, diferentes dos sistemas operacionais em que estes se baseiam no modelo tradicional entidade-relacionamento (E-R). O esquema dimensional possui uma estrutura mais intuitiva, permitindo uma consulta mais fácil para utilizadores não especializados (Kimball et al., 2002).

Com a conclusão da modelação e da carga do DW, os dados ficam disponíveis para serem tratados pelas ferramentas OLAP, que permitem a manipulação destes através criação de cubos de dados e do uso de operações específicas de manipulação desses cubos, como *Slice*, *Dice*, *Drill down* e *Drill up* (Kimball et al., 2002).

## 2.6 Cubos

A estrutura multidimensional resultante da combinação da tabela de factos com as tabelas de dimensões é referenciada na literatura como cubo de dados (Machado, 2004). Um cubo de dados representa um conjunto de métricas que compartilham o mesmo conjunto de dimensões (Kimball et al., 2008).

O cubo de dados é um formato lógico que permite modelar e visualizar dados sob várias perspectivas. É formado pela tabela de factos e pelas tabelas de dimensão que a rodeiam e representam possíveis formas de visualizar e consultar os dados. O cubo é assim um aglomerado de informações, gerado por assuntos relacionados entre si, permitindo que sejam associadas em várias combinações, resultando na extracção de várias visões sobre o mesmo tema.

A noção de dimensão é um conceito essencial e diferenciador em dados multidimensionais. As dimensões normalmente não possuem atributos numéricos, pois são meros descritores dos atributos das tabelas de facto, e são utilizadas com dois propósitos: a selecção de dados e o respectivo agrupamento a um dado nível. Cada dimensão pode estar organizada sob a forma de uma ou várias hierarquias.

Cada hierarquia é composta por um número de níveis, cada qual representando um nível de detalhe que pode interessar às análises a serem executadas. Também, para cada par dimensão/nível, existe um conjunto de instâncias que se denominam tipicamente por membros da dimensão ou valores da dimensão.

Para estruturas multidimensionais até três dimensões, é simples a observação do formato de um cubo, como ilustra a Figura 2.4, no entanto o modelo multidimensional possui mais de três dimensões, variando sempre de acordo com a complexidade do negócio, assumindo a designação de hipercubo.

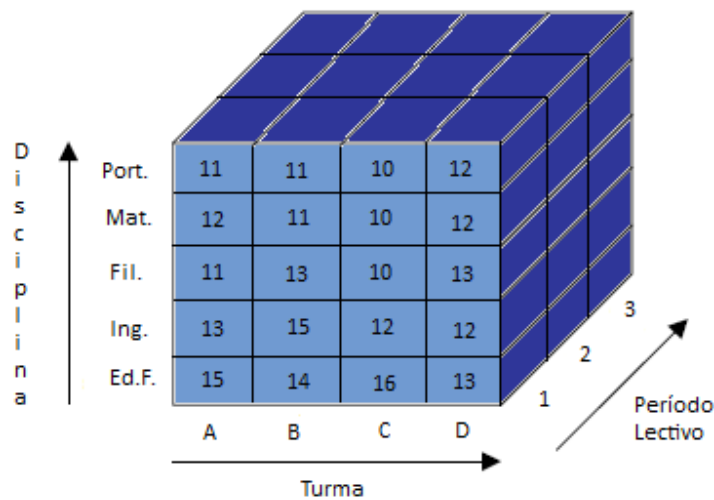


Figura 2.10 - Representação visual de um cubo. Fonte: Elaboração própria.

Um cubo de dados, não é então mais do que uma agregação de dados a um nível determinado de granularidade, entendida esta como o maior ou menor nível de detalhe.

## 2.7 Técnicas de análise multidimensional

A análise multidimensional é um meio natural e intuitivo de o utilizador obter entendimento nas informações contidas num DW. De forma simples, com ferramentas de navegação OLAP, é possível navegar entre diferentes níveis de granularidades de um cubo de dados de um DW. É assim possível o acesso a determinado grupo de dados, focalizando os dados numa orientação em particular e definindo cálculos analíticos. Atendendo a que os dados estão fisicamente armazenados numa estrutura multidimensional, a velocidade dessas operações é muito mais rápida e mais consistente do que em outras estruturas de base de dados. Esta combinação de rapidez, consistência e simplicidade é um dos benefícios fulcrais da análise multidimensional.

Os dois processos de análise multidimensional mais clássicos são comumente conhecidos como *Slice and Dice* e *Drill*. Através do processo chamado *Drill* o utilizador pode aumentar - *Drill down*, ou diminuir - *Drill up*, o nível de detalhe dos dados. *Drill down* e *Drill up* são operações realizadas dentro dos níveis hierárquicos de cada dimensão. Um *Drill down* aumenta o nível de detalhe da informação, ou seja, parte de um nível menos detalhado para um mais detalhado. *Drill up* é a operação inversa, o nível de detalhe é diminuído.

Usando o processo *Slice e Dice* conseguem-se criar visões dos dados através da sua reorganização, para que eles possam ser examinados sob diferentes perspectivas, isto é, possibilita a rotação, em qualquer sentido, dos lados do cubo de dados, permitindo a combinação de qualquer

dimensão para a obtenção de informações. Serve assim para modificar a visão de uma informação, trocando linhas por colunas de forma a facilitar a compreensão dos utilizadores e girar o cubo sempre que houver necessidade. *Slice* é o equivalente a um corte no cubo mantendo a mesma perspectiva na visualização dos dados, apresentando uma fatia dos dados. A operação *Dice*, é uma mudança na perspectiva da visualização no cubo (Inmon, 2005) (Machado, 2004).

## 2.8 Metadados

Os metadados, geralmente conhecidos como dados sobre os dados, podendo ainda ser considerados como uma abstracção dos dados, ou ainda dados de um nível superior que definem dados de um nível inferior, servem de apoio à automatização de tarefas repetitivas relacionadas com os dados do sistema.

Atendendo ao volume de dados armazenados num DW, é importante reter informação sobre a origem dos dados, o histórico da extracção e transformação que esses dados sofreram, entre outras informações pertinentes para melhor conhecimento do DW. Segundo Kimball, os metadados são dados sobre os próprios dados, que contêm informações diversas provenientes de vários componentes da arquitectura do DW (Kimball et al., 2002).

No desenvolvimento de um DW, vários são os metadados gerados, podendo estes ser classificados como metadados técnicos, que traduzem informações produzidas ao nível da definição das interrogações, descrição dos campos calculados, configurações do DW e acessos, alterações ou transformações dos dados aquando do ETL, bem como informação referente ao desempenho ou relatórios de execução. Outro tipo de metadados é os metadados de negócio, que possuem informação base sobre o negócio e traduzem algumas terminologias em uso no DW para formas simples entendíveis pelos utilizadores.

Os metadados devem ser gerados em todas as etapas de *Data Warehousing*, de forma a garantir que todas as informações base sejam incluídas (Inmon, 2005). São portanto um recurso elementar de um DW, desde a construção, passando pelo acesso dos utilizadores até à manutenção e actualização dos dados armazenados.

Numa organização são vários os metadados técnicos e de negócio, oriundos de várias ferramentas e existindo numa grande variedade de formatos, como se constata na Tabela 2.2. Estas fontes de metadados devem fluir directamente para o repositório para aí serem integrados.

Proveniência comum de Metadados	Tipos de Metadados
Ferramentas e Processos de ETL	Regras de transformação de dados Dependências do Programa Processos de carga de dados
Ferramentas de modelação de dados	Modelos lógicos e físicos dos dados Definições técnicas das entidades Definições técnicas dos atributos Domínio dos valores
Documentos	Políticas de negócio
Funcionários	Políticas de negócios Definição de entidades do negócio Definições de atributos do negócio Administração de dados
Ferramentas OLAP e de relatórios	Padrões de acesso do utilizador Relatório dos tempos de execução
Ferramentas da qualidade de dados	Controlo e auditoria

Tabela 2.2 - Tipos de metadados e sua proveniência. Fonte: Elaboração própria.

## 2.9 Arquitecturas, abordagens e metodologias de implementação de DW

A implementação de DW's é um tema que não reúne consenso por parte dos dois grandes autores, Inmon e Kimball, que norteiam o estudo de *Data Warehouses*, no que diz respeito à arquitectura, abordagem e metodologia a usar.

Inicialmente Inmon era defensor de uma abordagem *top-down* com uma estrutura orientada aos dados, sendo que a construção da base de dados do DW era construída como um todo, procurando uma forte integração entre todos os dados da organização num modelo corporativo e só posteriormente se abordariam os temas. Actualmente tal abordagem não fará muito sentido, tornando-se incompatível com a visão estratégica das organizações bem como com a necessidade premente de um sistema de DW de construção rápida e funcional, pelo que Inmon tenha vindo a modificar a sua visão inicial do desenvolvimento de sistemas de DW. No entanto, a metodologia de desenvolvimento apresentada por Inmon é bastante completa. Apresenta desde a metodologia aplicada ao desenvolvimento dos sistemas operacionais, a metodologia de implementação do DW, bem como as ferramentas a aplicar com descrição de tarefas, a sua ordem de execução, possíveis resultados esperados e o tempo necessário à sua conclusão.

Kimball apresenta uma abordagem mais simples e incremental. Como mentor do conceito de *star schema*, a sua metodologia incide na construção de DM's separados, com uma estrutura orientada aos objectivos que no decorrer do tempo serão integrados. Trata-se pois de uma abordagem *bottom-up*, tratando projectos de menor âmbito, com assuntos específicos, independentes, que se integrarão posteriormente.

Uma das arquitecturas propostas defende então a utilização de uma estrutura central normalizada, o DW, que poderá derivar em sucessivos *Data Mart's* (Inmon, 2005). É conhecida como arquitectura *Corporate Information Factory* ou *Hub-and-spoke*. Outra das arquitecturas defende em oposição a esta, a construção de *Data Mart's* multidimensionais, concebidos por assuntos específicos, cruzando informação através da partilha de dimensões comuns as *conformed dimensions* (Kimball et al., 2002). Esta arquitectura é conhecida como arquitectura em *Bus* ou *Data Mart Bus Architecture*.

Quanto ao tipo de abordagem na implementação, há quem defenda uma abordagem *top-down* que se traduz num modelo bastante ambicioso e complexo, atendendo a que se pretende englobar toda a informação da organização no DW. Este é o tipo de abordagem usado numa arquitectura *Corporate Information Factory*. Numa abordagem *bottom-up*, é dada especial ênfase à construção de DM's multidimensionais, construídos com o propósito de se integrarem e unificarem num DW, abordagem implementada em arquitecturas *Bus*.

Verifica-se assim que, a abordagem de Inmon se centra num estilo mais tradicional da construção de bases de dados, onde se procura uma forte integração entre todos os dados das diferentes áreas funcionais da organização e a abordagem de Kimball se centra num estilo incremental, apontando para projectos de DM's separados que deverão ser integrados à medida da sua evolução.

Em nenhuma das metodologias apresentadas os seus autores abordam com a profundidade desejada a gestão dos metadados.



## 3 Metodologias e ferramentas utilizadas

---

### 3.1 Introdução

Muitas são as opções existentes para a implementação de um Sistema de *Data Warehouse*, desde arquitectura, método e abordagens propostas e defendidas por autores, como outras oriundas dos fornecedores de soluções no mercado. A implementação do protótipo de DW requer um caso de estudo, onde as diferentes fases do projecto de DW fossem analisadas e executadas, de forma a identificar os problemas e definir soluções. Desta forma, neste capítulo é efectuada uma apresentação da Escola Secundária/3 de Amato Lusitano (ESAL) e uma introdução a alguns conceitos de gestão estratégica aplicados à instituição de ensino, caso de estudo para implementação do protótipo de DW.

O capítulo termina com a referência aos componentes da infra-estrutura tecnológica usada na construção do protótipo de DW, apresentando uma justificação para a sua escolha.

### 3.2 Apresentação do caso de estudo

A Escola Secundária/3 de Amato Lusitano (ESAL) é uma escola do 3º Ciclo e Ensino Secundário, sita na Avenida Pedro Álvares Cabral, perto do centro da cidade de Castelo Branco.

A ESAL (Figura 3.1), antiga Escola Industrial e Comercial de Castelo Branco, tem uma história e construiu uma tradição enquanto estabelecimento de ensino predominantemente vocacionado para a formação e vida activa.



Figura 3.1 - Vista aérea da ESAL. Fonte: ESAL.

No quadro da estrutura do Ensino Secundário, a ESAL reparte a sua oferta pelos CURSOS CIENTÍFICO-HUMANÍSTICOS, TECNOLÓGICOS, PROFISSIONAIS e CEF's. A oferta de CURSOS PROFISSIONAIS tem mantido a tradição e faz da ESAL uma das escolas que, a nível nacional, oferece um maior leque de escolhas nesta área de ensino (ESAL, 2009a).

O Projecto Educativo da ESAL reafirma a oferta, de forma equilibrada, desses cursos, de modo a proporcionar percursos diversificados que, segundo a escolha dos alunos, permitam o prosseguimento de estudos ou a integração imediata no mundo do trabalho.

A ESAL dispõe de espaço físico que suporta a concretização da orientação curricular definida anteriormente, existindo áreas específicas que asseguram o funcionamento dos cursos que constituem a sua oferta.

A orientação curricular assim definida exige, contudo, a reorganização e o restauro dos espaços existentes e a adaptação e requalificação de outras áreas em resposta à necessidade de criação de condições de bem-estar e de trabalho e de humanização geral do espaço.

Os equipamentos técnico-científicos e de apoio pedagógico-didáctico, devem ser sujeitos a permanente esforço de manutenção, actualização e modernização de modo a garantir a formação dos alunos com vista ao prosseguimento de estudos e à minimização do hiato entre a formação escolar e as exigências do mercado de trabalho.

Os cursos oferta da ESAL devem organizar-se e funcionar em interacção estreita com os agentes económicos, as autarquias, as associações e as colectividades nas diversas áreas, nomeadamente no que se refere ao ordenamento do espaço físico, aos equipamentos a instalar e à sua actualização e manutenção (ESAL, 2007).

Para os efeitos previstos, a ESAL celebrará protocolos de cooperação com entidades e organizações vocacionadas nas diferentes áreas.

A ESAL é uma comunidade de vivência colectiva de todos os actores educativos internos: alunos, professores, corpos administrativo e auxiliar.

As relações interpessoais institucionais, a todos os níveis da estrutura escolar, devem reger-se pela cordialidade, afectividade, compreensão e abertura ao diálogo no plano do exercício das competências próprias.

A ESAL, no cumprimento da sua função educativa-formativa, garante o respeito pelas diferenças individuais e a promoção da diversidade, desenvolvendo práticas efectivas na defesa desses valores, designadamente através das opções pedagógicas e metodológicas definidas no seu currículo-projecto (ESAL, 2007).

O espaço físico da ESAL ordena-se e organiza-se de modo a proporcionar a todos os elementos da comunidade escolar condições de bem-estar e de trabalho, devendo ser criados locais específicos para o trabalho individual e de grupo de todos os professores.

No ordenamento e organização do espaço físico, nomeadamente da sala de aula, têm-se em conta as exigências logísticas colocadas pelas opções curriculares fixadas neste projecto: pela natureza dos cursos; pelas orientações pedagógicas e didácticas; pelas actividades culturais e de lazer.

Os equipamentos educativos, em todo o espaço escolar, visam a criação de condições de bem-estar, de trabalho e de lazer para todos os elementos da comunidade escolar.

Os cursos em funcionamento na ESAL e que são resultado das opções definidas no Projecto Educativo da ESAL devem dispor dos equipamentos técnicos específicos e de apoio que permitam a concretização das orientações pedagógico-didácticas-metodológicas fixadas e a efectiva formação científico-técnica dos alunos que os frequentam (ESAL, 2007).

O número de docentes da escola tem-se mantido estável com um pequeno acréscimo nos dois últimos anos, sendo actualmente 149 o total de docentes. No geral, são docentes do quadro de escola, com mais de 15 anos de serviço, a grande maioria com mais de 20 anos de serviço e predominantemente do género feminino (ESAL, 2009b).

O número de pessoal não docente tem-se mantido relativamente estável ao longo dos últimos anos, perfazendo um total de 40, entre assistentes operacionais e assistentes administrativos.

No ano lectivo a que este estudo se reporta, 2009/2010, a escola tem uma diversidade significativa na sua oferta educativa em conformidade com a filosofia do seu Projecto Educativo (*A importância do aluno na centralidade da organização*). Os alunos estão distribuídos por 51 turmas de vários cursos e anos de escolaridade, num total de 965 alunos (ESAL, 2009b).

Um dos documentos essenciais, que na escola deve consubstanciar os aspectos estratégicos, é o projecto educativo, porque é nele que se definem as ambições, os fins e os objectivos, se pressupõe um diagnóstico e uma avaliação das estratégias, se exprime a decisão estratégica e as prioridades de desenvolvimento.

Assim concebido, o projecto educativo constitui-se, de facto, num instrumento institucional de organização/gestão de médio e longo prazo, devendo incluir, por conseguinte, o diagnóstico interno e externo da situação da escola, expressar as decisões estratégicas colectivamente assumidas e os contornos da identidade procurada, sistematizar os fins e objectivos estratégicos da instituição escolar, assegurando-lhe ao mesmo tempo coerência interna e externa.

Uma das implicações desta perspectiva de projecto educativo é que ele deve emergir como resultado de um processo participativo e negociado entre os diferentes actores sobre metas, valores, princípios e prioridades, enfim, sobre um futuro que se ambiciona construir, procurando reflectir deste modo uma dinâmica essencialmente política, globalizada e flexível. Ou seja, a construção do projecto educativo vai reclamar uma outra lógica, que não coincide propriamente com a lógica da racionalidade técnica da gestão estratégica *stricto sensu*, mas que apela a múltiplas referências, a critérios de natureza política e democrática (Estevão, 1999).

### **3.3 Gestão estratégica**

Segundo Kaplan e Norton, o primeiro princípio observado nas implementações de sucesso do *Balanced Scorecard* (BSC) consiste na tradução da estratégia em termos operacionais (Kaplan & Norton, 2001).

Após a apresentação e enquadramento da ESAL enquanto organização, optou-se neste estudo por efectuar uma caracterização um pouco mais pormenorizada, pelo que se elaborou uma matriz SWOT, baseada entre outros documentos, nos da avaliação externa que a ESAL foi sujeita em Abril de 2010 (IGE, 2010), com o intuito de promover o alinhamento das forças e fraquezas com as oportunidades e ameaças, conforme indicado na Tabela 3.1.

Pontos Fortes F	<p>F1. Quadro de professores estável;</p> <p>F2. Currículo generalista - Cursos Profissionais, CEF's, Tecnológicos e C.H;</p> <p>F3. Qualificação docente;</p> <p>F4. Qualificação discente;</p> <p>F5. Actuação diversificada;</p> <p>F6. Eficiência dos serviços prestados;</p> <p>F7. Boa estrutura administrativa;</p> <p>F8. Escola aberta ao exterior / comunidade.</p>
Oportunidades O	<p>O1. Processos administrativos orientados a processos de negócio;</p> <p>O2. Infra-estrutura (salas, laboratório, área de convivência social);</p> <p>O3. Gestão do espaço físico;</p> <p>O4. Implementação de Projectos inovadores e investimento na melhoria da qualidade de ensino;</p> <p>O5. Procedimentos e formulários suportados num sistema de informação;</p> <p>O6. Atendimento ao aluno / Pais e Encarregados de Educação.</p>
Pontos Fracos N	<p>N1. Taxa de insucesso nos exames nacionais e algum abandono escolar;</p> <p>N2. Pouca participação dos Pais e Encarregados de Educação na vida da escola;</p> <p>N3. Espaços exteriores (acessibilidade e segurança);</p> <p>N4. Escassez de espaços de trabalho para os docentes;</p> <p>N5. Falta de acompanhamento e envolvimento no processo educativo dos filhos / educandos;</p> <p>N6. Falta de visão sistémica.</p>
Ameaças A	<p>A1. Fraco rendimento escolar;</p> <p>A2. Falta de alunos.</p>

Tabela 3.1 - Análise SWOT. Fonte: Elaboração própria.

Sendo a ESAL uma instituição onde se concretiza o direito à educação, deve procurar a permanente optimização do seu funcionamento, bem como sustentar-se num modelo que permita o planeamento, a comunicação e a gestão dos factores estratégicos fundamentais. Para tal, e tendo por base a revisão bibliográfica efectuada, surge o presente projecto apoiando-se no *Balanced Scorecard* (BSC), considerado como uma ferramenta para uma melhor gestão, permitindo analisar o impacto das suas decisões estratégicas em várias perspectivas independentes entre si, mas que, por razões metodológicas se autonomizaram do seguinte modo: dos clientes, dos processos, da inovação e aprendizagem e da financeira, atendendo ainda a várias medidas relevantes como a análise do

meio envolvente da instituição, visão comum, *feedback* e aprendizagem, comunicação, envolvimento e planeamento.

O conceito do *Balanced Scorecard*, como se ilustra na Figura 3.2, coloca a estratégia no centro da gestão dos processos, permitindo a análise do desempenho da missão através das perspectivas do cliente, financeira, de processos internos de crescimento e aprendizagem, para termos uma visão equilibrada do desenvolvimento da instituição (Kaplan & Norton, 1992).

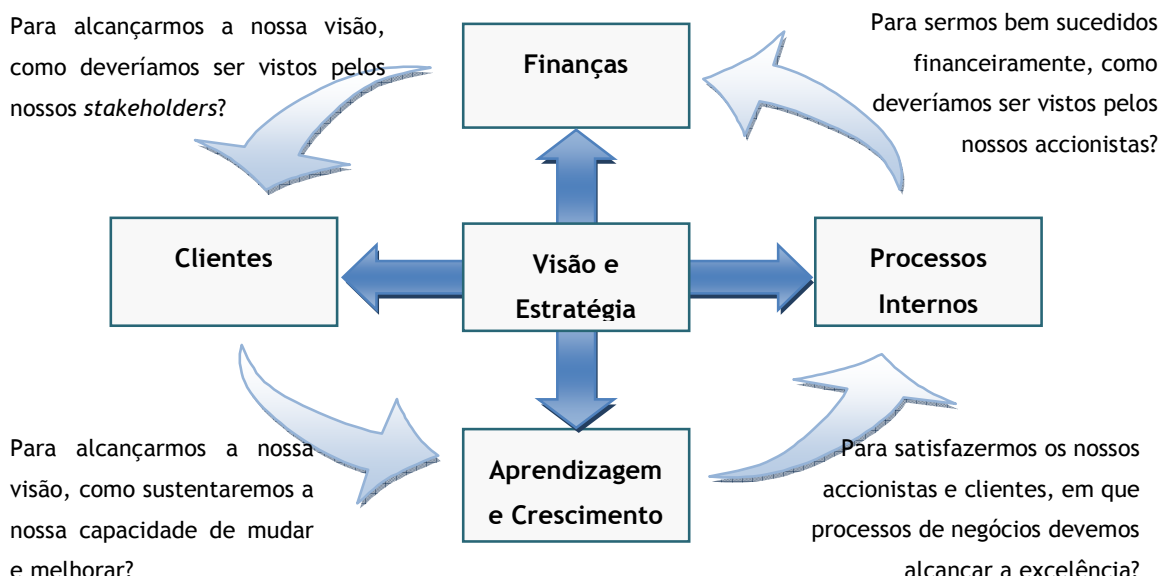


Figura 3.2 - Perspectivas de desempenho do BSC. Fonte: Adaptado de (Kaplan & Norton, 1992).

O modelo básico do BSC tem por pressuposto uma organização com fins lucrativos. Nesse sentido, a perspectiva financeira é a última e mais importante para este tipo de organização. Esse não é o caso de organizações sem fins lucrativos como organizações públicas de serviços, como do caso em estudo. Kaplan e Norton sugerem que uma organização sem fins lucrativos coloque em igual importância as perspectivas financeira e do cliente, subordinadas estas à missão da organização (Kaplan & Norton, 2000; Kaplan & Norton, 2001). A Figura 3.3 apresenta este modelo de perspectivas do BSC.

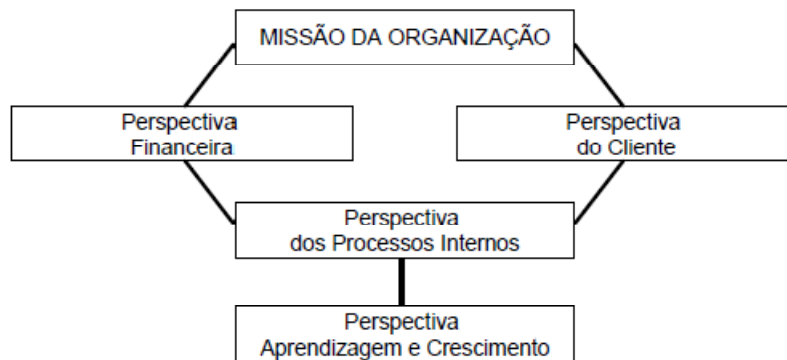


Figura 3.3 - Modelo de BSC para organizações sem fins lucrativos. Fonte: (Kaplan & Norton, 2000).

Na definição das perspectivas para instituições de ensino, as preocupações são genericamente as mesmas que para o sector público em geral. No entanto, a criação de uma perspectiva da educação seria pertinente, dado ser o principal objectivo dessas organizações (Rocha, 2000). Este aspecto será redundante, atendendo a que os objectivos educacionais dependem do desenvolvimento das restantes perspectivas e não são um fim em si próprio.

A definição dos objectivos estratégicos para o BSC surgiu da análise conjunta dos objectivos gerais à luz das metas globais estipuladas no Projecto Educativo da ESAL e, fundamentalmente, das estratégias elaboradas a partir da matriz SWOT. Dessa forma, resultou o BSC indicado na Figura 3.4.

Após leitura e reflexão sobre o Projecto Educativo da ESAL, poderá considerar-se a seguinte definição da **MISSÃO** da ESAL:

*“Prestar serviços educacionais de qualidade de modo a desenvolver as potencialidades do educando, assegurando a formação para o exercício da cidadania, o progresso no trabalho e na vida académica, e a consciência do seu papel como agente da história.”*

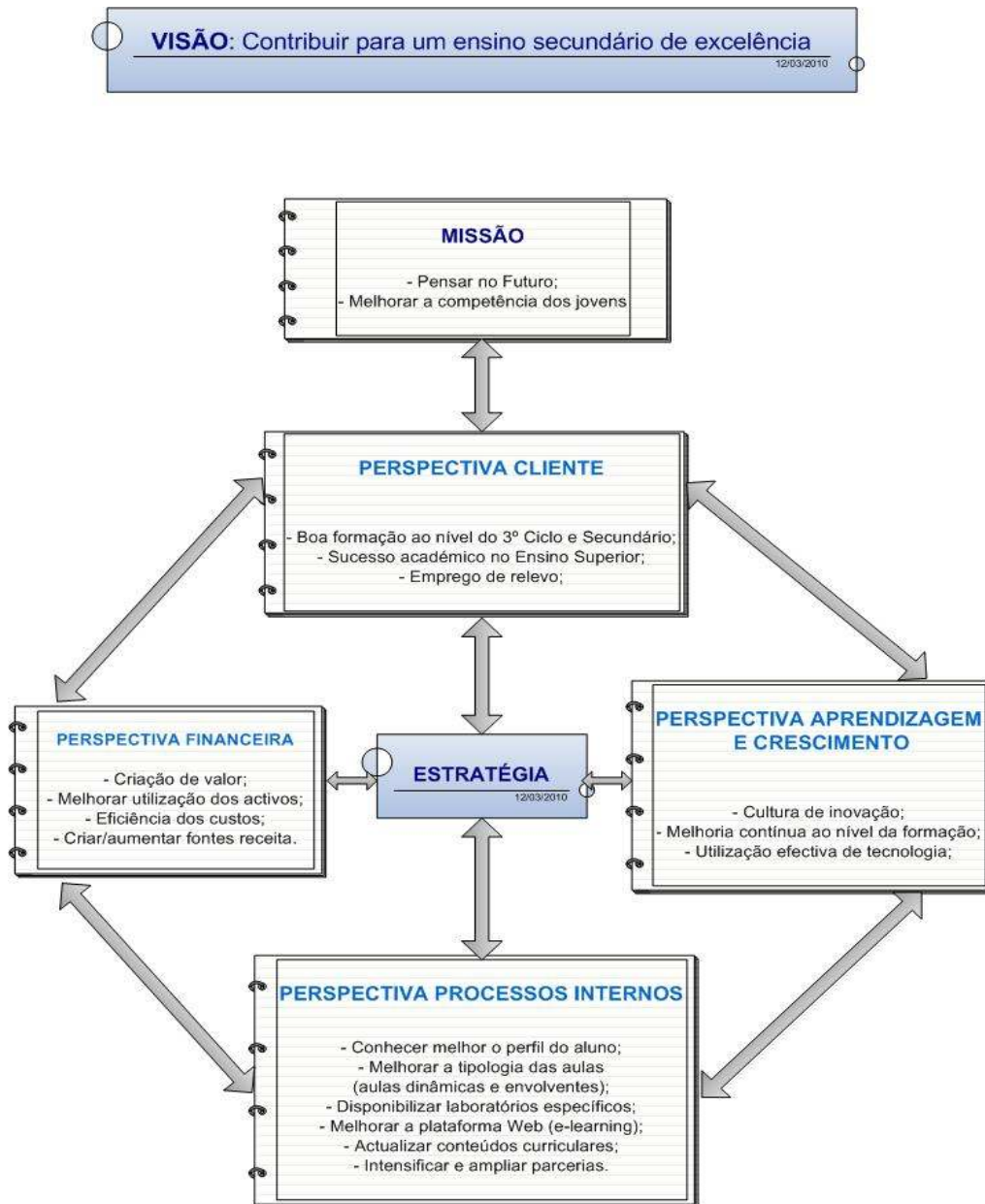


Figura 3.4 - Perspectivas BSC da ESAL. Fonte: Elaboração própria.

Como se pode verificar na figura anterior, o *Balanced Scorecard* da ESAL tem como preocupação a satisfação das necessidades dos alunos bem como da sociedade.

Partindo do princípio que a visão e os objectivos já foram definidos, a escolha dos indicadores estratégicos devem reflectir os objectivos estratégicos desejados (Tabela 3.2).

Objectivos Estratégicos	Indicadores Estratégicos	Medidas
<b>Perspectiva dos Clientes</b>		
<ul style="list-style-type: none"> <li>• Ser conhecido pela qualidade dos cursos;</li> <li>• Minimizar o nível de reclamações recebidas;</li> <li>• Diminuir transferências para outras escolas;</li> <li>• Diminuir anulações de matrículas.</li> </ul>	<ul style="list-style-type: none"> <li>• Qualidade de ensino;</li> <li>• Volume de reclamações recebidas;</li> <li>• Índice de transferências;</li> <li>• Índice de anulação de matrículas.</li> </ul>	<ul style="list-style-type: none"> <li>• Pesquisa/Avaliação;</li> <li>• Número de reclamações em cada curso;</li> <li>• Número de transferências;</li> <li>• Número de anulação de matrículas.</li> </ul>
<b>Perspectiva Financeira</b>		
<ul style="list-style-type: none"> <li>• Maximizar o retorno sobre o investimento;</li> <li>• Melhorar a utilização dos activos;</li> <li>• Melhorar a estrutura de custos;</li> <li>• Criar e aumentar fontes de receita.</li> </ul>	<ul style="list-style-type: none"> <li>• Retorno sobre os Investimentos;</li> <li>• Índice de horas de utilização das salas de aula;</li> <li>• Índice de aumento de despesas conservação / manutenção;</li> <li>• Novas receitas.</li> </ul>	<ul style="list-style-type: none"> <li>• Taxa mínima de retorno;</li> <li>• Número de horas de utilização das salas de aula;</li> <li>• Taxa de aumento de despesas de conservação / manutenção;</li> <li>• Taxa de aumento das receitas próprias.</li> </ul>
<b>Perspectiva dos Processos Internos</b>		
<ul style="list-style-type: none"> <li>• Manter a estrutura integrada no tratamento das operações;</li> <li>• Ter rapidez para se adequar as mudanças;</li> <li>• Melhorar serviços de apoio e infra-estrutura;</li> <li>• Investir em tecnologia;</li> <li>• Aumentar a utilização dos meios multimédia no processo ensino-aprendizagem;</li> <li>• Ampliar número parcerias.</li> </ul>	<ul style="list-style-type: none"> <li>• Tempo gasto no tratamento das operações;</li> <li>• Tempo médio de desenvolvimento de sistemas;</li> <li>• Índice de satisfação discente;</li> <li>• Novos investimentos em tecnologia;</li> <li>• Índice de requisições dos meios multimédia;</li> <li>• Novas parcerias.</li> </ul>	<ul style="list-style-type: none"> <li>• Tempo médio entre o requerimento e o tratamento oferecido;</li> <li>• Acompanhamento do tempo gasto no desenvolvimento;</li> <li>• Avaliação da satisfação;</li> <li>• Aquisição de novas tecnologias;</li> <li>• Número de requisições dos meios multimédia;</li> <li>• Contactos com parceiros.</li> </ul>
<b>Perspectiva Aprendizagem e Crescimento</b>		
<ul style="list-style-type: none"> <li>• Manter uma equipe motivada;</li> <li>• Treino contínuo ao nível da formação;</li> <li>• Desenvolver habilidades de relacionamento com os <i>stakeholders</i>/sociedade.</li> </ul>	<ul style="list-style-type: none"> <li>• Avaliação de desempenho;</li> <li>• Número de acções de formação realizadas;</li> <li>• Índice de satisfação discente.</li> </ul>	<ul style="list-style-type: none"> <li>• Pontuação média da avaliação;</li> <li>• Percentagem de cumprimento de metas;</li> <li>• Avaliação da satisfação.</li> </ul>

Tabela 3.2 - Objectivos Estratégicos e Indicadores. Fonte: Elaboração Própria.

Assim o mais importante, é entender o conteúdo das informações como sendo indicadores de tendência, ou seja, como esta série de informações recolhidas podem ser utilizadas no sentido de evidenciar a direcção que deve ser dada à instituição, no caso a ESAL, e mais, se esta direcção está de acordo com a missão definida.

### 3.4 Infra-estrutura tecnológica

A concretização do protótipo envolveu uma análise exaustiva sobre as diferentes soluções que iriam ser implementadas, pelo que uma grande parte do tempo nele despendido centrou-se na execução da tarefa de análise das tecnologias, tornando-se numa grande componente didáctica, tendo sido necessário despende de bastante tempo para a realização do trabalho de pesquisa e de investigação, permitindo assim, aprofundar os conhecimentos.

Na análise de requisitos feita à ESAL, verificou-se a existência de várias bases de dados operacionais, bem como dados registados em folha de cálculo Excel, tendo sido este o ponto de partida para o estudo da modelação a implementar. Para a modelação de qualquer sistema é fundamental estudar a sua realidade, tal como foi previamente feito.

A infra-estrutura tecnológica sobre a qual foi implementado o DW, assenta na tecnologia *Microsoft* tendo por sistema operativo o *Windows Server 2003*, sistema operativo para servidores de aplicações da *Microsoft*, plataforma de *software* concebida para simplificar a concepção, desenvolvimento e execução de aplicações multiutilizadores. Sobre este sistema operativo foi instalado o *Microsoft Office Visio Professional 2007*, onde foi modelado o projecto de *Data Warehouse*, a plataforma *Microsoft SQL Server 2008* e o *Microsoft Visual Studio 2008*, que permitiram a criação, implementação, análise e construção de relatórios do DW que através do *Microsoft Windows SharePoint Services 3.0* serão difundidos a toda a comunidade escolar (Figura 3.5).

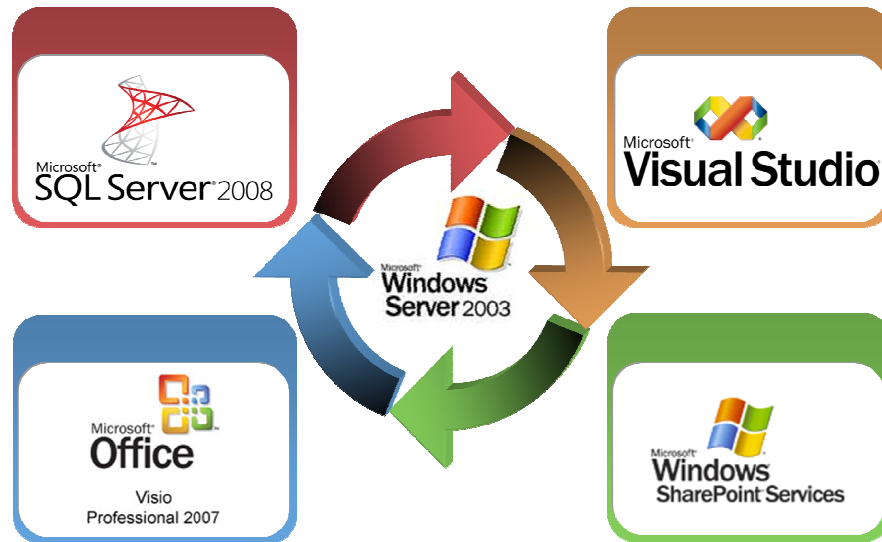


Figura 3.5 - Infra-estrutura tecnológica do DW. Fonte: Elaboração própria.

Todo o processo de desenvolvimento da plataforma analítica do protótipo de *Data Warehouse* foi feito com recurso à ferramenta *Business Intelligence* da *Microsoft*, o *SQL Server 2008*, que contém todas as funcionalidades necessárias à construção da plataforma analítica (Figura 3.5). A ferramenta permite a integração de múltiplas fontes de dados através do módulo *Integration Services*, o enriquecimento de dados e a criação de complexos processos de análise através do *Analysis Services*, e finalmente a elaboração, gestão e disponibilização de relatórios otimizados através do módulo *Reporting Services* (Magalhães, 2009). Desta forma obteve-se um protótipo, que se pretende que seja considerado como um apoio fundamental para a formulação de estratégias, realização de tarefas de análise, correlação de dados e elaboração de relatórios que permitam reportar a evolução da ESAL entre períodos de tempo diferentes.

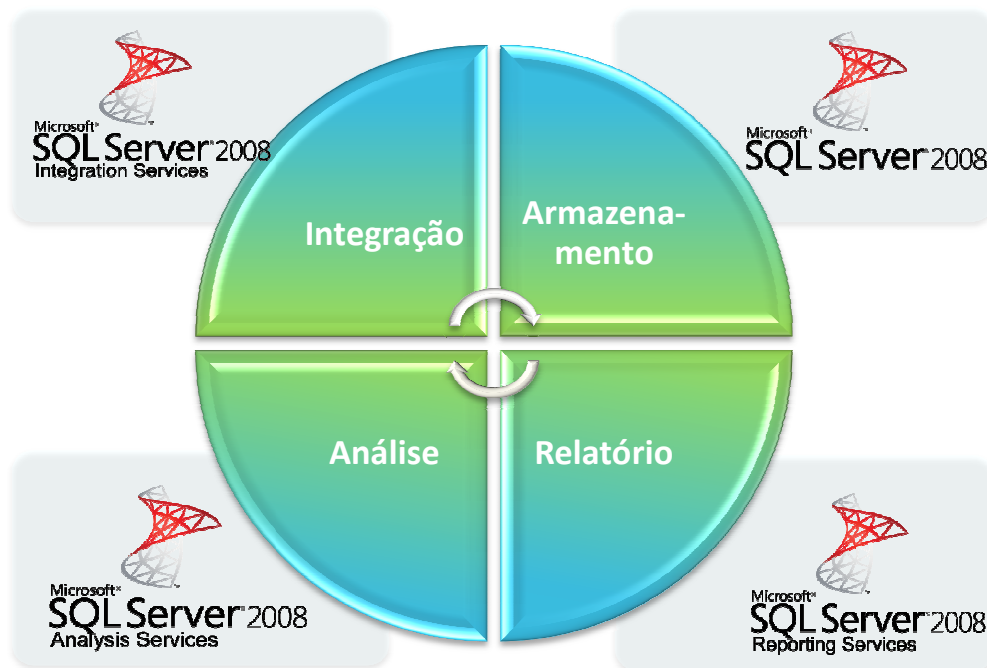


Figura 3.6 - Plataforma *Business Intelligence* da *Microsoft*. Fonte: Elaboração própria.

### 3.5 Outras ferramentas de *Business Intelligence*

Diversas são as ferramentas do âmbito de *Business Intelligence* disponíveis no mercado, nomeadamente a plataforma de alto desempenho da IBM da qual a COGNOS (IBM, 2010) faz parte, sendo considerada como uma das mais completas soluções de *Business Intelligence & Performance Management*. Outra plataforma muito completa é a produzida pela Oracle Corporation (Oracle, 2010), com aplicações de BI passíveis de funcionar sob diversos ambientes computacionais. O Oracle *Business Intelligence* torna-se um produto completo, com ferramentas OLAP, acesso integrado a base de dados, e sem necessidade de integração entre componentes de BI de fornecedores diferentes. A concorrente plataforma tecnológica desenvolvida pela SAP, a SAP NetWeaver BI (SAP, 2010), ou até a Pentaho Solutions (Pentaho, 2010) ferramenta *open source* que permite gerar relatórios, integração com o DW, análise de informações em tempo real, *data mining* e modelos estatísticos, poderiam ser alternativas à escolha feita.

A escolha da plataforma de *Business Intelligence* da *Microsoft*, que no momento oferece uma importante e completa plataforma para a administração de bases de dados e criação de arquitecturas de *Data Warehouse*, prendeu-se numa primeira instância pela disponibilidade de obtenção da mesma, bem como restante infra-estrutura tecnológica, no Centro de *Software* do

MSDN *Academic Alliance* através do acordo com o IPCB<sup>1</sup>. Numa segunda instância e após todo o trabalho de pesquisa e de investigação para aprofundar os conhecimentos na área de BI, esta plataforma revelou-se uma interessante opção por responder a quase todos os requisitos ambicionados para a concretização do projecto.

Na pesquisa efectuada sobre cada uma destas soluções de *software* de BI verificou-se que muitas ferramentas e metodologias são comuns entre elas, havendo também diferenças relevantes entre algumas das plataformas. No entanto a diferença a ser considerada centra-se no momento da implementação da plataforma escolhida pela empresa, que deverá efectuar uma análise detalhada verificando se a solução de BI pretendida responde a todos os seus requisitos, bem como ao impacto que causará na filosofia das suas actividades. Esta análise foi efectuada através do estudo comparativo dos folhetos informativos de cada ferramenta disponibilizados na *internet*, sendo restrita à informação aí disponibilizada.

---

<sup>1</sup> *Microsoft Developer Network Academic Alliance* (MSDN) - Este programa possibilita o acesso e utilização legal de *software Microsoft* relacionado com sistemas operativos, servidores e ferramentas de desenvolvimento ao universo do Instituto Politécnico de Castelo Branco.



## 4 Modelação

---

### 4.1 Introdução

Com o intuito de melhor visualizar todo o sistema de *Data Warehouse*, o seu comportamento, bem como a especificação das estruturas base, este capítulo faz uma primeira abordagem à modelação do protótipo de DW em *Unified Modeling Language*, ajudando à compreensão do sistema desenvolvido.

Neste capítulo é feita a modelação das funções base do protótipo de DW, evidenciando-se num diagrama de componentes a infra-estrutura tecnológica do projecto. Segue-se a modelação conceptual dos dados, desde a sua primeira abordagem até ao modelo conceptual final, passando-se para a elaboração do modelo lógico com a representação das entidades, relacionamentos, atributos e chaves intervenientes no DW.

É apresentado em seguida o diagrama de classes do Modelo Multidimensional e o capítulo termina com uma descrição de cada uma das tabelas de factos e de dimensão que constituem o Modelo Multidimensional.

A modelação apresentada baseia-se apenas nos dados disponíveis no momento, daí este ser apenas um protótipo de DW, não sendo considerados os indicadores estratégicos referidos no BSC anteriormente mencionado (segundo capítulo). Esta mesma modelação será pois considerada como trabalho futuro.

### 4.2 Nível mais elevado de descrição do projecto

Neste protótipo de *Data Warehouse* pretende-se que se efectuem consultas à informação nele contidas de uma forma simples. O DW contém um conjunto de visões multidimensionais associadas aos cubos gerados, sendo disponibilizado um conjunto de relatórios produzidos pela análise multidimensional.

De uma forma global pensou-se no protótipo de DW como a seguir se apresenta na Figura 4.1 no diagrama de caso de uso.

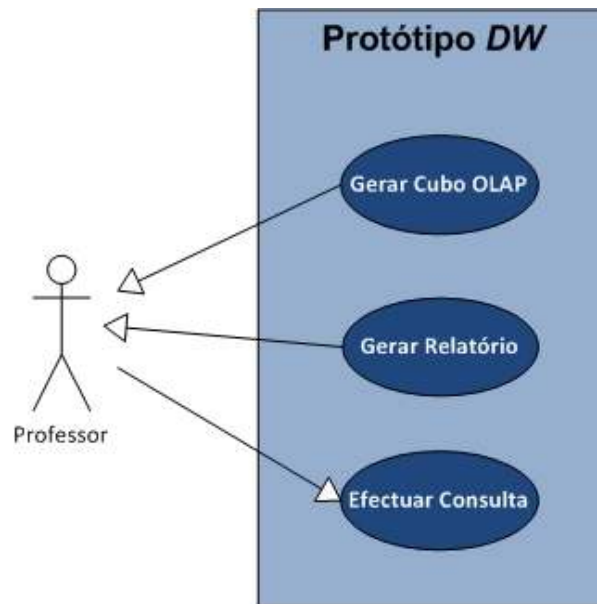


Figura 4.1 - Diagrama de caso de uso do protótipo DW. Fonte: Elaboração própria.

Num nível mais elevado de modelação do projecto e ainda sem qualquer nível de detalhe, foi pensada cada uma das funções que se pretendiam ver detalhadas posteriormente (Figura 4.2).

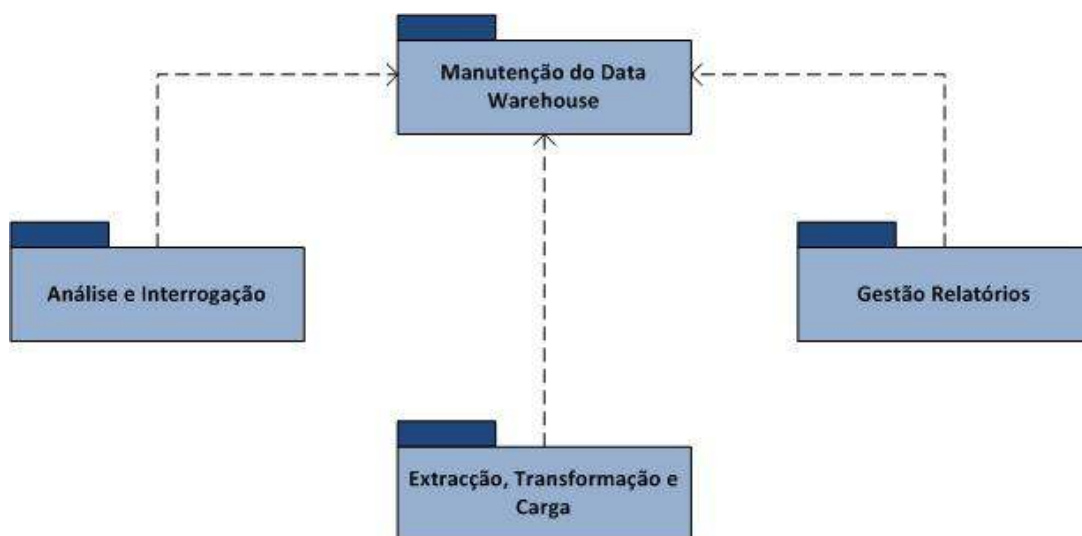


Figura 4.2 - Diagrama de pacotes. Fonte: Elaboração própria.

Associado ao DW, o processo inerente à extracção, transformação e carga de dados dos sistemas transaccionais e restantes fontes de dados para o DW foi um dos elementos preponderantes a constar no modelo, seguido da análise multidimensional e produção dos respectivos relatórios. A manutenção do *Data Warehouse*, representa o injectar incremental periódico dos dados.

O protótipo de DW está alinhado com o objectivo primordial de permitir o acesso a toda a informação nele contida para a tomada de decisões, de uma forma rápida, flexível e completa.

### **4.3 Componentes básicos**

Em paralelo à modelação do sistema de DW, foi elaborada a modelação da infra-estrutura tecnológica. Desta faz parte um servidor de dados, onde foi concebido o DW e um servidor *Web*, que serve como visualizador, ou *front-end*, dos relatórios produzidos para consulta (Figura 4.3).

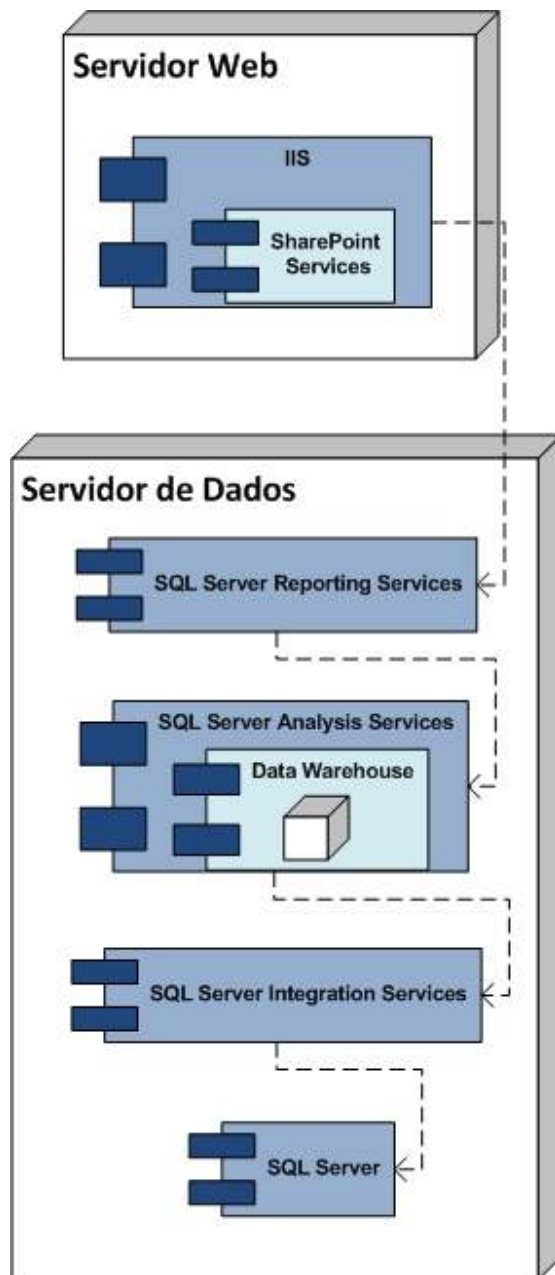


Figura 4.3 - Diagrama de componentes. Fonte: Elaboração própria.

#### 4.4 Modelo conceptual de dados

Esta fase de modelação, em que se construiu o modelo conceptual de dados (Figura 4.4), teve por objectivo a construção de uma primeira descrição ainda abstracta mas completa da estrutura do DW a partir dos requisitos analisados e dos objectivos pretendidos para as consultas, inerentes à fase de análise.

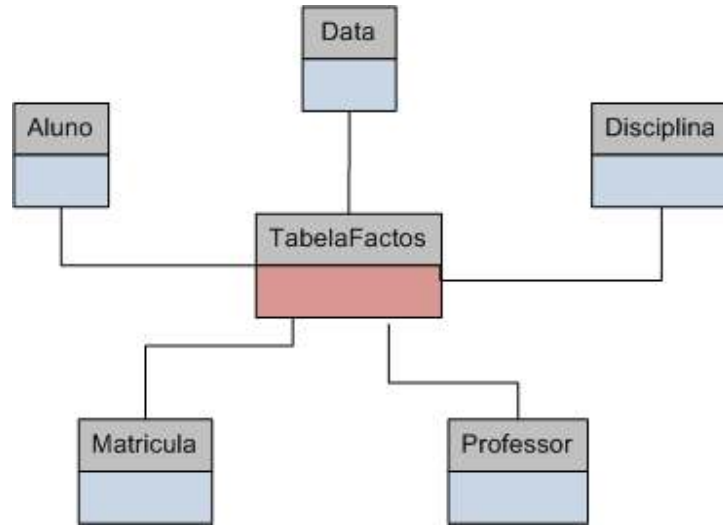


Figura 4.4 - Modelo conceitual inicial do Data Warehouse. Fonte: Elaboração própria.

Foram esboçadas as tabelas de dimensão e a tabela de factos que iriam fazer parte do modelo, ainda sem o pormenor de todos os seus atributos e medidas.

Com a implementação deste modelo, a informação referente aos alunos da ESAL do 3ºCiclo do Ensino Básico, do Ensino Secundário Regular e dos Cursos de Educação e Formação fica passível de ser usada para a construção dos futuros relatórios. As regras usadas na avaliação destes três tipos de ensino são similares, pelo que as medidas a constar da tabela de factos vão ao encontro deste requisito. Com o intuito de os Cursos Profissionais e o Curso Regular Nocturno, que a ESAL ministra, também fossem contemplados, foi criada uma segunda tabela de factos associada às dimensões *DimAluno*, *DimMatricula* e *DimData*, que veio a alterar o modelo inicialmente concebido (Figura 4.5).

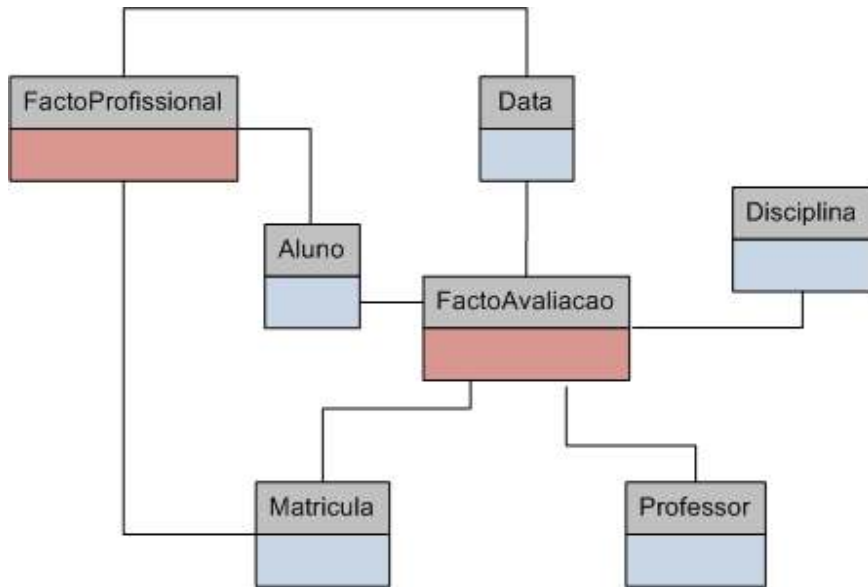


Figura 4.5 - Modelo conceitual final do Data Warehouse. Fonte: Elaboração própria.

Neste modelo conceitual passámos assim a ter dois esquemas em estrela, sendo este representado como se indica na Figura 4.6.

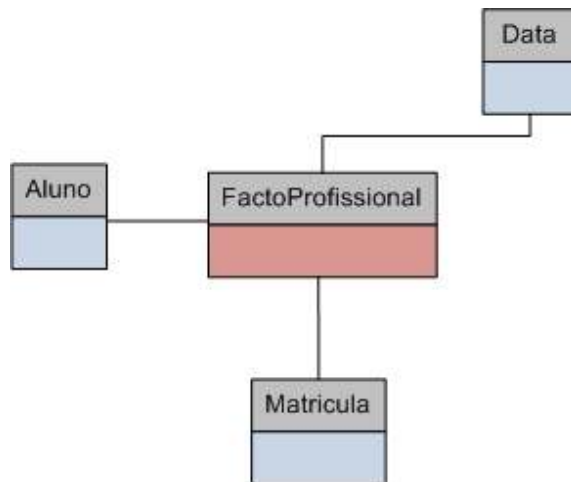


Figura 4.6 - Esquema em estrela dos dados dos Cursos Profissionais. Fonte: Elaboração própria.

## 4.5 Modelo lógico de dados

Após a fase anterior, donde resultou o modelo conceitual, foi elaborado o modelo lógico com todas as entidades intervenientes no DW e os relacionamentos entre elas; os atributos de cada entidade e respectivos tipos de dados a eles associados; a especificação das chaves primárias para cada entidade bem como as chaves estrangeiras (Figura 4.7).

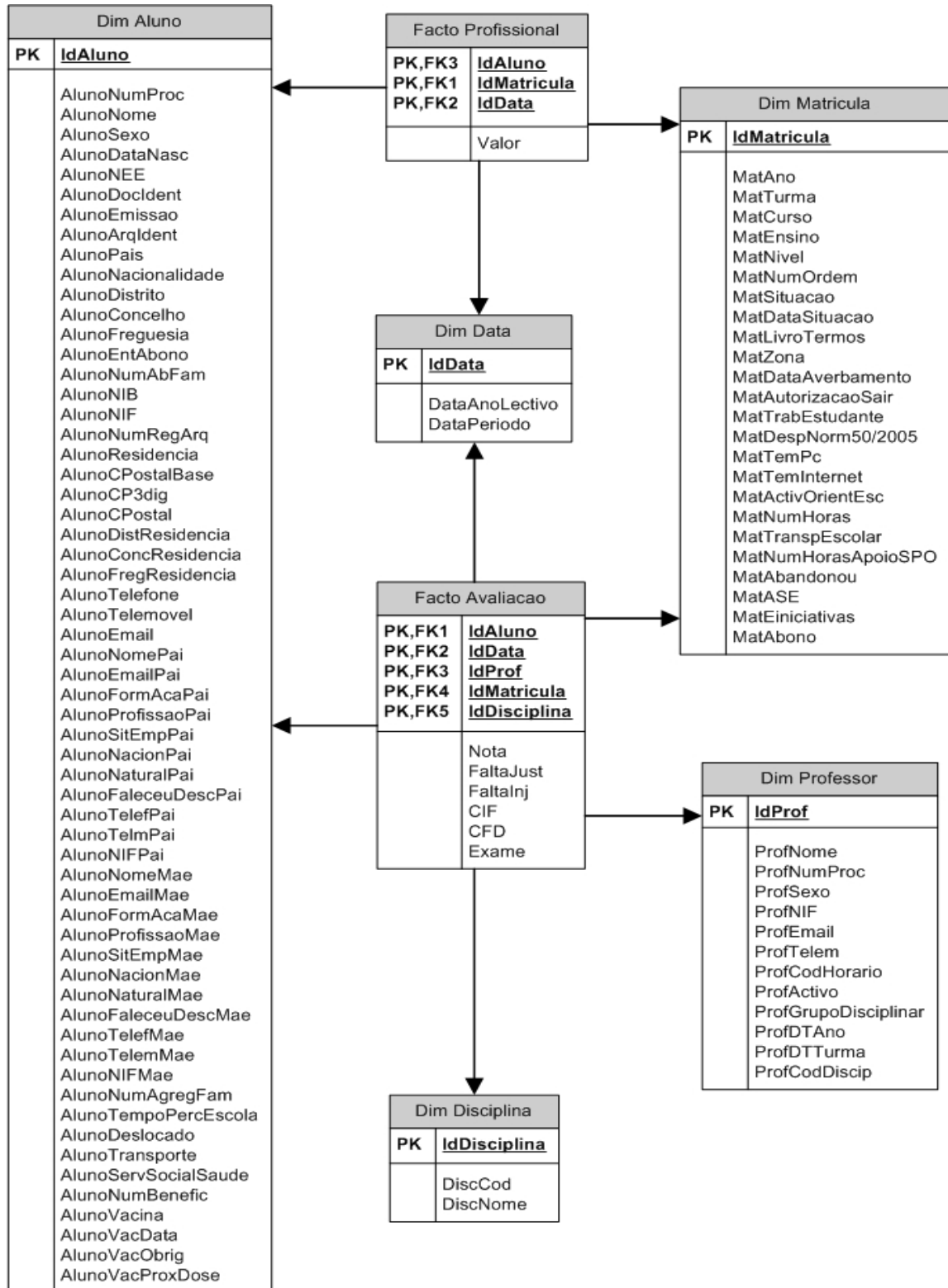


Figura 4.7 - Modelo lógico do Data Warehouse. Fonte: Elaboração própria.

O objectivo do modelo lógico de dados foi o de descrever os dados com o máximo detalhe, sem a preocupação em como seriam implementados fisicamente na base de dados. No seguimento da construção do modelo anterior, foi elaborado o diagrama de classes (Figura 4.8), onde cada classe representa uma tabela do DW, com os respectivos atributos e as operações que suporta.

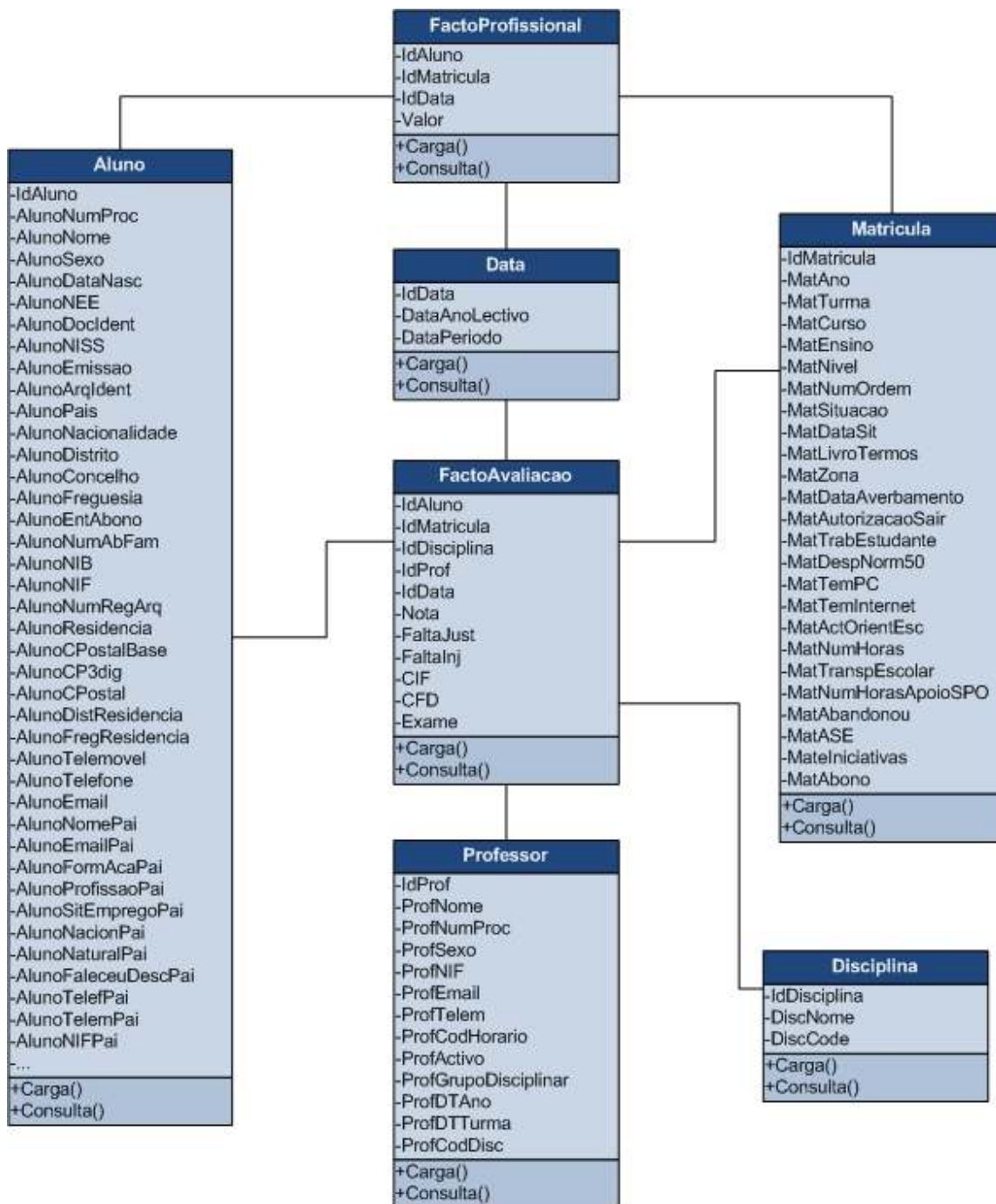


Figura 4.8 - Diagrama de classes do modelo dimensional. Fonte: Elaboração própria.

## 4.6 Descrição das tabelas

Nesta secção serão apresentados e descritos os atributos das tabelas de factos e de dimensão que compõem o modelo multidimensional criado para atender a solução proposta. Cada dimensão é definida pela sua única chave primária, que serve como base para manter a integridade referencial quando se liga à tabela de factos.

### 4.6.1 Tabela de dimensão *DimData*

A dimensão *DimData* expressa as diferentes granularidades temporais em que uma análise temporal pode ser efectuada. A dimensão *DimData* apresenta uma estrutura hierárquica com três níveis de análise, onde o atributo *DataAnoLectivo* é o atributo que identifica o ano lectivo a que se reportam os dados e o atributo *DataPeriodo*, descreve o período lectivo do respectivo ano (Figura 4.9).

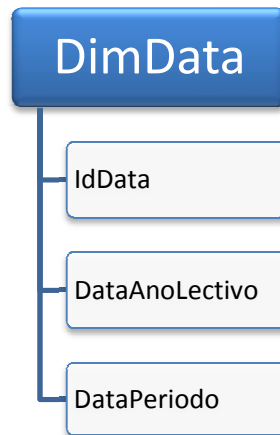


Figura 4.9 - Dimensão *DimData*. Fonte: Elaboração própria.

A representação do atributo descritor *DataPeriodo* no modelo multidimensional é a que se apresenta na Figura 4.10.

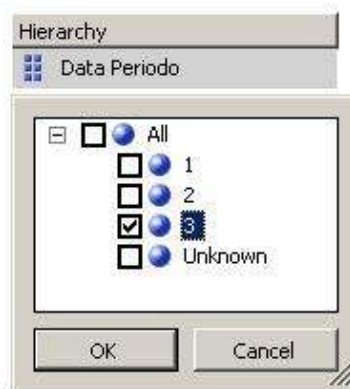


Figura 4.10 - Representação do *DataPeriodo* no modelo multidimensional. Fonte: Elaboração própria.

#### 4.6.2 Tabela de dimensão *DimAluno*

A dimensão *DimAluno* contém a informação referente aos alunos da ESAL. Esta dimensão é composta por um vasto conjunto de atributos, dos quais apresentamos apenas alguns (Figura 4.11).

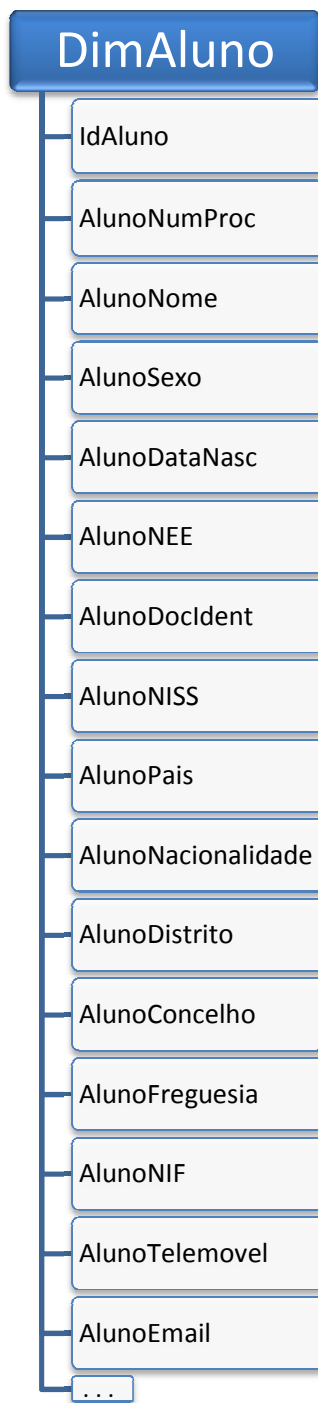


Figura 4.11 - Parte dos atributos da dimensão *DimAluno*. Fonte: Elaboração própria.

#### 4.6.3 Tabela de dimensão *DimMatricula*

A tabela dimensão *DimMatricula* armazena as informações sobre as matrículas de cada aluno. Esta dimensão é composta por vários atributos, onde os mais relevantes constam da Figura 4.12.

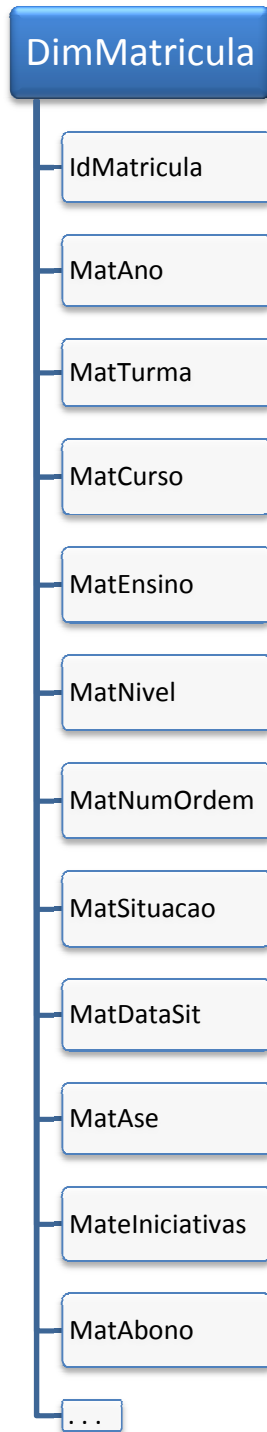


Figura 4.12 - Parte dos atributos da dimensão *DimMatricula*. Fonte: Elaboração própria.

Nesta dimensão uma das hierarquias criada apresenta quatro níveis de detalhe, sendo o nível de menor detalhe o representado pelo atributo *MatNivel*, contendo o atributo *MatEnsino*, dentro deste o atributo *MatAno* e por último o atributo *MatTurma* (Figura 4.13).

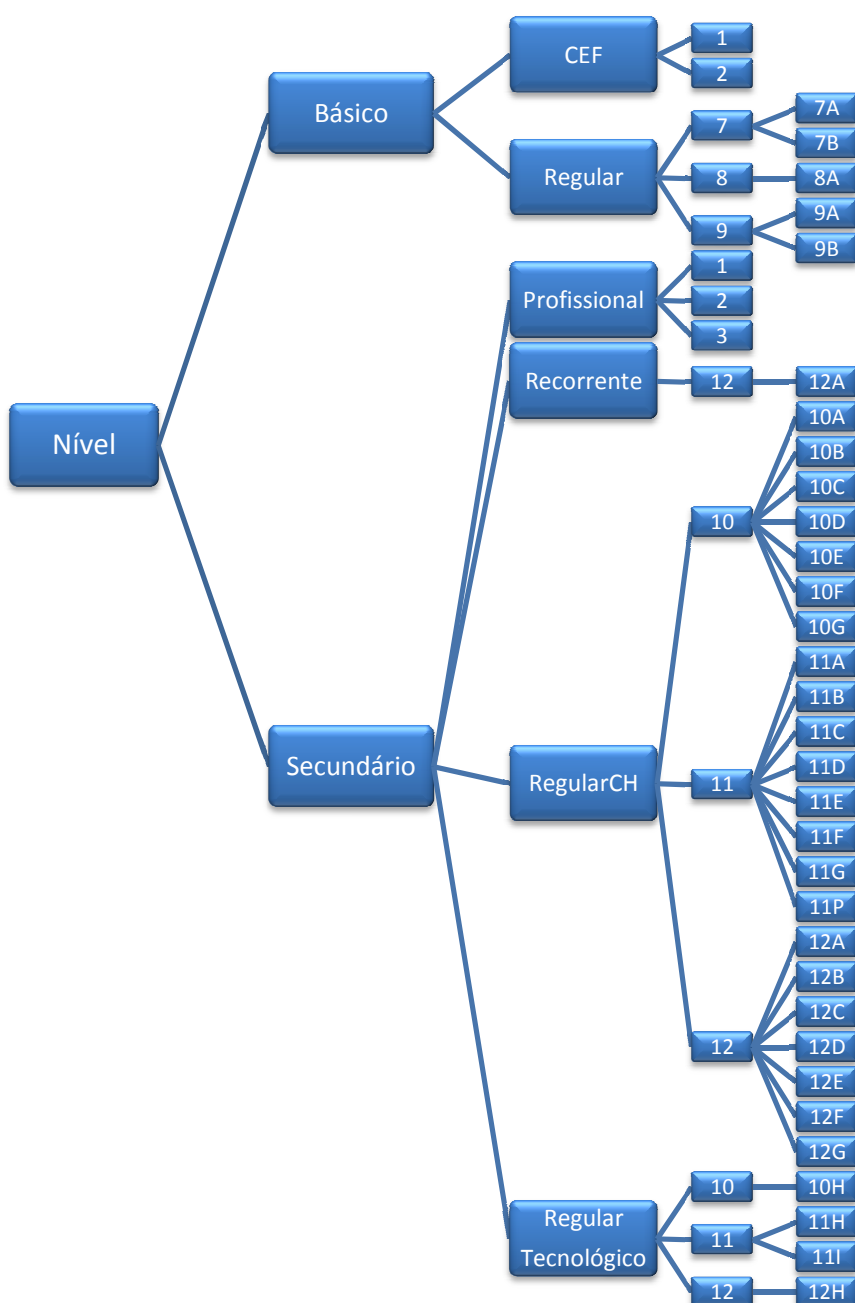


Figura 4.13 - Representação gráfica de uma hierarquia. Fonte: Elaboração própria.

No modelo multidimensional apresenta a estrutura representada na Figura 4.14.

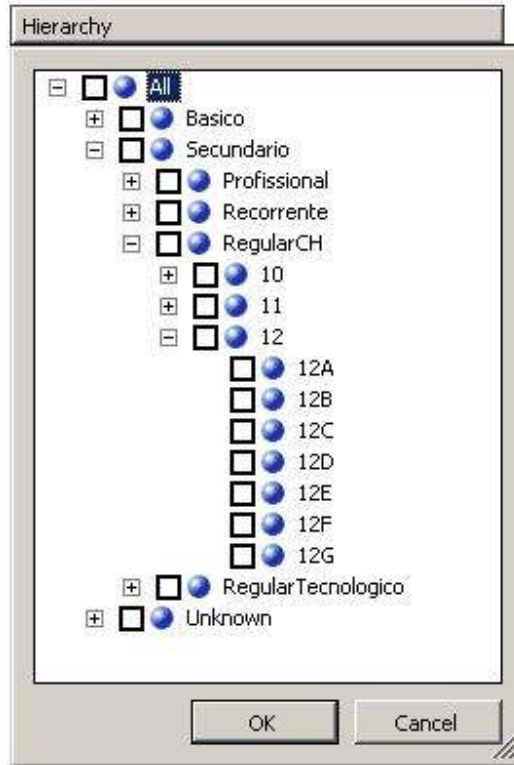


Figura 4.14 - Representação da informação de uma hierarquia. Fonte: Elaboração própria.

#### 4.6.4 Tabela de dimensão *DimDisciplina*

Esta tabela representa as disciplinas dos cursos ministrados na ESAL. Tem por atributos o *IdDisciplina*, o *DiscCod* e *DiscNome*, onde estes dois últimos caracterizam as disciplinas quer pelo seu código, quer pelo seu nome (Figura 4.15).

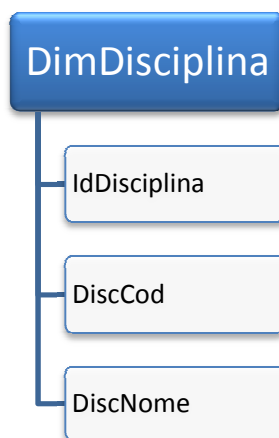


Figura 4.15 - Dimensão *DimDisciplina*. Fonte: Elaboração própria.

#### 4.6.5 Tabela de dimensão *DimProfessor*

A tabela dimensão *DimProfessor* armazena a informação referente aos professores da ESAL. Esta dimensão é composta por vários atributos representados na Figura 4.16.

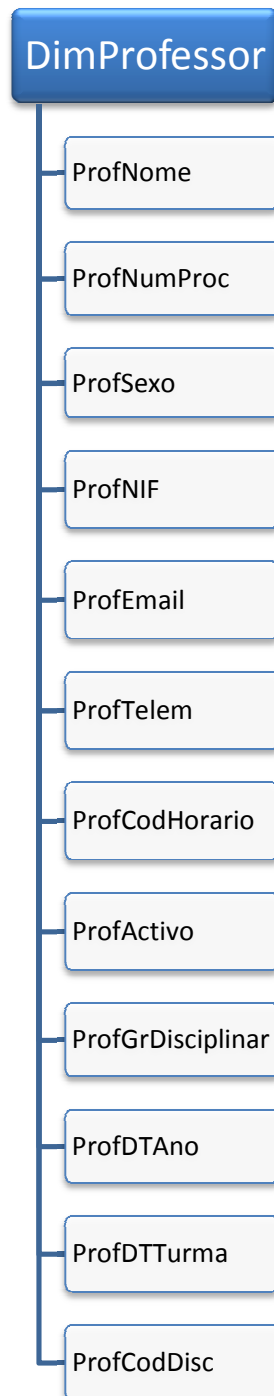


Figura 4.16 - Dimensão *DimProfessor*. Fonte: Elaboração própria.

#### 4.6.6 Tabela de facto *FactoAvaliacao*

A tabela *FactoAvaliacao* contém dois tipos de atributos, as métricas ou factos do modelo multidimensional e as chaves estrangeiras que a ligam a cada uma das tabelas de dimensão. Esta é a tabela dominante de um dos esquemas em estrela e tem como principal característica a presença de dados altamente redundantes com intuito de se obter um melhor desempenho (Figura 4.17).

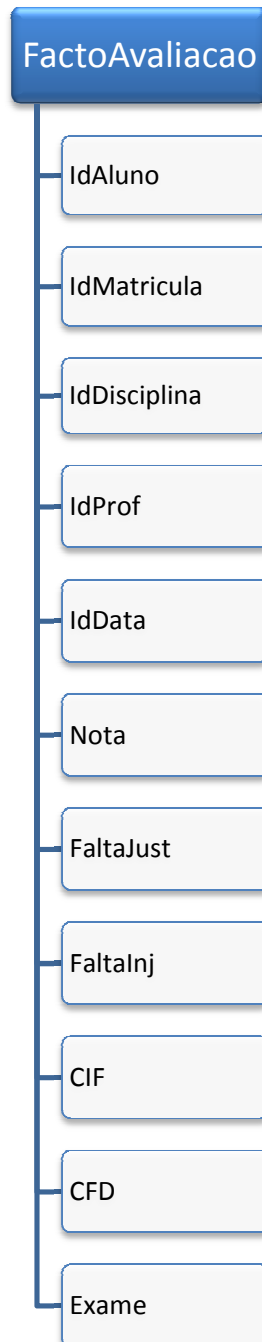


Figura 4.17 - Atributos da tabela facto *FactoAvaliacao*. Fonte: Elaboração própria.

As chaves estrangeiras, para além de estabelecerem regras de integridade referencial com as tabelas de dimensão, servem para utilizar os atributos das dimensões para caracterizar as métricas registadas e para parametrizar as consultas aos dados. As medidas que foram implementadas neste modelo multidimensional são resultado das necessidades de análise sentidas ao longo do desenvolvimento do projecto, estando representadas na Tabela 4.1.

Métricas	Descrição
Nota	Classificação obtida no final de um período lectivo
FaltaJust	Total de faltas justificadas obtidas a uma disciplina
FaltaInj	Total de faltas injustificadas obtidas a uma disciplina
CIF	Classificação Interna Final de uma disciplina
CFD	Classificação Final de Disciplina
Exame	Classificação de Exame Nacional

Tabela 4.1 - Descrição das métricas. Fonte: Elaboração própria.

#### 4.6.7 Tabela de facto *FactoProfissional*

A tabela *FactoProfissional* foi criada dado que as regras usadas na avaliação dos cursos profissionais são distintas das que se praticam nos cursos regulares ministrados na ESAL. Com esta tabela o que se pretendeu foi poder saber a situação do aluno do ensino profissional num qualquer instante do ano lectivo, sem entrar no pormenor das classificações obtidas nos módulos das várias disciplinas dos cursos (Figura 4.18).

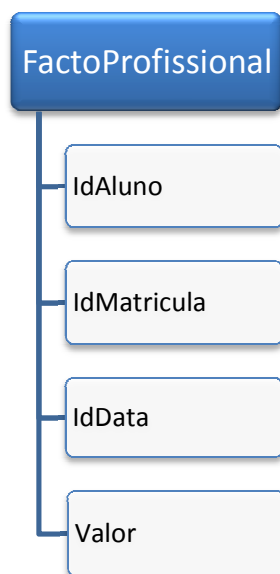


Figura 4.18 - Atributos da tabela facto *FactoProfissional*. Fonte: Elaboração própria.



## 5 Implementação do Protótipo de DW

---

### 5.1 Introdução

Este capítulo descreve a arquitectura e os aspectos de implementação da solução proposta, começando por identificar a abordagem *top-down* como a usada na construção do protótipo de DW, apresentando o ciclo de vida do mesmo. Em seguida é explicado como intervêm os componentes da plataforma de BI da *Microsoft* no processo de implementação do DW, quais os que ainda lhe são acoplados para que seja possível a construção de relatórios por parte dos utilizadores e futuramente partilhados com toda a comunidade escolar através da plataforma colaborativa *Windows SharePoint Services 3.0*, *front-end* do protótipo de DW.

Segue-se uma exposição de como foi feita a integração de dados no protótipo de DW, os processos de ETL, e dos recursos usados da plataforma de BI para tal, bem como do tratamento dos metadados. Em seguida descreve-se a forma como se fez a análise multidimensional, a criação dos cubos de dados, a elaboração dos relatórios e sua disponibilização, bem como dos recursos da infraestrutura tecnológica.

### 5.2 Ciclo de vida do projecto

No desenvolvimento deste projecto de protótipo de DW, iniciou-se por uma abordagem *top-down* depois de reunidos os requisitos necessários para a construção do DW e de ter uma visão geral da ESAL. Após identificação das diferentes fontes de dados, relacionadas com os requisitos estabelecidos como prioritários para o sistema de DW, procedeu-se à sua extracção e de acordo com as regras delineadas para o DW fez-se a limpeza, transformação e agregação dos dados, como se descreverá em detalhe posteriormente neste capítulo.

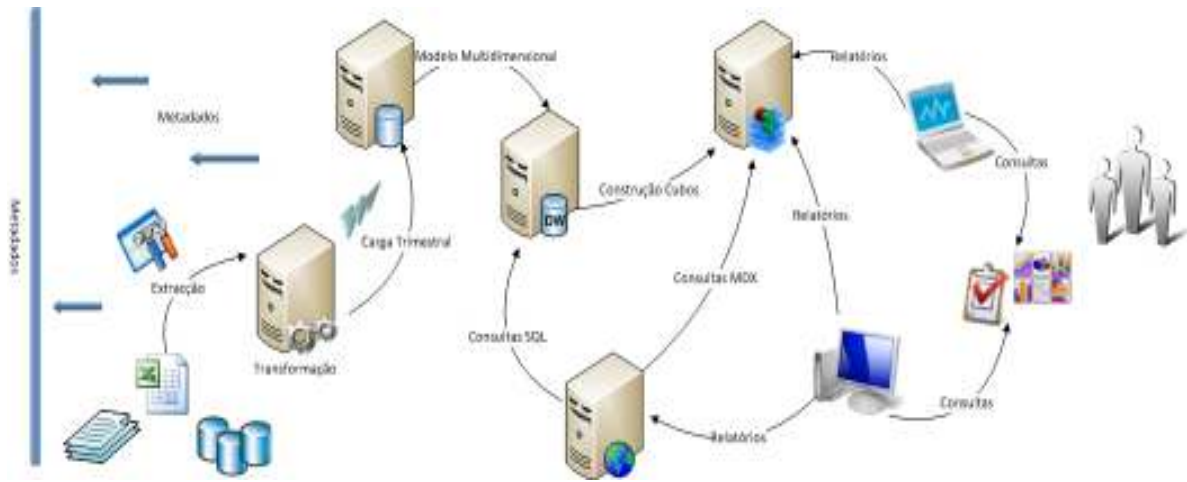


Figura 5.1 - Ciclo de vida do protótipo de DW. Fonte: Elaboração própria.

Foi criado o modelo multidimensional, para o qual foi efectuada a carga dos dados já integrados, povoando assim o DW. Com toda a informação no DW, geraram-se os cubos de acordo com as visões pretendidas, procedendo-se à análise multidimensional através de interrogações SQL à base de dados e *query's* MDX aos cubos gerados. Destas consultas produziram-se os relatórios que se partilham com a comunidade escolar, de forma a serem testadas as funcionalidades do protótipo, mas também para do lado dos utilizadores se verificarem as críticas à sua implementação, tendo em vista o ciclo de melhoria do projecto.

Durante as várias etapas de implementação, foram sendo produzidos e registados alguns metadados, em especial na etapa de ETL, como representado na Figura 5.1.

### 5.3 A plataforma *Business Intelligence* da *Microsoft*

O *Microsoft SQL Server 2008* possui muitos recursos na área de BI, de entre os quais se destacam os seguintes: *Analysis Services*, *Business Intelligence Development Studio*, *Integration Services* e *Reporting Services* (Magalhães, 2009), (Microsoft, 2010a). Para além destes recursos, o *Microsoft SQL Server 2008*, fornece ainda dois ambientes distintos para a construção, implementação e gestão do recurso de *Analysis Services*: O *Business Intelligence Development Studio* e o *SQL Server Management Studio* (Magalhães, 2009).

Para a construção do protótipo de DW a infra-estrutura que se aplicou foi a referida, tendo sido acoplado à mesma plataforma de *Business Intelligence* duas ferramentas de construção de relatórios, o *Excel Services* e o *Report Builder 2.0*, onde o utilizador final terá à disposição o manuseamento de relatórios já existentes, ou a permissão de através destas, construir novos relatórios por ele parametrizados. O *Excel Services* é uma tecnologia do *Microsoft Office SharePoint*

Server 2007 que simplifica a utilização, partilha, segurança e gestão dos livros do *Microsoft Office Excel 2007* como relatórios interactivos de uma forma consistente em toda a organização (Microsoft, 2010c).

Esta interacção com o modelo multidimensional do DW é possível de uma forma simples e directa, através da plataforma colaborativa *Windows SharePoint Services 3.0* que também a este protótipo se acoplou, sendo o *front-end* para os agentes escolares.

Toda esta infra-estrutura está representada na Figura 5.2.

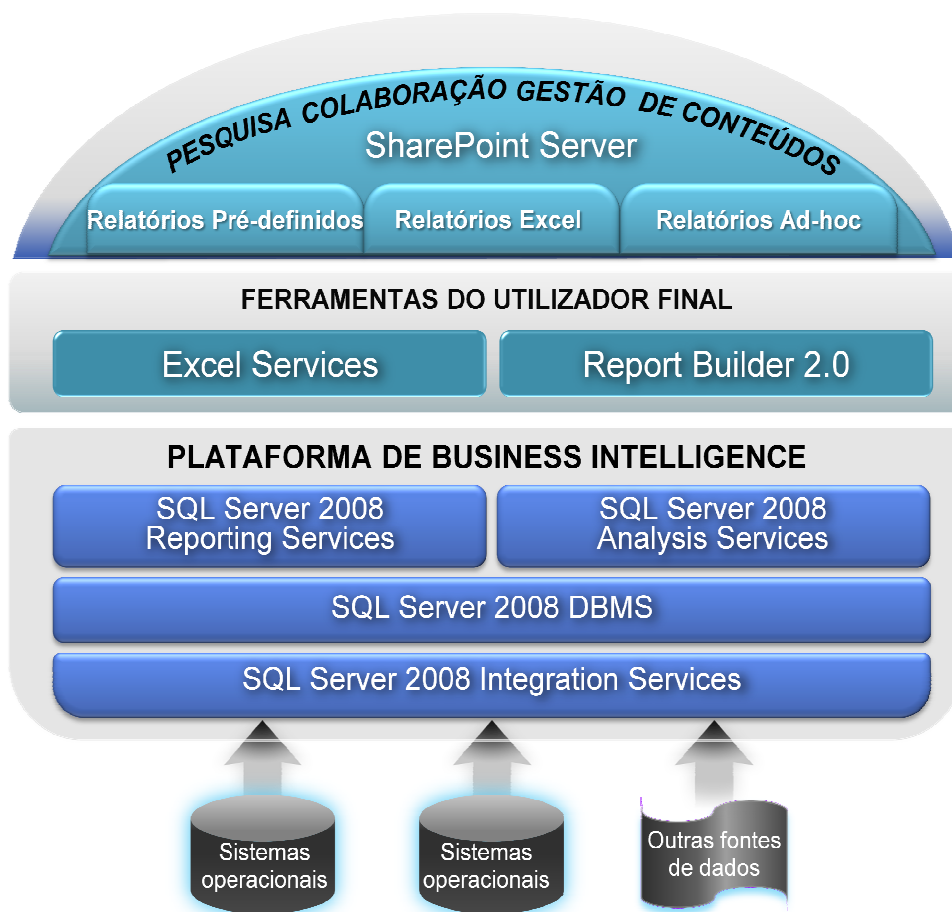


Figura 5.2 - Infra-estrutura tecnológica do protótipo de DW. Fonte: Elaboração própria.

O *SQL Server Database Engine*, através do ambiente *SQL Server Management Studio*, permitiu criar a base de dados multidimensional do DW, como em exemplo se apresenta nas Figuras 5.3 e 5.4. Atendendo ao seu mecanismo de armazenamento de dados escalonável e de alto desempenho para

grandes volumes de dados, foi a escolha ideal para a consolidação dos dados da ESAL num DW central para análise e construção de relatórios.

Através do ambiente de desenvolvimento integrado, o *Business Intelligence Development Studio*, baseado no *Microsoft Visual Studio 2008*, desenvolveu-se a análise multidimensional com a construção dos cubos e as respectivas interrogações em SQL e MDX incluídas na ferramenta *Analysis Services*; a construção dos pacotes com os respectivos manifestos que foram instalados no DW através do *Integration Services* e a criação dos vários relatórios incluídos no *Reporting Services*.

O *Business Intelligence Development Studio* é destinado ao desenvolvimento e distribuição de soluções baseadas na plataforma de *Business Intelligence*, melhorando a produtividade, a performance e o desenvolvimento de soluções baseadas no *SQL Server 2008* (Microsoft, 2010b; Magalhães, 2009).

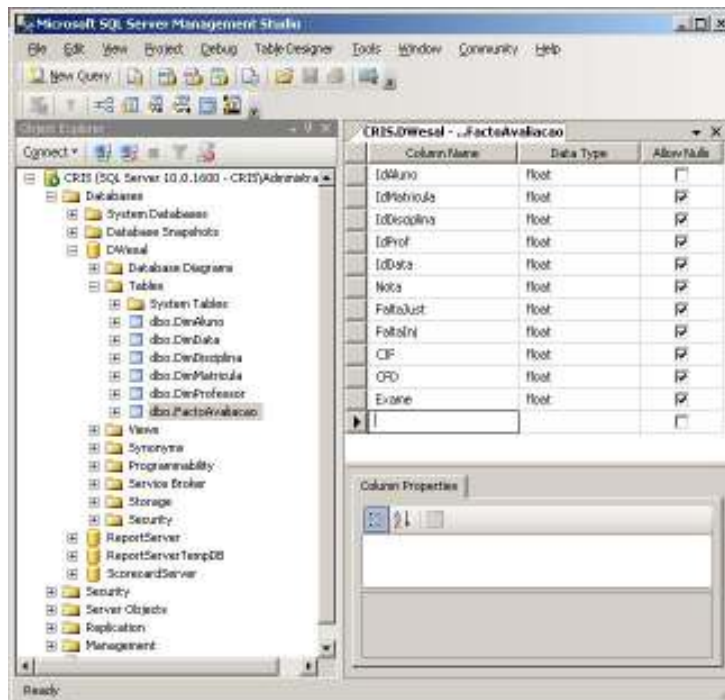


Figura 5.3 - Criação da tabela *FactoAvaliacao* do DW. Fonte: Elaboração própria.



extrair e que tipos de filtros aplicar, sendo esta actividade, uma das que mais tempo consumiu na construção do DW.



Figura 5.5 - Processo de ETL. Fonte: Elaboração própria.

O mecanismo de integração dos dados do *SQL Server 2008* é o *Integration Services* (Microsoft, 2010b). Com esta ferramenta realizou-se a transformação e extracção de dados da *Data Staging Area*, construindo um vasto conjunto de pacotes que permitiram povoar as dimensões e a tabela de factos. Depois de gerados os pacotes no *Business Intelligence Development Studio*, foram criados os respectivos ficheiros de manifesto, utilizados para instalar os pacotes de dados no destino, no *Integration Services* do *SQL Server 2008*, onde como última operação se executaram (Magalhães, 2009).

O *SQL Server Integration Services* (Figura 5.6) possui ligação com várias fontes de dados, agilizando o processo de extracção das informações a serem carregadas no DW. A origem de dados envolvida neste protótipo de DW e onde foram filtrados e tratados os dados, constituindo a *Data Staging Area*, foi o *Excel*. A carga dos dados no *Data Warehouse*, em particular na tabela de factos *FactoAvaliacao*, foi realizada de forma incremental, ou seja, foram inicialmente carregados os dados referentes ao final da avaliação do primeiro período, depois os dados referentes à avaliação final do segundo período e por fim os dados referentes ao final do terceiro período onde já se incluíram os resultados dos exames nacionais.

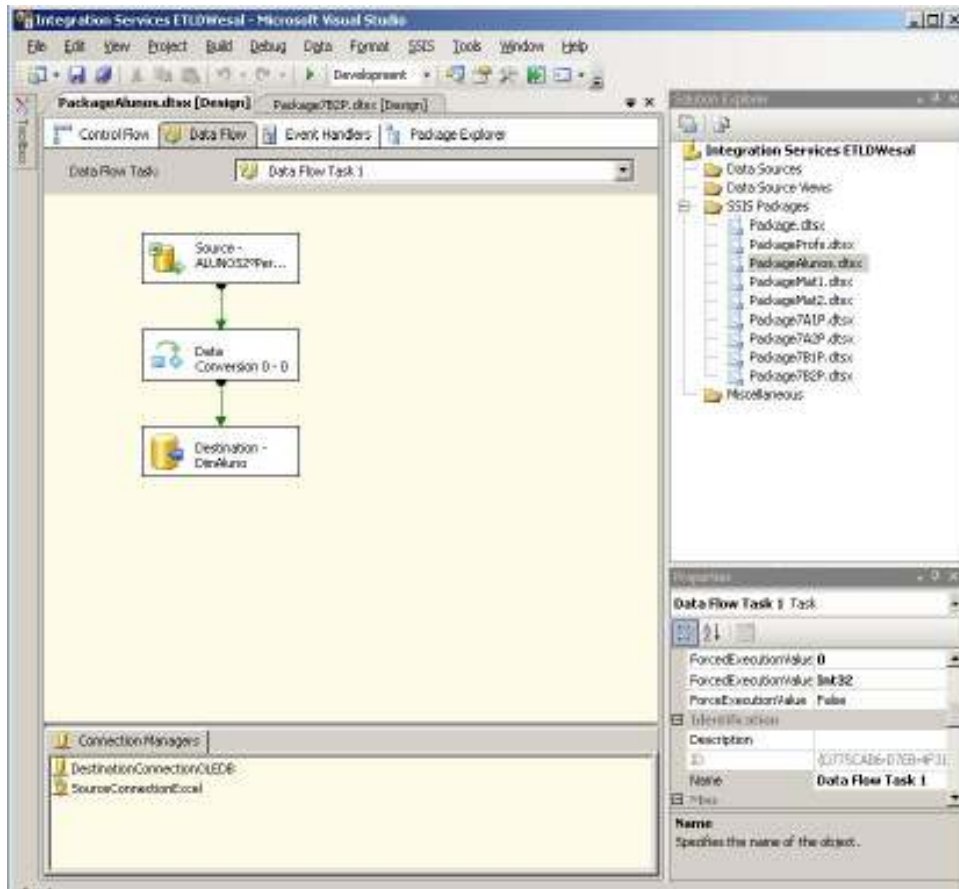


Figura 5.6 - Criação de pacotes de dados no *Business Intelligence Development Studio*. Fonte: Elaboração própria.

Para a tabela de dimensão *DimMatricula*, também foi feita uma carga de dados incremental e com a mesma periodicidade que a da tabela de factos *FactoAvaliacao*, pois contém atributos que poderão ser alterados no decorrer do tempo, como seja o atributo *MatSituacao* que descreve a situação actual da matrícula do aluno.

As restantes dimensões sofreram uma única carga de dados, pois estes só se alteram com o recomeço de um novo ano lectivo.

## 5.5 Gestão de metadados

Na fase de investigação e pesquisa bibliográfica, encontrou-se unanimidade nos autores consultados na referência ao quão importante se não mesmo indispensáveis são os metadados (Inmon, 2005; Kimball et al., 2002), sendo que no entanto pouca bibliografia se encontra sobre como implementar um modelo válido quando este não integra a solução de BI onde se desenvolve o projecto de DW.

Na plataforma de BI da *Microsoft*, a adoptada para este projecto, não existe um gestor de metadados proprietário, pelo que os metadados do protótipo de DW foram registados numa folha de cálculo adaptada (Thornthwaite et al., 2006), onde consta informação sobre os dados e as fontes origem, o mapeamento dos dados, a explicitação de algumas regras de conversão física dos dados, procedimentos e regras associadas ao processo de ETL, a periodicidade com que os dados são carregados, o volume de dados existente e ainda as tabelas, atributos e chaves do DW.

## 5.6 Análise multidimensional e criação dos cubos de dados

Um cubo de dados não é mais do que uma estrutura composta por tabelas de dimensões e por uma tabela de factos. As dimensões incorporam os eixos do cubo correspondendo aos atributos descritores do âmbito analisado sendo associados os respectivos dados/medidas da tabela de factos.

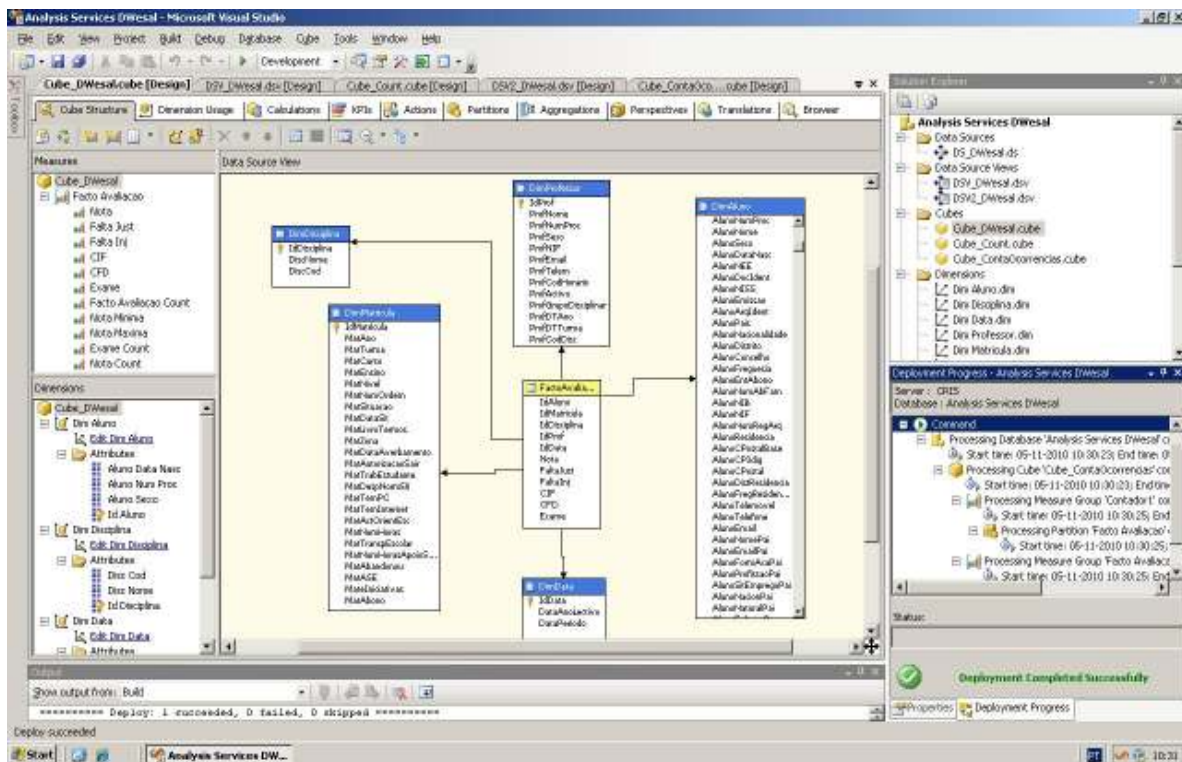


Figura 5.7 - Cubo de dados gerado no *Business Intelligence Development Studio*. Fonte: Elaboração própria.

Um cubo armazena todas as informações relacionadas com um determinado assunto, de forma a possibilitar que sejam realizadas várias combinações entre elas, resultando na extração de várias visões sobre o mesmo tema.

Assim, através do *SQL Server Analysis Services* (Figuras 5.7 e 5.8) e da pré-definição das consultas pretendidas geraram-se vários cubos para futura análise, definindo-se numa primeira fase as fontes de dados dos cubos, gerando-se as diferentes visões dessas mesmas fontes. Em seguida passou-se à selecção das tabelas de dimensão e facto que iriam constar dos cubos. Depois de criados e implementados, os cubos ficam disponíveis para serem acedidos para análise das suas informações, não sem antes serem testados no próprio *SQL Server Analysis Services* apenas com o simples “arrastar” com o rato de atributos das dimensões e das medidas para dentro da área do cubo, gerando-os no próprio ambiente de trabalho.

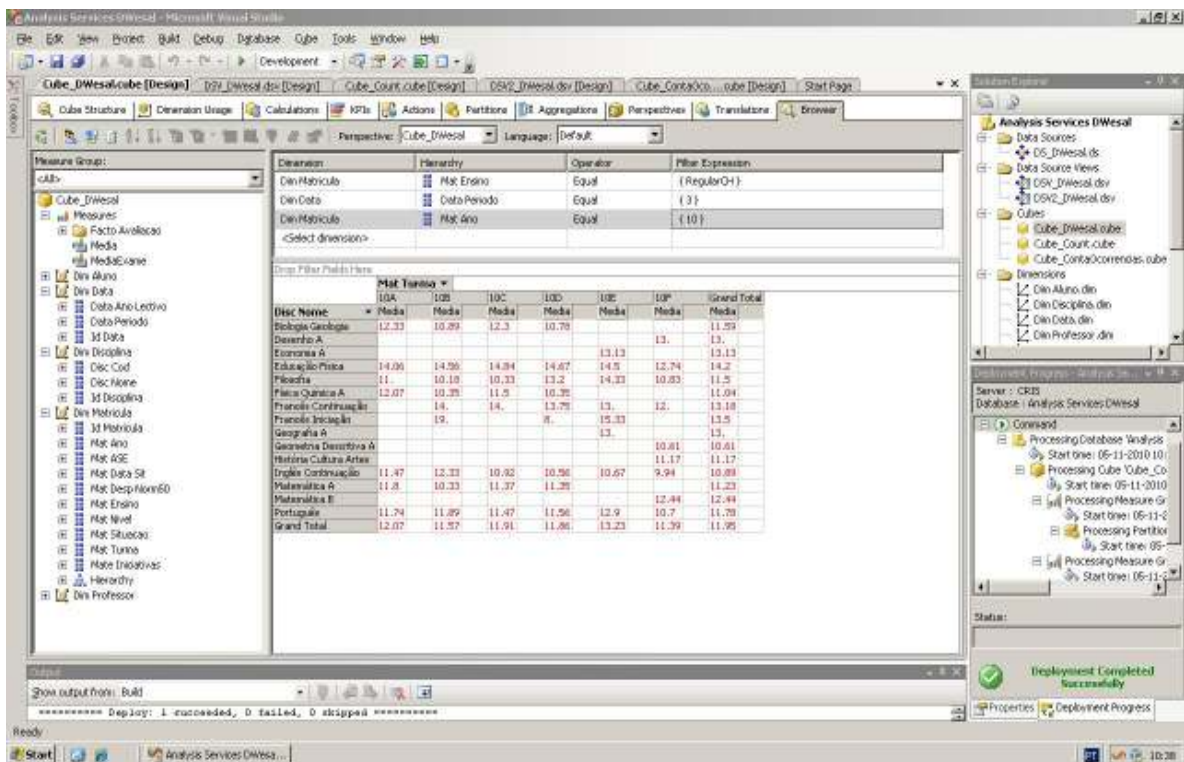


Figura 5.8 - Teste do Cubo de dados no ambiente de trabalho. Fonte: Elaboração própria.

## 5.7 Elaboração de relatórios

De entre os diversos benefícios que um DW traz ao quotidiano de uma organização, o mais imediato de constatar é a facilidade de extracção de relatórios integrados bem como a diminuição de tempo despendido na sua construção. A construção de relatórios sem um DW, numa organização que possui várias aplicações operacionais e ainda dados dispersos por outros tipos de suporte, tem de em primeiro lugar localizar a origem dos dados; recuperar os dados e extraí-los de cada um dos sistemas operacionais ou da origem identificada; transformar/integrar os dados padronizando-os;

reconciliar os dados; modificar as estruturas de dados; reunir os dados e, finalmente, construir o relatório. Usando um DW, todo este processo se dilui, atendendo a que a maioria dos processos descritos será apenas executado uma única vez, aquando do processo de ETL.

O *SQL Server 2008* permite a construção de relatórios através da ferramenta *Reporting Services*, no formato que se pretenda e com um *layout* que faça sentido com os dados nele contidos (Figura 5.9).

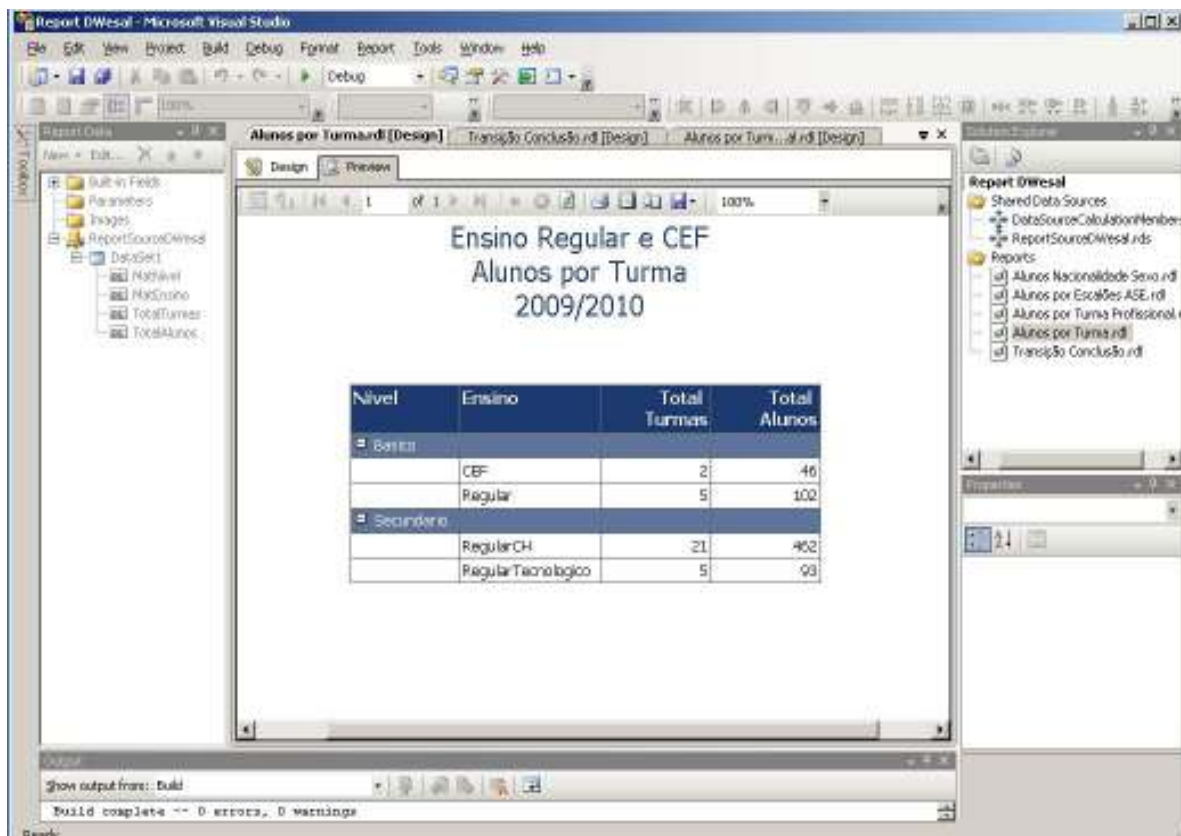


Figura 5.9 - Construção de relatórios no ambiente de trabalho. Fonte: Elaboração própria.

## 5.8 Disponibilização dos relatórios

Nos dias de hoje há uma necessidade básica em manter a informação organizada e centralizada num local e não em pessoas, independentemente do tamanho da organização. Não é tarefa fácil administrar a partilha de ficheiros online, seja ela através de páginas internet, pela dificuldade de actualização e criação das mesmas, seja através de correio electrónico que além de não ser uma boa prática para a partilha de informação, possui ainda alguma limitação no tamanho dos ficheiros. O *Windows SharePoint Services 3.0* (Figura 5.10) foi a plataforma escolhida para *front-*

*end* com a comunidade escolar. Trata-se de um componente que foi integrado ao *Windows Server 2003*, que com a sua magnitude permite a partilha de informação de forma simples, através de um portal de conteúdos onde os ficheiros são sincronizados pela *internet* no momento em que o mesmo é aberto ou fechado, possuindo uma grande capacidade de colaboração, permitindo que todos contribuam sem necessitar de suporte ou apoio directo.

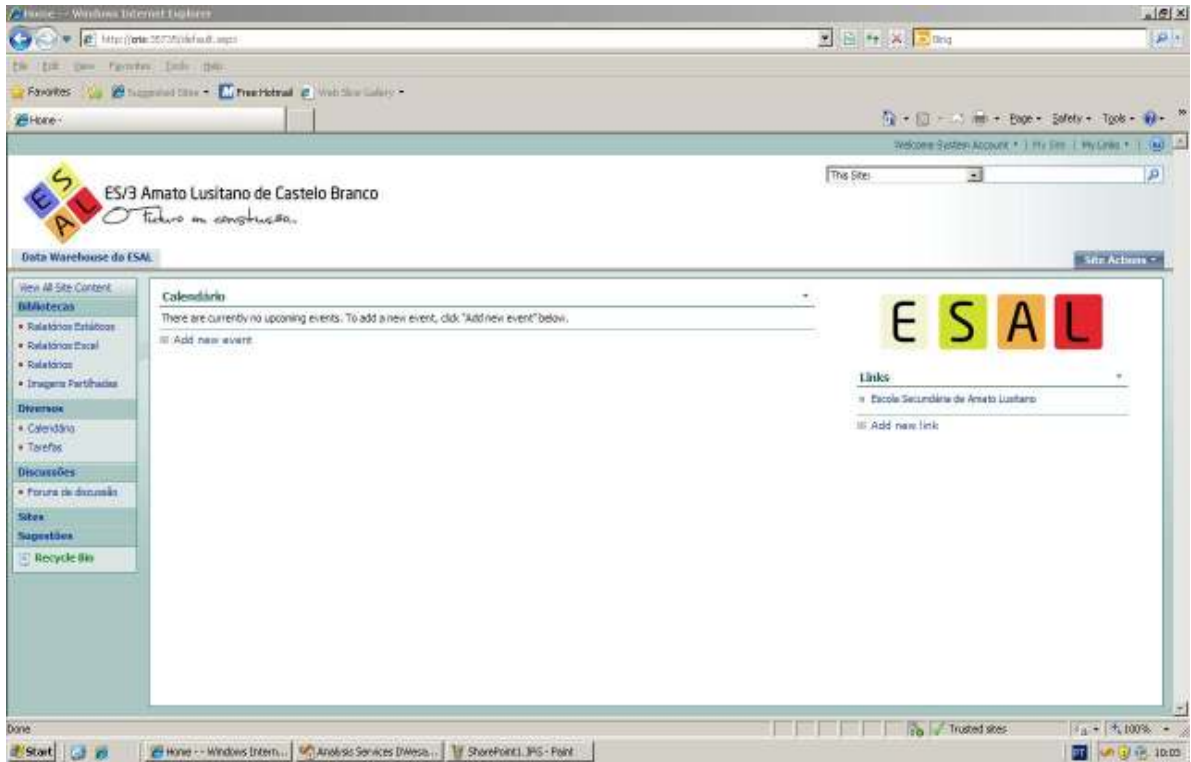


Figura 5.10 - Interface do *front-end* de disponibilização de relatórios. Fonte: Elaboração própria.

Quanto à usabilidade, disponibilizaram-se dois tipos de relatórios: os estruturais, com uma estrutura fixa pré-definida e os *Ad-Hoc*, relatórios construídos a pedido, com objectivos específicos e uma utilização casual, parametrizada pelo próprio utilizador.

Assim, com o intuito de facilitar todo o processo de análise aos dados do DW, foram desenvolvidas algumas consultas pré-definidas a partir das informações dos cubos de dados gerados pela análise multidimensional feita no *SQL Server Analysis Services*. Desta forma, os utilizadores poderão realizar consultas analíticas para controlar o desempenho dos alunos e professores da ESAL e desenvolver estratégias para as melhorar.

Um exemplo de relatório pré-definido que se pode encontrar na secção Relatórios Estáticos é o que se apresenta na Figura 5.11.

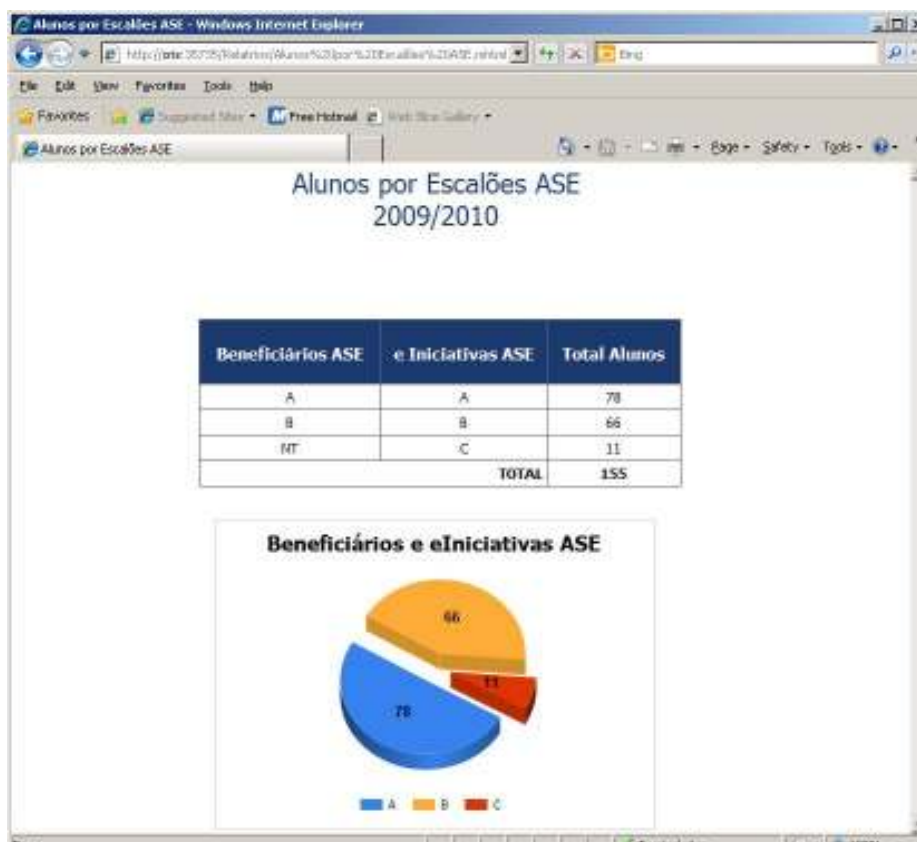


Figura 5.11 - Relatório pré-definido da secção Relatórios Estáticos. Fonte: Elaboração própria.

Além das consultas pré-definidas, é ainda disponibilizado ao utilizador a opção de criar consultas personalizadas a partir de qualquer um dos cubos de dados gerados. Estas consultas poderão ser gravadas e acedidas posteriormente, funcionalidade que aumenta ainda mais a agilidade na tomada de decisões, pois permite que os próprios utilizadores criem novas consultas sempre que surgir a necessidade de obter determinada informação não disponível nas consultas já existentes. Os utilizadores poderão assim criar tantas consultas quantas forem necessárias para satisfazer as suas necessidades de informação, pois não há limite para a criação das mesmas.

Através do *Report Builder 2.0* acoplado ao *SQL Server Reporting Services* deu-se permissão aos utilizadores para criarem relatórios *Ad-Hoc* de uma forma simples, bem como através do *Excel Services*, ferramenta que acede aos dados do *SQL Server 2008* quando integrado com o *Microsoft Office SharePoint Server* solução de *Business Intelligence* escolhida para a interacção com os cubos de dados no servidor, de forma rápida, simples e eficiente, permitindo a publicação de análises e relatórios efectuados ao DW deixando-os disponíveis a toda a comunidade escolar.

Um exemplo de relatório *Ad-Hoc* elaborado através do *Report Builder 2.0* é o que a seguir se apresenta na Figura 5.12.

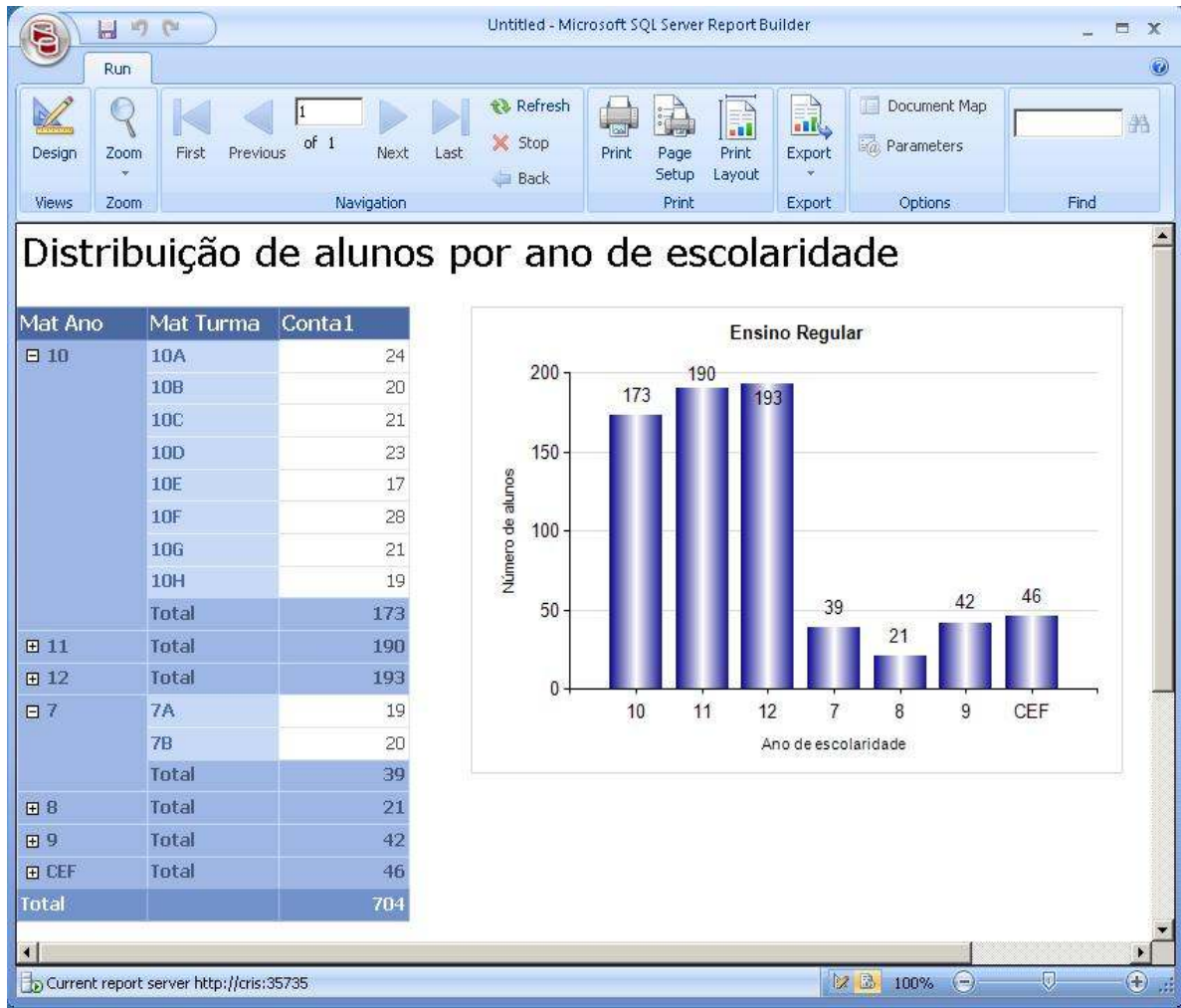


Figura 5.12 - Relatório *Ad-Hoc* elaborado no *Report Builder 2.0*. Fonte: Elaboração própria.

Depois de construído o relatório, este poderá passar a constar da lista de relatórios a disponibilizar à comunidade escolar, tornando-se uma mais-valia a plataforma colaborativa através da qual se acede aos dados do DW e se partilham os relatórios.

Todos os relatórios podem ser exportados para o formato que o utilizador pretenda, para que os possa analisar em pormenor e incluir noutros documentos, tal como se representa na Figura 5.13.

The screenshot shows a web browser window displaying a report titled "Ensino Regular e CEF Alunos por Turma 9/2010". The report is presented in a table with two columns: "Total Turmas" and "Total Alunos". The "Export" menu is open, showing options for exporting the report to various formats: Arquivo XML com dados de relatórios, CSV (delimitado por vírgulas), Arquivo Acrobat (PDF), MHTML (arquivo da Web), Excel, Arquivo TIFF, and Word. The table data is as follows:

	Total Turmas	Total Alunos
CEF	2	46
Regular	5	102
<b>Secundario</b>		
RegularCH	21	462
RegularTecnologico	5	93

Figura 5.13 - Exportação de relatórios para outros formatos. Fonte: Elaboração própria.

Outra solução implementada neste protótipo de DW para a disponibilização de relatórios foi efectuada através da integração do *Excel Services* à plataforma de BI, ferramenta para o acompanhamento de resultados e tomadas de decisão quando ligada ao *Analysis Services* onde as informações são actualizadas automaticamente baseando-se no processamento de cubos OLAP. O *Excel Services* faz parte de uma colecção de funcionalidades do *Microsoft Office SharePoint Server* designadas em conjunto por *Business Intelligence* e que podem ser utilizadas por um indivíduo, uma equipa ou uma organização (Microsoft, 2010c).

É facultado aos utilizadores o acesso em tempo real a partir de um *browser* a folhas de cálculo interactivas do *Microsoft Office Excel 2007* através da execução do *Excel Services* no *Microsoft Office SharePoint Server*, onde estes podem usar relatórios existentes ou construir novos, atendendo

a que se pré-estabeleceram ligações aos vários cubos de dados gerados no *SQL Server Analysis Services* (Figura 5.14).

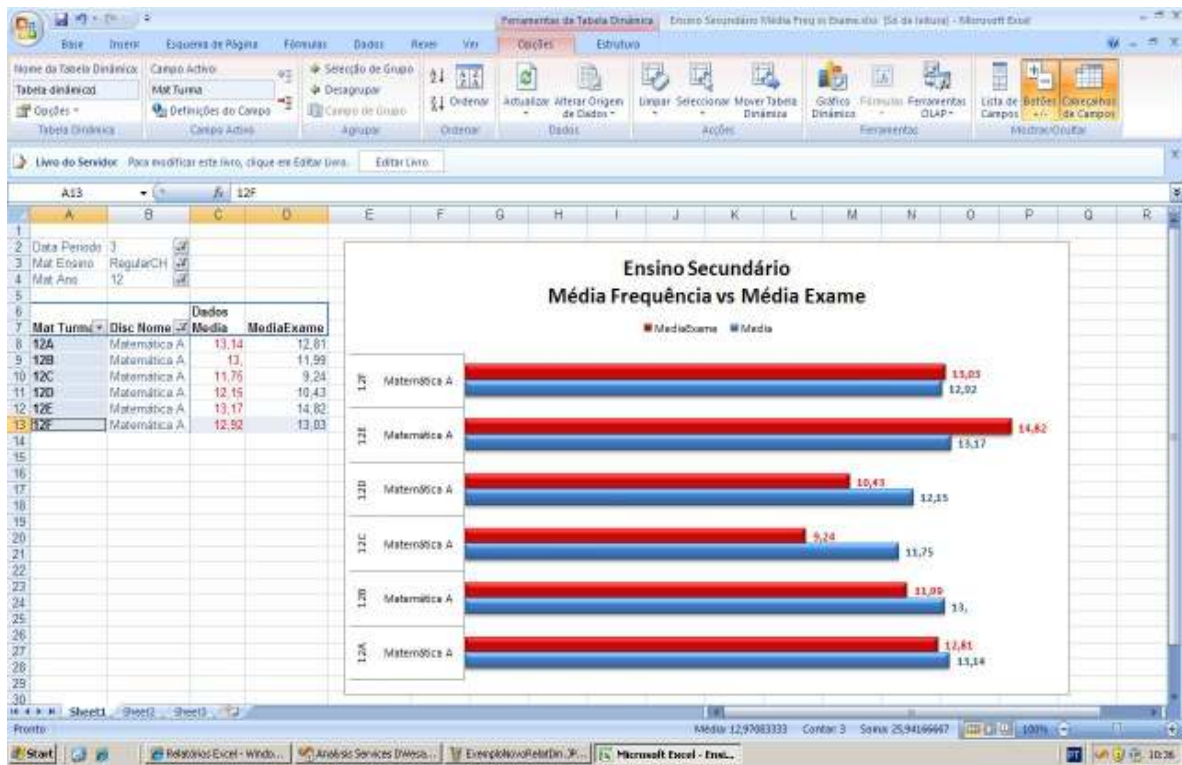


Figura 5.14 - Relatório produzido através do *Excel Services*. Fonte: Elaboração própria.



## 6 Conclusões

---

### 6.1 Reflexão crítica

Este trabalho resultou da percepção da importância e capacidade dos Sistemas de *Data Warehouse* na ajuda à tomada de decisão nas organizações. Dado que uma escola é uma organização com características particulares, onde o investimento dos agentes escolares é mensurável, surgiu a ideia deste projecto. Este protótipo de DW surgiu assim com o intuito de permitir analisar a informação da ESAL segundo várias parametrizações, de forma rápida e coerente sobre todos os dados, disponibilizando quer os resultados quer a capacidade de análise a toda a comunidade escolar.

Grande parte da informação desejada pela comunidade escolar não está disponível nos sistemas operacionais de uma escola, tendo em consideração que tais sistemas não foram projectados para atender necessidades analíticas. Durante o desenvolvimento do projecto, constatou-se que um ambiente de DW é o indicado quando existe necessidade de agregar dados de vários sistemas com vista a fornecer informações exactas, indo de encontro às expectativas do desenvolvimento dum protótipo de DW para apoio à decisão. Uma das vantagens da implementação deste protótipo de DW é propiciar a todos os elementos da ESAL uma economia de tempo e esforço no processo de análise e tomada de decisão. Outra vantagem que se deve evidenciar é a facilidade em comprovar os indicadores pré-estabelecidos na ESAL relativamente a cada uma das disciplinas dos cursos que ministra. Importa ressaltar, que a utilização deste ambiente não implica necessariamente melhores decisões, mas permite que as decisões se baseiem em melhores informações.

Na fase de investigação e após a revisão da literatura observa-se que há muitas dúvidas em relação às melhores práticas de construção e implementação de sistemas de DW que suportem as decisões estratégicas de uma organização, tal como foi abordado no segundo capítulo com a apresentação de metodologias de autores conceituados na área em causa. Os resultados da revisão

bibliográfica efectuada permitiram obter o conhecimento necessário, servindo de fundamentação ao processo de estruturação, modelação e implementação do protótipo de DW.

Para a criação e implementação do protótipo de DW, foi elaborado um caso de estudo aplicado à ESAL, através do qual as diferentes fases do projecto de DW foram analisadas e executadas. Futuramente, poderá ser consolidado este protótipo, modelando-se novos cubos de dados para os indicadores estratégicos referidos no BSC do segundo capítulo.

Do desenho do modelo conceptual e do modelo lógico multidimensional do DW e com o estudo prévio da plataforma de BI da *Microsoft*, implementou-se a infra-estrutura tecnológica, nem sempre com a facilidade que se desejava, pois surgiram incompatibilidades entre componentes que tiveram de ser ultrapassadas. Superadas as dificuldades e solucionados os problemas da infra-estrutura tecnológica, implementou-se o modelo físico do DW através do *Business Intelligence Development Studio*. Neste modelo multidimensional de dados optou-se por um esquema em estrela, dado que face às suas características permite interfaces mais compreensíveis e um processamento mais eficiente e onde novos elementos podem ser facilmente adicionados, não sendo necessário mudanças nas consultas e nos relatórios existentes, nem na recarga de dados.

No processo de ETL, também nem tudo foi simples, pois as fontes de dados eram várias, heterogéneas e redundantes, com a agravante de a mesma informação se encontrar especificada de forma diferente. Depois de extraídos os dados, estes foram transformados e limpos no *Data Staging Area*. No caso em particular deste protótipo, o *Excel* foi o *Data Staging Area* usado, onde os dados foram preparados para a carga no DW, feita através do *SQL Server Integration Services*. Foi de todos, o processo mais crítico e moroso deste projecto.

Com o DW povoado, procedeu-se à construção dos cubos de dados e respectiva análise multidimensional, através de interrogações SQL e MDX no *SQL Server Analysis Services*, processo que não envolveu grandes dificuldades. Através do *SQL Server Reporting Services* geraram-se os relatórios que, após integração e configuração do *Windows SharePoint Services* à restante plataforma, aí se disponibilizaram. Fez-se ainda a integração e configuração no *Windows SharePoint Services* de dois componentes que permitissem por parte do utilizador a produção e parametrização dos seus próprios relatórios, o *Report Builder 2.0* e o *Excel Services*. Desta forma, os utilizadores poderão realizar consultas analíticas para controlar o desempenho dos alunos e professores da ESAL e desenvolver estratégias para as melhorar.

Criou-se assim um armazém de dados integrado, coerente, de fácil acesso, estruturado de forma a suportar análises complexas, permitindo vastos critérios de pesquisa e que permite responder a questões de entidades externas e internas da organização.

Outra contribuição a considerar, fruto deste trabalho, é a verificação da aplicabilidade da ferramenta de *Business Intelligence* da *Microsoft* no desenvolvimento e implementação de *Data Warehouse*, revelando-se bastante completa e requerendo equipamentos com poucos recursos.

## 6.2 Avaliação do trabalho realizado

O desafio fundamental proposto nesta dissertação passava pelo desenvolvimento e criação de um protótipo de *Data Warehouse* solução que pudesse colmatar as necessidades de análise da ESAL. Esse desafio foi cumprido, atendendo a que todos os objectivos inicialmente propostos foram executados, sendo eles:

- Projectar e desenvolver um protótipo de *Data Warehouse*;
- Aplicar as tecnologias de modelação multidimensional;
- Seleccionar, preparar, limpar e carregar o *Data Warehouse* com os dados extraídos dos sistemas transaccionais da Escola Secundária/3 de Amato Lusitano. Os dados extraídos são provenientes dos programas de gestão de alunos WinGA e InovarProfissional; dos programas dos exames nacionais do ensino secundário, o ENES e do básico, o ENEB; do PortalSige; do programa de acção social escolar ASE; do programa InfoPonto e FaltasDiarias; e ainda de pautas e mapas estatísticos do *Excel* produzidos em Conselho de Turma. Os cubos foram desenvolvidos de acordo com o previsto, representando uma eficiência e estabilidade acima do inicialmente previsto, que contempla toda a informação disponível no DW.
- Produzir e disponibilizar relatórios que possibilitem extrair informação coerente. A construção dos relatórios tornou-se uma tarefa pacífica dada a estrutura que a suporta, tendo sido dada especial atenção à forma como os utilizadores irão interagir com a informação do DW, o portal colaborativo. Neste, encontram-se disponíveis as seguintes análises: relatórios de análise comparativa entre exames nacionais e classificações internas de frequência; relatórios de análise da distribuição de alunos por escalões de Acção Social Escolar; relatórios de análise comparativa das classificações internas de frequência por disciplina e por turma; relatórios de análise do total de alunos transferidos, anulação de matrícula e excluídos por faltas por ano e sexo; relatórios de análise de total alunos transitados e aprovados por ano, nível e ensino; relatórios de análise da distribuição de alunos com Necessidades Educativas Especiais entre outros.

Este projecto limitou-se a concretizar os objectivos estabelecidos. Dentro desta limitação não foram detectados problemas que reduzissem ou inviabilizassem a sua realização. Ter-se-á concretizado um pouco mais do que o que inicialmente foi delineado, já que foi implementado um

*front-end* para disponibilização e construção dos relatórios, através da plataforma colaborativa *SharePoint*. Por outro lado faltou implementar a vertente estratégica do BSC, que requer dados ainda não disponíveis na ESAL, podendo ser objecto de trabalho futuro.

Dos testes efectuados ao protótipo de DW com alguns docentes e direcção da ESAL, as críticas foram positivas. A possibilidade de gerar relatórios de uma forma simples e quase intuitiva foi o mais realçado, bem como a capacidade de poderem ser gerados para *Excel*, ferramenta de domínio generalizado, muito conhecida e utilizada por todos. Também o facto de os relatórios serem disponibilizados através da plataforma colaborativa *SharePoint*, acessível via *Web* a partir de qualquer computador da escola, foi mais um factor destacado, permitindo um melhor acompanhamento do desempenho dos alunos e professores. Neste ambiente de análise de dados é mais simples, eficiente e eficaz dar resposta às questões que, por exemplo, são colocadas pelo Ministério da Educação (IGE, 2008), aquando da realização das avaliações externas (Lei 31/2002), facilitando todo o processo de recolha e análise de dados até aqui usado.

O Director da ESAL, da demonstração efectuada, reforçou a crítica à capacidade do sistema responder com aparente simplicidade, dinamismo e rapidez às questões pretendidas, facilitando o planeamento estratégico e a tomada de decisões, mais ainda com a possibilidade de num trabalho futuro ser passível de se incluírem novos modelos multidimensionais com outros assuntos que não restringidos aos da avaliação.

As principais dificuldades encontradas neste projecto, que preferimos encarar como desafios e não dificuldades, foi a integração de todas as ferramentas da infra-estrutura tecnológica utilizadas para o desenvolvimento e implementação do protótipo de DW, atendendo a que as exigências e configurações a considerar, por vezes eram distintas de ferramenta para ferramenta. Outro problema, e que confirma as dificuldades de projectos de DW, prende-se com processos de ETL que se revelaram morosos e muito trabalhosos, tendo sido os dados preparados e analisados manualmente, podendo ser crítico se uma informação é carregada de forma errada, comprometendo a qualidade dos dados e posterior fiabilidade do DW.

### **6.3 Trabalho futuro**

O trabalho desenvolvido, apesar de consistir um sistema completamente funcional e que se adapta às necessidades da ESAL, tem ainda espaço para algumas melhorias, bem como a supressão de algumas limitações. Aponta-se como limitação o facto de apenas terem sido implementados modelos multidimensionais relacionados com a avaliação, objectivo base para o protótipo em causa.

A problemática da qualidade dos dados em sistemas de DW em geral e no caso em particular atendendo à diversificação de origens dos mesmos e à impossibilidade de acesso directo às bases de dados operacionais poderá ser resolvida em trabalhos futuros. Seria interessante assim, como

trabalho futuro a investigação em temas como, o desenvolvimento de mecanismos de ETL que prevejam a automatização ou semi-automatização de entrada dos dados no DW e que desse modo contribuam para a diminuição do principal factor de origem das deficiências nos dados.

A utilização de DW com disponibilização da informação via *internet*, apresenta também grandes desafios de segurança, colocando-se algumas questões éticas, sendo dever de qualquer organização proteger a privacidade e confidencialidade das informações nele contidas. Não foi este o âmbito nem objecto de estudo considerado neste projecto, embora os acessos ao portal colaborativo sejam feitos mediante *login* e *password*. Num trabalho futuro seria pertinente fazer um estudo relacionado com o controlo dos acessos à informação, para que políticas de privacidade, confidencialidade e segurança sejam implementadas, garantindo por exemplo, a inviabilidade de acessos não autorizados, bem como vulnerabilidades do *login/password*. Poderá ainda considerar-se uma monitorização adequada do uso do DW. A segurança deve ser considerada mais uma funcionalidade do DW e, embora se saiba que não há sistemas seguros, há sempre possibilidade de implementar medidas que previnam, detectem ou minimizem problemas ou falhas de segurança. Pelos motivos apresentados e por não ser um objectivo deste estudo, neste protótipo não foram considerados os dados da aplicação de gestão de recursos humanos GestRec da ESAL, podendo futuramente ser modelado um novo cubo neste âmbito.

Poderia ainda ser implementado em trabalho futuro, outros cubos de dados para responder a questões de tipologia diferente das do âmbito da avaliação, também elas pertinentes numa organização como a ESAL, como acima referido e no BSC apresentado. Para o efeito, o desenvolvimento de uma série de inquéritos ao meio onde a escola está inserida, pode aferir outras necessidades, que poderão protagonizar novos modelos multidimensionais, para posterior inserção e análise no protótipo de DW. Idem para inquéritos aos alunos, pais e sociedade, acerca da sua opinião da escola, dos professores, dos funcionários, dos órgãos de direcção e outros. Em resumo, um projecto de BI é um processo constante, os requisitos das organizações mudam constantemente e o mesmo acontece com as aplicações de BI. É por isso necessário estar preparado para a necessidade de ir actualizando as ferramentas às necessidades estratégicas da organização.

Por fim, porque não num trabalho futuro, implementar-se sistemas de DW semelhantes em outras instituições de ensino.

#### 6.4 Considerações finais

Nos tempos que correm, é já frequente a utilização de sistemas de DW, em áreas como bancos, seguradoras, comércio, *marketing*, finanças, instituições governamentais, entre outras, onde o volume de informação é considerável e as decisões a tomar fulcrais. Estes sistemas revelam-se uma mais-valia para a tomada de melhores e mais fundamentadas opções, com a integração dos

dados da organização num único sistema, permitindo a execução de análises detalhadas e dinâmicas da informação, na forma de relatórios com tabelas, gráficos e mapas, que permitem ter uma visão global na tomada de decisão, agilizando a eficiência operacional. Quanto mais são usados os DW, mais se percebe o quão útil são para a gestão estratégica das organizações.

O objectivo primordial das organizações é obter melhores resultados nos seus negócios, tornando-se mais eficientes e cada vez mais competitivas. Para tal, as organizações que melhor uso fizerem da informação e do conhecimento, mais se distinguirão ao nível competitivo.

No entanto, vários são os factores intervenientes no sucesso de sistemas de DW, como a qualidade dos dados, em que estes devem ser actuais, completos e fiáveis, a fim de possibilitar correctas análises às operações de negócios para as decisões operacionais e estratégicas a tomar. Os dados operacionais são armazenados de forma redundante e em formatos diferentes, podendo estar correctos numa base de dados e incorrectos ou desactualizados noutra. Só com uma correcção manual os dados poderão ser confiáveis, implicando custos e recursos humanos.

A dispersão de assuntos de negócio representados num DW, em que vários dados se cruzam, implica também um maior dispêndio de tempo. Por um lado no processo de ETL, na limpeza, transformação dos formatos de origem para os formatos pré-definidos no DW e integração dos dados. O processo de ETL revela-se o processo mais complexo e demorado de um projecto de DW, comprometendo o tempo global da sua construção. Por outro lado, todas as fontes de dados dos vários assuntos de negócio têm de ser integradas, atendendo às necessidades dos utilizadores da organização, que por vezes são ambíguas e pouco precisas, levando tempo na recolha dos requisitos e na decisão de que dados devem ser armazenados. A consciencialização de que os requisitos das organizações estão em constante mudança é imprescindível, sendo necessária a sua adequação e das ferramentas de BI ao longo do tempo às necessidades estratégicas da organização.

A forma como é disponibilizada a informação ao utilizador, deve ser simples e intuitiva, não dependendo da mudança de ferramentas associadas ao evoluir das tecnologias. Se estas soluções se apresentarem via internet, a procura de informação o mais actual possível vai ocorrer, situação que também ela pode comprometer o DW.

Como pontos fracos podemos assim apontar o custo, o tempo e o trabalho despendido na sua implementação e manutenção, não esquecendo que é fundamental incutir mudanças culturais na organização e no próprio perfil do utilizador, que deve passar a ter uma visão mais analítica dos dados.

Contraopondo aos pontos fracos, destacam-se como pontos fortes a possibilidade de um desenvolvimento gradual do DW, com uma abordagem incremental, priorizando metas solidamente planeadas, alcançando objectivos parciais, onde cada nova fase do projecto será melhor sucedida

que a anterior. Outro ponto forte, o acesso interactivo e dinâmico aos dados da organização, com capacidade de os manipular, permitindo as mais variadas análises para suportar as tomadas de decisão.

No actual ambiente competitivo, apesar da crise financeira a nível mundial, a aposta em tecnologias de BI tenderá a aumentar, pois estão na base do desenvolvimento de vantagens competitivas, onde o conhecimento criará valor para o negócio das organizações.



## REFERÊNCIAS BIBLIOGRÁFICAS

- Barry, D. K. *Common Warehouse Metamodel (CWM)*. [Online] Disponível em: [http://www.service-architecture.com/web-services/articles/common\\_warehouse\\_meta-model\\_cwm.html](http://www.service-architecture.com/web-services/articles/common_warehouse_meta-model_cwm.html) [Acedido em: 27 Julho de 2010].
- Bensberg, F., 2003. “Controlling the Data Warehouse - a Balanced Scorecard Approach” *Journal of Computing and Information Technology*. September, p.233-241, 2003.
- Breslin, M., 2004. *Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models*. *Business Intelligence Journal*. 9(1), 2004.
- Cortes, B., (2005). *Sistemas de Suporte à Decisão*. Lisboa: FCA - Editora Informática, Lda.
- ESAL, 2007. *Projecto Educativo 2007/2010*. [Pdf] Disponível em: [http://moodle.esal.edu.pt/~orion/esalsite/index.php?option=com\\_docman&task=cat\\_view&gid=115&Itemid=616](http://moodle.esal.edu.pt/~orion/esalsite/index.php?option=com_docman&task=cat_view&gid=115&Itemid=616) [Acedido em: 28 Fevereiro 2010].
- ESAL, 2009. *Projecto Curricular de Escola 2009/2010*. [Pdf] Disponível em: [http://moodle.esal.edu.pt/~orion/esalsite/index.php?option=com\\_docman&task=cat\\_view&gid=115&Itemid=616](http://moodle.esal.edu.pt/~orion/esalsite/index.php?option=com_docman&task=cat_view&gid=115&Itemid=616) [Acedido em: 28 Fevereiro 2010].
- ESAL, 2009. *PCE Mundo de Cores 2009/2010 Caracterização*. Anexo 0 - Projecto Curricular de Escola 2009/2010. [Pdf] Disponível em: [http://moodle.esal.edu.pt/~orion/esalsite/index.php?option=com\\_docman&task=cat\\_view&gid=115&Itemid=616](http://moodle.esal.edu.pt/~orion/esalsite/index.php?option=com_docman&task=cat_view&gid=115&Itemid=616) [Acedido em: 28 Fevereiro 2010].
- Estevão, C., 1999. *Gestão Estratégica nas escolas*. Cadernos de Organização e Gestão Escolar. Instituto de Inovação Educacional. [Pdf] Disponível em: <http://area.dgicd.min-edu.pt/inovbasic/biblioteca/ccoge09/caderno9.pdf> [Acedido em: 08 Março de 2010].
- Greenfield, L., 2010. *The Data Warehousing Information Center*. [Online] Disponível em: <http://www.dwinfocenter.org/> [Acedido em: 29 Maio de 2020].
- IBM, 2010. *Cognos Business Intelligence*. [Online] Disponível em: <http://www-01.ibm.com/software/analytics/cognos/business-intelligence/> [Acedido em: 01 Junho de 2010].
- IGE, 2008. *Tópicos para a apresentação da escola*. Inspecção Geral da Educação. [Pdf] Disponível em: [http://www.ige.min-edu.pt/upload/AEE\\_2011/AEE\\_10\\_11\\_Topicos.pdf](http://www.ige.min-edu.pt/upload/AEE_2011/AEE_10_11_Topicos.pdf) [Acedido em: 28 Janeiro de 2011].

IGE, 2010. *Avaliação Externa das Escolas. Relatório da Escola Secundária com 3º CEB de Amato Lusitano em Castelo Branco*. Delegação Regional do Centro da IGE. [Pdf] Disponível em: [http://www.ige.min-edu.pt/upload/AEE\\_2010\\_DRC/AEE\\_10\\_ES3\\_Amato\\_Lusitano\\_R.pdf](http://www.ige.min-edu.pt/upload/AEE_2010_DRC/AEE_10_ES3_Amato_Lusitano_R.pdf) [Acedido em: 26 Junho de 2010]

Inmon, W., 1996. *Building the Data Warehouse*. Second Edition. New York, EUA: John Wiley & Sons, Inc.

Inmon, W., 2005. *Building the Data Warehouse*. Fourth Edition. New York, EUA: John Wiley & Sons, Inc.

Kaplan, R. & Norton, D., 1992. *The Balanced Scorecard - Measures that drive performance*. Boston: Harvard Business School Press.

Kaplan, R. & Norton, D., 2000. *Having Trouble with Your Strategy? Then Map It*. Boston: Harvard Business Review.

Kaplan, R. & Norton, D., 2001. *The strategy-focused organization*. Boston: Harvard Business School Press. *Strategy & Leadership*, 2001, Vol. 29 Issue 3, p41, 2p.

Kimball, R., Reeves, L., Ross, M. & Thornthwaite, W., 1998. *The Data Warehouse Lifecycle Toolkit*. New York, EUA: John Wiley & Sons, Inc.

Kimball, R. & Ross, M., 2002. *The Data Warehouse Toolkit*. Second Edition. New York, EUA: John Wiley & Sons, Inc.

Lei nº 31/2002, de 20 Dezembro de 2002. *Diário da República N° 294 - I Série A*. Ministério da Educação. Lisboa.

Machado, F., 2004. *Tecnologia e Projeto de Data Warehouse*. 4ª Edição. São Paulo: Érica.

Magalhães, A., 2009. *SQL Server 2008. Curso Completo*. Lisboa: FCA - Editora Informática, Lda.

Microsoft Corporation, 2010. *Microsoft SQL Server 2008 Business Intelligence*. [Online] Disponível em: <http://www.microsoft.com/sqlserver/2008/en/us/business-intelligence.aspx> [Acedido em: 03 Julho de 2010].

Microsoft Corporation, 2010. *Integration Services*. [Online] Disponível em: <http://msdn.microsoft.com/pt-br/library/ms130214.aspx> [Acedido em: 20 Julho de 2010].

Microsoft Corporation, 2010. *Introdução ao Excel Services e Excel Web Access*. [Online] Disponível em: <http://office.microsoft.com/pt-pt/sharepoint-server-help/introducao-ao-excel-services-e-excel-web-access-HA010105476.aspx> [Acedido em: 06 Novembro de 2010].

Niedrite, L., Treimanis, M., Solodovnikova, D. & Grundmane, L., 2008. *Development of Data Warehouse Conceptual Models: Method Engineering Approach*. In *Data Warehousing and Mining Book Series*, IGI Global publisher.

Oracle, 2010. *Oracle Business Intelligence*. [Online] Disponível em: <http://www.oracle.com/us/solutions/ent-performance-bi/business-intelligence/index.html> [Acedido em: 01 Junho de 2010].

Pentaho, 2010. *Pentaho Open Source Business Intelligence*. [Online] Disponível em: <http://www.pentaho.com/products/> [Acedido em: 01 Junho de 2010].

Rocha, D., 2000. *Desenvolvimento do Balanced Scorecard para Instituição de Ensino Superior Privada - Estudo de Caso da Unidade de Negócios 4 da Universidade Gama*. Dissertação de Mestrado em Engenharia. Florianópolis: Universidade Federal de Santa Catarina.

Santos, A., 2008. *Gestão Estratégica - Conceitos, modelos e instrumentos*. Lisboa. Escolar Editora.

SAP Portugal, 2010. *SAP Business Intelligence*. [Online] Disponível em: <http://www.sap.com/portugal/platform/netweaver/businessintelligence/index.epx> [Acedido em: 01 Junho de 2010]

Silvers, F., 2008. *Building and Maintaining a Data Warehouse*. New York, EUA: Auerbach Publications. Taylor & Francis Group.

Singh, H., 1997. *Data Warehousing: Concepts, Technologies, Implementations, and Management* - Upper Saddle River. NJ: Prentice Hall.

Thornthwaite, W., Mundy, J. & Kimball, R., 2006. *The Microsoft Data Warehouse Toolkit: With SQL Server 2005 and the Microsoft Business Intelligence Toolset*. John Wiley & Sons. Ch.13. [Pdf] Disponível em: <http://www.kimballgroup.com/html/booksMDWTtools.html> [Acedido em 28 Julho de 2010].

Turban, E., Aronson, J. & Liang, T., 2005. *Decision Support Systems and Intelligent Systems*. Seventh Edition. Prentice Hall.

Turban, E., Sharda, R., Aronson, J. & King D. (2008). *Business Intelligence: a managerial approach*. Prentice Hall.