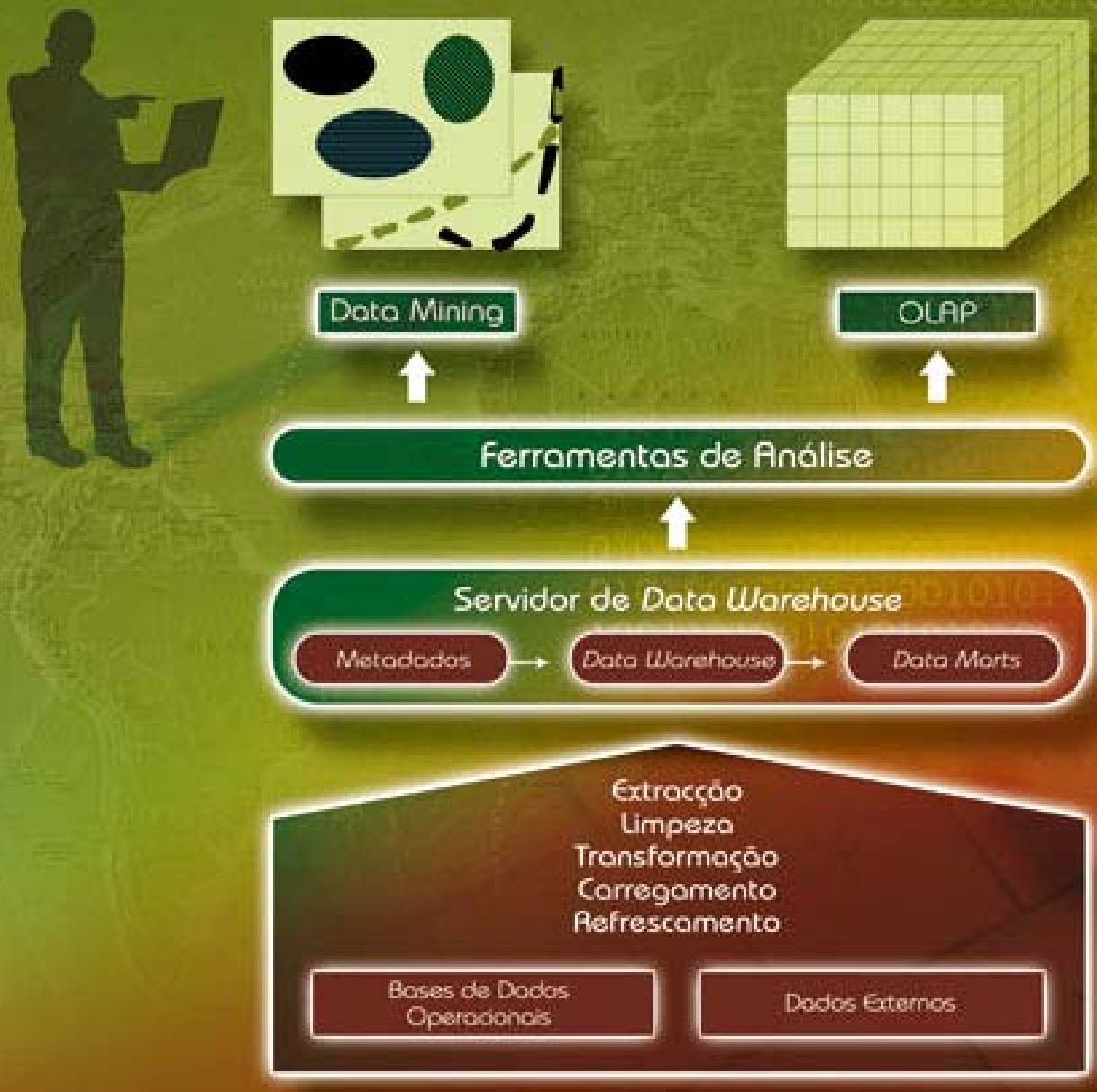
SISTEMAS DE INFORMAÇÃO

Business Intelligence

Tecnologias da Informação na Gestão de Conhecimento



Maribel Yasmina Santos • Isabel Ramos

FCA

ÍNDICE

PREFÁCIO ..... **I**

AGRADECIMENTOS ..... **III**

ÍNDICE ..... **V**

LISTA DE FIGURAS ..... **IX**

LISTA DE TABELAS ..... **XIII**

LISTA DE QUADROS ..... **XV**

1. INTRODUÇÃO **1**

1.1 Motivação 1
1.2 Estrutura do Livro 3

2. DA INFORMAÇÃO AO CONHECIMENTO **7**

2.1 Introdução 7
2.2 Várias Perspectivas sobre Informação e Conhecimento 8
2.3 Tipos de Conhecimento 12
2.4 O Processo de Construção de Conhecimento Organizacional 14
2.4.1 O Contexto Histórico e Sociocultural do Conhecimento 17

3. CONHECIMENTO ORGANIZACIONAL **21**

3.1 Introdução 21
3.2 Realidades Socialmente Construídas 22
3.3 Identidade Organizacional 24
3.3.1 Quatro Perspectivas sobre as Organizações: Estrutural, Social, Política, e Simbólica 25
3.3.1.1 Perspectiva Estrutural 27
3.3.1.2 Perspectiva Social 29
3.3.1.3 Perspectiva Política 30
3.3.1.4 Perspectiva Simbólica 32
3.4 Memória Organizacional 34
3.5 Aprendizagem Organizacional 36

© - FCA – EDITORA DE INFORMÁTICA

3.6 Criatividade e Inovação: Aprendizagem e Adaptação 39
3.6.1 Riscos Associados à Mudança Organizacional 40
3.6.2 Mudança e Cultura Organizacional 41
3.6.3 Inovação e as TI 43
3.6.4 As Dimensões Simbólica e Política da Mudança Organizacional 46

4. GESTÃO DE CONHECIMENTO ORGANIZACIONAL 49
4.1 Introdução 49
4.2 O Processo de Gestão de Conhecimento Organizacional 49
4.2.1 A Abordagem Sociológica ou Comportamental à Gestão de Conhecimento 53
4.2.2 A Abordagem Económica à Gestão de Conhecimento 53
4.2.3 A Abordagem Tecnológica à Gestão de Conhecimento 55
4.3 Os Sistemas Informáticos de Apoio à Gestão de Conhecimento 55
4.4 O Contributo da Gestão de Conhecimento para a Eficiência Organizacional 57

5. *BUSINESS INTELLIGENCE* 61
5.1 Introdução 61
5.2 O Conceito 61
5.3 *Data Warehousing* 62
5.3.1 Características de um *Data Warehouse* 64
5.3.2 Concepção de um *Data Warehouse* 66
5.3.2.1 Esquema em Estrela 66
5.3.2.2 Esquema em Floco de Neve 68
5.3.2.3 Esquema em Constelação 70
5.3.3 Carregamento e refrescamento de um *Data Warehouse* 71
5.4 Exploração de um *Data Warehouse* 72
5.5 Infra-estrutura Tecnológica de Apoio ao *Business Intelligence* 76

6. DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS 81
6.1 Introdução 81
6.2 Princípios 81
6.3 Fases do Processo 83
6.3.1 Selecção dos Dados 83
6.3.2 Tratamento dos Dados 83
6.3.3 Pré-processamento dos Dados 83
6.3.4 *Data Mining* 84
6.3.5 Interpretação de Resultados 85
6.4 A Metodologia CRISP-DM 86
6.4.1 Compreensão do Negócio 88
6.4.1.1 Identificar Objectivos do Negócio 89
6.4.1.2 Descrever o Contexto 90

6.4.1.3 Identificar os Objectivos do *Data Mining* 91
6.4.1.4 Produzir Plano do Projecto 91
6.4.2 Compreensão dos Dados 92
6.4.2.1 Recolher os Dados 92
6.4.2.2 Descrever os Dados 93
6.4.2.3 Explorar os Dados 93
6.4.2.4 Verificar a Qualidade dos Dados 93
6.4.3 Preparação dos Dados 94
6.4.3.1 Seleccionar Dados 95
6.4.3.2 Limpar Dados 95
6.4.3.3 Construir Dados 95
6.4.3.4 Integrar Dados 96
6.4.3.5 Formatar Dados 96
6.4.4 Modelação 96
6.4.4.1 Seleccionar Técnicas de Modelação 97
6.4.4.2 Definir Mecanismos de Teste 98
6.4.4.3 Construir Modelos 98
6.4.4.4 Analisar Modelos 98
6.4.5 Avaliação 99
6.4.5.1 Avaliar Resultados 99
6.4.5.2 Rever Processo 100
6.4.5.3 Identificar os Próximos Passos 100
6.4.6 Desenvolvimento 101
6.4.6.1 Planear o Desenvolvimento 101
6.4.6.2 Planear a Monitorização e Manutenção 101
6.4.6.3 Produzir Relatório Final 102
6.4.6.4 Rever o Projecto 102

7. *DATA MINING* 103

7.1 Introdução 103
7.2 Tarefas de *Data Mining* 103
7.2.1 Classificação 104
7.2.2 Segmentação 106
7.2.3 Sumariação 107
7.2.4 Modelação de Dependências 108
7.3 Técnicas de *Data Mining* 108
7.3.1 Árvores de Decisão 108
7.3.2 Regras de Associação 111
7.3.3 Regressão Linear 112
7.3.4 Redes Neuronais Artificiais 114
7.3.5 Algoritmos Genéticos 118
7.3.6 Vizinhos mais próximos 119

© - FCA – EDITORA DE INFORMÁTICA

8. ANÁLISE DE UMA BASE DE DADOS ORGANIZACIONAL COM *OLAP* 123

8.1 Introdução 123
8.2 O Domínio de Aplicação 123
8.3 Arquitectura do *Data Mart* 125
8.4 Análise do *Data Mart* com a Ferramenta OLAP 126
8.4.1 Cliente e Tipo de Bem Financiado 127
8.4.2 Cliente, Tipo de Bem Financiado e Tipo de Situação 130
8.4.3 Tipo de Bem Financiado, Tipo de Pagamento e Tipo de Situação 132
8.4.4 Tipo de Bem Financiado, Tipo de Contrato e Tipo de Situação 133
8.4.5 Tipo de Bem Financiado, Tipo de Duração e Tipo de Situação 135

9. ANÁLISE DE UMA BASE DE DADOS ORGANIZACIONAL COM *DATA MINING* 139

9.1 Introdução 139
9.2 Compreensão do Negócio e dos Dados 139
9.2.1 O Negócio 140
9.2.2 Os Dados 141
9.3 A Descoberta de Conhecimento 146
9.3.1 Selecção dos Dados 146
9.3.2 Tratamento dos Dados 148
9.3.3 Pré-processamento dos Dados 148
9.3.4 *Data Mining* 154
9.3.5 Interpretação de Resultados 163

10. A CONSTRUÇÃO SOCIAL DO CONHECIMENTO: ANÁLISE DA BASE DE DADOS ORGANIZACIONAL 167

10.1 Introdução 167
10.2 *Data Mining* e Construção de Conhecimento Organizacional 168
10.2.1 *Data Mining* e Construção Individual de Conhecimento 169
10.2.2 *Data Mining* e Construção Social de Conhecimento 172
10.3 Orientações Práticas para Facilitar a Partilha do Conhecimento Criado 173
10.4 Casos de Utilização das Ferramentas de *Data Mining* 175
10.4.1 Sector de Retalho 175
10.4.2 Sector Financeiro 177
10.4.3 Sector Automóvel 180
10.4.4 Sector das Telecomunicações 182

11. CONCLUSÕES 183

REFERÊNCIAS BIBLIOGRÁFICAS 185

ÍNDICE REMISSIVO 193

1

INTRODUÇÃO

1.1 MOTIVAÇÃO

As Tecnologias de Informação (TI) têm tido, desde há já vários anos, um papel extremamente importante nas organizações. Se a sua adopção se iniciou com o propósito de automatizar os processos operacionais da organização, aqueles que suportam as actividades do dia-a-dia, a verdade é que o papel das TI nas organizações tem evoluído consideravelmente.

Aplicações de suporte à gestão dos recursos humanos, gestão financeira e contabilística, gestão de stocks, gestão da produção, gestão de projectos, gestão logística, gestão de clientes e fornecedores, gestão da qualidade, entre outras, encontram-se perfeitamente enraizadas nas empresas. Além do suporte aos processos operacionais, estas aplicações geram uma quantidade avultada de informação, que deve ser utilizada pela organização para aumentar o conhecimento:

1. Sobre si própria - Quais os produtos e serviços que disponibilizamos e como essa disponibilização é feita? Quais os problemas que temos vindo a enfrentar e como os solucionamos? Quais as competências que temos? De que forma podemos usar essas competências para atingir as metas que definimos? Qual a probabilidade de sucesso da alteração do modelo de negócio actualmente implementado?
2. Sobre as entidades externas cuja colaboração é essencial para o sucesso do negócio - Quais são os nossos melhores clientes? Quais foram os produtos mais vendidos no mês passado? Quais os nossos fornecedores? Quais os fornecedores que, por norma, cumprem os prazos de entrega? Que organizações competem connosco? Que alterações políticas ou ambientais podem afectar o nosso negócio?
3. Sobre a sua capacidade para influenciar o ambiente externo em benefício próprio - Qual será o impacto nas vendas da diminuição do preço do artigo xpto? Que segmento de clientes deve ser visado pela campanha de

© - FCA – EDITORA DE INFORMÁTICA

marketing ‘20% de desconto em grandes marcas’? Qual a probabilidade de sucesso de um novo produto ou serviço? Como podemos vencer a concorrência?

Ao permitirem obter uma resposta para as questões acima indicadas e muitas outras que não é possível aqui detalhar, as aplicações informáticas usadas pela organização são elementos essenciais da construção e partilha de conhecimento sobre si própria, o negócio que realiza, o seu exterior e a sua possibilidade de interacção com as entidades externas relevantes para o negócio.

Todo este conhecimento é fundamental para apoiar o processo de tomada de decisão. A informação gerada pelas aplicações informáticas disponibiliza aos gestores um conjunto de indicadores sobre o negócio, que lhe dão indicações do que aconteceu no passado e lhe permitem traçar cenários para o futuro.

Neste livro é abordado o conceito de *Business Intelligence* e as TI a ele associadas. Os sistemas de *Business Intelligence* utilizam os dados disponíveis nas organizações para disponibilizar informação relevante para a tomada de decisão. Combinam um conjunto de ferramentas de interrogação e exploração dos dados com ferramentas que permitem a geração de relatórios, para produzir informação que será posteriormente utilizada pela gestão de topo das organizações, no suporte à tomada de decisão.

Os sistemas de *Business Intelligence* estão tradicionalmente associados a três tecnologias: *Data Warehouses*, *On-Line Analytical Processing* e *Data Mining*. Um *Data Warehouse* é um armazém de dados, um repositório integrado que permite o armazenamento de informação relevante para a tomada de decisão. Estes repositórios podem ser analisados utilizando ferramentas *On-Line Analytical Processing* e/ou ferramentas de *Data Mining*. No primeiro caso, estamos perante a análise multidimensional, que nos permite examinar a informação sob diferentes perspectivas. No segundo caso, algoritmos de exploração dos dados identificam padrões, relacionamentos, modelos, etc., que estão ocultos na grande quantidade de dados armazenados. Estes modelos podem, por exemplo, ser utilizados pela organização em tarefas de previsão.

A informação que estas tecnologias colocam ao dispor das organizações irá tornar-se um componente essencial do entendimento que os agentes organizacionais partilham sobre o negócio e a possibilidade que têm de agir para criar condições internas e externas favoráveis ao sucesso da organização. Desta forma, os sistemas de *Business Intelligence* contribuem para aumentar: (1) a inteligência colectiva da organização na medida em que facilitam a construção do conhecimento necessário para planear e implementar soluções para problemas e desafios percebidos como

podendo ameaçar a sobrevivência e bem-estar da organização; (2) a capacidade de aprendizagem da organização na medida em que contribuem para alterar e sofisticar a forma como colectivamente os agentes organizacionais captam, entendem e se comportam com o intuito de se acomodarem a uma experiência de trabalho em contínua remodelação, e (3) a criatividade organizacional ao apoiar a produção de novas ideias, e produtos ou serviços que permitam à organização adaptar-se de forma dinâmica aos desafios e oportunidades internos e externos.

Queremos, no entanto, realçar que estes sistemas devem ser vistos apenas como ferramentas de apoio que dentro de circunstâncias organizacionais e de gestão adequadas favorecem a inteligência, aprendizagem e criatividade organizacional. Sem políticas adequadas de gestão de recursos humanos, gestão estratégica e operacional, sem a existência de ambientes de trabalho que favoreçam a colaboração, comunicação, a aceitação do risco e a tolerância ao erro, estes sistemas podem nunca cumprir os objectivos para que foram adoptados e tornarem-se encargos financeiros demasiado pesados para a maioria das organizações.

A motivação deste livro está assim associada à junção dos conceitos de gestão de conhecimento e *Business Intelligence*, evidenciando como é que o conhecimento existente na organização pode ser conjugado com a utilização destes sistemas de forma a ser convenientemente utilizado no processo de decisão.

Este livro começa por apresentar os conceitos associados ao conhecimento, conhecimento organizacional e gestão de conhecimento, que são posteriormente complementados com os princípios associados aos sistemas de *Business Intelligence*, e às tecnologias associadas, nomeadamente *Data Warehousing*, *On-Line Analytical Processing* e *Data Mining*. Para evidenciar o processo de consolidação do conhecimento obtido pelos sistemas de *Business Intelligence*, é apresentada a análise de uma base de dados organizacional, com o objectivo de por um lado mostrar como é que estas aplicações funcionam (*On-Line Analytical Processing* e *Data Mining*) e que tipo de informação disponibilizam, e por outro, como é que este conhecimento pode ser directamente utilizado pela organização.

1.2 ESTRUTURA DO LIVRO

A estrutura deste livro reflecte, por um lado, o conjunto de conceitos com que o leitor deve estar familiarizado em cada instante, fornecendo, capítulo a capítulo, esses conceitos, e por outro lado, a sequência com que idealmente as diversas TI abordadas neste livro devem ser introduzidas na organização.

© - FCA – EDITORA DE INFORMÁTICA

Além deste capítulo introdutório, que apresenta a motivação e a estrutura do livro, esta obra integra outros 10 capítulos.

No capítulo 2, são apresentadas e discutidas as definições de informação e conhecimento mais divulgadas na literatura académica e de negócios, ligando-as a formas específicas de gerir informação e conhecimento. Neste capítulo, é ainda indicada a definição de conhecimento que orienta as discussões teóricas e aplicações de tecnologia contidas neste livro. Finalmente é descrito o processo de construção de conhecimento que serve de contexto à análise da tecnologia normalmente associada à inteligência de negócio.

No capítulo 3, o conhecimento organizacional é apresentado como uma construção social em contínua reformulação e patente em todos os momentos da vida da organização. Neste capítulo, são abordados os conceitos de identidade, memória e aprendizagem organizacional.

No capítulo 4, é apresentado o processo de gestão de conhecimento organizacional e o seu contributo para a eficiência das organizações.

No capítulo 5, é apresentado o conceito de *Business Intelligence*, caracterizando-o globalmente, e descrevendo duas das tecnologias que lhe estão associadas: *Data Warehousing* e *On-line Analytical Processing*. Em relação aos sistemas de *Data Warehousing*, apresentam-se as características destes repositórios, descreve-se a modelação multidimensional de dados na concepção da arquitectura destes sistemas, assim como são contextualizadas as funções das ferramentas de extracção, transformação e carregamento de dados. As ferramentas *On-Line Analytical Processing* são também caracterizadas, apresentando as diversas funções de manipulação dos cubos. Este capítulo termina com a apresentação de uma arquitectura de uma infraestrutura tecnológica de apoio ao *Business Intelligence*, que define como é que estas diferentes tecnologias podem ser integradas na organização.

O capítulo 6 é dedicado ao processo de descoberta de conhecimento em bases de dados. Para este processo são descritas as diversas fases que o integram, assim como é apresentada a metodologia *Cross Industry Standard Process for Data Mining*, de suporte a projectos de *Data Mining*. Esta metodologia descreve um modelo de referência que define as fases a seguir, as tarefas a executar em cada uma das fases, e os resultados esperados pela execução das mesmas.

No capítulo 7, são apresentadas as diversas tarefas associadas ao *Data Mining*, assim como são descritas as principais técnicas utilizadas na sua concretização. Para as diversas técnicas são apresentadas explicações genéricas sobre o modo de funcionamento das mesmas. Estas explicações, não muito

detalhadas, visam permitir a compreensão das técnicas por leitores com diferentes formações académicas.

No capítulo 8, é analisada uma base de dados organizacional com uma ferramenta *On-Line Analytical Processing*. Neste capítulo, e antes da análise propriamente dita, são descritos os dados disponíveis para análise e definida a estrutura de um pequeno armazém de dados para onde os dados serão carregados. Os dados são depois analisados recorrendo a diversos cubos com vistas multidimensionais, que permitem aumentar o conhecimento da organização acerca dos seus clientes.

No capítulo 9, são utilizadas técnicas de *Data Mining* para analisar o conjunto de dados introduzido e explorado no capítulo 8. Na concretização desta tarefa são executadas as diversas fases do processo de descoberta de conhecimento em bases de dados, desde a selecção dos dados até à interpretação de resultados, as quais permitiram a identificação de modelos que descrevem o perfil dos clientes que integram o conjunto de dados analisado.

No capítulo 10, é analisado o caso descrito no capítulo 9, à luz dos conceitos teóricos apresentados nos capítulos iniciais do livro. Neste capítulo, são ainda apresentadas algumas orientações práticas para a partilha de conhecimento construído com o apoio de ferramentas de *Data Mining*. Finalmente, é apresentada evidência prática da relevância daquelas orientações.

O capítulo 11 culmina este livro com uma síntese dos aspectos mais relevantes na adopção dos sistemas de *Business Intelligence*, e na consolidação do conhecimento gerado a partir dos mesmos.

© - FCA – EDITORA DE INFORMÁTICA