

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE INFORMÁTICA
CURSO TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

RAFAEL LUCIANO GONÇALVES CORREIA

TÉCNICAS DE AQUISIÇÃO DE CONHECIMENTO NA COLETA DE
REQUISITOS NA MODELAGEM DE UM BANCO DATA
WAREHOUSE

TRABALHO DE CONCLUSÃO DE CURSO

PONTA GROSSA

2017

RAFAEL LUCIANO GONÇALVES CORREIA

**TÉCNICAS DE AQUISIÇÃO DE CONHECIMENTO NA COLETA DE
REQUISITOS NA MODELAGEM DE UM BANCO DATA
WAREHOUSE**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas, do Departamento Acadêmico de Informática, da Universidade Tecnológica Federal do Paraná.

Orientador: Profa. Dr^a Simone de Almeida

PONTA GROSSA

2017



TERMO DE APROVAÇÃO

TÉCNICAS DE AQUISIÇÃO DE CONHECIMENTO NA COLETA DE REQUISITOS NA MODELAGEM DE UM BANCO DATA WAREHOUSE

por

RAFAEL LUCIANO GOLÇALVES CORREIA

Este Trabalho de Conclusão de Curso (TCC) foi apresentado em 20 de novembro de 2017 como requisito parcial para a obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Simone de Almeida
Prof.(a) Orientador(a)

Geraldo Ranthum
Membro titular

Rogério Ranthum
Membro titular

Prof^a. Helyane Bronoski Borges
Responsável pelo Trabalho de
Conclusão de Curso

Prof^a. Mauren Louise Sguario
Coordenadora do curso

- O Termo de Aprovação assinado encontra-se na Coordenação do Curso -

AGRADECIMENTOS

Agradeço a Deus por tudo.

Agradeço a Profa. Dra. Simone de Almeida, uma amiga que tanto me auxiliou.

Agradeço a Daniela, pelo carinho e apoio.

Agradeço a minha família, em especial aos que deixaram saudade.

A nada dar mais valor que a Cristo.
(São Cipriano de Cartago)

RESUMO

CORREIA, Rafael Luciano G. **Técnicas de Aquisição de Conhecimento na Coleta de Requisitos na Modelagem de Um Banco Data Warehouse**. 2017. 38f. Trabalho de Conclusão de Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas - Universidade Tecnológica Federal do Paraná. Ponta Grossa, 2017.

A crescente necessidade de geração de conhecimento em cenários compostos por significativo número de variáveis oriundas de grandes volumes de dados com o objetivo de apoiar o processo de tomada de decisão dentro das organizações, tem como solução o *Data Warehouse*. Entretanto, cada organização apresenta características próprias fazendo com que o desenvolvimento de um *Data Warehouse* seja específico e completamente influenciado pelos requisitos de negócios dos usuários. Desta forma, este trabalho descreve os elementos que compõe um *Data Warehouse*, sua arquitetura, modelagem, seus processos. O trabalho também apresenta as técnicas de aquisição de conhecimento, as abordagens para levantamento de requisitos, pesquisas recentes na área e as metodologias para desenvolvimento de um *Data Warehouse* com enfoque nos requisitos.

Palavras-chave: Apoio a Decisão. Aquisição de Conhecimento. Coleta de Requisitos. *Data Warehouse*.

ABSTRACT

CORREIA, Rafael Luciano G. **Knowledge Acquisition Techniques in Requirements Collection in the Modeling of a Data Warehouse Database.** 2017. 38p. Conclusion Work of Higher Course of Technology in Analysis and Development of Systems - Federal Technological University of Paraná. Ponta Grossa, 2017.

The growing need to generate knowledge in scenarios composed by significant number of variables from large volumes of data to support the decision-making process within organizations, has as solution the Data Warehouse. However, each organization has its own characteristics making the development of a Data Warehouse specific and completely influenced by the business requirements of users. In this way, this work describes the elements that make up a Data Warehouse, its architecture, modeling, it's processes. The paper also presents the knowledge acquisition techniques, the approaches to requirements gathering and the methodologies for developing a Data Warehouse with a focus on requirements.

Keywords: Support the Decision. Acquisition of Knowledge. Requirements Collection. Data Warehouse.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de MER.....	19
Figura 2 – Exemplo do esquema Estrela	20
Figura 3 – Exemplo do esquema Floco de Neve	21
Figura 4 – <i>Framework</i> do método <i>Iterations</i>	30

LISTA DE QUADROS

Quadro 1 – Diferenças entre Gestão da Informação e Gestão do Conhecimento	14
Quadro 2 – Comparação do nível de atendimento a práticas de engenharia de requisitos pelas metodologias de desenvolvimento de DW	31

LISTA DE SIGLAS

BDL	<i>Business Development Lifecycle</i>
BI	<i>Business Intelligent</i>
CDWD	<i>Conceptual Data Warehouse Design</i>
DDL	<i>Data Definition Language</i>
DM	<i>Data Mart</i>
DW	<i>Data Warehouse</i>
ETL	<i>Extract, Transform, Load</i>
FWD	<i>Foreign Data Wrappers</i>
GC	Gestão do Conhecimento
GI	Gestão da Informação
MER	Modelagem Entidade-Relacionamento
SAD	Sistema de Apoio a Decisão
SR	Sistema de Recomendação
SSD	<i>Solid State Disk</i>

SUMÁRIO

1 INTRODUÇÃO	10
1.1 OBJETIVOS.....	10
1.1.1 Objetivo Geral.....	11
1.1.2 Objetivos Específicos.....	11
1.2 PROBLEMA.....	11
1.3 ORGANIZAÇÃO DO TRABALHO.....	11
2 DATA WAREHOUSE	13
2.1 DADOS, INFORMAÇÃO E CONHECIMENTO.....	13
2.2 DIFERENÇAS ENTRE GESTÃO DO CONHECIMENTO E da INFORMAÇÃO.....	13
2.3 TÉCNICAS PARA AQUISIÇÃO DE CONHECIMENTO.....	14
2.4 TOMADA DE DECISÃO NO SISTEMA DE INFORMAÇÃO.....	16
2.5 DEFINIÇÃO DE <i>DATA WAREHOUSE</i>	16
2.6 CARACTERÍSTICAS DO <i>DATA WAREHOUSE</i>	17
2.6.1 Orientado a Assunto.....	17
2.6.2 Variação de Tempo.....	17
2.6.3 Não-Volátil.....	18
2.6.4 Integração.....	18
2.7 MODELAGEM DOS DADOS.....	18
2.7.1 Modelagem Entidade-Relacionamento (MER).....	19
2.7.2 Modelagem Multidimensional.....	20
2.8 METADADOS.....	21
2.9 DATA MART.....	21
2.10 GRANULARIDADE.....	22
2.11 ARQUITETURAS DE UM DATA WAREHOUSE.....	22
2.12 PROCESSO DE DATA WAREHOUSING.....	23
3 REQUISITOS PARA DATA WAREHOUSE	25
3.1 ABORDAGENS DE LEVANTAMENTO DE REQUISITOS PARA DW.....	25
3.2 PESQUISAS RECENTES EM DW.....	26
4 PRINCIPAIS METODOLOGIAS PARA DESENVOLVIMENTO DE SISTEMAS DATA WAREHOUSE COM ENFOQUE NOS REQUISITOS	28
4.1 O MÉTODO HADDEN-KELLY.....	28
4.2 BUSINESS DEVELOPMENT LIFECYCLE (BDL).....	29
4.3 GOLFARELLI E RIZZI.....	29
4.4 CONCEPTUAL DATA WAREHOUSE DESIGN (CDWD).....	30
4.5 ITERATIONS.....	30
4.6 QUADRO COMPARATIVO.....	31
5 CONSIDERAÇÕES FINAIS	32
5.1 CONCLUSÃO.....	32
5.2 TRABALHOS FUTUROS.....	33

1 INTRODUÇÃO

Com a evolução da tecnologia, a capacidade competitiva das organizações se dá, em parte, pela habilidade de seus gestores em tomar decisões de forma rápida e precisa. A correta interpretação de dados nos mais variados cenários se apresenta como uma grande vantagem neste contexto (PAIM, 2003).

Calazans e Raslan (2014, p. 29) afirmam que o gerenciamento dos dados se tornou uma atividade fundamental porque o conhecimento que uma empresa tem sobre si permite que ela aperfeiçoe ou crie novos processos e estratégias, promovendo o aumento de sua lucratividade e possibilitando o seu ingresso em novos mercados.

Para auxiliar nessa necessidade, surgiu o conceito de *Data Warehouse* (Armazém de Dados), o qual não se restringe à simples extração de dados com finalidades diversas, e sim ao negócio como um todo (WAGNER, 2003).

Para implantar um *Data Warehouse* (DW) deve-se primeiro conhecer os requisitos dos decisores. Estes requisitos possuem natureza diferente daqueles especificados para os sistemas de informações convencionais, que executam atividades operacionais da empresa, estão associados aos processos de tomada de decisão, o que é uma função mais complexa do que a execução de procedimentos diários e operacionais, e dependem do conhecimento e experiência da pessoa que o executa (SIDREIRA, 2009). Sendo assim, deve-se utilizar técnicas mais adequadas para a coleta dos requisitos para um desenvolvimento de um DW.

A coleta de requisitos para um DW não tem sido muito valorizada pelos autores da área. Historicamente, a construção de DW foi identificada como a etapa mais desafiadora na sua implantação (KIMBALL, 1996). O processo de localizar a origem dos dados, integrá-los e organizá-los de maneira a tornar eficiente o acesso a um grande volume dados, despertou o interesse em pesquisa-lo.

Desta forma, este trabalho descreve os elementos que compõe um *Data Warehouse*, sua arquitetura, modelagem, seus processos. O trabalho também apresenta as técnicas de aquisição de conhecimento, as abordagens para levantamento de requisitos, pesquisas recentes na área e as metodologias para desenvolvimento de um *Data Warehouse* com enfoque nos requisitos

1.1 OBJETIVOS

Os objetivos deste trabalho estão descritos a seguir.

1.1.1 Objetivo Geral

Levantamento de técnicas de aquisição do conhecimento para a coleta de requisitos e modelagem de um banco *Data Warehouse*.

1.1.2 Objetivos Específicos

- Apresentar as principais arquiteturas e modelos de um banco de dados *Data Warehouse*;
- Pesquisar as formas atuais de levantamento dos requisitos para um *Data Warehouse*;
- Levantar as técnicas de aquisição de conhecimento;
- Definir a ferramenta para o projeto de *Data Warehouse*;

1.2 PROBLEMA

Atualmente a coleta de requisitos para DW vem sendo realizada aplicando técnicas como entrevistas e sessões de facilitação. Estas técnicas são muito dependentes de fatores humanos, como a habilidade de comunicação e precisam de uma equipe com talentos especiais para obterem sucesso. A coleta de requisitos em DW está mais para arte do que ciência (KIMBALL, 1996).

As atividades dos decisores são intensivas em conhecimento e envolve geralmente aspectos como incertezas e *feeling*. Em um contexto como este, as limitações das equipes de coleta de requisitos em fazer perguntas certas e as dificuldades do usuário em expressar todos os fatores envolvidos na tomada de decisão se somam e prejudicam o sucesso de um projeto dessa natureza (KIMBALL, 1996).

É preciso contornar estas dificuldades utilizando técnicas adequadas para entender como se dá o processo de tomada de decisão e que informações serão efetivamente aplicadas. Tais técnicas são oriundas da Engenharia do Conhecimento.

1.3 ORGANIZAÇÃO DO TRABALHO

O estudo sobre Técnicas de Aquisição do Conhecimento para a coleta de requisitos e modelagem de um banco *Data Warehouse* foi organizado em 5 capítulos. O Capítulo 2 aborda

e define, por meio da literatura especializada, o que é um *Data Warehouse*, apresentando suas principais características e conceitos de utilização e desenvolvimento.

O Capítulo 3 apresenta abordagens específicas para o levantamento de requisitos de um *Data Warehouse*.

O Capítulo 4 apresenta as principais metodologias para desenvolvimento de sistemas *Data Warehouse* com enfoque nos requisitos. O Capítulo 5 finaliza o trabalho apresentando a conclusão.

2 DATA WAREHOUSE

Este capítulo traz informações acerca dos conceitos fundamentais para a compreensão de um DW. A seção 2.1 apresenta o significado de dados, informações e conhecimento. A seção 2.2 aborda as diferenças entre a gestão da informação e a gestão do conhecimento. A seção 2.3 justifica a utilização de sistemas de apoio à tomada de decisão. A seção 2.4 apresenta algumas definições sobre DW. A seção 2.5 descreve as principais características de um DW. Na seção 2.6 define-se a modelagem de dados. Na seção 2.7 define-se metadados. Na seção 2.8 define-se *Data Mart*. Na seção 2.9 define-se o conceito de granularidade de dados. Na seção 2.10 descreve-se o processo de DW.

2.1 DADOS, INFORMAÇÃO E CONHECIMENTO

Segundo Silva et al (2015) dados apresentam significados específicos que quando isolados não possuem a capacidade de transmitir uma mensagem, mas que podem ser agrupados e ordenados com o propósito de constituir informações operacionais ou gerenciais.

Os dados são os registros de eventos em seu maior grau de detalhamento possível, como por exemplo os valores individuais relacionados às vendas de uma organização, informações são dados processados, como por exemplo a totalização do número de vendas de cada produto em um determinado período produz a informação sobre quais foram os produtos mais vendidos neste período (FAVARETTO, 2005).

Para Bonfadini (2014) o conhecimento é a capacidade de compreender o relacionamento entre as informações trazendo com isso o saber necessário para apoiar a tomada de decisão.

2.2 DIFERENÇAS ENTRE GESTÃO DO CONHECIMENTO E DA INFORMAÇÃO

De acordo com Barbosa (2008), enquanto a Gestão da Informação (GI) tem por insumo o registro e por finalidade o processamento da informação explícita, a Gestão do Conhecimento (GC) tem por foco o conhecimento pessoal que precisa ser descoberto e socializado. O Quadro 1 resume as principais diferenças entre GC e GI.

Quadro 1 – Diferenças entre Gestão da Informação e Gestão do Conhecimento

Critério	Gestão da Informação	Gestão do Conhecimento
1. Fenômenos centrais	Informação ou conhecimento Explícito	Conhecimento tácito, competências pessoais
2. Visibilidade dos fenômenos	Baixa	Muito baixa
3. Processos críticos	Organização e tratamento da Informação	Descoberta e compartilhamento do conhecimento
4. Nível de centralidade para gestão estratégica	Mediana	Alta
5. Influência da cultura organizacional sobre processos e resultados	Mediana	Alta
6. Possibilidade de gerenciamento	Baixa ou mediana	Baixa ou muito baixa
7. Outros conceitos relacionados	Sistemas de informação, gestão eletrônica de documentos	Capital intelectual, ativos intangíveis, aprendizagem organizacional

Fonte: Barbosa (2008)

No Quadro 1 é possível verificar a diferença central entre GI e GC e algumas características oriundas dessa diferença. Na GI ocorre a gerência da informação explícita e, portanto, apresenta como principal dificuldade a organização e o tratamento dos dados com os quais indicadores serão gerados para auxiliar no processo de tomada de decisão. A GC tem por objetivo extrair e relacionar os conhecimentos pessoais para a promoção de novas ideias e soluções, apresentando como principal dificuldade a descoberta e compartilhamento destes conhecimentos.

A gestão da informação objetiva dar suporte a execução de tarefas operacionais cotidianas e a tomada de decisão, enquanto a gestão do conhecimento tem por objetivo promover a concepção de ideias, a solução de problemas e a tomada de decisão (BENCK e GURA, 2011).

2.3 TÉCNICAS PARA AQUISIÇÃO DE CONHECIMENTO

A aquisição do conhecimento se dá mediante a compreensão e organização do conhecimento proveniente de fontes diversas, como por exemplo livros, diagramas, manuais, computadores e especialistas humanos, profissionais que possuem conhecimento profundo e específico sobre determinado assunto (SILVA, 2007).

De acordo com Ziulkoski (2003), a Engenharia do Conhecimento promove a construção de sistemas de conhecimento por meio de uma série de metodologias, técnicas e formalismos, dentre elas residem as Técnicas de Aquisição de Conhecimento das quais o autor descreve:

- **Análise de Protocolos:** O especialista deve falar seus pensamentos enquanto executa um processo, decompondo este processo em subprocessos.
- **Focalizando Contextos ou Cenários:** O especialista deve executar testes que demonstrem detalhes sobre a execução do processo.
- **Observação:** Observar a realização de uma tarefa pelo especialista, o que permite a visualização da execução do processo no ambiente em que ocorre, entretanto se faz necessário a realização de entrevistas posteriores para o entendimento das decisões tomadas.
- **Recuperação de Eventos:** O especialista deverá descrever as soluções utilizadas para a resolução das exceções que poderão ocorrer durante o processo.
- **Entrevistas:** Podem ocorrer de forma não estruturada, sem a definição de uma pauta nas etapas iniciais para estabelecer um bom relacionamento com o especialista, obter uma visão ampla a respeito dos processos a serem abordados e conhecer as prioridades desse processo. Também podem ocorrer de forma estruturada, dirigidas por uma pauta previamente planejada para a aquisição de conhecimentos específicos sobre o domínio. Segundo Silva (2007) e execução de entrevistas não estruturadas são recomendáveis apenas nas etapas iniciais, pois podem se estender a tópicos irrelevantes.
- **Classificação de Termos:** Processo iterativo que consiste em entregar ao especialista um conjunto de fichas contendo um conceito cada, pedir para ele agrupe-as de acordo com seus próprios critérios e posteriormente questionar o critério e os rótulos que ele daria a cada conjunto, promovendo com isso o reconhecimento dos relacionamentos entre estes termos, das diferentes formas de visualizar o domínio e a obtenção de termos não verificados anteriormente.

As tarefas executadas para aquisição de conhecimento podem apresentar complicações como, a falta de domínio de questões técnicas necessárias para o entendimento dos processos que criará dificuldades de compreensão do conteúdo de documentos e na comunicação com os especialistas, promovendo uma comunicação simplificada e menos precisa, a incapacidade do especialista expressar todas as informações necessárias, a existência de conhecimentos tácitos e difíceis de se explicitar, a possibilidade do especialista não estar comprometido com o projeto e omitir deliberadamente as informações (SILVA, 2007).

2.4 TOMADA DE DECISÃO NO SISTEMA DE INFORMAÇÃO

A amplitude dos problemas e o nível de complexidade das variáveis que compõe o processo decisório em uma organização, sobretudo em um mercado globalizado, faz com que este processo deva ser estruturado e resolvido de modo formal, detalhado, consistente e transparente, para assim manter a competitividade reduzindo os custos e aumentando os lucros mediante a rapidez, assertividade e a correta abrangência das ações executadas, tornando, para isso, indispensável a utilização de Sistemas de Apoio à Decisão (BONFADINI, 2014).

Sistemas de Apoio a Decisão (SAD) objetivam fornecer ferramentas que permitam comparar, analisar, simular e apoiar a escolha de alternativas pertencentes ao domínio de um processo decisório, convertendo volumosas bases de dados em conhecimento e auxiliando no processo cognitivo relacionado à análise desse conhecimento, minimizando dificuldades inerentes a capacidade de percepção humana para considerar e tratar a totalidade das variáveis disponíveis nesse cenário complexo (HEINZLE, 2011).

A implementação e utilização dessa tecnologia depende da disponibilidade de uma fonte de dados não transacional que possua todos os dados relevantes ao domínio da aplicação do SAD (HEINZLE, 2011).

2.5 DEFINIÇÃO DE *DATA WAREHOUSE*

Segundo Calazans et al (2014) o conceito de DW é a organização integrada dos dados corporativos de modo a gerar uma única fonte histórica de dados, possibilitando assim a identificação de tendências que irão indicar um melhor posicionamento estratégico, aumentando a competitividade e o lucro.

De acordo com Ramez et al (2011) pode-se definir o conceito de DW de forma mais ampla, como uma coleção de tecnologias de apoio a decisão que permite ao gestor tomar decisões melhores e mais rapidamente.

É necessário diferenciar DW de bancos de dados transacionais, pois ele não tem por finalidade o controle do negócio e sim a gestão estratégica. Bancos de dados transacionais podem executar operações que permitem a alteração dos dados, os DW devem manter um histórico de dados e para isso possuem dados sumarizados, agregados e consolidados, que não são alterados, ele é uma forma de analisar o grande volume de dados espalhados por diferentes sistemas de uma organização (CALAZANS; RASLAN, 2014).

Segundo Benck et al (2011) sistemas que necessitam processar transações consistentemente, gerenciar a concorrência de acesso e apresentam a possibilidade de recuperação e atualização constante dos dados são conhecidos por OLTP (*Online Transaction Processing* ou *Processamento de Transações Online*), enquanto sistemas que tem por finalidade manter um histórico de dados estáticos e completos são conhecidos por OLAP (*Online Analytical Processing* ou *Processamento Analítico Online*).

Ferramentas OLAP surgem para atuarem como mediadoras entre usuário e bases de dados de DW. Sistemas OLAP permitem a análise de grandes quantidades de dados consolidados, além do processamento de consultas para descobrir tendências e analisar fatores importantes para o negócio (FILHO, 2014). Estes dados podem ser acessados sob uma larga possibilidade de visões da informação, de forma rápida, consistente e interativa.

2.6 CARACTERÍSTICAS DO DATA WAREHOUSE

O DW integra e consolida os dados de diferentes fontes, permitindo a realização de análises e com isso a criação de oportunidades de negócios (FAVARETTO, 2005). Para tanto, o DW apresenta as características a seguir.

2.6.1 Orientado a Assunto

O DW é orientado por assuntos relacionados às atividades importantes realizadas na organização e a escolha correta desses assuntos determina o sucesso do projeto (BENCK; GURA, 2011). De acordo com Paim (2003) os dados são organizados conforme os usuários se referem a eles em seu dia a dia e essa forma de orientação promove flexibilidade para a análise gerencial dos dados, por estruturá-los segundo as áreas de atuação e objetivos estratégicos da empresa.

2.6.2 Variação de Tempo

O DW deve manter o histórico de todas as operações realizadas pela organização, por isso os seus dados não são modificados, todas as variações dos dados são inseridas, permitindo a realização de consultas por períodos de tempo (PAIM, 2003). Segundo Benck et al (2011) estes dados históricos são denominados *snapshots*, ou fotos instantâneas e eles podem sofrer

alterações se estiverem incorretos, mas assim que são considerados corretos não podem mais ser atualizados.

2.6.3 Não-Volátil

A manutenção de um histórico objetiva produzir indicadores que representem o desenvolvimento ao longo do tempo, por este motivo não se pode excluir ou alterar os dados no DW, os quais apresentam somente operações de carga e consulta a esses dados (CALAZANS; RASLAN, 2014).

2.6.4 Integração

Os dados advindos de diferentes bases de dados devem sofrer uma padronização ao serem movidos para o ambiente analítico, havendo assim uma convenção consistente de nomes, forma das variáveis, atributos, etc (CALAZANS; RASLAN, 2014). Para realizar esta integração, os dados são traduzidos para um formato em concordância com a análise estratégica e posteriormente são tratados para eliminar inconsistências, incompatibilidades e ausência de conteúdo essencial (PAIM, 2003).

2.7 MODELAGEM DOS DADOS

A modelagem é uma forma de representação do ambiente observado por meio de um modelo, permitindo a análise dos relacionamentos entre os objetos de estudo, formação de conceitos sobre esses objetos, abstração de suas características comuns e seu agrupamento em conjuntos, assim dando condições para a validação da organização dos dados e gerando maior estabilidade a essa organização (CALAZANS; RASLAN, 2014).

Segundo Benck et al (2011) a modelagem é a exposição da forma como as informações se relacionam através de diagramas e notações, com a qual pode-se eliminar redundâncias e inconsistências de dados e a não utilização dessa ferramenta implica em maiores custos e dificuldades de desenvolvimento devido ao crescimento desorganizado dos dados.

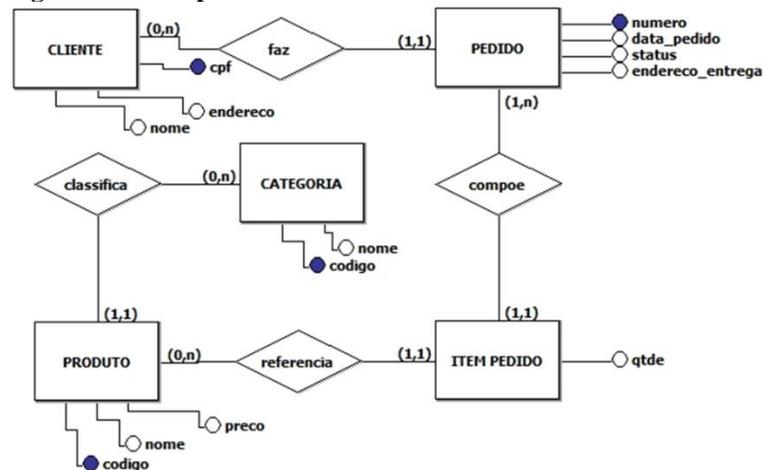
Para Ramez et al (2011) um modelo de dados é formado por uma coleção de conceitos que podem ser usados para descrever a estrutura de um banco de dados. Todas as operações que podem ser realizadas com os dados dependem da forma como eles estão armazenados, essa arquitetura pode ser relacional ou dimensional (CALAZANS; RASLAN, 2014). Segundo

Benck et al (2011) deve-se definir a modelagem dos dados para realizar a construção do DW, e esta pode ser a Modelagem Entidade-Relacionamento ou a Modelagem Multidimensional.

2.7.1 Modelagem Entidade-Relacionamento (MER)

Forma-se uma representação global do problema observado por meio da técnica diagramática Entidade-Relacionamento, identificando-se as entidades e relacionamentos ignorando ponderações sobre o armazenamento e eficiência e representando o problema através dessas entidades e das relações entre elas, as quais podem possuir atributos qualificadores dessa realidade. Esse modelo futuramente será traduzido para um modelo lógico (CHEN, 1990).

Figura 1 – Exemplo de MER



Fonte: Siqueira (entre 1999 e 2017)

O MER emprega três noções básicas (SILBERSCHATZ et al., 1999):

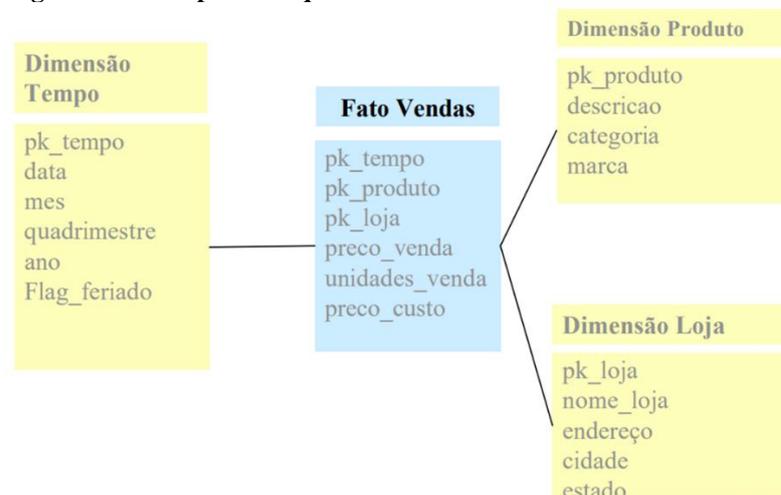
- Entidade: elemento identificado no mundo real e único em relação aos demais elementos.
- Atributo: propriedades de uma entidade. Podem ser simples ou compostos e podem ser monovalorados, recebendo um único valor, ou multivalorados, podendo receber um conjunto de valores.
- Relacionamento: associação entre entidades, que pode ser denominado unário ao associar uma entidade com ela própria, binário ao ocorrer entre duas entidades distintas ou ternário quando ocorre entre três entidades. A função que uma entidade exerce em um relacionamento é dá-se o nome de papel.

2.7.2 Modelagem Multidimensional

Segundo Ramez et al (2011) o modelo multidimensional armazena os dados em matrizes multidimensionais de acordo com seus relacionamentos inerentes e este modelo é constituído por dois tipos de tabela: tabelas de fatos e tabelas de dimensões.

- **Tabelas de Fatos:** Armazenam valores numéricos denominadas medidas ou métricas, constituídos por indicadores matemáticos como participação, evolução, índices, não possui identificador percentual e não indica inexistência do fato, é resultado de uma análise incompleta (BENCK; GURA, 2011).
- **Tabelas de Dimensões:** Informações que contextualizam os fatos, que podem ser caracterizadas por quando foi realizada, onde foi realizada, quem realizou e o que foi. As dimensões podem ser divididas em membros hierárquicos, por exemplo a dimensão localização pode se subdividir em: cidade, estado e país (BENCK; GURA, 2011). Segundo Calazans et al (2014) o modelo dimensional pode ser representado pelo esquema estrela ou pelo esquema floco de neve, os quais são coleções de elementos de bancos de dados (tabelas, visões e índices).
- **Estrela:** Seu diagrama de entidade-relação assemelha-se com uma estrela, é constituído em seu centro por uma tabela de fatos e suas pontas por tabelas de dimensão, as quais se relacionam à tabela de fatos por meio de sua chave primária e não existem relacionamentos entre as tabelas dimensões.

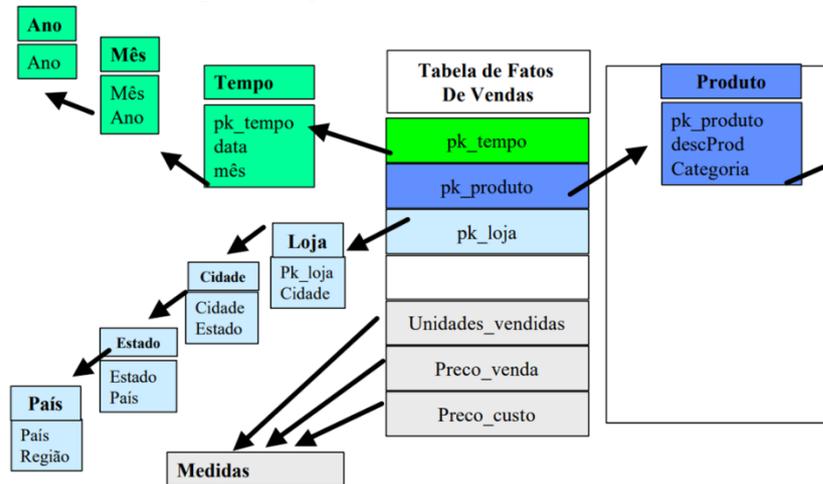
Figura 2 – Exemplo do esquema Estrela



Fonte: Tanaka (entre 1997 e 2017)

- **Floco de neve:** Nesse esquema os dados das dimensões são agrupados em várias tabelas com o objetivo de eliminar a redundância, assim aumentando o número de tabelas de dimensão.

Figura 3 – Exemplo do esquema Floco de Neve



Fonte: Tanaka (entre 1997 e 2017)

2.8 METADADOS

Segundo Sell (2001) metadados são os dados que descrevem a estrutura dos dados armazenados, que tem por finalidade o atendimento de necessidades do sistema de DW e que podem ser classificados como metadados técnicos quando propiciam a manutenção de um DW ou são classificados como metadados de negócios quando auxiliam a interpretação da informação.

2.9 DATA MART

De acordo com Batista et al (2014) os *Data Marts* (DM) são subconjuntos de um DW, no qual cada DM armazena dados referentes a assuntos específicos, como por exemplo os dados de departamentos individuais de uma organização, formando deste modo repositórios de dados de pequeno porte, em relação ao sistema completo, apresentando, portanto, maior agilidade na execução de consultas.

2.10 GRANULARIDADE

A granularidade é o nível de detalhe ou de resumo em que os dados são armazenados no DW, afetando assim o volume de dados armazenados e a qualidade das consultas que poderão ser executadas, desta forma se faz necessário defini-la apropriadamente para que o sistema apresente o correto desempenho (CAMARGO, 2004).

Quanto menor for o nível de granularidade, maior quantidade de dados serão armazenados e conseqüentemente um número maior de informações poderá ser extraído destes dados ao custo de um maior tempo de processamento, quanto maior for o nível de granularidade menor será a riqueza de informações possíveis, porém maior será a agilidade do sistema, assim a granularidade deve ser definida almejando-se o equilíbrio entre os recursos do sistema e as necessidades do usuário (SELL, 2001).

Uma maior eficiência de armazenamento pode ser obtida por meio da utilização de dois, ou mais níveis de granularidade, como por exemplo manter um alto nível de detalhamento para os dados recentes e resumir os dados mais antigos, assim um número maior de informações relevantes poderá ser extraído a um menor custo de processamento (CAMARGO, 2004).

2.11 ARQUITETURAS DE UM DATA WAREHOUSE

Segundo Domenico (2001), um DW deve possuir uma arquitetura que o permita responder, de forma rápida e precisa, consultas complexas a dados provenientes de sistemas distintos, destacando-se as seguintes abordagens:

- **Arquitetura *Top-Down*:** Inicia pela extração, transformação e integração dos dados provenientes das fontes externas, os dados e metadados resultantes desse processo são carregados no DW.
- **Arquitetura *Bottom-Up*:** Implementação de DMs independentes com o armazenamento de dados em baixo nível de granularidade e sem a realização de um DW, que é construído de forma incremental ao longo do tempo. O processo tem início pela extração, transformação e integração dos dados provenientes das fontes externas, os dados e metadados resultantes desse processo são carregados para um ou mais DMs independentes.

2.12 PROCESSO DE *DATA WAREHOUSING*

O processo de *Data Warehousing* se inicia pela extração dos dados de suas variadas fontes de armazenamento, posteriormente ocorre a integração e a transformação de modo a promover uma padronização destes dados, e por fim a carga no DW e sua utilização (CAMARGO, 2004).

- **Extração:** O mapeamento para extração é realizado por meio de metadados que indicam a localização dos dados e sua possível forma de extração e a definição de quais dados devem ser extraídos se dá a partir do modelo dimensional (FAVARETTO, 2005). O processo de extração é responsável então por selecionar e copiar os dados provenientes de diferentes fontes para a área de transformação.
- **Transformação:** Os dados coletados pelo processo de extração podem apresentar valores incompletos ou alheios aos limites esperados, assim se faz necessária a execução de um processo de limpeza e padronização (FAVARETTO, 2005). A transformação dos dados coletados é realizada em uma área intermediária, área de estágio, por uma ferramenta capaz de trabalhar com dados heterogêneos. (SELL, 2001).

Segundo Camargo et al (2004) os processos de transformação mais utilizados são:

- ✓ Limpeza: objetiva eliminar inconsistências, duplicidades, a existência de valores incorretos e promover a uniformidade;
 - ✓ Eliminação: trata-se de desconsiderar dados irrelevantes ao DW;
 - ✓ Combinação: trata-se de mesclar registros iguais ou complementares;
 - ✓ Desnormalização e normalização: reunir em uma única dimensão as hierarquias de dados previamente separadas pelo processo de normalização, ou normalizar partes de registros totalmente desnormalizados;
 - ✓ Cálculos, derivação e alocação: são transformações a serem executadas de acordo com as regras de negócio previamente identificadas.
- **Carga:** Após o processo de transformação, os dados eles devem ser carregados para o DW, levando-se em consideração a sua integridade, se a carga deverá ser incremental ou total, a capacidade de atender a exigências prévias ou posteriores à carga como a eliminação e recriação de índices e o particionamento físico de tabelas e índices (CAMARGO, 2004).

Por fim os dados tornam-se acessíveis para as aplicações de apoio a tomada de decisão (CAMARGO, 2004). Estas aplicações dependem da adaptação da ferramenta OLAP para as necessidades específicas do usuário (PAIM, 2003).

3 REQUISITOS PARA DATA WAREHOUSE

Este capítulo apresenta as principais abordagens de levantamento de requisitos para a construção de um DW, como também cita algumas pesquisas recentes na área.

3.1 ABORDAGENS DE LEVANTAMENTO DE REQUISITOS PARA DW

Os requisitos são as capacidades operacionais que um sistema ou processo deve possuir para atender as necessidades do negócio, por isso, as atividades de levantamento e análise de requisitos servem como base para o desenvolvimento de qualquer sistema (CARVALHO, 2009).

Segundo Moreira (2015), a definição inadequada dos requisitos impacta o desenvolvimento de um DW gerando problemas como a ausência de objetivos que justifiquem a sua implementação, dificuldade de compreensão das informações requeridas e a implementação baseada em suposições incorretas. Ainda de acordo com Moreira (2015), existem três abordagens específicas para o levantamento de requisitos de um DW:

- **Abordagem Orientada a Dados:** Consiste em levantar os requisitos por meio de análise direta dos sistemas de origem, sem ocorrer muita interação com os usuários que irão utilizar o sistema e, portanto, não priorizando o atendimento aos requisitos de negócio.
- **Abordagem Orientada a Requisitos:** Por meio das técnicas de aquisição de conhecimento determina-se os requisitos de informação dos usuários. O desenvolvimento do ETL (*Extract, Transform, Load*) deve se adequar aos requisitos levantados podendo ocorrer de os sistemas de origem não possuírem dados suficientes para a composição das informações solicitadas, pode-se gerar duplicidade de informações e conseqüente aumento no consumo dos recursos da infraestrutura do sistema.
- **Abordagem Orientada a Metas:** Abordagem que visa atender aos processos chave do negócio por meio da obtenção dos requisitos mediante entrevistas com a alta gerência da organização, requerendo para isso que a equipe de projeto já possua prévio conhecimento a respeito de questões técnicas necessárias para o entendimento das necessidades.

3.2 PESQUISAS RECENTES EM DW

A área da estratégia empresarial que engloba os processos que auxiliam os decisores por meio de informações é conhecida como *Business Intelligence* (BI) (BENCK; GURA, 2011).

Em decorrência de uma série de fatores, é possível (e inclusive bastante comum) que o desempenho de soluções de BI seja tão aquém do esperado que acabem inviabilizando seu uso. Diante deste cenário, várias são as pesquisas que objetivam aperfeiçoar os tradicionais DW. Uma das técnicas mais utilizadas para melhorar o desempenho de DW é a criação de tabelas agregadas (MACHADO, 2000).

Uma forma de criação de tabelas agregadas é através do uso de visões materializadas (FORLANI, 2006). Sidi (2016) utilizou duas técnicas diferentes em conjunto e obteve uma melhora de 20% em alguns casos: particionamento da tabela fato e utilização de índice *bitmap* de junção - IBJ.

Moskado (2016) sugere uma implantação ágil em ambientes de BI que utilizem bancos de dados relacionais: a Agile ROLAP. O referido pesquisador usa *Foreign Data Wrappers* (FDW), dispensando, assim, a necessidade de implementação de DW físicos nem de processos de ETL pois, através do uso de *Foreign Tables*, o FDW fica vinculado às tabelas dos sistemas de origem. Além disso, FDW permite a implementação de uma solução de BI porque as ferramentas OLAP são compatíveis com *Foreign Tables*.

Carniel (2012), utilizou o BJIn OLAP (um servidor OLAP baseado em IBJ) e, através da ferramenta *Star Schema Benchmark*, aferiu uma melhora de até 98% no tempo de resposta para consultas se comparado às mesmas consultas realizadas em um DW equivalente implementado em um servidor OLAP Mondrian.

Carniel e Sá (2012) analisaram o desempenho de bancos de dados relacionais e NoSQL no contexto de DW. Foram analisados o tempo de resposta para consultas, o uso de processador, o uso de memória RAM e o consumo de disco de um DW implementado no banco PostgreSQL combinado de diferentes formas com as seguintes tecnologias: MongoDB, FastBit e LucidDB.

A pesquisa mostrou que o LucidDB usa menos de 50% do espaço em disco utilizado pelo PostgreSQL e menos de 10% que o utilizado pelo MongoDB. Com a pesquisa foi verificado que o FastBit permitiu uma melhora de até 99,8% no tempo de resposta de consultas.

Moreira (2013) propõe uma maneira de integrar dados estruturados e dados não estruturados em um mesmo DW, armazenando-os em DM distintos. A integração destes diferentes DM é feita através de um processo denominado *Linkage*: técnica que consiste na

criação de uma nova tabela, a qual fica responsável por armazenar relacionamentos entre as tabelas fato dos diferentes DM em questão.

Walter (2016) trabalha com a utilização de operadores OLAP em bancos de dados NoSQL. Sua pesquisa mostra a utilização do banco de dados MongoDB com o MR OLAP: ferramenta responsável pela geração de operadores OLAP no MongoDB.

Brainer (2016) analisa o impacto que o uso de diferentes mídias de armazenamento ocasiona no desempenho de bases de dados no que se refere a tempo de resposta. Utilizando a ferramenta *TPCH Benchmark*, o pesquisador mostra que o emprego de *Solid State Disk (SSD)* como mídia de armazenamento para bancos de dados não proporciona grandes ganhos se comparado com os obtidos com o uso de associações de HDDs de boa qualidade. O pesquisador aponta a gestão de buffer dos bancos de dados como o gargalho que prejudica o desempenho de SSD no contexto de bancos de dados.

Santos (2016) propõe um tratamento diferenciado para dados estacionários (os que devem ser mantidos no DW por mais tempo) e dados transitórios (os que apenas servem para produzir um insight para a gestão da organização). Apesar de integrados na visão, cada uma dessas classes de dados deve ser armazenada separadamente no DW, de acordo com a pesquisadora.

A pesquisa de Santos (2016), combina o conceito de Sistema de Banco de Dados Ativo e o de Sistema de Recomendação (SR), propondo um monitoramento de eventos de *Data Definition Language (DDL)* disparados contra um banco de dados para, com base neles, propor uma série de recomendações para otimização do banco de dados. Reconhecendo os impactos negativos que eventos DDL disparados contra um banco de dados podem implicar em processos de ETL, por exemplo, percebe-se a relevância da pesquisa de Santos no cenário de *Data Warehousing*.

Percebe-se, enfim, o esforço da comunidade científica no sentido de buscar soluções e aprimoramentos para a área de BI. São várias as possíveis melhorias que podem aprimorar o desempenho de ambientes de DW. Essa série de possíveis melhorias podem resultar em consultas OLAP ágeis e tomadas de decisões estratégicas *on-time*.

4 PRINCIPAIS METODOLOGIAS PARA DESENVOLVIMENTO DE SISTEMAS DATA WAREHOUSE COM ENFOQUE NOS REQUISITOS

O desenvolvimento de um sistema *Data Warehouse* é completamente influenciado pelos requisitos de negócios dos usuários, pois eles determinam quais informações deverão ser geradas e em qual periodicidade, além de como estas informações deverão ser organizadas. Por este motivo mudanças nos requisitos de negócios podem afetar seriamente o sistema e, portanto, é necessário realizar corretamente a especificação de requisitos (PAIM, 2003).

Segundo Paim (2003), o emprego de uma metodologia organizada que favoreça a compreensão dos requisitos promoverá o estabelecimento de um esquema conceitual formal, completo e livre de ambiguidades, a redução de problemas futuros relacionados a incorreta alocação de requisitos por meio da detecção antecipada de erros na modelagem, a possibilidade de construção de um sistema escalável e a redução no tempo de desenvolvimento deste sistema.

Neste capítulo serão apresentadas cinco metodologias para o desenvolvimento de sistemas DW selecionadas a partir do levantamento bibliográfico.

4.1 O MÉTODO HADDEN-KELLY

Segundo Moreira (2000) esse método objetiva o rápido desenvolvimento de um DW por meio de uma construção incremental, com a obrigatoriedade de apresentar benefícios significativos e mensuráveis a cada refinamento e é composto por quatro fases:

- **Preparação:** Etapa na qual se deve compreender as regras de negócio, averiguar recursos necessários e o ambiente técnico, com isso criar um Plano de Ação para a organização (PAIM, 2003).
- **Planejamento:** Etapa em que ocorre a identificação e documentação das informações que cada área específica de negócio necessita, determinando deste modo os objetivos a serem alcançados, seus indicadores e prioridades e construindo o Plano de Implementação por meio da averiguação dos dados necessários para obtenção das as informações requeridas, da decisão das regras para a implementação e a definição da infraestrutura de armazenamento de dados (MOREIRA, 2000).
- **Construção:** Construção de um *Data Mart*, com a implementação da infraestrutura, sua modelagem e mapeamento dos dados operacionais, o desenvolvimento e testes de programas para ETL (PAIM, 2003).

- **Operação:** Implantação no cliente final, procedimentos de configuração, segurança, monitoramento de desempenho (MOREIRA, 2000).

Esta abordagem demonstra uma maior ênfase na gerência de projeto do que na gerência de requisitos, não apresentando uma sistemática específica para a coleta e validação dos requisitos, entretanto a construção incremental facilita o gerenciamento de mudanças de requisitos, além disso, suas fases de preparação e planejamento promovem a coleta de um escopo bem definido do projeto (PAIM, 2003).

4.2 *BUSINESS DEVELOPMENT LIFECYCLE* (BDL)

Segundo Vaz (2009) este método é iniciado pela etapa de planejamento do projeto, na qual ocorre a avaliação das características do negócio que justifiquem a implementação de um DW, dos riscos e do retorno de investimento associados a essa implementação e a familiarização com a cultura organizacional. Posteriormente, de acordo com o mesmo autor, ocorre a etapa de compreensão e definição dos requisitos do negócio e, quando em posse destes requisitos, ocorre a etapa de construção do modelo multidimensional.

As etapas subsequentes dizem respeito à construção deste modelo, das ferramentas para ETL e a implantação do DW (PAIM, 2003). Em relação à aplicação de práticas de engenharia de requisitos neste método ocorre a separação dos requisitos de negócio e do projeto físico da aplicação, a busca pela conformidade entre os requisitos multidimensionais e o modelo corporativo, a realização de entrevistas e reuniões para o levantamento dos requisitos do sistema (PAIM, 2003).

4.3 GOLFARELLI E RIZZI

Metodologia estruturada em seis fases, das quais podem ser citadas como relevantes ao processo de levantamento de requisitos a primeira fase que diz respeito ao estudo de prévia documentação a respeito das fontes de dados com a finalidade de definir esquemas de integração iniciais e a segunda fase em que ocorre a especificação de requisitos na qual os requisitos dos usuários são filtrados de acordo com os esquemas definidos na primeira fase, produzindo uma especificação preliminar de fatos e dimensões (PAIM, 2003).

4.4 CONCEPTUAL DATA WAREHOUSE DESIGN (CDWD)

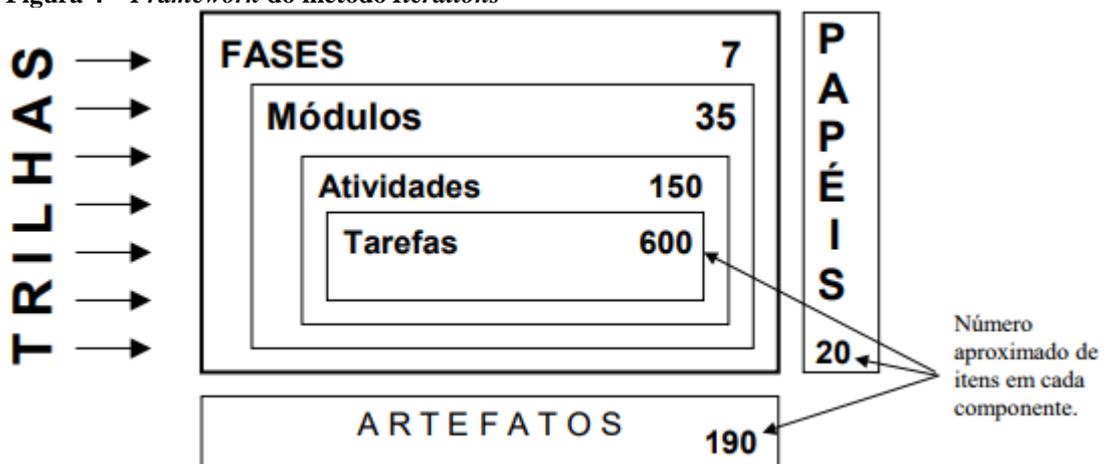
Metodologia para o desenvolvimento de um projeto multidimensional conduzido por uma sequência de etapas e exigindo grande atenção ao levantamento de requisitos e a elaboração do projeto conceitual que tem por finalidade derivar o esquema de dimensões, métricas e hierarquias que formará a base da aplicação apresentando, para isso, diretrizes para a modelagem do esquema de dados, um formalismo gráfico que captura apropriadamente a distinção entre atributos e propriedades num nível dimensional e a aplicação de formas normais multidimensionais (PAIM, 2003).

4.5 ITERATIONS

Segundo Sá (2009), o método é constituído por cinco trilhas, sete fases e trinta e cinco módulos com atividades específicas subdivididas em tarefas, apresentando vinte papéis a serem desempenhados pela equipe de desenvolvimento, conforme exposto na Figura 1.

Trilhas são conjuntos distintos de módulos que devem ocorrer em paralelo, fases representam um agrupamento progressivo destes módulos, os quais são concluídos ao termino da fase (PAIM, 2003).

Figura 4 – Framework do método Iterations



Fonte: Paim (2003)

Segundo Sá (2009), a primeira trilha do método agrupa módulos relacionados a gestão do projeto, a segunda agrupa módulos relacionados a modelagem dos requisitos de negócio e sua aprovação pelos usuários finais, a terceira agrupa módulos relacionados diretamente a construção e população do DW, a quarta agrupa módulos relacionados a avaliação de recursos

técnicos, seleção, integração, preparação do ambiente e testes e, por fim, a quinta trilha agrupa módulos relacionados a integração dos metadados técnicos com os metadados de negócios e o desenvolvimento do acesso aos metadados.

Como primeira fase do desenvolvimento temos a definição de qual arquitetura do DW será implementada, a segunda fase diz respeito a organização da gestão do projeto, na terceira fase se dá análise do projeto em si, na quarta fase ocorre a modelagem do DW, na quinta fase o desenvolvimento do DW, a sexta fase é composta pelos testes e na sétima fase se dá a implantação para a utilização dos clientes finais, além do monitoramento de desempenho para a execução de otimizações (SÁ, 2009).

Este método apresenta características importantes em relação a especificação de requisitos como possuir uma trilha dedicada à análise e aprovação dos requisitos de negócios, papéis bem definidos, *templates* de artefatos para registro de requisitos multidimensionais identificados; *checklists* para cada uma das atividades (PAIM, 2003).

4.6 QUADRO COMPARATIVO

O Quadro 2 expõe o nível de atendimento a práticas de engenharia de requisitos em cada uma das metodologias abordadas nesse capítulo.

Quadro 2 – Comparação do nível de atendimento a práticas de engenharia de requisitos pelas metodologias de desenvolvimento de DW

Boas Práticas em Engenharia de Requisitos	Hadden-kelly	Bdl	Golfarelli E rizzi	Cdwd	Iterations
Processo Iterativo com refinamentos sucessivos	C	C			C
Gerenciamento de Mudanças em Requisitos	P				
Modelagem de Aspectos Organizacionais	P				C
Técnicas de Elicitação de Requisitos		P	P	P	P
Técnicas de Análise e Negociação dos Requisitos		P			P
Técnicas de Documentação dos Requisitos		P		P	P
Técnicas de Validação dos Requisitos			P		P
Especificação do Sistema em Alto Nível		P			
Modelo de Requisitos Detalhado					
Separação entre Requisitos e Aspectos Implementacionais	P	C		P	P
Planejamento da Gerência de Requisitos	P	P			P
Modelagem de Hierarquias Múltiplas			C	C	
Modelagem da Aditividade dos Fatos		C	C	C	
Modelagem da Cardinalidade das dimensões		C			
Conformidade de Requisitos Multidimensionais		C			
Conformidade de Requisitos Funcionais					
Legenda: C= atende completamente à prática, P= atende parcialmente à prática, Brancos= não tende a prática.					

Fonte: Paim (2003)

5 CONSIDERAÇÕES FINAIS

Este capítulo está dividido em duas seções, a primeira seção aborda as conclusões sobre o presente trabalho. A segunda seção apresenta os possíveis trabalhos futuros complementando o trabalho atual.

5.1 CONCLUSÃO

Atualmente, a realidade de um mercado globalizado e, por consequência, altamente competitivo impôs aos gestores um grande aumento na complexidade do processo decisório, pois cada decisão a ser tomada dentro do domínio de um negócio específico pode ser composta por um número significativo de variáveis presentes em um grande volume de dados.

A especificidade existente para a tomada de decisão dentro de cada domínio de negócio de um departamento de cada empresa que atua em algum setor distinto e que, além disso, possui uma cultura organizacional e um conjunto de sistemas próprios, e por consequência, geram e armazenam os seus dados das mais variadas formas possíveis, exige que a implementação de uma solução para o apoio a tomada de decisão seja específica para cada empresa.

Por isso, nota-se que o sucesso do desenvolvimento de um DW está intrinsecamente relacionado ao entendimento da cultura organizacional e dos requisitos de negócio dos usuários. A dificuldade de identificação das necessidades de informação dos usuários pode impactar de forma profunda a implementação do sistema de DW, por isso se faz necessário desenvolver cuidadosamente a especificação de requisitos.

A falta de domínio de questões técnicas necessárias para o entendimento dos processos que estão sendo analisados poderá comprometer a compreensão dos documentos avaliados e das informações obtidas dos profissionais especialistas, podendo gerar interpretações incompletas ou incorretas, além disso, a falta de comprometimento dos especialistas com o projeto poderá gerar prejuízos.

Um DW por executar consultas complexas em grandes volumes de dados, provenientes de sistemas distintos também exige o desenvolvimento cuidadoso de seus requisitos no que tange a infraestrutura, arquitetura, modelagem dos dados. A decisão incorreta do nível de granularidade, por exemplo, poderá afetar drasticamente o desempenho do sistema.

5.2 TRABALHOS FUTUROS

Para trabalhos futuros pode ser mencionada a aplicação das Técnicas de Aquisição do Conhecimento para a especificação de um banco *Data Warehouse*, e dos demais conteúdos apresentados neste trabalho, em um estudo de caso prático.

REFERÊNCIAS

- BARBOSA, Ricardo R. Gestão da informação e do conhecimento: origens, polêmicas e perspectivas. **Informação & Informação**, Londrina, v.13, n. esp., 2008. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/download/1843/1556>>. Acesso em: 25 out. 2017.
- BATISTA, Cleisson F. L.; NETO, Mário G.; SOUZA, Ellen P. R. Adequação do processo de desenvolvimento de data warehouse para empresa varejista que utiliza software COTS. **Revista Brasileira de Administração Científica**, Aquidabã, v.5, n.2, out. 2014. Disponível em <<http://sustenere.co/journals/index.php/rbadm/article/view/SPC2179-684X.2014.002.0012/473>>. Acesso em: 25 out. 2017.
- BENCK, Larissa L. N.; GURA, Emanoély F. **Construção de um Data Warehouse, Aliado a uma Ferramenta Open Source Ireport na Geração de Informações para Tomada de Decisão**. Universidade Tecnológica Federal Do Paraná, 2011. 88 f. Trabalho de Conclusão de Curso (Graduação) – Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas. Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2011.
- BONFADINI, Gerson J. ; MARTINS, Bárbara. **Tomada de Decisão: Analisando o Uso de Sistemas de Informação na Empresa Joagro Ferragens de Estrela/Rs**. Centro Universitário UNIVATES, 2014. Trabalho de Conclusão de Curso (Graduação) – Curso Superior de Administração. Centro Universitário UNIVATES, Lajeado, 2014. Disponível em <<http://www.univates.br/revistas/index.php/destaques/article/download/453/445>> Acesso em: 28 out. 2017.
- BRAINER, Angelo. Hardware-adware Database Systems: A new era for database is coming. **In: Simpósio Brasileiro de Banco de Dados**, v 31, p. 187192, 2016.
- CAMARGO, Denis; HOKAMA, Daniele D. B.; FOGLIENE, João L. V.; FUJITA, Francine. **A Modelagem de Dados no Ambiente Data Warehouse**. Universidade Presbiteriana Mackenzie, 2004. Trabalho de Conclusão de Curso (Graduação) – Curso Superior de Bacharelado em Sistemas de Informação. Universidade Presbiteriana Mackenzie, São Paulo, 2004. Disponível em <<http://homes.dcc.ufba.br/~mauricio052/Topicos%20em%20BD/Material%20Didático/Monografias/A%20Modelagem%20de%20Dados%20no%20Ambiente%20DW.pdf>> Acesso em: 25 out. 2017.
- CARVALHO, Elaine A. de. **Heurísticas para Identificação de Requisitos de Data Warehouse a partir de Indicadores de Desempenho**. Pontifícia Universidade Católica do Rio de Janeiro PUC-RJ, 2009. Disponível em <https://www.maxwell.vrac.puc-rio.br/15136/15136_4.PDF> Acesso em: 30 out. 2017.

CARNIEL, Anderson Chaves. **BJIN OLAP**: uma ferramenta OLAP baseada no índice bitmap de junção. 2012. 37 p. Trabalho de conclusão de curso (Tecnólogo em Análise e Desenvolvimento de Sistemas) – Instituto Federal de São Paulo – IFSP. São Paulo, 2012.

CARNIEL, Anderson Chaves; SÁ, Aried de Aguiar et al. Análise experimental de dados relacionais e noSQL no processamento de consultas sobre Data Warehouse. **In: Simpósio Brasileiro de Banco de Dados**, v 27, p. 113-120, 2012.

DOMENICO, Jorge A. D. **Definição de um Ambiente Data Warehouse em Uma Instituição de Ensino Superior**. Universidade Federal de Santa Catarina, 2001. Disponível em <<https://repositorio.ufsc.br/bitstream/handle/123456789/79520/179215.pdf?sequence=1&isAllowed=y>> Acesso em: 30 out. 2017.

FAVARETTO, Fábio. Ambiente de Data Warehouse para análise de algumas medidas utilizadas na administração de produção. **Revista da FAE**, Curitiba, v. 8, n. 2, jul. 2005. Disponível em <<https://revistafae.fae.edu/revistafae/article/download/387/271>> Acesso em: 25 out. 2017.

FILHO, Bruno Edson Martins de Albuquerque. **OBAS**: um benchmark para serviços de análises. 2014. 161 p. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Pernambuco. Recife, 2014.

FORLANI, Diogo Tuler. **Fragmentação vertical de dados em data warehouses no sistema WEBD²W**. 2006, 140 p. Dissertação (mestrado em Ciência da Computação) – UEM. Paraná. 2006.

HEINZLE, Roberto. **Um Modelo de Engenharia do Conhecimento para Sistemas de Apoio a Decisão com Recursos para Raciocínio Abduativo**. Universidade Federal de Santa Catarina UFSC, 2011. Disponível em <<https://repositorio.ufsc.br/bitstream/handle/123456789/95417/289391.pdf?sequence=1&isAllowed=y>> Acesso em: 30 out. 2017.

KIMBALL, R. **The Data Warehouse Toolkit**. New York, John Wiley & Sons. 1996.

MACHADO, Felipe Nery Rodrigues. **Projeto de data warehouse: uma visão multidimensional**. 2ª Ed. São Paulo: Érica, 2000. 318 p.

MOREIRA, João P. C. L. M. **A construção de um sistema de Armazenamento de Dados no âmbito do sistema GIST98/EUROBUS**. Escola de Gestão do Porto, 2000. Disponível em

<<https://repositorio-aberto.up.pt/bitstream/10216/6818/2/25727.pdf>> Acesso em: 29 out. 2017.

MOREIRA, João Luiz et al. JointOLAP - sistema de informação para exploração conjunta de dados estruturados e textuais: um estudo de caso no setor elétrico. **In: Simpósio Brasileiro de Banco de Dados**, v 28, p. 756-767, 2013.

MOREIRA, Walber R. **Estudo Comparativo de Métodos para Levantamento de Requisitos de Data Warehouse**. Instituto de Pesquisas Tecnológicas do Estado de São Paulo, 2015. Disponível em <http://cassiopea.ipt.br/teses/2015_EC_Walber_Moreira.pdf> Acesso em: 30 out. 2017.

MOSKADO, Felipe Igawa et al. Utilização e integração de bases de dados relacionais por meio de foreign tables para utilização em ferramentas OLAP. **In: Workshop de Iniciação Científica em Sistemas de Informação**, v 3, p. 52-55, 2016.

PAIM, Fábio R. S. **Uma Metodologia para Definição de Requisitos em Sistemas Data Warehouse**. Universidade Federal de Pernambuco UF-PE, 2003. Disponível em <http://www.cin.ufpe.br/~if696/referencias/warehousing/Uma_Metodologia_para_Analise_de_Requisitos_em_Sistemas_Data_Warehouse.pdf> Acesso em: 8 mar. 2016.

RAMEZ, Elmasri; SHAMKANT, B. Navathe **Sistemas de Banco de Dados**, 6. ed. São Paulo:Pearson Addison Wesley, 2011.

SÁ, Jorge V. de O. **Metodologia de Sistemas de Data Warehouse**. Universidade do Ninho, 2009. Disponível em <https://repositorium.sdum.uminho.pt/bitstream/1822/10663/4/Tese%20de%20doutoramento_Jorge%20Vaz%20de%20Oliveira%20e%20Sá_2009.pdf> Acesso em: 29 out. 2017.

SANTOS, Paulo Henrique. **Uma arquitetura para monitoramento de banco de dados e recomendações utilizando sistema de banco de dados ativos**. 2016. 67 p. Dissertação (Mestrado em Computação Aplicada) – UTFPR. Curitiba, 2016.

SELL, Denilson. **Uma Arquitetura para Distribuição de Componentes Tecnológicos de Sistemas de Informações Baseados em Data Warehouse**. Universidade Federal de Santa Catarina UFSC-SC, 2001. Disponível em <<https://repositorio.ufsc.br/bitstream/handle/123456789/82258/185086.pdf?sequence=1&isAllowed=y>> Acesso em: 26 out. 2017.

SIDI, Emany. Star schema advantages on data warehouse: using bitmap index and partitioned fact tables. In: **International Journal of Computer Applications**, v 134, n 13, 11-13, 2016.

SIDREIRA, Ana Cristina do V. **Data Warehouse: A Importância dos Sistema Data Warehouse na Tomada de Decisão da Empresa**. Universidade Candido Mendes UCAM-RJ, 2009. Disponível em <http://www.avm.edu.br/docpdf/monografias_publicadas/k211439.pdf> Acesso em: 11 mar. 2016.

SILVA, Franciny Q. da ; TORRES, Juliano R. **Proposta de Data Warehouse para ser Consumido por Gestores de Produção**. Faculdade de Tecnologia de Bauru FATEC-SP, 2015. Disponível em <<http://fatecbauru.edu.br/mtg/source/Proposta%20de%20Data%20warehouse.pdf>> Acesso em: 25 out. 2017.

SILVA, Shirly C. M.; COSTA, Welbson S. Aquisição de Conhecimento: O Grande Desafio Na Concepção de Sistemas Especialistas. **HOLOS**, [S.l.], v. 2, p. 37-46, dez. 2007. ISSN 1807-1600. Disponível em: <<http://www2.ifrn.edu.br/ojs/index.php/HOLOS/article/view/71/77>>. Acesso em: 30 out. 2017. doi:<http://dx.doi.org/10.15628/holos.2005.71>.

SIQUEIRA, Fernando de. **Banco de Dados I**. Centro Universitário Plínio Leite UNIPLI, [entre 1999 e 2017]. Disponível em <<https://sites.google.com/site/uniplibancodedados1/exercicios/exercicio-16---sql>> Acesso em: 30 out 2017.

TANAKA, Asterio K. **Banco de Dados Distribuídos e Datawarehousing**. Centro Universidade Federal do Estado do Rio de Janeiro UNIRIO, [entre 1997 e 2017]. Disponível em <<http://www.uniriotec.br/~tanaka/TIN0036/10-BDDDW-Mod-Dimens-Basicol.pdf>> Acesso em: 30 out 2017.

WAGNER, Cláudio A. **Estudo para Implantação de um Data Warehouse em um Ambiente Empresarial**. Universidade Federal de Santa Catarina, 2003. Disponível em <<https://repositorio.ufsc.br/bitstream/handle/123456789/84989/196929.pdf?sequence=1>> Acesso em: 12 mar 2016.

WALTER, Roberto et al. Implementação de operadores OLAP utilizando o modelo de programação Map Reduce no MongoDB. In: **Escola Regional de Informática de Banco de Dados**, v 12, p. 68-77, 2016.

ZIULKOSKI, Luís C. C. Coleta de Requisitos e Modelagem de Dados para Data Warehouse: Um Estudo de Caso Utilizando Técnicas de Aquisição de Conhecimento. Universidade Federal do Rio Grande do Sul UFRS-RS, 2003. Disponível em <<http://www.inf.ufrgs.br/bdi/wp-content/uploads/Monografia-LuisZiulkoski.pdf>> Acesso em: 30 out 2017.