

Daniela Andrade Raslan¹
 Angélica Toffano Seidel Calazans²

Resumo

O gerenciamento de dados tornou-se uma das atividades mais importantes nas organizações. À medida que nossa sociedade se orienta cada vez mais para a informação, organizar dados para maximizar sua utilidade mostra-se um desafio para o profissional que lida com essa área. Embora os sistemas de banco de dados simplifiquem a tarefa de manter e recuperar uma grande quantidade de dados, ainda são encontradas inúmeras dificuldades na organização dos dados para utilizar a capacidade total de um banco de dados. Atualmente, as empresas que mantêm um volume muito grande de dados distribuídos em diversos sistemas pela organização têm optado por uma forma de gerir suas informações: o *Data Warehouse*. O objetivo desta pesquisa é identificar os principais conceitos relacionados a *Data warehouse*, e verificar sua disseminação em algumas empresas brasileiras e as vantagens de sua aplicação. Por meio do processo de revisão sistemática — um modo de identificar e avaliar todas as pesquisas relevantes a respeito de uma questão particular, bem como identificar lacunas capazes de direcionar novas pesquisas e novas atividades de investigação —, verificou-se, inicialmente, a disseminação de conceitos relativos à modelagem de dados, a banco de dados e a *Data Warehouse*. Com essa investigação, foi possível verificar como se desenvolveu e se consolidou o processo de *Data Warehouse*. Para verificar a aceitação desse conceito nas empresas da área de tecnologia da informação, investigou-se junto a algumas empresas de Brasília a aplicabilidade dos conceitos desse processo no mercado, o que permitiu relacionar as dificuldades enfrentadas e as desvantagens encontradas para a implantação do processo, bem como as vantagens alcançadas com o término dos projetos.

Palavras-chave: Revisão Sistemática. Banco de dados. *Data Warehouse*.

Abstract

Data management has become one of the most important activities in organizations. As our society is geared increasingly to the information, organize data to maximize its utility proves a challenge for the professional who deal with this area. Although database systems simplify the task of maintaining and retrieving a large amount of data, many difficulties are still encountered in data organization to utilize the full capacity of a database. Currently, companies that maintain a very large amount of data distributed across many systems the organization has opted for a way to manage your information: the *Data Warehouse*. The objective of this research is to identify the main concepts related to data warehouse, and check its spread in some Brazilian companies and the advantages of their application. Through the systematic review process - a way to identify and evaluate all relevant regarding a particular research question, as well as identify gaps capable of driving new research and new research activities - initially it appeared the spread of concepts relating to data modeling, the database and the data Warehouse. With this research, we observed how it developed and consolidated the process of *Data Warehouse*. To verify the acceptance of this concept in the business field of information technology, was investigated with some companies Brasilia the applicability of the concepts of this process on the market, allowing relate the difficulties and disadvantages encountered in the implementation process, and the advantages achieved with the completion of the projects.

Keywords: Systematic Review. Database. *Data Warehouse*.

* Recebido em: 23/09/2013

Aprovado em: 09/04/2014

¹ Bacharela e Licenciada em Física pela Universidade de Brasília (2006). Graduada em Análise e Desenvolvimento de Sistemas pelo UniCEUB (2013). Atualmente é professora de Física e Matemática da Secretaria de Educação do Distrito Federal.

² Doutorado em Ciência da Informação pela Universidade de Brasília (2008) e mestrado em Gestão do conhecimento e TI pela Universidade Católica de Brasília (2003). Atuou 28 anos como especialista da Caixa Econômica Federal e atualmente é professora titular do Centro Universitário de Brasília-UniCEUB.

1 Introdução

Conforme nos movemos para uma sociedade orientada para a informação, a determinação de como organizar os dados para maximizar sua utilidade torna-se um problema muito importante. Por essa razão, o gerenciamento de dados tornou-se uma atividade de extrema importância nas organizações, uma vez que a correta manipulação de dados, além de dinamizar os diversos setores da organização, pode também fornecer indicadores que auxiliam a gestão estratégica das corporações.

Sistemas de banco de dados simplificam a tarefa de manter e recuperar uma grande quantidade de dados. Contudo, o problema de como organizar os dados para utilizar a capacidade total de um banco de dados não é tarefa simples. Para simplificar essa organização, surgiram as técnicas de modelagens de dados que permitem a construção de abstrações do mundo real para relacionar dados de maneira eficiente. Paralelamente ao desenvolvimento das técnicas de modelagem de dados, aprimoraram-se também as abordagens para os modelos de Banco de Dados (BD) e os Sistemas Gerenciadores de Banco de Dados (SGBD).

As empresas que mantêm um volume muito grande de dados espalhados em diversos sistemas têm optado por uma nova forma de obtenção de informações: o *Data Warehouse* (DW). Esse conceito consiste em organizar os dados corporativos de maneira integrada, com uma única versão da verdade, histórico variável com o tempo e uma única fonte de dados. Tal integração possibilita à empresa identificar tendências em sua corporação, o que a possibilita posicionar-se estrategicamente no mercado de forma mais competitiva e, conseqüentemente, aumentar os lucros.

Assim, identificar os principais conceitos relacionados a *Data warehouse* e verificar sua disseminação em algumas empresas brasileiras e as vantagens de sua aplicação são os objetivos gerais desta pesquisa. São objetivos específicos: identificar seus principais conceitos e verificar sua disseminação nas empresas, averiguando as vantagens e desvantagens de sua adoção.

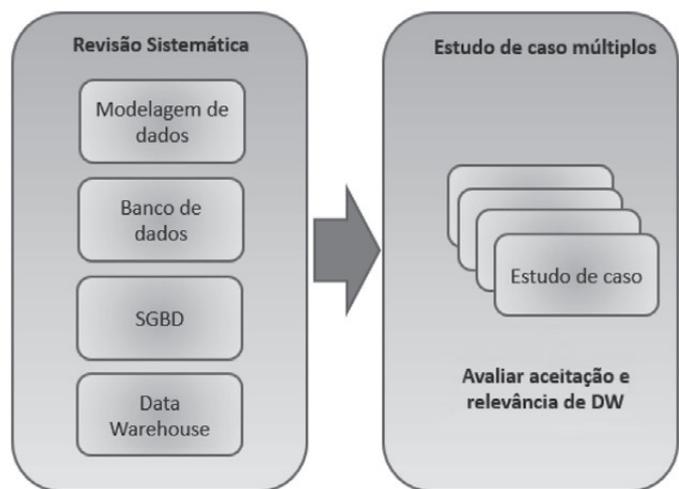
Para tanto, este trabalho está organizado do seguinte modo: 1) inicialmente descreve-se a metodologia de revisão sistemática utilizada para pesquisa dos conceitos ligados a *Data Warehouse*; 2) conceituam-se modelagem de dados, banco de dados e SGBD, assuntos necessários para o entendimento do objeto de pesquisa; 3) conceitua-se *Data Warehouse*, salientando-se as van-

tagens e desvantagens da adoção da tecnologia, segundo os artigos e livros pesquisados, e também os processos de construção de um *data warehouse*; e 4) verifica-se a aceitabilidade desse conceito pelas corporações por meio um instrumento de coleta de dados.

2 Metodologia

O processo definido para a pesquisa está apresentado na Figura 1. Inicialmente, realizou-se uma pesquisa documental, por meio do processo de revisão sistemática, sobre os conceitos de modelagem de dados, banco de dados, sistemas gerenciadores de banco de dados e *Data Warehouse*. Em seguida, fez-se um estudo de casos múltiplo com os profissionais da tecnologia da informação de empresas que utilizam ou utilizaram o *Data Warehouse*. Por fim, para avaliar a aceitação e relevância do *Data Warehouse* no mercado, compararam-se os resultados obtidos no estudo de caso e os da pesquisa documental.

Figura 1 – Modelo da pesquisa



Fonte: Próprio autor

2.1 Revisão Sistemática

A revisão sistemática é um modo de identificar e avaliar todas as pesquisas relevantes a respeito de uma questão particular. Ela pode ser útil para identificar lacunas que podem direcionar novas pesquisas e novas atividades de investigação. É um modo de os pesquisadores resumirem toda a informação existente sobre determinado fenômeno de forma completa e imparcial (KITCHENHAM, 2004).

As revisões sistemáticas iniciam-se pela definição de um protocolo de revisão que estabelecerá uma estratégia de pesquisa capaz de detectar toda literatura relevante

para determinada questão. A estratégia deve estabelecer os critérios para julgar qual literatura é concernente ao tema da pesquisa. São três os estágios da revisão sistemática: 1) Planejamento da Revisão; 2) Realização da Revisão; e 3) Relato da Revisão.

O Planejamento da Revisão envolve a identificação de o que é necessário para se iniciar a revisão e o desenvolvimento do protocolo de revisão. São definidos nessa fase o objetivo da pesquisa; as fontes de pesquisa; os critérios de inclusão e exclusão dos artigos de pesquisa; os critérios de avaliação dos estudos e os critérios de extração de dados de estudo. O desenvolvimento do protocolo de pesquisa é de suma importância, pois é uma maneira de impedir que a seleção de estudos do pesquisador seja unicamente dirigida às suas expectativas, o que implica delimitar as razões e as questões de pesquisa, as quais a revisão sistemática tem a intenção de responder. O protocolo de pesquisa possibilita que esta seja direcionada de modo imparcial. Com o protocolo, são estabelecidos também os termos que serão pesquisados e as fontes de dados da pesquisa.

A Realização da Revisão deve registrar as estratégias de busca e seus resultados. As estratégias de busca são geralmente iterativas e podem se beneficiar de combinações de termos de pesquisas que são derivados das questões de pesquisa. Deve ser redigida uma lista de sinônimos e abreviaturas que auxiliem as estratégias de busca. “O processo de realização de uma revisão sistemática deve ser transparente e replicável. A revisão deve ser documentada em detalhes suficientes para que os leitores sejam capazes de avaliar o rigor da pesquisa.” (KITCHENHAM, 2004)

O Relato da Revisão deve reunir a síntese dos dados pesquisados, a qual pode ser descritiva, não quantitativa. No entanto, é possível completar essa síntese com um resumo quantitativo utilizando-se técnicas de estatística para a elaboração de uma visão mais completa do trabalho de pesquisa.

A revisão sistemática foi feita de modo a explorar os conceitos de modelagem de dados, banco de dados e *Data Warehouse* com objetivo de responder aos seguintes objetivos:

Quais os conceitos de modelagem de dados, banco de dados e SGBD?

Qual o conceito de *Data Warehouse*?

Quais os processos de construção de um *Data Warehouse*?

Qual a aceitabilidade do conceito de *Data Warehouse* pelas corporações e pelos profissionais da área, relacionando as vantagens e desvantagens dessa tecnologia?

Desse modo, os artigos pesquisados na revisão sistemática responderam às seguintes questões de pesquisas:

O que é a modelagem de dados?

Quais as vantagens e desvantagens do uso da modelagem de dados?

Qual o tipo de modelagem de dados mais utilizado?

O que são banco de dados?

Quais os tipos de bancos de dados existentes?

Qual o tipo de banco de dados mais utilizados?

O que são os SGBDs e quais suas funções e importância?

Quais são os tipos de SGBD mais largamente utilizados atualmente no mercado?

O que é o *Data Warehouse*?

Quais os objetivos da criação de um *Data Warehouse*?

Quais as etapas de construção de um *Data Warehouse* e quais os desafios a serem vencidos em cada etapa?

Quais as vantagens e desvantagens da adoção de um *Data Warehouse*?

Qual a aceitabilidade do conceito de *Data Warehouse* no mercado?

Para estabelecimento do protocolo de pesquisa da revisão sistemática, adotaram-se com base nos termos chaves das questões de pesquisa e nos objetivos as seguintes *strings* para buscar os artigos: modelagem de dados, *data modeling*, banco de dados, *data base*, software de banco de dados, sistema gerenciador de banco de dados, SGBD, *Data Warehouse* e *Data Warehousing*.

As fontes de pesquisa utilizadas para a pesquisa foram:

Google Acadêmico: página eletrônica de pesquisas de artigos revisados por especialistas (*peer-reviewed*), teses, livros, resumos e artigos de editoras acadêmicas, organizações profissionais, bibliotecas de pré-publicações, universidades e outras entidades acadêmicas (GOOGLE, 2014);

Referências: Livros citados nas referências e escolhidos a partir dos livros mais vendidos nas livrarias virtuais Cultura, Saraiva e Siciliano e mais citados por especialistas da área de tecnologia da informação (TI);

Portal da CAPES: uma biblioteca virtual que reúne e disponibiliza a instituições de ensino e pesquisa no Brasil o melhor da produção científica internacional. Ele conta atualmente com um acervo de mais de 35 mil periódicos com texto completo, 130 bases referenciais, onze bases dedicadas exclusivamente a patentes, além de livros, enciclopédias e obras de referência, normas técnicas, estatísticas e conteúdo audiovisual. (CAPES, 2014);

SciELO - Scientific Electronic Library Online: modelo para a publicação eletrônica cooperativa de periódicos científicos na Internet. Foi desenvolvido para responder às necessidades da comunicação científica nos países em desenvolvimento e particularmente na América Latina e Caribe. (SCIELO, 2014]) Nesse site foi pesquisada a revista JISTEM (Journal of Information Systems and Technology Management);

Journal of Computer Science: revista de Ciência da Computação que tem como objetivo publicar artigos de pesquisa sobre fundamentos teóricos da informação

e computação e de técnicas para a sua implementação e aplicação em sistemas de computador. (JOURNAL OF COMPUTER SCIENCE, 2014]).

Como critérios de inclusão e exclusão de documentos de pesquisa, estabeleceu-se que: 1) os artigos deveriam estar disponíveis gratuitamente na *web* ou impressos, no caso de livros; 2) eles deveriam ser redigidos em português ou inglês; 3) os artigos, dissertações e teses deveriam datar entre os anos de 2011 e 2013; 4) a primeira seleção dos documentos seria feita pelo título que deveria fazer referência a alguma das questões de pesquisa; 5) a segunda seleção dos documentos seria efetuada após a leitura dos resumos dos documentos em que pelo menos 30% das questões de pesquisa deveriam ser respondidas.

A revisão sistemática foi conduzida no segundo semestre de 2012 e no início de 2013, entre os meses de setembro e fevereiro. A tabela abaixo demonstra os resultados encontrados:

Tabela 1 - Números da Pesquisa de Revisão Sistemática

Fonte	String de busca	Número de documentos como resultados da busca	Número de documentos após primeira seleção	Número de documentos após segunda seleção
Google Acadêmico	modelagem de dados (conceito, histórico, tipos)	165	6	4
Google Acadêmico	Data Modeling (concepts, types)	1340	9	4
Google Acadêmico	banco de dados, SGBD (conceitos, tipos)	231	18	1
Google Acadêmico	data base, SGBD (concepts, types)	18	18	2
Google Acadêmico	Data Warehouse, Data Warehousing (conceitos, concepts)	15	2	2
Portal da CAPES	modelagem de dados (conceito, histórico, tipos)	794	6	3
Portal da CAPES	Data Modeling (concepts, types)	1204	16	2
Portal da CAPES	banco de dados, SGBD (conceitos, tipos)	0	0	0
Portal da CAPES	data base, SGBD (concepts, types)	7810
Portal da CAPES	Data Warehouse, Data Warehousing (conceitos, concepts)	148	5	2
JISTEM	modelagem de dados (conceito, histórico, tipos)	1	1	1
JISTEM	Data Modeling (concepts, types)	3	2	2
JISTEM	banco de dados, SGBD (conceitos, tipos)	1	-	-
JISTEM	data base, SGBD (concepts, types)	2	2	2
JISTEM	Data Warehouse, Data Warehousing (conceitos, concepts)	0	0	0
Journal of Computer Science	modelagem de dados (conceito, histórico, tipos)	0	0	0
Journal of Computer Science	Data Modeling (concepts, types)	0	0	0
Journal of Computer Science	banco de dados, SGBD (conceitos, tipos)	8	2	2
Journal of Computer Science	data base, SGBD (concepts, types)	2	2	2
Journal of Computer Science	Data Warehouse, Data Warehousing (conceitos, concepts)	3	2	2

Fonte: Próprio autor

As buscas nas fontes de pesquisas citadas permitiram a seleção de trinta documentos que respondiam apenas parcialmente a algumas das questões de pesquisa. Após a leitura integral dos documentos, selecionaram-se dezoito documentos que respondiam adequadamente a, pelo menos, 30% das questões de pesquisa e que eram disponíveis gratuitamente na *web*. Não foi encontrado nenhum artigo que respondesse a todas as questões de pesquisa e fizesse a relação entre os temas modelagem de dados, banco de dados e *Data Warehouse*. Os livros pesquisados traziam em sua maioria explicações mais completas a respeito dos temas. Após a leitura dos documentos e livros, retiraram-se dos textos os trechos que respondiam às questões propostas e posteriormente foi feito um compêndio das definições encontradas.

2.2 Estudos de Caso

Segundo Yin e Martins (2001, 2006 apud CALAZANS, 2007), o Estudo de Caso é uma forma de investigação empírica que permite o estudo de fenômenos sociais complexos ao analisar o fenômeno contemporâneo dentro de seu contexto da vida real. De acordo com Martin (2006 apud CALAZANS, 2007, p. 40), “O estudo de caso é uma estratégia que pede avaliação quantitativa”, e por isso, de acordo com Cassel & Simon (1994 apud CALAZANS, 2007), a pesquisa deve ter foco na interpretação e não na quantificação. Ao realizar os estudos de caso, o pesquisador não tem controle sobre os eventos e deve descrever, compreender e interpretar o seu resultado (CALAZANS, 2007).

Para a definição dos estudos de caso, é necessário especificar o objeto de pesquisa, o fenômeno específico no tempo em que se deseja estudar e a unidade de análise em que se procedida à investigação. Os estudos de caso múltiplos são aqueles que utilizam unidades de análise diferentes. Sobre os estudos de casos múltiplos, Calazans (2007, p. 46) descreve:

Os Estudos de Casos Múltiplos são utilizados em mais de uma unidade de análise e a provas resultantes destes estudos são consideradas mais convincentes, pois normalmente são estudos mais robustos e rigorosos e utilizam uma abordagem mais completa. Estudos de Casos Múltiplos utilizam a lógica da replicação e não da amostragem [...]

Com intuito de verificar a aceitação dos conceitos de *Data Warehouse* pelos profissionais da área de tecnologia da informação, fez-se um estudo de casos múlti-

plos, por meio da aplicação de um questionário (conferir *Apêndice*), que permitiu aos participantes expressar sua avaliação a respeito da elaboração de um *Data Warehouse*, das dificuldades encontradas nessa elaboração, os casos de sucesso, as vantagens, desvantagens etc. O objetivo foi avaliar se os dados colhidos da revisão sistemática sobre *Data Warehouse* são confirmados na prática pelos profissionais que utilizam a teoria.

Na primeira parte do questionário, verificou-se o perfil do profissional de TI participante para averiguar sua formação e seu entendimento sobre o conteúdo avaliado. A segunda parte do questionário permitiu ao participante expressar suas experiências ao desenvolver projetos ligados a *Data Warehouse*.

A interpretação dos resultados da pesquisa de casos múltiplos foi feita por meio da condensação de significados, isto é, a abreviação do texto dos participantes em partes mais sucintas identificando pontos comuns entre as diversas falas dos entrevistados.

3 Resultados

Os resultados obtidos por meio da revisão sistemática permitiram responder aos questionamentos elaborados. Assim, a seguir apresentam-se sucintamente os conceitos de Modelagem de dados, Banco de dados, Sistemas Gerenciadores de Banco de Dados e *Data Warehouse*.

3.1 Modelagem de Dados

Pode-se afirmar que as informações, que uma empresa possui, são sua fonte de alimentação para novos processos, novas decisões, novas estratégias. Somente a partir do conhecimento que a empresa tem sobre si, sobre seus clientes, sobre seus fornecedores e sobre seus processos, é que ela pode avançar rumo a novos mercados e maiores lucratividades. Dessa forma, o gerenciamento de dados tornou-se uma atividade de extrema importância nas organizações, uma vez que a correta manipulação de dados, além de dinamizar os diversos setores de uma organização, pode também fornecer indicadores de preferências que auxiliam a gestão estratégica das corporações.

A modelagem de dados objetiva a construção de modelos que representem de maneira simples e abstrata o ambiente observado. Por meio dela, é possível observar com nitidez os processos de relacionamento entre os objetos em estudo e desenvolver uma documentação que

forneça mecanismos de validação da descrição do ambiente observado.

Define-se modelagem de dados como uma técnica que visa aperfeiçoar e dar maior estabilidade à organização dos dados. Para executar essa técnica, passos como a observação dos elementos de um ambiente, a elaboração de conceitos sobre esses elementos, a abstração de características comuns aos elementos e união desses elementos em conjuntos, que representarão a realidade, devem ser percorridos. A modelagem de dados possibilita a criação de um modelo de dados, isto é, de um conjunto de conceitos que podem ser usados para descrever a estrutura de um banco de dados.

Para que a modelagem de dados seja feita de forma a retratar a realidade em estudo, o primeiro passo deve ser o levantamento e análise dos requisitos com objetivo de entender e documentar como os dados de interesse para o problema são obtidos. Na sequência cria-se um esquema conceitual de alto nível. Segundo Elmasri e Navathe (2004), o modelo de dados conceitual é uma descrição concisa dos requisitos de dados e inclui descrições detalhadas de tipos de entidades, relacionamentos e descrições. A etapa seguinte é a implementação real do projeto no banco de dados com o auxílio de um SGBD.

Ainda sobre modelagem de dados, Heuser (2004) descreve essas etapas de construção mais detalhadamente:

Modelo Descritivo: descrição textual das estruturas de dados independente da abordagem do tipo de banco de dados.

Modelo conceitual: modelo abstrato de dados que descreve a estrutura de um banco de dados independente de sua classificação. Os objetos, suas características e relacionamentos entre elementos da realidade têm a representação fiel do ambiente observado, independente de quaisquer limitações impostas por tecnologias, técnicas de implementação ou dispositivos físicos.

Modelo lógico: modelo de dados que representa a estrutura de um banco de dados conforme vista pelo usuário do SGBD. Leva em conta limites impostos por algum tipo de banco de dados como banco de dados hierárquico, banco de dados em rede, banco de dados relacional, banco de dados orientado a objetos etc.

Modelo físico: é uma representação gráfica do detalhamento físico das estruturas dos dados que o banco de dados irá armazenar. Esse modelo é dependente do SGBD que será implementado. Leva em consideração limites impostos pelo SGBD e pelos requisitos não funcio-

nais dos programas que acessam os dados.

Uma das abordagens de modelagem mais comum é a Entidade-Relacionamento, um modelo de alto nível e independente do SGBD. Ela foi criada em 1976 por Peter Chen e tem como conceito essencial a percepção do mundo real representada por entidades e relacionamentos. A notação que será utilizada para a representação deste modelo é o Diagrama Entidade-Relacionamento (DER), exemplificado na Figura 2, em que os retângulos representam as entidades (elementos do domínio do problema) e os losangos representam os relacionamentos entre essas entidades. Naturalmente, existem outros elementos utilizados na construção desse diagrama, como cardinalidade, agregação, relacionamento ternário (ou de maior grau), autorrelacionamento e generalização/especialização.

Figura 2 - Notação do Diagrama Entidade Relacionamento



Fonte: (ELMASRI; NAVATHE, 2005).

3.2 Banco de dados e sistemas gerenciadores de bancos de dados

Um banco de dados é uma coleção de dados organizados de modo a satisfazer as necessidades dos usuários. Os bancos de dados são usados por meio de softwares chamados de Sistemas de Gerenciadores de Banco de Dados que consistem em um grupo de programas que executam manipulações reais sobre o banco de dados e fornecem uma interface entre os bancos de dados e o usuário ou entre os bancos de dados e os programas aplicativos.

O SGBD é o responsável pela manutenção e armazenamento dos dados de um BD. Ele contém, portanto, um catálogo com descrição completa do banco que contém informações sobre a estrutura de cada arquivo, o tipo e o formato de armazenamento de cada tipo de dado, restrições, entre outros. A informação armazenada no catálogo é denominada metadados. Os SGBD permitem a correta utilização e manipulação dos dados em um BD ao possibilitar que critérios como integridade, compartilhamento, segurança, controle da redundância e relacionamento dos dados sejam obedecidos. (ELMASRI; NAVATHE, 2005)

Os SGBD mais utilizados hoje foram concebi-

dos com base no modelo relacional, criado por Codd em 1970, derivado da teoria matemática dos conjuntos. Nesse modelo, os dados do banco são representados por meio de suas relações que consistem em uma coleção de tabelas, cada uma designada por um nome único.

A grande quantidade de dados nos bancos de dados informatizados pode produzir conhecimentos valiosos e úteis para a tomada de decisão. A utilidade desses dados, no entanto, irá depender da forma com que eles estão armazenados e organizados e do modo como serão acessados. Esse é, pois, o objetivo da modelagem de dados: permitir o entendimento do problema em estudo para que os dados relativos a esse problema possam ser armazenados e organizados seguramente no banco de dados por meio dos SGBD, de modo que possibilitem o acesso à informação útil na tomada de decisões.

A geração das informações geradas pelos bancos de dados necessárias para uma organização pode ser feita por sistemas transacionais ou por sistemas analíticos. Os sistemas transacionais são aqueles que se caracterizam pela alta taxa de atualização nos bancos de dados, pelo grande volume de dados e acessos pontuais e por pesquisas cujos resultados sejam de pequeno volume, até milhares de consultas. Esses sistemas são também conhecidos como sintéticos ou ainda OLTP – *Online Transactional Processing* -, pois, como o nome sugere, baseiam-se em transações. Esses sistemas são utilizados no dia a dia da organização e desenvolvidos com base nos modelos de negócio da organização, de acordo com Gardner (1998, apud CALAZANS, 2003).

Os sistemas analíticos, ou *Online Analytical Processing* (OLAP) caracterizam-se por fornecer subsídio para tomadas de decisão a partir de análises realizadas sobre bases de dados históricas com grande volume de dados, por vezes com milhões de registros a serem totalizados. São sistemas compostos de informações de vários departamentos ou unidades da empresa e que, após o cruzamento de informações, geram informações gerenciais para análises de negócios, de forma a garantir uma vantagem competitiva (CALAZANS, 2003). Os *Data Warehouse* são exemplos de sistemas analíticos.

3.3 Data Warehouse

De acordo com Moody e Kortnink (2000, apud CALAZANS, 2003), a tecnologia de *Data Warehouse* surgiu em meados da década de 1990 como uma solução para satisfazer à necessidade de informações gerenciais

da organização. À época já havia um aumento considerável nos sistemas de gestão empresarial, e, como consequência, a quantidade de dados também cresceu. Os bancos de dados tiveram, então, de evoluir para atender a esse crescimento tecnológico e a toda atmosfera de gestão informatizada. Para essa nova realidade, era preciso trabalhar em um contexto de dados distintos para uni-los externamente. Foi, então, que um projeto acadêmico conceitual da década de 1980 possibilitou, no início dos anos 1990, os primeiros testes com *Data Warehouse* – do-ravante DW.

O DW deve ser entendido como uma estratégia de organização de dados voltada para a gestão estratégica das corporações, e não para o controle do negócio. Nesse ponto faz-se importante diferenciar *Data Warehouse* de bancos de dados. Os bancos de dados possuem operações que são transações repetitivas, estruturadas, isoladas, detalhadas, com atualização ou leitura de dados, e os registros são acessados normalmente por chaves primárias (CHAUDHURI; DAYAL, 1997, apud CALAZANS, 2003). Os DW possuem dados sumarizados, agregados e consolidados que são armazenados por longos períodos de tempo (CALAZANS, 2003, p.7). DW é uma forma de gerir volumes muito grande de dados que se encontram, geralmente, espalhados em diversos sistemas de uma organização. Ele possibilita a análise de grandes volumes de dados coletados dos sistemas transacionais (OLTP). Em outras palavras, os DW das corporações são construídos a partir dos diferentes bancos de dados gerenciais de uma organização.

O conceito de DW consiste em organizar os dados corporativos de maneira integrada, com uma única versão da verdade, em um histórico variável no tempo, gerando uma única fonte de dados. Essa estrutura possibilita à empresa identificar tendências em sua corporação para posicionar-se estrategicamente no mercado tornando-se mais competitiva e, conseqüentemente, aumentando os seus lucros.

De acordo com Machado (2000), o DW é um sistema de computação utilizado para armazenar informações relativas às atividades de uma organização em bancos de dados, de forma consolidada. São as chamadas séries históricas que possibilitam uma melhor análise de eventos passados, oferecendo suporte às tomadas de decisões presentes e à previsão de eventos futuros.

O DW não é um produto que possa ser comprado como um software de banco de dados; nem aprendido

ou codificado como uma linguagem; nem é somente um modelo de banco de dados ou a constituição de vários modelos. O sistema de *data warehouse* deve ser pensado como um processo que está sempre em crescimento para disponibilizar informações que apoiem as decisões estratégicas da organização.

Os DW proporcionam a integração sólida e concisa dos dados para a realização de análises gerenciais estratégicas. É uma arquitetura que possibilita a organização dos dados, como os de clientes e fornecedores e os de operações, de modo que eles possibilitem a obtenção de informações estratégicas a partir da análise de dados sob diversas perspectivas, e da capacidade de navegação no nível de detalhe da informação. O DW organiza informações de fontes internas da organização, fontes que em grande parte dos casos são heterogêneas, e fontes externas filtrando, limpando esses dados e preparando-os para a análise e suporte à decisão (MACHADO, 2000).

Os DW são uma boa solução para organizações que têm várias plataformas de *hardware* e *software* e que sofrem com a falta de padronização e integração dos dados e com numerosas alterações nos sistemas transacionais. Com os DW, pode-se fornecer um modelo de dados comum para diferentes áreas de interesse, independentemente da fonte de dados. Dessa forma, torna-se mais fácil de reportar e analisar informações. As inconsistências das diferentes fontes de dados são resolvidas antes de serem carregadas as informações no DW, o que torna o processo de elaboração de relatórios e análise muito simples.

Uma das funcionalidades do DW é fazer com que as informações de uma empresa sejam facilmente acessadas. De acordo com Kimball (2002), as informações devem ser compreensíveis aos usuários com dados intuitivos e óbvios e, ao mesmo tempo, devem possuir diferentes formas de combinação de dados para a geração das informações. Além disso, as informações devem ser consistentes e com dados confiáveis. Ainda de acordo com Kimball (2002), há outras características do DW, como sua adaptabilidade e flexibilidade com relação a mudanças e um nível de segurança que proteja os dados.

Dentre as vantagens de construção de um DW, Machado (2000) cita:

- Disponibilidade de informações para a gestão das corporações;
- Geração de curvas de comportamento;
- Agilidade no uso de ferramentas de apoio à decisão;

- Segurança de informações para decisão;
- Maior abrangência na visão de indicadores.

Singh e Malhotra (2011) destacam também como vantagens da adoção dos DW o fato de que o armazenamento de dados revela-se útil na prestação de informações coletivas a todos os seus usuários e que a informação pode ser facilmente e armazenada de forma segura por um período de tempo mais longo.

Para que os DW sejam considerados um sucesso, a comunidade do negócio deve aceitá-lo. Os indicadores recolhidos a partir do DW devem atender às demandas da gestão estratégica empresarial. Essa é comumente a desvantagem mais destacada pelos profissionais usuários dos conceitos de DW: falta de apoio da alta gerência para o desenvolvimento do projeto de DW. Outra desvantagem costumeiramente descrita é a de que o DW deve ser cuidadosamente modelado para que de fato atendam às expectativas de seus patronos, fornecendo-lhes informações estratégicas, caso contrário eles tornam-se meras repetições dos sistemas transacionais já existentes na corporação.

Boateng et al. (2013) cita ainda como desvantagens o fato de que os armazéns de dados não são o ambiente ideal para dados não estruturados. Além disso, como no processo de *Data Warehousing*, os dados devem ser extraídos, transformados e carregados no armazém, há um elemento de latência de dados no DW. E durante sua vida, os DW podem ter custos elevados e ficarem obsoletos de forma relativamente rápida, o que pode causar grandes prejuízos.

Os DW são projetados para tamanhos de centenas de gigabytes até terabytes e podem armazenar grandes quantidades de informação, às vezes, divididas em unidades lógicas menores que são chamadas de *Data Marts* (DM).

Os *data marts* representam um subconjunto dos DM que permitem o acesso descentralizado à informação. Os DM podem ser direcionados a um departamento ou área específica do negócio (MACHADO, 2000). O esquema de dados mais utilizado na construção de um DM é o esquema estrela ("*Star Schema*"), também conhecido como Modelagem Multidimensional. (KIMBALL, 2002).

Data Warehousing é nome dado ao processo sistemático de construção de um sistema de DW. Esse processo é composto por uma coleção de tecnologias, algoritmos, ferramentas, técnicas e por uma arquitetura concebida para facilitar o armazenamento e o gerenciamento desses grandes volumes de dados e de várias origens, com o objetivo de proporcionar ao trabalhador do

conhecimento (executivos, gerentes e analistas) a visão do todo ou de parte do negócio. (CHAUDHURI; DAYAL, 1997; GARDNER, 1998 apud CALAZANS, 2003).

3.4 Características DW

A primeira característica de um DW é que ele é orientado por assunto. Outro aspecto importante é que os dados de um DW são precisos com relação ao tempo, isto é, são inseridos com posições históricas das atividades no tempo. E por isso não se modificam. Os dados de DW não podem, portanto, ser excluídos ou modificados, somente inseridos. Por isso, os DW são não voláteis e possuem apenas as operações básicas de carga de dados e de leitura de dados. Esse histórico de armazenamento dos dados tem como finalidade gerar indicadores que representem uma evolução ao longo de grande espaço de tempo. Por fim, pode-se afirmar que o mais importante aspecto do ambiente de DW é que seus dados são integrados. Essa integração mostra-se em diferentes maneiras: na convenção consistente de nomes, na forma consistente das variáveis, na estrutura consistente de códigos, nos atributos físicos consistente dos dados etc.

Os dados a serem inseridos em um DW passam por um processo de filtragem, limpeza e sumarização – o processo de *data warehousing*. Esse processo corresponde à retirada dos dados de origem dos sistemas transacionais e à concentração desses dados no *Operational Data Storage* (ODS) ou *Staging Area*, um local de armazenamento temporário dos dados (MACHADO, 2000). No ODS os dados originais serão ajustados para que ocorra a integração das informações. Esse é um processo que aperfeiçoa a criação do *data warehouse* e do *data mart* porque possibilita uma única fonte de dados sincronizada. Segundo Calazans (2010), os dados na *staging area* sofrem as seguintes transformações:

Integração: geração de chaves substitutas para cada registro, de modo a evitar a dependência de chaves definidas no sistema legado;

Limpeza: correção de códigos e caracteres especiais, resolvendo problemas de domínios, tratando dados perdidos e corrigindo valores duplicados ou errados;

Eliminação: de campos e dados provenientes dos sistemas legados que não serão úteis ao DW ou ao DM;

Combinação: unificação de dados que possuem os mesmos valores;

Verificação de integridade referencial: verificação da correspondência entre os dados de duas diferentes tabelas;

Desnormalização e renormalização: reunificação das hierarquias de dados, separadas pela normalização dentro de uma tabela desnormalizada;

Conversão de tipo de dados: transformação de dados de baixo nível em tipo de dado em outro formato;

Cálculos, derivação e alocação: transformações a serem aplicadas sobre as regras de negócio identificadas durante o processo de levantamento de requisitos;

Auditoria: no conteúdo dos dados.

Os meios utilizados para transformar e proceder à carga dos dados, para recuperar, analisar e extrair os dados e gerenciar dados de dicionário são considerados componentes essenciais de um DW. Essas operações são dependentes da forma como os dados são armazenados. Nos bancos de dados do DW, os dados são armazenados em cubos. Existem duas abordagens principais para o armazenamento de dados: a abordagem normalizada e a abordagem dimensional (MUHEET; QUADRI; ZAMAN, 2012).

Na abordagem normalizada, os dados no DW são armazenados seguindo, de certa forma, a regra de normalização Codd. As tabelas são agrupadas por áreas temáticas que refletem categorias de dados gerais. A principal vantagem dessa abordagem consiste na facilidade, que ela permite, ao adicionar informações na base de dados. E a desvantagem refere-se à dificuldade que os usuários possuem para unirem os dados provenientes de diferentes fontes de informação significativa.

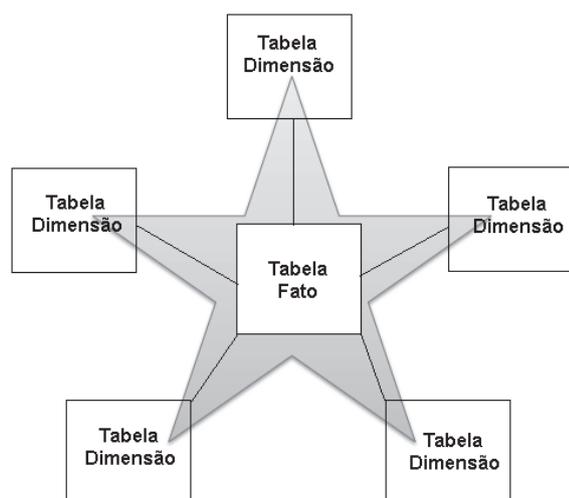
Na abordagem dimensional, os dados da transação são divididos em fatos e dimensões. Os fatos são os dados numéricos da transação e as dimensões são as informações de referência que contextualizam os fatos. Essa abordagem possibilita ao modelador do DW uma liberdade de organização dos dados e possibilita aos usuários o fácil entendimento do banco e a grande usabilidade. A recuperação dos dados do DW nessa abordagem também tende a operar muito rapidamente. Como desvantagens da abordagem pode-se citar: a complexidade no carregamento de dados a partir de diferentes sistemas operacionais e a dificuldade de modificar a estrutura de armazenamento de dados. Esse modelo dimensional pode ser

representado pelos esquemas estrela ou flocos de neve.

Os esquemas de um DW são uma coleção de objetos de banco de dados, incluindo tabelas, visões, índices e sinônimos. O esquema em estrela, ilustrado na figura 3, é, talvez, o esquema de armazenamento de dados mais simples. Ele é chamado de esquema em estrela porque o diagrama de entidade-relação desse esquema se assemelha a uma estrela, com pontos de irradiação de um ponto central. O centro da estrela consiste em uma grande tabela de fatos e as pontas da estrela são a dimensão. A estrela é uma associação entre uma tabela fato e uma série de tabelas de dimensão. Cada tabela de dimensão está associada à tabela de fatos usando uma chave primária para se juntar a uma chave estrangeira. As tabelas de dimensões não são unidas umas às outras. As principais vantagens dos esquemas em estrela são o fornecimento de um mapeamento direto e intuitivo entre as entidades de negócios, sendo analisados pelos usuários finais, o design do esquema e o desempenho altamente otimizado para consultas típicas de estrelas devido ao grande número de ferramentas de buscas já desenvolvidas. (MUHEET; QUADRI; ZAMAN, 2012).

Outro modelo de esquema é o floco de neve, um modelo mais complexo que esquema em estrela. Nesse esquema as dimensões visam eliminar a redundância, isto é, dados da dimensão são agrupados em várias tabelas. Esse esquema embora permita uma economia de espaço, aumenta o número de tabelas de dimensão e requer mais chaves estrangeiras. Como consequência, o resultado é consultas mais complexo e o desempenho é reduzido (MUHEET; QUADRI; ZAMAN, 2012).

Figura 3 - Esquema Estrela



Fonte: Baseado em Muheet, Quadri e Zaman (2012).

Independentemente de quão inteligente e produtivo seja o sistema de gerenciamento de informações planejado, construído e operado, ele é basicamente um repositório de dados. O valor a esses dados só é acrescido por meio das aplicações analíticas que possibilitam o acesso e o processo de apresentação dos dados, informações e conhecimentos para sustentar investigação e resolução de problemas necessidades. Esse é o processo de mineração de dados ou *data mining*. Sua função principal é a varredura de grande quantidade de dados a procura de padrões e detecção de relacionamentos entre informações gerando novos subgrupos de dados. *Data Mining* é como um agregador e organizador de dados. É gigantesco o resultado desse processamento como, por exemplo, a formação de hipóteses e principalmente regras de dados que respondem às questões do negócio.

3.5 Estudo de Casos Múltiplos: resultados do questionário

Foi realizada a aplicação de questionários a oito profissionais da área de TI que trabalhavam em quatro diferentes empresas. A aplicação dos questionários para a elaboração do estudo de casos múltiplos objetivava verificar se havia uma correspondência entre os conteúdos obtidos sobre DW com a revisão sistemática e o uso cotidiano desses conceitos pelos profissionais no mercado.

Dentre os participantes, cinco deles trabalham na área de tecnologia da informação há mais de vinte anos, dois participantes trabalham entre dez e vinte anos e um participante possui experiência na área de TI entre cinco e dez anos. Desses, três estão aposentados e os outros ainda desenvolvem atividades ligadas a DW. Quatro profissionais são especialistas na área de gestão da tecnologia e três possuem nível de escolaridade *stricto sensu*. Todos afirmam conhecer os conceitos de DW e classificam seu conhecimento sobre o assunto como ótimo ou bom. Os entrevistados declararam terem desenvolvido mais de dois projetos de *data warehouse* a partir do ano de 1997 e afirmaram em sua maioria que, durante o desenvolvimento dos projetos, tinham 100% do seu tempo envolvido com assuntos correlatos de DW.

Para averiguar a percepção que os entrevistados têm a respeito dos objetivos de implantação de um DW, perguntou-se: quais são, em sua opinião, os principais objetivos para a implantação de um projeto de *Data Warehouse*? Enumere em ordem de importância. As respostas dos participantes estão de acordo com o que foi

verificado com a revisão sistemática que a elaboração de um projeto de DW deve servir à gestão estratégica do negócio. A seguir, parte das falas dos entrevistados:

- Facilitar o acesso a informações gerenciais com a resolução de problemas [...]
- Resolver a necessidade de informações gerenciais para um problema específico [...]
- Permitir a execução de consultas gerenciais envolvendo milhares/milhões de registros com desempenho aceitável [...]

Ainda com objetivo de analisar a percepção dos profissionais da área de TI sobre DW, perguntou-se: quais os benefícios dos projetos de Data Warehouse? Os benefícios descritos com a implantação de um DW expressados pelos participantes também corroboram o resultado alcançado pela revisão sistemática:

- O benefício da cultura da informação dentro de uma corporação.
- Oferecer dados consistentes e confiáveis de suporte à decisão.
- “Racionalidade no processo. Economia no parque computacional.”
- Divulgação da informação dentro da corporação. Conhecimento informacional de maneira corporativa.
- Possibilidade de recuperar dados históricos

A média de tempo de desenvolvimento de um projeto de DW relatada pelos participantes foi de dois anos e as equipes de desenvolvimento em sua maioria eram compostas por mais de cinco pessoas.

Para alcançar o objetivo específico de verificação das vantagens e desvantagens da adoção de um DW perguntou-se aos participantes:

1) Você acha que a adoção do processo de *Data Warehouse*, pela sua empresa, foi vantajosa? Em caso afirmativo, em sua opinião quais foram os ganhos para a organização com a adoção do processo? Se foram vários projetos, por favor, discrimine as vantagens para cada um;

2) Os objetivos do projeto foram atingidos? Se foram vários projetos, por favor, discrimine cada um; e

3) Quais as principais dificuldades encontradas para a implantação do projeto de *Data Warehouse*? Como elas foram contornadas?

Quando questionados se acharam vantajosa a implantação do DW em sua empresa e se houve o atingimento dos objetos do projeto de DW, os participantes descreveram que de modo geral foi vantajoso à organização porque os DW possibilitavam a obtenção de estra-

tégias que tornavam a empresa mais competitiva. Essas descrições revelaram-se nas seguintes falas:

- Em minha opinião a utilização do BI, com aplicação adequada da informação, o retorno é sempre vantajoso. (sic)
- Sim foi muito vantajoso. A empresa necessitava urgentemente ter uma ferramenta ágil, como o DW, para entrar no mercado em igualdade de condições com as concorrentes.
- A empresa onde trabalham atualmente faz uso extremamente vantajoso das tecnologias de Data Warehouse.

Os participantes que relataram não ter sido vantajosa a adoção da arquitetura de *data warehouse* descreveram com principal impedimento a falta de preparo e maturidade da organização para a adoção dessa abordagem. Como destacado nas falas a seguir:

- O processo de data warehousing é muito importante para uma organização, entretanto nem todas as organizações estão preparadas para utilizá-lo.
- Não foi vantajoso, pois o mesmo nunca foi colocado em produção devido a problemas políticos dentro da empresa. (sic)

Os profissionais entrevistados descreveram que a falta de participação corporativa em torno da solução, a falta de integração dos dados que compõem o DW, as questões políticas, a falta de maturidade da organização, a gestão sem objetividade e com pouco conhecimento de suas necessidades foram as principais dificuldades enfrentadas no processo de implantação de um DW. Devido a isso, todos são unânimes em afirmar que já participaram de projetos que foram abandonados antes de sua conclusão devido a problemas em seu desenvolvimento.

Se incentivados a fazerem os comentários que julgassem pertinentes sobre DW, os participantes descreveram suas dificuldades no processo:

- Um comentário que julgo importante seria fazer com que os diretores tivessem maior conhecimento dos benefícios de um DW, mesmo sendo de difícil implantação, não pelos técnicos, mas quase sempre por fatores extras, como por exemplo a falta de um patrono. (sic)
- A falta de um planejamento adequado para os projetos de BI, geram problemas [...](sic)
- Penso que a maior dificuldade em um projeto de warehouse, limita-se aos seguintes pontos:
 - Não é um objetivo comum, quando deveria ser: boa parte da equipe de negócios não tem ideia do que necessita [...]
 - O Data Warehouse é uma ótima ferramenta para ser utilizada nas organizações, porém é

preciso que os usuários tenham conhecimento do que é o DW e para que serve, além de um conhecimento grande do negócio em que o DW foi construído. Isso é necessário para o sucesso do uso da ferramenta [...]

A avaliação dos questionários em comparação com os resultados da revisão sistemática descritos está em consonância, o que expressa há uma integração entre a literatura sobre DW e a atuação dos profissionais da área de tecnologia da informação no mercado. Os benefícios descritos, as vantagens alcançadas com implantação do DW e dificuldades enfrentadas no processo de implementação do DW relatados pelos participantes das entrevistas complementam a visão acadêmica alcançada pelo processo de revisão sistemática.

4 Considerações Finais

O objetivo desta pesquisa foi identificar os conceitos relacionados a *Data Warehouse* (DW) descrevendo as vantagens e desvantagens da adoção dessa arquitetura. Outro objetivo da pesquisa foi avaliar, junto a empresas que utilizam bancos de DW, a percepção que os profissionais da área de tecnologia da informação tinham sobre o assunto.

A metodologia de revisão sistemática, utilizada para pesquisa documental sobre definições e conceitos de modelagem de dados, banco de dados, sistemas gerenciadores de banco de dados (SGBD) e *Data Warehouse*, permitiu avaliar todas as pesquisas relevantes a respeito desses assuntos. A revisão sistemática é útil para identificar lacunas que podem direcionar novas pesquisas e novas atividades de investigação. É um modo de os pesquisadores resumirem toda a informação existente sobre determinado fenômeno de forma completa e imparcial. Por meio da pesquisa de revisão sistemática, foi possível averiguar a escassez de literatura que relacione de modo completo os conceitos de modelagem de dados, banco de dados e *Data Warehouse* abordados e relacionados nessa pesquisa.

Com intuito de verificar a aceitação dos conceitos de *Data Warehouse* pelos profissionais da área de tecnologia da informação, fez-se um estudo de casos múltiplos, por meio da aplicação de um questionário, que permitiu aos participantes expressar sua avaliação a respeito da elaboração de um *Data Warehouse*, das dificuldades encontradas nessa elaboração, os casos de sucesso, as vantagens, desvantagens etc.

Os resultados colhidos da metodologia de revisão

sistemática sobre os conceitos de DW foram complementados pelo estudo de casos múltiplos realizado com os profissionais de TI. Foi possível verificar que os profissionais confirmam a ideia de que um DW consiste em organizar os dados corporativos de maneira integrada, com uma única versão da verdade, em um histórico variável no tempo, gerando uma única fonte de dados. Em suas descrições, os participantes da pesquisa afirmaram que, entre as principais vantagens da adoção de DW, está a possibilidade de a empresa identificar tendências em sua corporação para posicionar-se estrategicamente no mercado tornando-se mais competitiva e, conseqüentemente, aumentando os seus lucros.

Os resultados mostram ainda que a aquisição da arquitetura de DW é vantajosa e, por essa razão, de ampla divulgação no mercado. Esse sucesso se deve ao fato de o DW possibilitar às empresas a obtenção de formação estratégica, o que os bancos de dados também faziam, mas de forma ineficiente.

Referências

- ADHIKARI, A.; JAIN L. C.; RAMANNA, S. Analysing effect of database grouping on multi-database mining. *IEEE Intelligent Informatics Bulletin*, Hong Kong, v. 12, n. 1, p. 25-32, dez. 2011.
- AKINTOLA, K.G.; ADETUNMBI, A. O.; ADEOLA, O. S. Building data warehousing and data mining from course management systems: a case study of FUTA course management information systems. *International Journal of Database Theory and Application*, Australia, v. 4, n. 3, p. 13-24, set. 2011.
- BOATENG, O.; GREESHMA, J. S.; GREESHMA, P. S. Data warehousing. *Business Intelligence Journal*, Renton, WA, v. 5, n. 2, p. 224-234, jul. 2012.
- BUTT, M. A.; QUADRI, S. M. K.; ZAMAN, M. Data warehouse implementation of examination databases. *International Journal of Computer Applications*, New York, USA, v. 44, n.5, p. 18-23, abr. 2012.
- CALAZANS, A. T. S. *Medição de Tamanho para Sistemas de Data Mart*. 2003.138 f. Dissertação (Mestrado em Gestão do Conhecimento e Tecnologia da Informação) – Universidade Católica de Brasília. Brasília: UCB, 2003.
- CALAZANS, A. *Estudo de caso: uma estratégia de pesquisa in Métodos para a pesquisa em Ciência da Informação*. Brasília: Thesaurus, 2007
- CHEN, P. *Modelagem de dados: a abordagem entidade relacionamento para projeto lógico*. São Paulo: Makron Books, McGraw-Hill, 1990.

- DHIMAN, A. *Importance of metadata in data warehousing*. 2012. 52 f. Dissertação (Mestrado em Ciência da Computação) – San Diego State University, San Diego, 2012.
- DUMITRITA, B. M. *Business intelligence*. Disponível em: <<http://www.rebe.rau.ro/RePEc/rau/jisomg/SP11/JISOM-SP11-A20.pdf>>. Acesso em: 26 jul. 2013.
- ELMASRI R., NAVATHE, S. *Sistemas de banco de dados*. São Paulo: Pearson Addison Wesley, 2005.
- GOOGLE ACADÊMICO. *Sobre o Google Acadêmico*. Disponível em: <<http://scholar.google.com.br/intl/pt-BR/scholar/about.html>>. Acesso em: 19 ago. 2013.
- HALPIN, T. *Structural aspects of data modeling languages*. Malaysia: LogicBlox, Australia, INTI International University, 2011.
- HEUSER, C. A. *Projeto de banco de dados*. 5. ed. Porto Alegre: Sagra Luzzato, 2004.
- KIMBALL, R. and ROSS, M. *The Data warehouse toolkit: guia completo para modelagem dimensional*. 4. ed. Rio de Janeiro: Campus, 2002.
- KITCHENHAM, B. *Procedures for performing systematic reviews*. United Kingdom: Keele University Technical Report, 2004.
- MACHADO, F. *Projeto de data warehouse: uma visão multidimensional*. São Paulo: Erica, 2000.
- MACHADO, F. *Banco de dados projeto de implementação*. São Paulo: Érica, 2004.
- MARWA, S. F.; MOHAMED E. M.; LAILA, M. Transforming conceptual model into logical model for temporal data warehouse security: a case study. *International Journal of Advanced Computer Science and Applications*, England, v. 3, n. 3, p. 115-122, 2012.
- PERIÓDICOS CAPES. *Institucional*. 2014. Disponível em: <http://www.periodicos.capes.gov.br/index.php?option=com_painstitucional&Itemid=103>. Acesso em: 19 ago. 2013.
- OLIVEIRA, M. *Sumário*. 2009. Disponível em: <http://www.datawarehouse.inf.br/Academicos/A%20PUBLICAR_DATA_WAREHOUSE_MARCELL_OLIVEIRA.pdf>. Acesso em: 10 jan. 2013.
- RAHMAN, N.; MARZ, J.; AKHTER, S. An ETL metadata model for data warehousing. *Journal of Computing and Information Technology*, Croácia, v. 20, n. 2, p. 95-111, Trimestral 2012.
- SANTOS, J. G. dos. *Conceitos de banco de dados*. 2011. Disponível em: <<http://www.pusivus.com.br/apostilas-doc/ConceitosBancoDados.pdf>>. Acesso em: 10 ago. 2013.
- SANTOS, I. S. et al. *Testes de aplicações móveis: uma análise das pesquisas científicas via revisão sistemática*. Ceará: Universidade Federal do Ceará, 2013.
- SCIENCE PUBLICATION. *Home*. Disponível em: <<http://thescipub.com/journals/jcs>>. Acesso em: 19 ago. 2013.
- SCIENTIFIC ELECTRONIC LIBRARY ONLINE (SIELO). *Serial browsing*. Disponível em: <<http://www.scielo.br/#about>>. Acesso em: 19 ago. 2013.
- SHARMA, M.; DORA, S. Efficient Approach for Compression in Data Warehouse. *Internacional Journal of Computing and Information Technology*, USA, v. 53, n. 9, p. 1-4, set. 2012.
- SILVA, L. S. S. Metodologia de Desenvolvimento de Sistemas: Análise Essencial, Estruturada e Orientada a Objetos. *Revista Olhar Científico, Faculdades Associadas de Ariquemes*, Ariquemes, RO, v. 1, n. 2, p. 257-266, ago./dez. 2010.
- SINGH, S.; MALHOTRA, S. Data Warehouse and its methods. *Journal of Global Research in Computer Science*, India, v. 2, n. 5, p. 113-115, maio 2011.