

**UNIVERSIDADE DE TAUBATÉ**  
**Ramon Modesto Pacheco**

**GUIA DE BOAS PRÁTICAS PARA  
IMPLANTAÇÃO DE DATA WAREHOUSE**

**Taubaté**  
**2019**

**Ramon Modesto Pacheco**

**GUIA DE BOAS PRÁTICAS PARA  
IMPLANTAÇÃO DE DATA WAREHOUSE**

Trabalho de Monografia para obtenção do Certificado de Especialização pelo Curso de Gestão de Projeto em Business Intelligence do Departamento de Informática da Universidade de Taubaté.

Orientador: Prof. Me. Antônio Ricardo Mendrot

**Taubaté - SP  
2019**

**SIBi – Sistema Integrado de Bibliotecas / UNITAU**

P116g Pacheco, Ramon Modesto  
Guia de boas práticas para implantação de Data Warehouse / Ramon Modesto Pacheco. -- 2019.  
64 f. : il.

Monografia (Especialização) – Universidade de Taubaté, Departamento de Informática, Matemática e Física, 2020.

Orientação: Prof. Me. Antônio Ricardo Mendrot, Departamento de Informática.

1. Data Warehouse. 2. Boas Práticas. 3. Business Intelligence. I. Título. II. Pós-graduação em Gestão de Projeto em Business Intelligence.

CDD – 005.7

**RAMON MODESTO PACHECO**

**GUIA DE BOAS PRÁTICAS PARA  
IMPLANTAÇÃO DE DATA WAREHOUSE**

Monografia apresentada para a obtenção do Certificado de Especialização pelo Curso de Gestão de Projetos em Business Intelligence do Departamento de Informática da Universidade de Taubaté.

**Data:** \_\_\_\_/\_\_\_\_/\_\_\_\_

**Resultado:** \_\_\_\_\_

**BANCA EXAMINADORA**

Prof. <> \_\_\_\_\_

Assinatura \_\_\_\_\_

Prof. <> \_\_\_\_\_

Assinatura \_\_\_\_\_

Prof. <> \_\_\_\_\_

Assinatura \_\_\_\_\_

A minha família pelo apoio e compreensão.

## **AGRADECIMENTOS**

A todos que de forma direta e indireta contribuíram para o sucesso dessa monografia.

“O temor do Senhor é o princípio da sabedoria, e o conhecimento do Santo a prudência”

Provérbios 9:10

## RESUMO

*Business Intelligence* tem se mostrado cada vez mais relevante diante do cenário altamente competitivo nos dias hodiernos, onde o volume de dados gerado cresce rapidamente e a necessidade de extrair informações dos mesmos aumenta conforme as organizações competem pela primazia do mercado de trabalho. Dentre as muitas tecnologias utilizadas no processo de BI tem-se o *Data Warehouse (DW)* que consiste em um banco de dados onde o processo histórico dos dados da organização será armazenado, é nele, e nas suas vertentes (*Data Marts*) que serão realizadas todas as consultas e minerações com o intuito de fornecer informações que darão suporte à tomada de decisão pelo nível gerencial. Já que Tecnologia da informação (TI) é uma área bastante maleável um *DW* pode ser implementado de formas diferentes acarretando em problemas como gargalos e intermitências, frustrando os usuários finais e no pior cenário, levando a empresa a ter prejuízo financeiro. A presente monografia se dispõe a apresentar um checklist de boas práticas a serem utilizadas durante o processo de implementação de um *Data Warehouse* com a finalidade de proporcionar qualidade ao processo, mitigar riscos ocasionados pelo esquecimento e uma implementação sólida e consistente. O método utilizado caracteriza-se por uma pesquisa bibliográfica e documental. Com a pesquisa realizada adquiriu-se informações valiosas tornando possível a criação do checklist que poderá ser utilizado durante o processo de implementação do *Data Warehouse*. Após a criação do *checklist* foi possível ter uma visão panorâmica sobre as fases de criação de um *DW*, também foi reafirmada a importância que *BI* tem no cenário organizacional e que o seu uso não deve ser negligenciado pelos que almejam estar alinhados com o mundo competitivo atual. A pesquisa verifica que um guia de boas práticas para a implementação do *Data Warehouse*, se seguido, resultará em uma implementação sólida e consistente gerando um sistema mais robusto e estável.

**Palavras chave:** Data Warehouse. Boas Práticas. Business Intelligence.

## ABSTRACT

Business Intelligence has proven increasingly relevant in today's highly competitive scenario, where the volume of data generated grows rapidly and the need to extract information from it grows as organizations compete for the primacy of the labor market. Among the many technologies used in the BI process we have the Data Warehouse (DW) which consists of a database where the historical process of the organization's data will be stored, it is in it, and in it is strands (Data Marts) where will be performed all queries and data mining to provide information that will support decision making at the managerial level. Since Information Technology (IT) is a very malleable area, a DW can be implemented in different ways leading to problems such as bottlenecks and intermittences, frustrating end users and in the worst-case scenario, leading to financial loss. This paper presents a checklist of good practices to be used during the process of implementing a Data Warehouse in order to provide quality to the process, mitigate risks caused by forgetting and a solid and consistent implementation. The method used is characterized by a bibliographic and documentary research. With this research, valuable information was acquired making it possible to create the checklist that can be used during the Data Warehouse implementation process. After the creation of the checklist it was possible to have a panoramic view on the phases of creating a DW, it was also reaffirmed the importance that BI has in the organizational scenario and that its use should not be neglected by those who wish to be aligned with the current competitive world. The Research finds that a good practice guide for Data Warehouse implementation, if followed, will result in a solid and consistent implementation generating a more robust and stable system.

**Keywords:** Data warehouse. Good Practices. Business Intelligence.

## LISTA DE FIGURAS

Figura 1 - Visão geral dos passos do processo de KDD.....	15
Figura 2 – Planejamento – Início.....	26
Figura 3 – Planejamento – Infraestrutura.....	29
Figura 4 – Planejamento – Backup e recuperação de desastres.....	31
Figura 5 – Desenvolvimento – Início.....	33
Figura 6 – Desenvolvimento – ETL.....	36
Figura 7 – Desenvolvimento – Testes.....	38
Figura 8 – Desenvolvimento – Visualização.....	40
Figura 9 – Implementação.....	42
Figura 10 – Pós-Implementação.....	44

## LISTA DE SIGLAS

BI	<i>Business Intelligence</i>
DW	<i>Data warehouse</i>
KDD	<i>Knowledge Discovery in Database</i>
OLAP	<i>Online Analytical Processing</i>
OLTP	<i>Online Transaction Processing</i>
TI	Tecnologia de Informação
DBMS	<i>Database Management System</i>
PMI	<i>Project Management Institute</i>
ISO	<i>International Organization for Standardization</i>
ETL	<i>Extract Transform and Load</i>
TCC	Trabalho de Conclusão de Curso
PDCA	Plan, Do, Check and Act

## SUMÁRIO

1 INTRODUÇÃO	11
1.1 OBJETIVO GERAL	12
1.2 OBJETIVOS ESPECÍFICOS	12
1.3 RELEVÂNCIA	12
1.4 ESTRUTURA DO TRABALHO	12
2 REVISÃO DE LITERATURA	14
2.1 BUSINESS INTELLIGENCE	14
2.2 OLTP	15
2.3 KDD	15
2.3.1 ETAPAS DO KDD	16
2.3.1.1 SELEÇÃO	16
2.3.1.2 LIMPEZA DOS DADOS	17
2.3.1.4 REDUÇÃO DOS DADOS	18
2.3.1.5 MINERAÇÃO DOS DADOS	18
2.3.1.6 INTERPRETAÇÃO	19
2.3.1.7 USANDO O CONHECIMENTO DESCOBERTO	19
2.4 DATA WAREHOUSE	20
2.4.1 DATA MART	21
2.5 OLAP	21
2.6 BOAS PRÁTICAS	22
3 METODOLOGIA	24
4 RESULTADOS E DISCUSSÃO	25
4.1 PLANEJAMENTO – INÍCIO	26
4.2 PLANEJAMENTO – INFRAESTRUTURA	29
4.3 PLANEJAMENTO – BACK E RECUPERAÇÃO DE DESASTRES	31
4.4 DESENVOLVIMENTO – INÍCIO	33
4.5 DESENVOLVIMENTO – ETL	36
4.6 DESENVOLVIMENTO – TESTES	38
4.7 DESENVOLVIMENTO – VISUALIZAÇÃO	40
4.8 IMPLEMENTAÇÃO	42
4.4 PÓS-IMPLEMENTAÇÃO	44
5 CONCLUSÃO	46
REFERÊNCIAS	47
APÊNDICE A – <i>CHECKLIST</i> DA FASE DE PLANEJAMENTO.	50
APÊNDICE B – <i>CHECKLIST</i> DA INFRAESTRUTURA.	53
APÊNDICE C – <i>CHECKLIST</i> DO <i>BACKUP</i> E RECUPERAÇÃO DE DESASTRES	55
APÊNDICE D – <i>CHECKLIST</i> DO DESENVOLVIMENTO	57
APÊNDICE E – <i>CHECKLIST</i> DO ETL	58
APÊNDICE F – <i>CHECKLIST</i> DOS TESTES	59
APÊNDICE G – <i>CHECKLIST</i> DA VISUALIZAÇÃO	61
APÊNDICE H – <i>CHECKLIST</i> DA IMPLEMENTAÇÃO	62
APÊNDICE I – <i>CHECKLIST</i> DA PÓS-IMPLEMENTAÇÃO	63

# 1 INTRODUÇÃO

*Business Intelligence* (BI) se tornou indispensável para as empresas que querem continuar competitivas e ofertando um serviço a cada dia mais alinhado com as demandas dos seus clientes.

Através da implementação do *Business Intelligence* é garantido a alta gerência que as decisões de negócio sejam tomadas consubstanciadas em informações que muitas das vezes passariam despercebidas ao olhar do profissional mais dedicado e sensitivo, garantindo assim que a sinergia com os clientes seja mantida e os concorrentes não estejam a frente no tocante ao mercado.

A implementação de um ambiente de BI abarca o uso de diversas ferramentas e tecnologias, segundo Turban (2009, p. 27) "BI é um termo guarda-chuva que inclui arquiteturas, ferramentas, banco de dados, aplicações e metodologias", dentre as tecnologias citados por Turban têm-se o *Data Warehouse* (DW) que configura uma das partes mais essenciais de todo ambiente, porque é nele, e nas suas vertentes (*Data Marts*) que serão realizadas todas as consultas e minerações que fornecerão suporte à tomada de decisão, Turban (2009, p. 27) informa que "o processo de BI baseia-se na transformação de dados em informações, depois em decisões e finalmente em ações".

Um projeto de BI bem-sucedido é aquele que consegue fornecer em tempo hábil as informações necessárias para as consultas de seus usuários, em geral essas consultas retornam dados que serão usados para plotar gráficos e criação de relatórios, quanto mais veloz a resposta da consulta for tão cedo serão obtidos indicadores que ajudarão na tomada de decisões.

Um projeto de *BI* mal implementado, em especial o *DW*, invariavelmente é o resultado da falta do uso de boas práticas durante o processo, gerando problemas como um ecossistema com gargalos e intermitências, frustrando os usuários finais e no pior cenário levando a empresa a ter prejuízo financeiro.

Já que Tecnologia da informação (TI) é uma área bastante maleável um *DW* pode ser implementado de formas diferentes acarretando nos problemas supracitados, as boas práticas, expressas em padrões e métodos, são importantes para que algo seja implementado da melhor maneira possível, garantindo assim a

qualidade uma ou n vezes; mitigando o risco de algo ser esquecido durante o processo.

### **1.1 Objetivo Geral**

Desenvolver, com base no estado da arte da literatura e estudos de caso de implantação de B.I um guia de boas práticas para implementação de *Data Warehouses* garantindo a qualidade do processo, mitigação dos riscos ocasionados pelo esquecimento e uma implementação sólida e consistente.

### **1.2 Objetivos Específicos**

- Levantar o estado da arte da literatura sobre B.I e, especificamente, *Data Warehouses*;
- Analisar estudos de caso de implementação de soluções B.I disponíveis na base acadêmica.
- Elaborar com base na literatura consultada um guia de boas práticas para implementação de *Data Warehouses*.

### **1.3 Relevância**

Um guia de boas práticas será de grande valia para a implementação do *Data Warehouse* garantindo qualidade ao processo e mitigando os riscos de ao final da implementação ter-se um sistema com gargalos e intermitências, proporcionando a organização ter um sistema de suporte a tomada de decisão eficiente. Soma-se a isso o fato de encontrar-se poucos trabalhos que abordam boas praticas na construção *do Data Warehouse*.

### **1.4 Estrutura do Trabalho**

O trabalho está estrutura em 5 capítulos descritos a seguir:

No capítulo 1 encontra-se a introdução, objetivo geral, objetivo específico e relevância.

No capítulo 2 encontra-se a revisão da literatura;

No capítulo 3 é apresentada a metodologia;

No capítulo 4 é apresentado os resultados e discussão;

No capítulo 5 cinco tem-se as considerações finais.

## 2 REVISÃO DE LITERATURA

A seguir apresenta-se os conceitos dos assuntos que balizaram esse trabalho, sendo eles sobre o *Business Intelligence*; *OLAP (Online Transaction Processing)*; *KDD (Knowledge Discovery in Database)* e todos o seus processos; *Data Warehouse* e *Data Marts*, *OLAP (Online Analytical Processing)* e Boas práticas.

### 2.1 Business Intelligence

Segundo a literatura, *Business Intelligence* é constituído pela realização de um conjunto de processos que irão transformar os dados em informações que por sua vez serão usados no processo de tomada de decisão.

“Os principais objetivos de *BI* são permitir o acesso interativo aos dados (às vezes, em tempo real), proporcionar a manipulação desses dados e fornecer aos gerentes e analistas de negócios a capacidade de realizar a análise adequada. Ao analisarem dados, situações e desempenhos históricos e atuais, os tomadores de decisão conseguem valiosos *insights* que podem servir como base para decisões melhores e mais informadas” (Turban, 2009, p. 27).

Segundo PRIMAK (2008) Na década de 80 o termo *Business Intelligence* foi cunhado pelo Gartner Group referindo-se ao processo inteligente de coleta, organização, análise, compartilhamento e monitoração de dados. Esses dados podem estar contidos em *Data Warehouses* ou *Data Marts*, sua análise gera informações que serão usadas para corroborar na tomada de decisão no ambiente de negócios.

Rainer e Cegielski (2011) diz que *BI* se trata de aplicações e tecnologias que tem o intuito de analisar e oferecer acesso a grandes quantidades de dados ajudando os usuários a tomarem melhores decisões empresariais e estratégicas. Aplicações de *BI* também oferecem visões históricas, atuais e previsíveis das operações de negócio.

PRIMAK (2008) mostra o fato interessante que no início da década de 90 a maioria das empresas de grande porte já faziam uso de Centro de Processamento de Dados (CPD) que embora mantivessem os dados proporcionavam pouquíssima disponibilidade de informações.

## 2.2 OLTP

OLTP "*Online Transaction Processing*" refere-se aos sistemas transacionais das organizações, esses sistemas são usados para gerenciamento dos dados transacionais gerados pela mesma, estão fortemente relacionados com as operações diárias e ao mundo relacional.

Segundo Microsoft (2018) "Dados transacionais são informações que acompanham as interações relacionadas às atividades de uma organização. Normalmente, essas interações são transações comerciais..."

Segundo Henrique (2012) o OLTP (*On-line Transaction Processing*) consiste em capturar as transações de negócio de um sistema e armazenar no banco de dados. Esse banco de dados é utilizado em sistemas que realizam pequenas transações (*INSERT, UPDATE, DELETE*) em tempo real, que ocorrem constantemente e de forma rápida.

Segundo Turban (2009, p. 36) "...lidam com os negócios rotineiros no andamento de uma empresa".

Turban (2009) ainda diz que o processo de OLTP oferece uma solução eficaz baseada em atividades repetitivas e de rotina para processamento de transações fazendo uso de um ambiente de banco de dados relacional distribuído com manipulação simples.

## 2.3 KDD

KDD é a abreviação para "*Knowledge Discovery in Database*", é um processo amplo que engloba desde a estruturação do banco de dados até *Data Mining*.

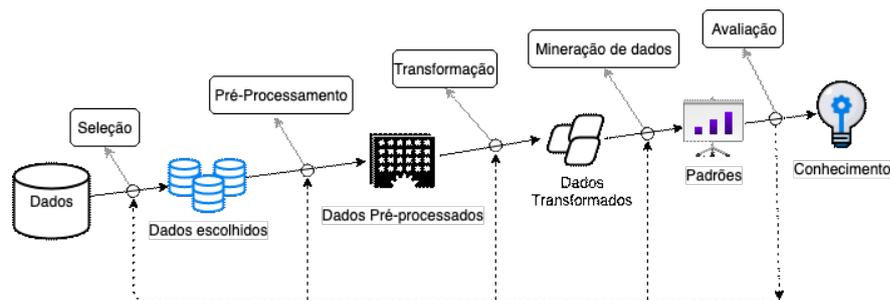


Figura 1 - Visão geral dos passos do processo de KDD

Fonte: Adaptada de Fayyad (1996)

Em cenários com grande volume de dados sendo gerados, como a Amazon por exemplo, uma abordagem de análise manual primeiro consistiria em gerar relatórios em arquivos impressos ou planilhas, depois esses relatórios seriam passados para um corpo de analistas que detém alto conhecimento do negócio, geralmente a gerência, então essa análise levaria um tempo considerável para ser completada; para daí então algum *insight* ser obtido. Essa abordagem se torna inviável levando em consideração o tempo gasto, o capital humano alocado para a mesma e que a quantidade de dados gerados tende sempre a crescer exigindo um crescimento exponencial no esforço de sua análise.

Ceci (2012) diz que o processo de KDD consiste fundamentalmente em: estruturação do banco de dados; na seleção, preparação e pré-processamento dos dados; bem como na transformação, adequação e redução da dimensionalidade dos dados; no processo de *Data Mining*; também nas análises, assimilações, interpretações e uso do conhecimento extraído do banco de dados através do uso do processo de *Data Mining*.

De acordo com Brachman e Anand (1994, tradução nossa) geralmente descobrir “conhecimento” em grandes quantidades de dados é atraente e intuitivo, porém no campo técnico é extremamente desafiador e difícil; O processo de KDD é considerado a extração não trivial de informações implícitas, previamente desconhecias e potencialmente úteis dos dados.

### **2.3.1 Etapas do KDD**

A seguir são apresentadas as etapas que compõem o processo de KDD.

#### **2.3.1.1 Seleção**

Na etapa de seleção, que é a primeira, tem-se maior interação humana onde os dados que serão analisados pelo processo de descoberta devem ser escolhidos, a origem desses dados pode ser de diversas fontes como planilhas e banco de dados.

Segundo Fayyad (1996, tradução nossa) essa primeira etapa se resume a "selecionando um conjunto de dados ou concentrando-se em um subconjunto de variáveis ou amostras de dados, dos quais a descoberta deve ser executada."<sup>1</sup>

### 2.3.1.2 Limpeza dos Dados

Como os dados, na maioria das vezes, são produzidos por pessoas; é possível que os mesmos possam conter algumas inconsistências como: informações faltantes causadas por má vontade ou esquecimento da pessoa que deveria inseri-las, erradas e duplicadas que levarão a resultados errados na análise.

É nessa etapa onde a limpeza dos dados é feita resultando em um *dataset* mais confiável para ser analisado aumentando a chances de uma descoberta verdadeira.

Segundo Fayyad (1996, tradução nossa) o processo de limpeza de dados inclui operações básicas a saber: remoção de ruído ou *outliers*, se apropriado, coleta de informações necessárias para modelar ou responsabilizar pelo ruído, decide estratégias para lidar com falta de dados e contabilizando informação de sequência de tempo e mudanças conhecidas, bem como resolver problemas de DBMS, como tipos de dados, esquemas e mapeamento de valores ausentes e desconhecidos.<sup>2</sup>

Brachman e Anand (1994, tradução nossa) compara a limpeza de dados a uma faca de dois gumes no tocante a dados que ocasionalmente poderiam ser o indicador mais crucial de um fenômeno de domínio serem removidos como se fossem *outliers* ou anomalias.<sup>3</sup>

---

<sup>1</sup> selecting a data set, or focusing on a subset of variables or data samples, on which discovery is to be performed.

<sup>2</sup> Data cleaning and preprocessing: includes basic operations, such as removing noise or outliers if appropriate, collecting the necessary information to model or account for noise, deciding on strategies for handling missing data fields, and accounting for time sequence information and known changes, as well as deciding DBMS issues, such as data types, schema, and mapping of missing and unknown values.

<sup>3</sup> Data cleaning is a double-edged sword. It is almost always necessary because of inevitably poor data quality, but occasionally what looks like an anomaly to be scrubbed away is in reality the most crucial indicator of an interesting domain phenomenon. In other words, what look like outliers to be dismissed can actually be the key data points worth focusing on.

#### 2.3.1.4 Redução dos Dados

Nessa etapa do processo a massa de dados, apesar de estar limpa pelo processo anterior, é reduzida para conter apenas variáveis que podem ajudar no objetivo da análise. Ao ter-se em uma fonte de dados que foi populada via *web scraping* invariavelmente haverá dados que não serão de serventia para o objetivo da análise, talvez esses mesmos dados possam ser úteis em outra ocasião, tudo depende do objetivo do momento.

Segundo Fayyad (1996, pg 30, tradução nossa) "Redução e projeção de dados: inclui encontrar recursos úteis para representar os dados, dependendo do objetivo da tarefa, e usar métodos de redução ou transformação de dimensionalidade para reduzir o número efetivo de variáveis[...]"<sup>4</sup>

#### 2.3.1.5 Mineração dos Dados

É nessa etapa onde são escolhidos e aplicados os algoritmos (e.g., classificação e aprendizado de máquina) a fim de analisar a base de dados, essa análise tem como objetivo encontrar padrões ou modelos que resultarão em um novo conhecimento ou confirmando o que é sabido de forma tácita, outra vez há uma maior participação nessa parte do processo; tanto na parte da escolha do algoritmo quando na avaliação do resultados.

Witten (2011, tradução nossa) diz que a mineração de dados se trata do processo de descoberta de padrões nos dados que mais comumente é feito de forma automática e que essa descoberta quando significativa promove alguma vantagem, na maioria das vezes econômica.<sup>5</sup>

Segundo Fayyad (1996, tradução nossa) a mineração de dados engloba a montagem de modelos ou a determinação de padrões oriundos da análise dos dados.

---

<sup>4</sup> Data reduction and projection: includes finding useful features to represent the data, depending on the goal of the task, and using dimensionality reduction or transformation methods to reduce the effective number of variables.

<sup>5</sup> Data mining is defined as the process of discovering patterns in data. The process must be automatic or (more usually) semiautomatic. The patterns discovered must be meaningful in that they lead to some advantage, usually an economic one. The data is invariably present in substantial quantities.

Esses modelos e padrões encontrados são subjetivos portanto necessitam de um julgamento onde um ser humano deve estar envolvido.<sup>6</sup>

Enquanto Fayyad cita modelo e padrão como sinônimos Kraker (2013) nos mostra uma diferenciação interessante: "...os modelos são como a equação geral de uma linha:

$$y = a + bx$$

enquanto os padrões são como uma equação específica:

$$\text{e.g. } y = 5 + 2x."$$
<sup>7</sup>

### 2.3.1.6 Interpretação

É a fase onde os padrões descobertos pela mineração são interpretados por uma pessoa e caso haja a necessidade de refinar algo nos processos anteriores, será feito; esse *loop* continuará até que o resultado pareça melhor se adequar aos critérios da análise.

Segundo Fayyad (1996, tradução nossa) consiste em interpretar os padrões descobertos e traduzi-los para termos mais compreensíveis pelos usuários, durante esse processo padrões redundantes e irrelevantes podem ser removidos.

### 2.3.1.7 Usando o Conhecimento Descoberto

Nesta última etapa do processo de KDD o conhecimento encontrado é usado em benefício do escopo em que a análise feita está inserida.

Tem-se a história, bem conhecida, que é a de que uma rede de varejo nos Estados Unidos descobriu que as vendas de fraldas descartáveis estavam associadas a venda de cerveja, os compradores eram homens que quando iam comprar fraldas a noite e aproveitavam para levar cerveja também. Após esse conhecimento ser adquirido a rede varejista resolveu colocar os dois produtos juntos e o resultado foi um crescimento significativo nas vendas.

---

<sup>6</sup> Data mining involves fitting models to or determining patterns from observed data. The fitted models play the role of inferred knowledge. Deciding whether or not the models reflect useful knowledge is a part of the overall interactive KDD process for which subjective human judgment is usually required.

<sup>7</sup> models are like the general equation of a line  $y = a + bx$ , while patterns are like a specific equation, e.g.  $y = 5 + 2x$ .

Fayyad (1996, tradução nossa) diz que o uso do conhecimento descoberto deve ser usado para aumento de desempenho, bem como tomar ações com base nesse novo conhecimento, também é possível simplesmente documentá-lo e informá-lo às partes interessadas. Essa nova descoberta pode ocasionar a atualização do que outrora era tido como certo.<sup>8</sup>

## 2.4 Data Warehouse

Trata-se de um repositório central e otimizado de dados históricos que serão usados para análises e criação de *dashboards* fornecendo suporte a tomada de decisão, esses dados podem vir de fontes como bancos de dados transacionais e planilhas.

Os dados salvos no *Data Warehouse* já passaram por um processo de limpeza e transformação na etapa de *ETL*, com isso as discrepâncias, duplicações e ausências nos dados foram eliminadas garantindo sua consistência.

Uma das características interessantes do *DW* é que ele possui um esquema desnormalizado enquanto os bancos de dados transacionais primam pela normalização, essa forma de ser permitirá com que as consultas retornem com maior rapidez; essas mesmas consultas em um banco de dados transacional teriam um desempenho aquém do aceitável em termos de tempo.

Turban (2009, pg. 57) diz que se trata de um conjunto de dados que é produzido para oferecer suporte a tomada de decisões; uma de suas características é guardar os dados históricos de possível interesse aos gerentes de toda a organização, esses dados são estruturados de forma otimizada para o processamento analítico dos mesmos.

Segundo Barbieri (2001, p. 49) "[...] definido como um banco de dados, destinado a sistemas de apoio à tomada de decisão e cujos dados foram armazenados em estruturas lógicas dimensionais, possibilitando o seu processamento analítico por ferramentas especiais [...]"

---

<sup>8</sup> Using discovered knowledge: includes incorporating this knowledge into the performance system, taking actions based on the knowledge, or simply documenting it and reporting it to interested parties, as well as checking for and resolving potential conflicts with previously believed (or extracted) knowledge.

Kimball (2004, pg. 22, tradução nossa) também nos diz que: "A missão do Data Warehouse é publicar os dados dos ativos da organização para dar suporte mais eficaz a tomada de decisões [...]"

Popular o DW é um processo demorado e custoso Kimball (2004, pg. 23, tradução nossa) relata que "o processo que envolve a transformação de dados de seu formato original a um repositório de dados dimensionais representa pelo menos 70% do tempo, esforço e despesas da maioria dos projetos de Data Warehouse."

#### 2.4.1 Data Mart

O *Data Mart* é parecido com *Data Warehouse* só que com menor abrangência, os dados armazenados no *Data Mart* representam uma área de negócio específica enquanto que o DW representa todo o negócio.

Talend (2019, tradução nossa) relata que *Data Marts* tem o objetivo de atender a uma divisão específica ou função de negócios, diferentemente dos *Data Warehouses* que tem um escopo maior quanto aos dados armazenados pois os mesmos são relacionados a todo o negocio <sup>9</sup>

#### 2.5 OLAP

*OLAP (Online Analytical Processing)* permite-se fazer análises complexas em grandes massas de dados, geralmente esses dados estão no *DW* e como estão à parte dos sistemas transacionais, pois estes não suportariam tais consultas, não produz instabilidade para os mesmos.

Microsoft (2018) relata que "...os bancos de dados *OLAP* são otimizados para cargas de trabalho com leitura intensa e com pouca gravação."

Turban (2009) diz que geralmente *OLAP* está relacionado as atividades normalmente executadas por usuários finais em sistemas *online*. Dentre essas

---

<sup>9</sup> Data marts and data warehouses are both highly structured repositories where data is stored and managed until it is needed. However, they differ in the scope of data stored: data warehouses are built to serve as the central store of data for the entire business, whereas a data mart fulfills the request of a specific division or business function.

atividades podem citar a geração de relatórios e gráficos *ad hoc*, análise estatística tradicionais ou modernas e construção de apresentações visuais.

Atualmente existem várias ferramentas *OLAP* com as quais realizar algumas consultas é facilitado inclusive para usuários com menos conhecimento técnico, já que a modelagem semântica (onde o nome já descreve o significado) das tabelas produz nomes mais intuitivos e fáceis de entender. A seguir algumas ferramentas *OLAP*:

- Cognos PowerPlay
- JasperReports
- Iccube
- Pentaho BI

Turban (2009) informa que as ferramentas relacionais *OLAP* permitem com que usuários de perfil avançado façam consultas investigativas no banco de dados em busca de qualquer resposta, é possível, realizar funções de *drill down*, que é ir a um nível mais detalhado dos dados.

## 2.6 Boas Práticas

Boas práticas consistem em ações realizadas que trarão um resultado positivo para quem as emprega. Essas ações são aceitas e validadas por outras pessoas quem também as praticam. Por exemplo, no tocante a higiene tem-se a boa prática, incentivada pelos profissionais da área da saúde e por quem já foi prejudicado por negligenciar essa ação, de lavar as mãos com maior frequência, essa boa pratica é incentivada porque as mãos sempre estão em contato com locais que podem estar contaminados; fazendo assim as mãos ficam higienizadas ajudando na prevenção de doenças como gripe por exemplo.

“Conjunto das técnicas, processos, procedimentos e atividades identificados, utilizados, comprovados e reconhecidos por diversas organizações, em determinada área do saber, como sendo os melhores quanto ao mérito, eficácia e sucesso alcançados pela sua aplicação na realização de uma tarefa” (Dicionário Infopédia, 2019).

Um outro exemplo de boa prática, agora em TI; principalmente na área de gestão da qualidade, é aplicação do ciclo *PDCA* (*Plan, Do, Check and Act*) que é uma metodologia para melhoria continua de processos. Através do uso desse método é

possível validar um processo e melhorá-lo, o ciclo é iniciado quantas vezes forem necessárias até que a qualidade desejada seja atingida. (Martins, 2018).

Segundo PMBOK (2008) as boas práticas estão relacionadas com a evolução do conhecimento. Esse é um ponto de vista interessante, ao pensar-se em como o conhecimento humano evolui tem-se que concordar que as boas práticas surgem da lapidação do que conhecido até o momento e esse conhecimento lapidado é passado para outras pessoas que o põem em prática melhorando o mesmo, se possível, e reiniciando o processo de passagem de conhecimento.

### 3 METODOLOGIA

Este trabalho utilizou-se das formas de pesquisa bibliográfica e documental.

Através da pesquisa realizada de estudo de casos foram aproveitados 16 materiais, estes variando entre artigos, tccs, monografia dissertações e teses. As principais áreas onde houve a implementação do DW foram instituições de ensino, instituições governamentais e saúde.

Os ambientes onde a pesquisa foi realizada foram, principalmente, as bases do Google Scholar<sup>10</sup> e SciELO<sup>11</sup> utilizando os critérios: “*estudo de caso ‘data warehouse’*” como texto base, com os filtros de data entre os anos 2000 e 2019. As outras fontes de pesquisas foram livros e sites sobre o assunto.

Visando tornar o processo metodológico mais organizado o mesmo foi dividido em três etapas subsequentes: levantamento de material de referência, análise do material levantado e proposição descritas abaixo:

Na primeira etapa deu-se a formação do arcabouço literário, tendo como alvo o *Data Warehouse* e tudo que o circunda, foram realizadas pesquisas sobre o mesmo com foco na sua implementação bem como suas características principais.

A segunda etapa caracterizou-se onde o material reunido na etapa anterior foi analisado de forma exaustiva com o objetivo de serem extraídos os pontos elencados a seguir:

- Levantamento de requisitos
- Criação dos modelos
- Segurança
- Implementação
- Boas práticas utilizadas antes, durante e depois da implementação
- Dificuldades enfrentadas antes durante e depois da implementação

Todos os pontos significantes para essa pesquisa foram registrados em um arquivo de log com o intuito de facilitar a compilação e cruzamento das ideias.

Na terceira e última etapa foi desenvolvido, baseado nas informações adquiridas na etapa que a antecede, um *checklist* com boas práticas para serem observadas no processo de implementação do um *Data Warehouse*.

---

<sup>10</sup> <https://scholar.google.com/>

<sup>11</sup> <https://www.scielo.org/>

## 4 RESULTADOS E DISCUSSÃO

Após a análise dos casos foram detectados os pontos em comum durante o processo de implementação do *Data Warehouse* que são: Planejamento, desenvolvimento, implementação e pós implementação.

Para facilitar o entendimento dos processos foram criados fluxogramas para cada ponto encontrado, a seguir são expostos esses fluxogramas seguidos da discussão:

## 4.1 Planejamento – Início

A implementação de um *DW*, mesmo se não encarado como, trata-se de um projeto, então nada mais correto do que termos a estrutura organizacional, ainda que simplificada, montada; Santos (2006) diz que definir uma estrutura com capacidade de coordenar e executará as tarefas é essencial.

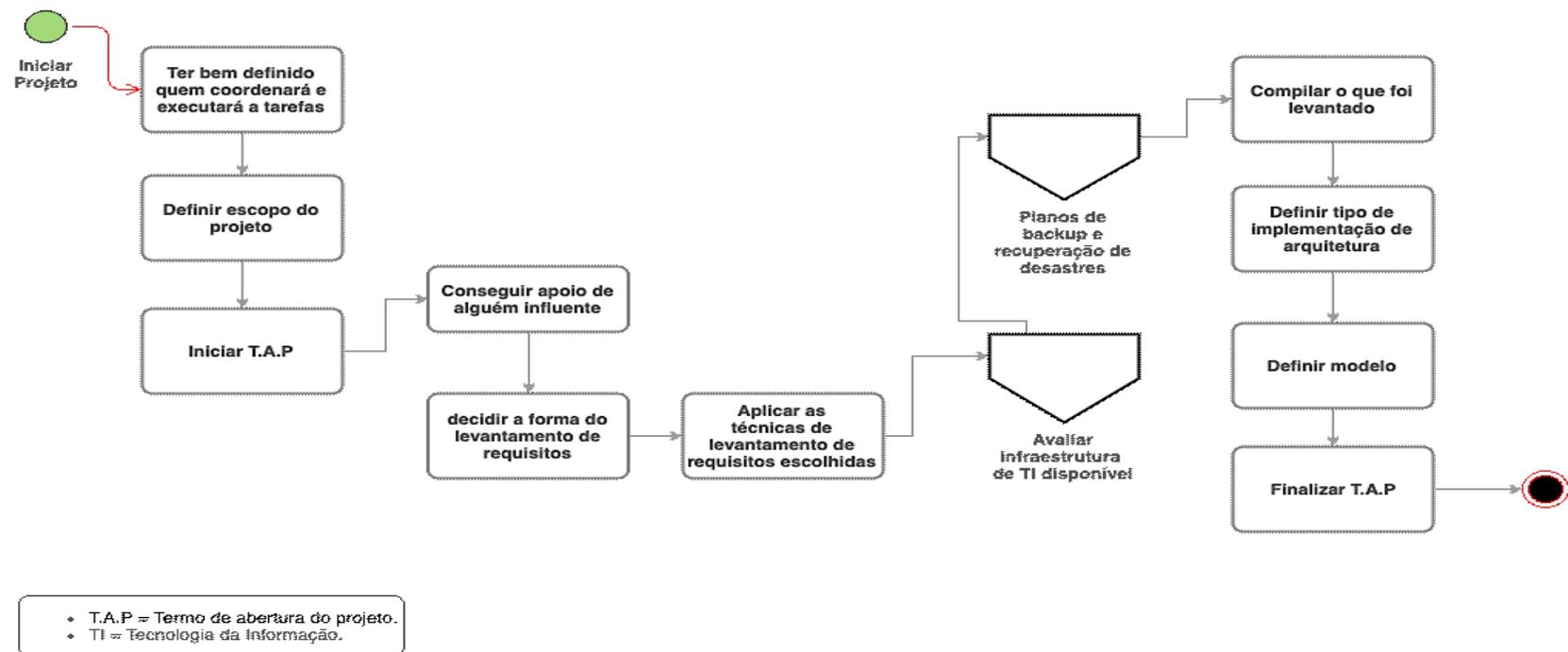


Figura 2 – Planejamento – Início.

Fonte: Própria autoria.

O escopo do projeto é fator inalienável, PMBOK (2008) diz que o escopo do projeto inclui os processos necessários para garantir que consta no projeto todo o trabalho necessário e somente ele.

O Termo de Abertura do Projeto abarca todas as necessidades do projeto bem como o escopo do projeto, ter essa fonte de documentação fornece segurança para medir o andamento das atividades e detectar se o projeto está se distanciando do que foi proposto; um empreendimento como a implementação de um *DW* que começa sem um norte se aproxima da linha do fracasso.

Clemes (2001) cita sobre a importância do apoio de pessoas influentes durante a implementação do projeto de *Data Warehouse* esse apoio será útil para contornar problemas e negociar prioridades. Pode-se ampliar o poder dessa influencia para quando o projeto for implementado, quando pessoas influentes aderem a novas tecnologias, gradativamente mais e mais usuários tendem a baixar as suas defesas e sentirão vontade de testar e aderir a solução, gerando assim um círculo virtuoso para a organização.

Os requisitos expressam as necessidades do que será implementado, são de extrema importância para o sucesso do projeto é através deles que será possível definir como o *DW* será construído e o que será possível extrair do mesmo. Almeida (2006) aponta que quando os requisitos errados são escolhidos tornam-se indicadores de possíveis fracassos, o que faz total sentido levando em consideração que com os requisitos errados, informações não tão relevantes serão produzidas das consultas incorrendo no abandono e descrédito do sistema. PMI (2008) salienta que os requisitos atenderão as expectativas das partes interessadas, então a escolha das pessoas e métodos que serão abordados deve ser feita com máxima atenção.

Há várias formas de levantamento de requisitos, PMI (2008) lista algumas delas que são:

- Entrevistas
- Discussão em grupo
- *Workshops*
- Técnicas de criação em grupo
- Técnicas de tomada de decisão em grupo
- Questionários
- Pesquisas observações

- Protótipos

Já Santos (2006) cita o método *Source-Driven* onde os requisitos serão identificados com base na observação dos dados fornecidos pelas fontes. As formas citadas pelo PMI são tradicionais e refletem a maioria dos levantamentos de requisitos conduzidos nos projetos. Essas opções ajudam a refinar e entender melhor as expectativas dos interessados produzindo uma visão mais apurada dos assuntos que devem e não devem ser levados em consideração.

A essa altura do projeto compile as informações obtidas para que fique mais fácil o seu manuseio, isso facilitará sua inserção e adequação ao TAP.

A definição da arquitetura da implementação deve ser baseada nos requisitos colhidos e até mesmo nas condições financeiras no momento. Almeida (2006) cita dois tipos mais famosos de arquitetura a *Top-Down defendida por Inmon* e a *Bottom-Up defendida por Kimball*.

A arquitetura *Top-down* consiste na criação de um DW central, ele será o centro analítico e também alimentará o *Data Marts*. Essa opção é a mais tanto financeiramente como na sua implementação.

A arquitetura *Bottom-Up* consiste na criação dos *Data Marts*, os quais atendem a departamentos específicos, eles podem ser implementados de forma gradativa, é a solução mais barata e a sua reunião dará origem a DW da organização.

O modelo trata de como as tabelas estarão relacionadas, isso tem relação direta no tempo em que as consultas levarão para retornarem os resultados, quanto mais níveis de junções (*joins*) existirem mais lenta a consulta será. Kimball (2013) relata que o modelo estrela corresponde a uma grande tabela central denominada “Tabela Fato” e pequena tabelas conectadas a mesma, denominadas de “Dimensões” formando o modelo dimensional. O modelo estrela é o mais utilizado.

Agora que todos os detalhes iniciais do projeto estão definidos o termo de abertura do projeto pode ser concluído e o prosseguimento para a fase de implementação deve ser iniciado.

## 4.2 Planejamento – Infraestrutura

Ainda na fase de planejamento tem-se a infraestrutura de TI que é parte fundamental para implantação do projeto de Data Warehouse.

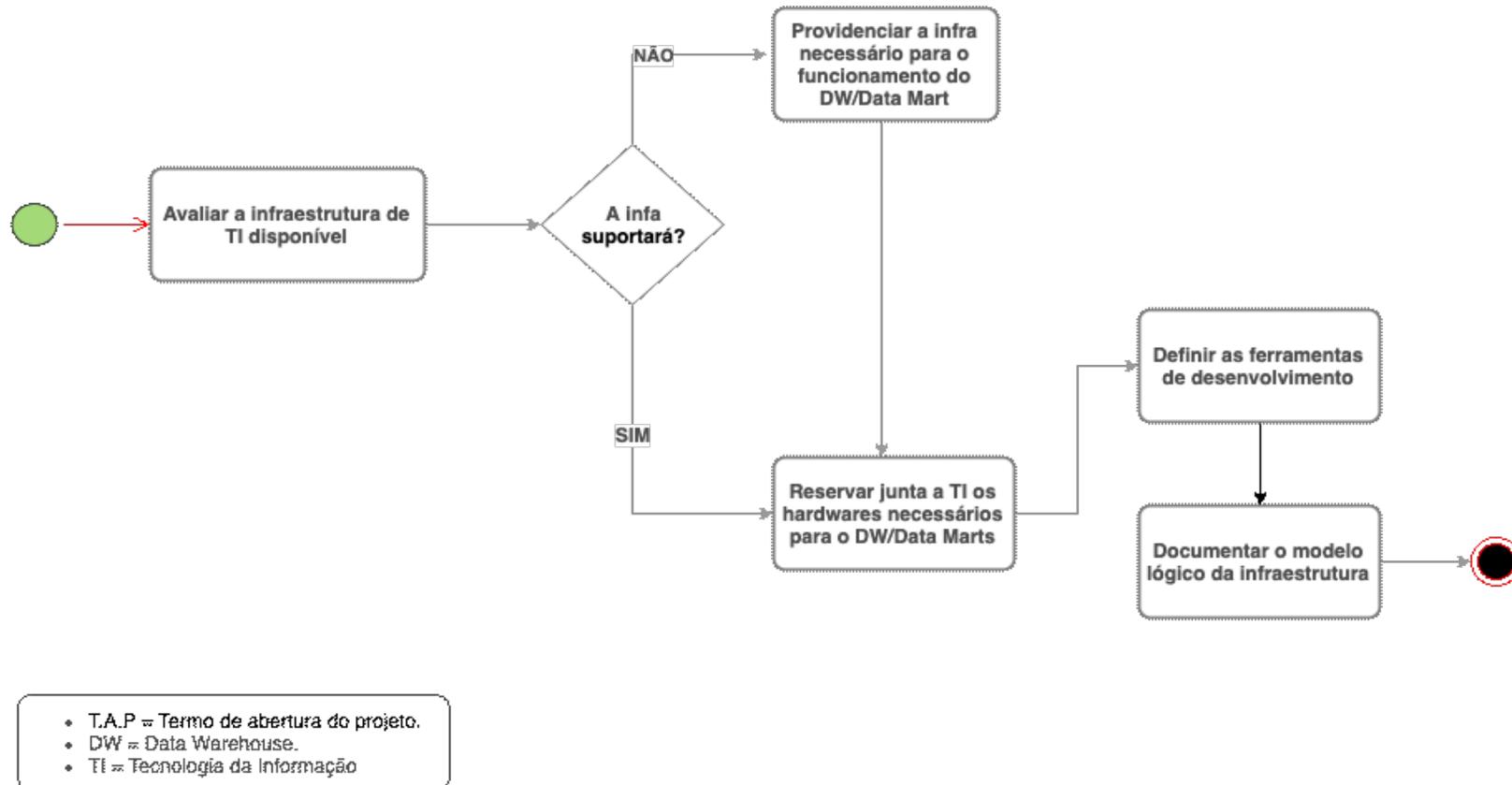


Figura 3 – Planejamento – Infraestrutura.

Fonte: Própria autoria.

Deve-se dispensar especial atenção para a infraestrutura de TI onde o *Data Warehouse* será implementado, essa atenção se redobra se a arquitetura escolhida for a *Top-down*, pois essa arquitetura abrange toda a organização, demandando mais hardware e demais recursos. IBM (apud Cledes 2001) informa que a escolha da implementação é influenciada pela infraestrutura disponível. Se não houver escalabilidade no escopo da infra atual da empresa o projeto sofrerá danos ou poderá se tornar inviável haja visto que quanto maior o volume de dados que serão processados maior se tornarão as exigências em termos de hardware.

Caso a infraestrutura atual não suporte a implementação do *DW* será preciso agir no sentido ou de adaptar o projeto para que se acomode ao atual cenário ou sugerir a compra dos ativos necessários para que o projeto se realize, nesse momento ter obtido o apoio de pessoas influentes será de grande serventia no tocante a liberação de verba para a compra. Singh (apud Cledes 2001) diz que a infraestrutura deve ser maleável para crescimento futuro. Essa é uma observação bastante pertinente já que o ambiente de TI é dinâmico, ter um projeto que não prevê possível expansão foge ao bom senso.

Caso a infraestrutura comporte a implementação do *DW* deve ser reservado junto a equipe de TI os recursos necessários para o projeto bem como a configuração dos ambientes, manter boas relações com o setor de Tecnologia da Informação pode garantir o bom andamento do projeto.

As ferramentas de desenvolvimento também devem ser escolhidas com cuidado, é comum encontrar profissionais desgostosos com certas escolhas de ferramentas que depois de algum tempo de uso revelaram-se inadequadas para as tarefas propostas, esse problema se agrava se as ferramentas forem pagas.

Pode-se categorizar as ferramentas, grosso modo, como não pagas e pagas, sem adentrarmos nas nuances de *free* e *open source*. Existem boas opções dos dois lados é preciso pesquisar e escolher a que se adequa melhor as necessidades do projeto.

Após a infraestrutura ser definida é importante que o modelo lógico seja documentado para facilitar as consultas e entendimento de todos os envolvidos.

### 4.3 Planejamento – Backup e Recuperação de Desastres

No planejamento também é verificado sobre as políticas de backup e recuperação de desastres.

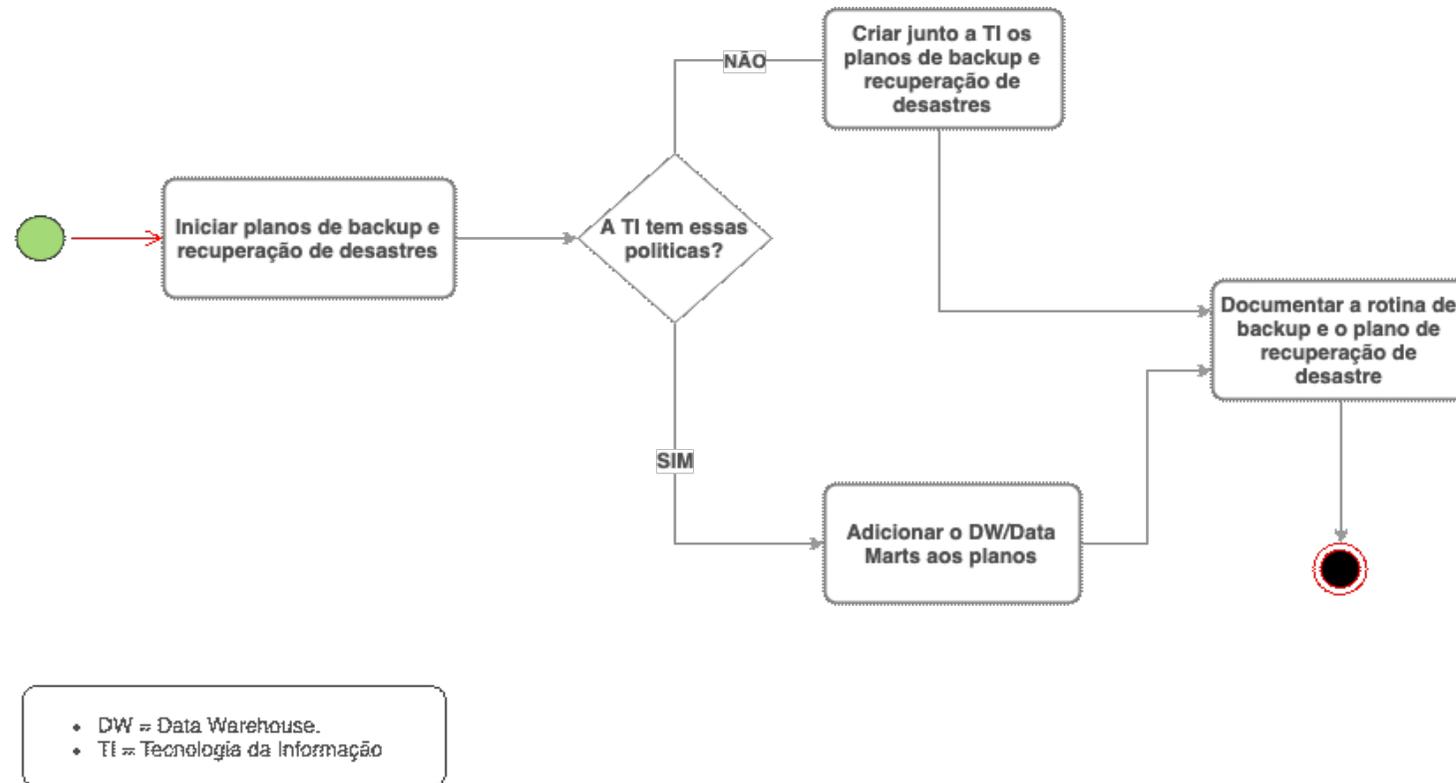


Figura 4 – Planejamento – Backup e recuperação de desastres.

Fonte: Própria autoria.

A necessidade de uma política de *backup* e recuperação de desastre é inerente a todo projeto que envolva dados sensíveis à organização e que possa trazer prejuízo financeiro a mesma, o *Data Warehouse* se enquadra nos dois pontos, os dados que o mesmo armazena são sensíveis a organização e impossibilidade da consulta ao *DW* para suporte a tomada de decisão pode levar a insucesso no que for decidido ou um sucesso menor no que poderia ter sido.

É preciso verificar se no departamento de informática já existem políticas voltadas para o *backup* e recuperação de desastre, caso não existam é extremamente importante que a criação de tais políticas seja realizada, o *DW* trabalha com dados históricos isso quer dizer que caso esses dados sejam perdidos por algum motivo e não existir *backup*, não será possível realizar quaisquer tipos de análise significativas apenas com os dados após o incidente, pelo menos por um bom tempo.

Meira (2003) observa que uns dos cuidados que se deve ter é com manutenção do *Data Warehouse* e que a manutenção envolve as rotinas de *backup* e restauração. Corroborando com que o Meira salienta é notável a falta de coerência entre o quão importante um *DW* é para a organização e a não adoção de uma política de continuidade do negócio para salvaguardar o uso a um ativo tão valioso.

Na maioria dos departamentos de TI já existem políticas voltadas aos tópicos de segurança da informação, *backup* e recuperação de desastre. Normalmente essas políticas são baseadas na norma ISO 27001 onde são especificadas boas práticas para implementá-las. Caso esse seja o cenário deve-se solicitar a adição do projeto nesse escopo.

Por fim é sempre bom angariar toda a documentação possível sobre essas políticas e planos para que todos possam estar a par do cenário geral do projeto.

#### 4.4 Desenvolvimento – Início

A fase de desenvolvimento do projeto segue após a fase de planejamento, nesse estágio um detalhamento mais aprofundado é feito no sentido de criar de fato os parâmetros, rotinas e o que mais for necessário para a fase a seguir que é a de implementação.

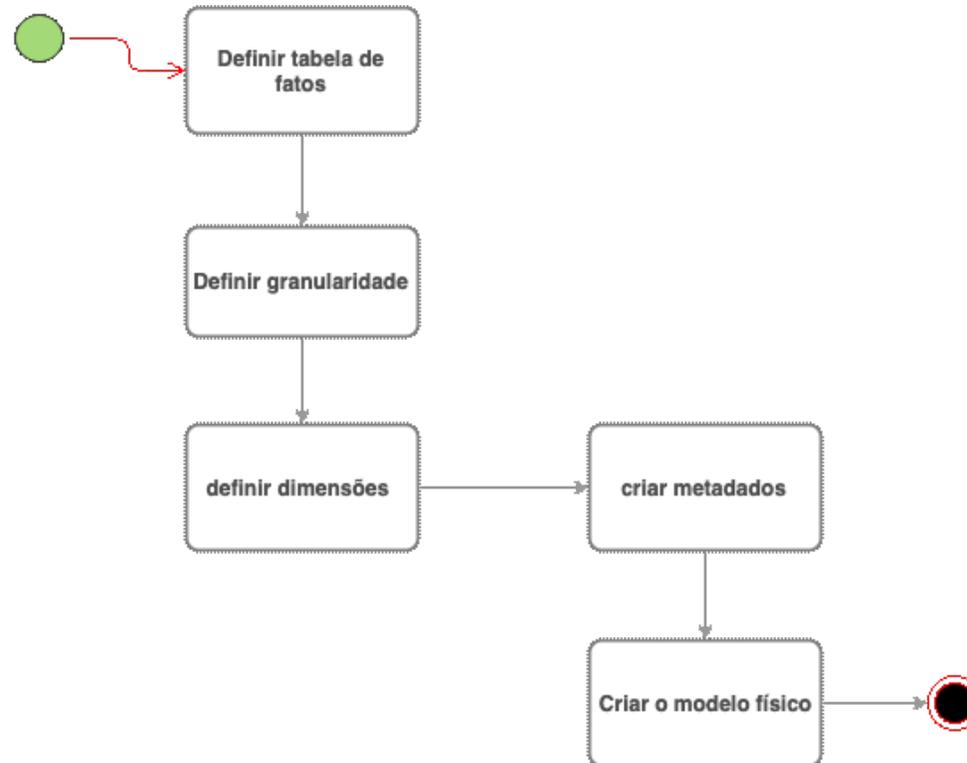


Figura 5 – Desenvolvimento – Início.

Fonte: Própria autoria.

A Tabela de Fatos é uma parte essencial do *DW*, é a tabela principal no esquema estrela. Essa tabela será criada com base no levantamento de requisitos pois ela abrigará dados que serão usados para produzir as informações esperadas pelos usuários.

Kimball (2013, tradução nossa) relata que as tabelas de fatos guardam as medições de desempenho originárias do resultado dos processos de negócios<sup>12</sup>. Pela razão citada por Kimball é que os fatos geralmente são valores numéricos.

A granularidade é fator determinante para o quão detalhista podem ser as informações obtidas das consultas ao *DW*. Inmon (2005, tradução nossa, pg 139) diz que o problema de *design* mais importante para o desenvolvimento de um *DW* acaba sendo decidir o nível de granularidade<sup>13</sup>. É preciso chegar a um equilíbrio entre a granularidade ideal e o volume de dados armazenados por conta disso, quando mais detalhes forem armazenados maior o banco de dados ficará. Ballard (1998) não é possível ir para um nível de detalhes maior do que o que foi decidido ser mantido, mas sempre é possível fazer *roll up* que quer dizer ir de um nível maior de detalhes para um menor<sup>14</sup>. Decidir a granularidade demanda tempo e foco para que as perguntas atuais possam respondidas bem como a possíveis futuras.

As dimensões são as tabelas que estarão ligadas à tabela de fatos com caráter descritivo e são conhecidas por terem poucas alterações nos seus atributos. Kimball (2013) observa que um *DW* é tão bom quanto os atributos das suas dimensões<sup>15</sup>. A afirmação de Kimball é verdadeira pois não será possível extrair informações valiosas se os dados necessários forem negligenciados, usando um exemplo simplista é como se em uma dimensão de nome DIM\_CLIENTE o campo "idade" não fosse considerado e durante a análise desejasse-se saber a faixa etária dos clientes mais ativos nos últimos 6 meses. Veja que a definição da granularidade ajudará também na criação das dimensões.

Os metadados são essenciais para um *Data Warehouse*, já que é nos metadados que encontraremos uma definição detalhada dos dados. Inmon (2002) diz

---

<sup>12</sup> The fact table in a dimensional model stores the performance measurements resulting from an organization's business process events.

<sup>13</sup> The single most important design issue facing the data warehouse developer is determining the proper level of granularity of the data that will reside in the data warehouse.

<sup>14</sup> This is because you cannot change data to a higher level than what you have decided to keep. You can, however, always roll up (summarize) the data to create a table with a lower level of granularity.

<sup>15</sup> In many ways, the data warehouse is only as good as the dimension attributes; the analytic power of the *DW/BI* environment is directly proportional to the quality and depth of the dimension attributes.

que se trata de dados sobre dados, conterà a descrição da estrutura conteúdo e etc.<sup>16</sup>, será mais fácil fazer uso do *DW* com o seu dicionário em mão, Inmon (2005) complementa que quando um usuário lida com um *Data Warehouse* que não tem metadados ele encontra dificuldade em saber por onde começar a analisar.<sup>17</sup> O ideal é que a criação do metadados seja feita no andamento do projeto e não seja deixada para ser feita no final.

A criação do modelo físico serve para produzir a documentação mais detalhada do *DW*.

---

<sup>16</sup> metadata (1) data about data; (2) the description of the structure, content, keys, indexes, etc., of data.

<sup>17</sup> Put differently, when a user approaches a data warehouse where there is no metadata, the user does not know where to begin the analysis.

## 4.5 Desenvolvimento – ETL

O processo de *ETL* (*Extract Transform and Load*) é o mais custoso em *BI*, é dispensado um tempo consideravelmente alto para extrair os dados de todas as fontes necessárias, transformá-los, isto é remover redundâncias, normalizar nomes, adaptar dados, dentre outras coisas e persistir os dados em todas as tabelas de Fatos e Dimensões pertinentes ao processo.

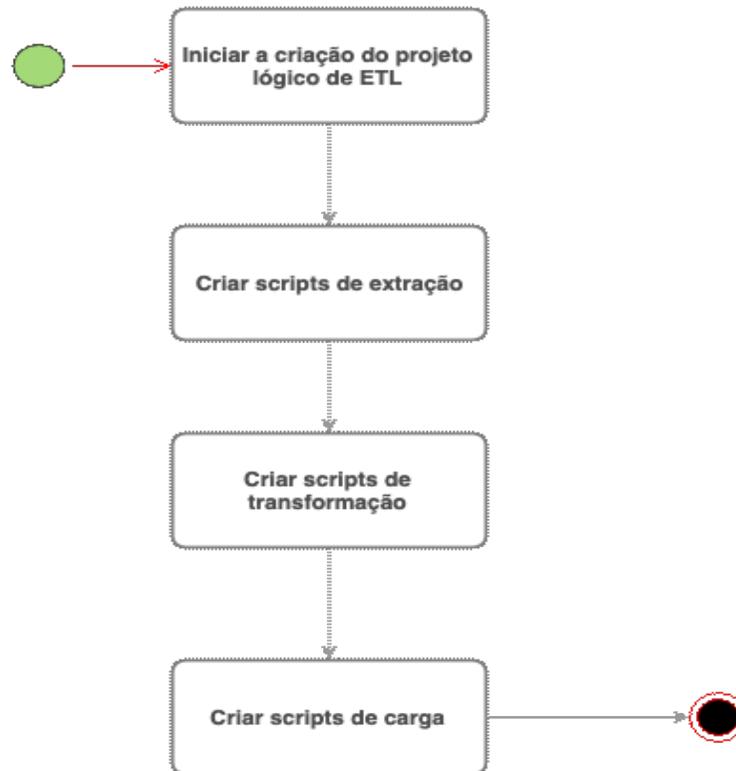


Figura 6 – Desenvolvimento – ETL.

Fonte: Própria autoria.

Existe uma frase do filósofo grego Aristóteles que é: “Bem começado, metade feito”, quando se inicia corretamente algum processo aumenta-se a chance de sucesso. Vieira (2009) diz que antes de iniciar a extração deve-se criar um projeto lógico. O que faz todo sentido, fazendo assim ficará mais fácil relembrar, implementar e mostrar para os interessados todo o fluxo do *ETL*. Kimball (2004) também sustenta a ideia do que o projeto lógico de ver ser feito antes de tudo<sup>18</sup>.

Os *scripts* de extração, transformação e carga devem ser criados voltados para o desempenho, cada minuto a menos na realização desses processos conta preciosos pontos no panorama geral.

---

<sup>18</sup> The ETL process must be figured out logically and documented.

## 4.6 Desenvolvimento – Testes

Depois que o projeto logico for criado é necessário iniciar os testes para experimentar tudo o que foi idealizado, pôr em prática tudo que até então estava no papel.

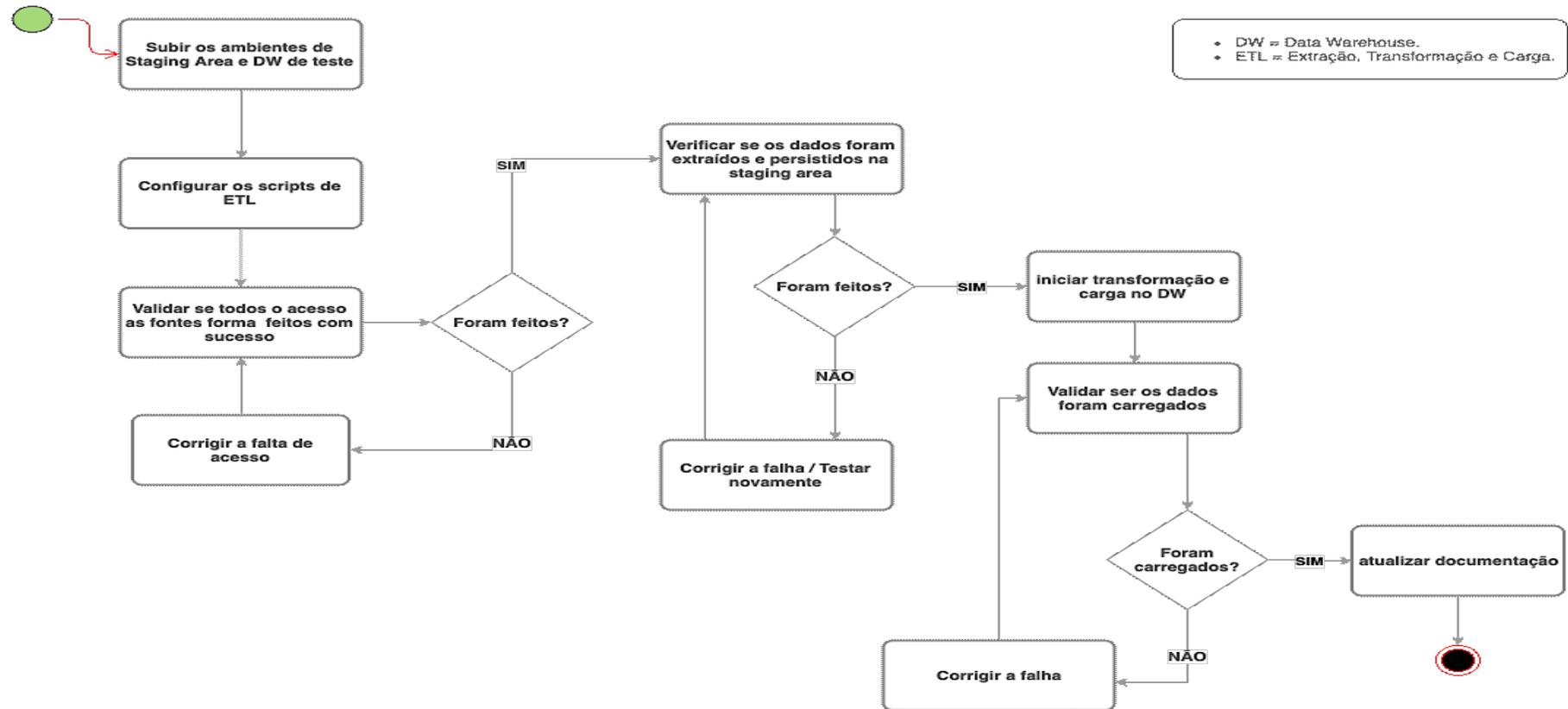


Figura 7 – Desenvolvimento – Testes.

Fonte: Própria autoria.

No início dos testes o acesso ao ambiente de desenvolvimento é restrito a poucos usuários, deverá ser iniciado o ambiente de *Staging Area* e *Data Warehouse*.

*Staging Area* é um local onde, os dados extraídos das fontes envolvidas em alimentar o *DW*, ficam armazenados de forma transitória para serem transformados e carregados de fato no *DW*. Kimball (2004, tradução nossa) se refere a *Staging Area* como a cozinha do *Data Warehouse* completamente além dos limites do usuário final<sup>19</sup>. A vantagem de ter uma *Staging Area* é que no processo de transformação, que pode ser demorado, ao invés de acessar os dados nas fontes o acesso é feito em um local dedicado para isso, dessa forma o risco de impactar negativamente o desempenho das fontes é reduzido drasticamente.

Os *scripts* de *ETL* devem ser configurados na plataforma escolhida para esse fim para serem executados.

Alguns acessos podem ter sido perdidos dependendo do tempo decorrido do planejamento até a fase de desenvolvimento ou se as fontes serão arquivos eles podem ter sofrido alterações de nome ou mudado de lugar, então conferir a permanência do acesso é ideal.

Ao final do processo de extração os dados devem ter sido persistidos na *Staging Area*, caso isso não ocorra o provavelmente o script *ETL* contém falhas e deve ajustado.

Os *scripts* de transformação e carga farão respectivamente a adequação dos dados para deixá-los prontos para as tabelas do *DW* e a persistência no *DW*. Ao final também é necessário validar se os dados esperados foram inseridos com sucesso, caso não forem correções devem ser feitas.

Após o ciclo de *ETL* ter sido concluído a documentação deve ser alterada para refletir as mudanças que foram feitas.

---

<sup>19</sup> The staging area is the kitchen of the data warehouse. It is a place accessible only to experienced data integration professionals. It is a back-room facility, completely off limits to end users

## 4.7 Desenvolvimento – Visualização

Todo o investimento na implementação de um *DW* seria inútil se não fosse possível realizar consultas aos dados armazenados nele e com isso obter informação para suporte à tomada de decisão.

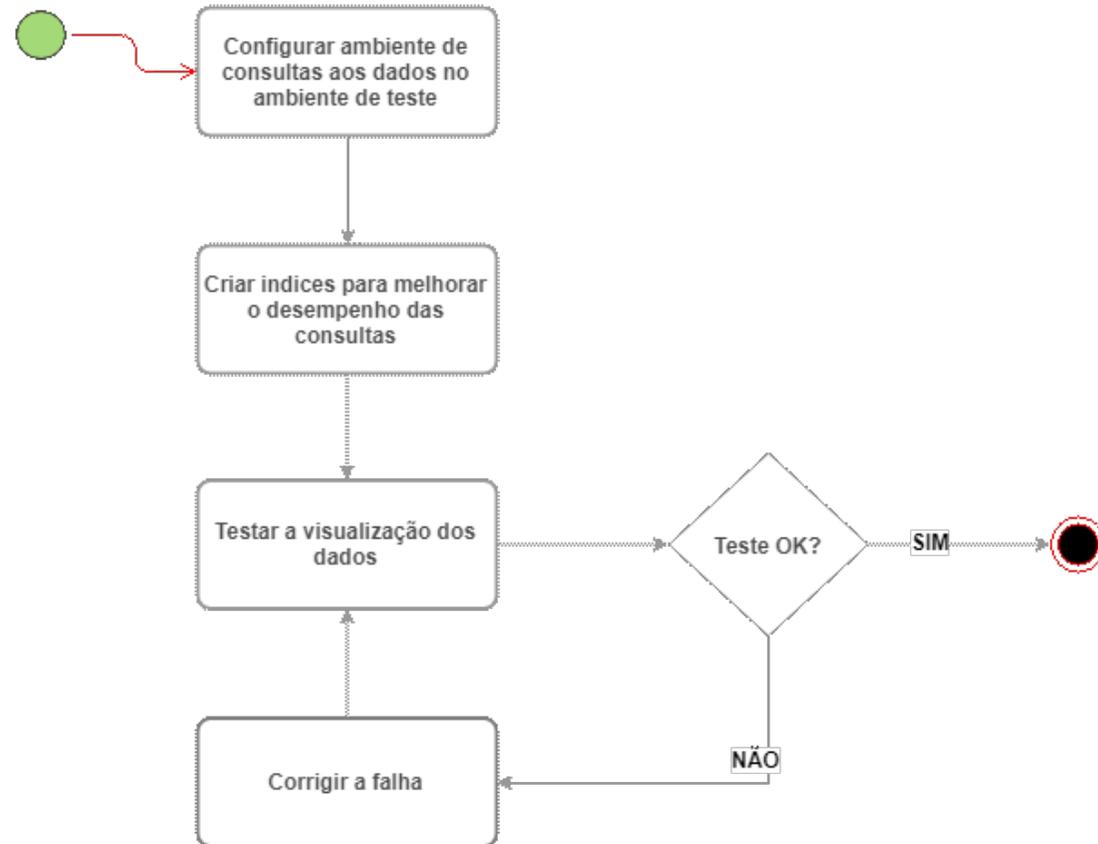


Figura 8 – Desenvolvimento – Visualização.

Fonte: Própria autoria.

Essas consultas são efetuadas, na maioria das vezes, em uma interface que facilita a sua criação, as interfaces podem ser web ou através de outra ferramenta obtida pela empresa. Kimball (2004, tradução nossa) salienta que o *DW* e a parte de visualização são lógica e fisicamente separados, que o acesso ao *DW* diretamente deve ser proibido, ou seja, toda consulta ao *Data Warehouse* deve ser feita por uma interface e nunca diretamente a ele<sup>20</sup>. Bloqueando o acesso direto a banco de dados, principalmente por parte dos usuários finais, pode evitar queries que podem ocasionar lentidão e travamento no mesmo.

A criação de índice é indispensável para melhorar o desempenho das consultas, nesse ambiente quanto mais rápida for a resposta que o usuário final obtém maiores as chances de se obter vantagens sobre os concorrentes. Golfarelli (1998, tradução nossa) diz que a escolha do índice tem um papel crucial quando ao desempenho do *DW*<sup>21</sup>. Pode-se concluir pelo que foi dito por Golfarelli que podem existir índices que prejudiquem o desempenho do banco, então a escolha deve ser feita com cuidado. Ao tomar-se como exemplo o banco de dados *SQL Server* tem-se os índices do tipo: *Hash*, Não clusterizado com otimização de memória, Clusterizado, Não clusterizado, Exclusivo, Columnstore, Índice em colunas computadas, Filtrado, Espacial, XML e Texto completo. (Microsoft, 2016).

Caso o retorno das consultas não esteja em sintonia com o que foi definido no levantamento de requisito, deve-se atuar no problema para normalizar a situação.

---

<sup>20</sup> The back room and the front room of the data warehouse are physically, logically, and administratively separate. Our approach to data warehousing assumes that data access is prohibited in the back room, and therefore the front room is dedicated to just this one purpose.

<sup>21</sup> Index selection has a crucial role in determining the DW performance...

## 4.8 Implementação

Uma das etapas mais importantes de todo projeto é o momento de liberar o uso ao usuário final. O uso do ambiente deve ser incentivado, mais uma vez o apoio de pessoas influente obtido no início do projeto deve ser explorado fazendo com que eles tenham papel importante na divulgação.

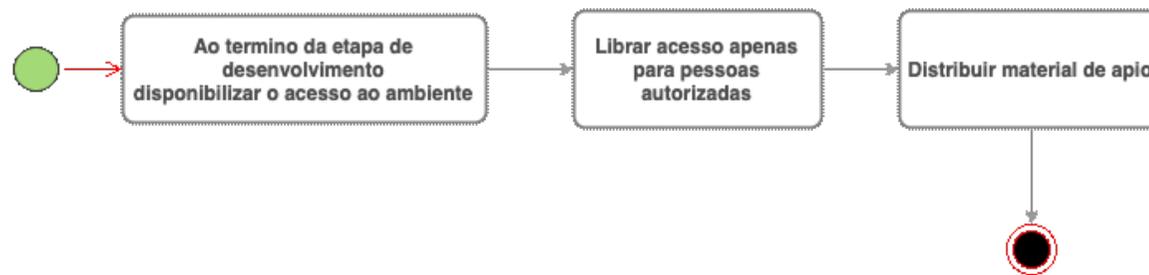


Figura 9 – Implementação.

Fonte: Própria autoria.

Pela alta sensibilidade dos dados que o *Data Warehouse* armazena, o acesso a ele deve ser protegido, o modo mais comum é solicitando *login* e senha para acesso, essa necessidade aumenta se o acesso ao ambiente estiver disponível pela internet, usuários mal-intencionados poderiam fazer uso dessa brecha para obter vantagens sobre as informações.

Distribuir um material para auxiliar o uso do ambiente é um ponto importante, com isso o usuário final sentirá maior confiança no uso do sistema. Treinamentos também podem ser oferecidos caso necessário.

#### 4.4 Pós-Implementação

Já com o sistema em completo funcionamento é o momento de tomar nota sobre como ele se comporta diante do uso diário.

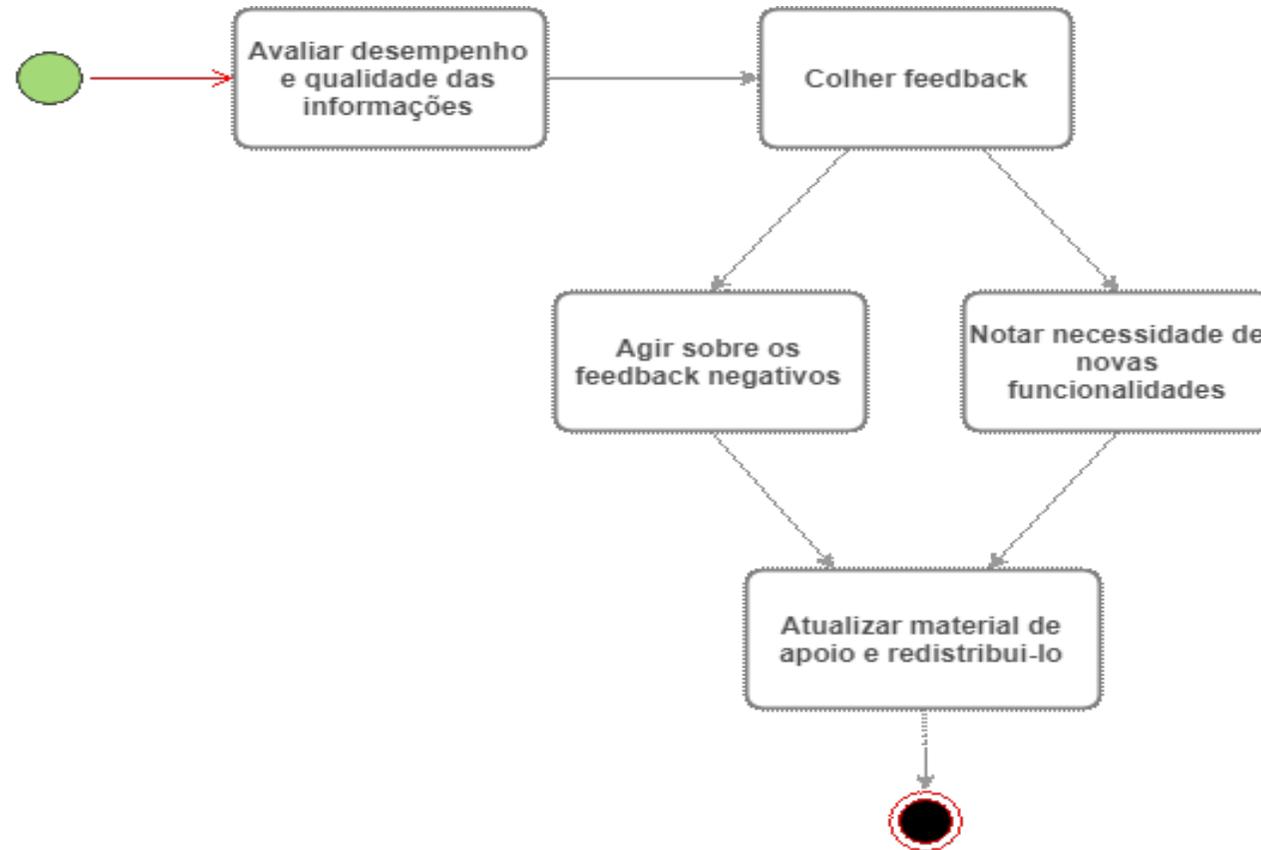


Figura 10 – Pós-Implementação.

Fonte: Própria autoria.

Clemes (2001) relata que o desempenho e a qualidade das informações produzidas devem ser avaliados. Se os resultados ainda estão de acordo com o que foi especificado no levantamento de requisitos é um bom sinal que o projeto feito devidamente implementado.

É muito importante colher *feedback* de todos os usuários do sistema com isso será possível melhorar o ambiente.

Os *feedbacks* negativos são ótimas oportunidades de adicionar melhorias no sistema, não desprezar esses *feedbacks* é uma sábia decisão.

No momento em que os usuários finais passam a utilizar de fato o sistema o entendimento é aprimorado e novas ideias podem surgir, essas ideias devem ser colhidas e submetidas à avaliação da equipe de implementação para possível adição ao sistema.

O material de apoio deve ser atualizado caso haja alguma mudança no ambiente ou *feedback* colhido aponte falta de clareza em seu conteúdo.

Não é uma tarefa fácil a implementação de um *DW*, é preciso análise minuciosa e entender o domínio do ambiente onde a mesma se dará, contudo quando faz-se uso de boas práticas o projeto atinge um nível elevado de qualidade; entre os materiais de pesquisa encontrados tem-se o exemplo da implementação de um *Data Warehouse* no Estaleiro Atlântico Sul Vieira (2009) diz que essa implementação “trouxe ganhos à qualidade da informação e, conseqüentemente ao processo de tomada de decisão”.

## 5 CONSIDERAÇÕES FINAIS

É inegável o quão inseparável *BI* está se tornando das empresas que querem permanecer competitivas no mercado, através desse trabalho foi possível constatar que o *Data Warehouse* é peça importantíssima para *BI* e que o zelo para seu funcionamento de forma otimizada não deve ser desprezado; sendo assim, um guia de boas praticas agrega valor a todo o processo, quando mais rápidas forem as respostas das consultas ao *DW* mais competitiva se tornará a organização que faz uso do mesmo.

O objetivo proposto nesta monografia foi desenvolver, com base no estado da arte da literatura e estudos de caso de implantação de *B.I* um guia de boas práticas para implementação de *Data Warehouses*.

Na pesquisa bibliográfica e documental os materiais de pesquisa serviram para a criação dos fluxogramas que melhoraram o entendimento de como todo processo se desenrola, com isso foi possível a apresentação de um guia de boas práticas, na forma de *checklist*, para ser usado no contexto de *Business Intelligence*, especificamente durante a fase de implementação do ambiente de *Data Warehouse*.

Durante o processo de análise dos estudos de caso foram encontradas valiosas informações sobre a implementação de um *Data Warehouse* em diversos cenários, pôde-se extrair as boas práticas e notar as más; isso contribuiu consideravelmente para a criação do *checklist*.

O *checklist* será de grande serventia para todos que pretendem implementar um *Data Warehouse*, especialmente para os iniciantes na área de *Business Intelligence*, levando em consideração que a implementação será conduzida por seres humanos os quais são fadados a erros e esquecimentos, ter um documento que o “lembre” do caminho a ser seguido é de extrema valia.

É possível concluir que um guia de boas práticas para a implementação do *Data Warehouse*, se seguido, proporcionará qualidade ao processo, mitigação de riscos ocasionados pelo esquecimento, robustez ao sistema e uma implementação sólida e consistente.

Como trabalhos futuros sugere-se o uso do *checklist* aqui apresentado para medir o nível de eficiência do mesmo e assim poder melhorá-lo caso necessário.

## REFERÊNCIAS

ALMEIDA, Alexandre Marques. **Proposição de indicadores para avaliação técnica de projetos de data warehouse: um estudo de caso no data warehouse da plataforma lattes**. 2006. 83f. Dissertação (Mestrado em Engenharia de Produção) UFSC, Florianópolis 2006. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/88222/228102.pdf?sequence=1&isAllowed=y>. Acesso em: 13 ago. 2019.

BALLARD, Chuck. et al. **Data Modeling Techniques for Data Warehousing**. San Jose, Califórnia, 1998.

BARBIERI, Carlos. **Business Intelligence: modelagem e tecnologia**. Rio de Janeiro, RJ, 2001

BRACHMAN, Ronald J. ANAND, Tej. **The process of knowledge discovery in databases: a first sketch**. 1994, Disponível em: <https://pdfs.semanticscholar.org/2db5/ec88e07974242eb8f8de867275bec8f29e3a.pdf>. Acesso em: 14 jul. 2019.

CECI, Flávio. **Business Intelligence**. Palhoça, SC. 2012.

CLEMES, Marcio. **Data warehouse como suporte ao sistema de informações gerenciais em uma instituição de ensino superior: estudo de caso na UFSC**. 2001. 118f. Dissertação (Mestrado em Engenharia de Produção) UFSC, Florianópolis 2001. Disponível em: <http://repositorio.ufsc.br/xmlui/handle/123456789/79664>. Acesso em: 06 ago. 2019.

FAYYAD, Usama. PIATETSKY-SHAPIRO, Gregory. SMYTH, Padhraic. **Knowledge discovery and data mining: towards a unifying framework**. 1996. Disponível em: <https://www.aaai.org/Papers/KDD/1996/KDD96-014.pdf>. Acesso em: 14 jul. 2019

GOLFARELLI, Matteo; RIZZI, Steffano. **A methodological framework for Data warehouse Design**. DOLAP, 1998. Disponível em: <http://www.few.vu.nl/~kgr700/A%20methodological%20framework%20for%20data%20warehouse%20design%20-golfarelli.pdf>. Acesso em: 13 set. 2019.

HENRIQUE, Ozimar. **OLTP X OLAP**. 2012. Disponível em: <https://social.technet.microsoft.com/wiki/pt-br/contents/articles/12576.oltp-x-olap.aspx>. Acesso em: 13 jul. 2019

INMON, William B. **Building the Data Warehouse**, 3 ed. New York, NY, 2002.

INMON, William B. **Building the Data Warehouse**, 4 ed. Indianapolis, Indiana, 2005.

KIMBALL, Ralph., CASERTA, Joe. **The Data Warehouse ETL Toolkit: practical techniques for extracting, cleaning, conforming and delivering data**. Indianapolis, Indiana, 2004.

KIMBALL, Ralph. **The Data Warehouse Toolkit: the definitive guide to dimensional modeling**. 3 ed. Indianapolis, Indiana, 2013.

MEIRA, Carlos Alberto Alves. **Projeto físico de Data Warehouse no Oracle: um estudo de caso**. Embrapa, 2003. Disponível em: <https://www.infoteca.cnptia.embrapa.br/bitstream/doc/8838/1/doc35.pdf>. Acesso em: 16 ago. 2019.

Índices. **Microsoft**, 2016. Disponível em: <<https://docs.microsoft.com/pt-br/sql/relational-databases/indexes/indexes?view=sql-server-2017>>. Acesso em: 19 ago. 2019.

OLTP (processamento de transações online). **Microsoft**, 2018. Disponível em: <<https://docs.microsoft.com/pt-br/azure/architecture/data-guide/relational-data/online-transaction-processing>>. Acesso em: 13 jul. 2019

PRIMAK, Fábio Vinícius. **Decisões com BI**. Rio de Janeiro, RJ, 2008.

PROJECT MANAGEMENT INSTITUTE (PMBOK). **Guia PMBOK: um guia do conhecimento em gerenciamento de projetos**. 4 ed. PMI, 2008.

SANTOS, Ricardo S. et al. **Data warehouse para a saúde pública: estudo de caso SES-SP**. Academia.edu, 2006. Disponível em: <[https://www.academia.edu/4997943/Data\\_Warehouse\\_para\\_a\\_Sa%C3%BAdede\\_P%C3%BAblica\\_Estudo\\_de\\_Caso\\_SES-SP](https://www.academia.edu/4997943/Data_Warehouse_para_a_Sa%C3%BAdede_P%C3%BAblica_Estudo_de_Caso_SES-SP)>. Acesso em: 05 ago. 2019.

RAINER, R. K., CEGIELSKI, C. G. **Introdução a sistemas de informação**, 3 ed. Rio de Janeiro, RJ, 2011.

What is a data warehouse and why does it matter to your business? **Talend**, 30 de abr. de 2019. Disponível em: <<https://www.talend.com/resources/what-is-data-warehouse/>>. Acesso em: 21 de jul. de 2019.

TURBAN, Efrain. et al. **Business Intelligence – Um enfoque gerencial para a inteligência de negócio**. São Paulo, SP, 2009

VIEIRA, Inês Beltrão Gama. **Um processo para tomada de decisão apoiado em Data Warehouse – Um estudo de caso no Estaleiro Atlântico Sul**. 2009. 112f.

Dissertação (Mestrado em Ciência da Computação) UFPE, Recife, 2009. Disponível em: [https://repositorio.ufpe.br/bitstream/123456789/1895/1/arquivo1896\\_1.pdf](https://repositorio.ufpe.br/bitstream/123456789/1895/1/arquivo1896_1.pdf). Acesso em: 19 ago. 2019.

WITTEN, Ian H. EIBE, Frank. HALL, Mark A. **Data Mining: practical machine learning tools and techniques**. 3 ed. Indianapolis, Indiana, 2011.

Dicionário Infopédia da Língua Portuguesa, **Infopédia**, 21 de nov. de 2019. Disponível em: < <https://www.infopedia.pt/dicionarios/lingua-portuguesa/boas-praticas>>. Acesso em: 21 de nov. de 2019.

DILL, Sérgio Luis. **Uma metodologia para desenvolvimento de data Warehouse e estudo de caso**. 2002. 131f. Dissertação (Mestrado em Ciência da Computação) UFSC, Santa Catarina, 2002. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/82897/188747.pdf?sequence=1&isAllowed=y>. Acesso em: 19 ago. 2019.

Informações corporativas. **Amazon**, 2019. Disponível em: <[https://www.amazon.com.br/gp/feature.html?ie=UTF8&docId=5721910011&ref\\_=footer\\_gw\\_m\\_b\\_corporate](https://www.amazon.com.br/gp/feature.html?ie=UTF8&docId=5721910011&ref_=footer_gw_m_b_corporate)>. Acesso em: 16 dez. 2019.

Martins, Rosemary. **O Que é PDCA**. Blog da Qualidade, 2018. Disponível em: <<http://gestao-de-qualidade.info/ferramentas-da-qualidade/pdca.html>>. Acesso em: 16 dez. 2019.

## APÊNDICE A – Checklist da fase de planejamento.

PLANEJAMENTO			
Nº	Item	Feito?	Observação
1	Estrutura		A implementação de um <i>Data Warehouse</i> também é um projeto, e como tal deve ter bem definido os papéis: quem coordenará e que executará as tarefas. Santos (2006).
1.1	Quem coordenará?	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	
1.2	Quem executará?	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	
2	Escopo do projeto	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Inclui os processos necessários para assegurar que o projeto inclua todo o trabalho necessário, e apenas o necessário, para que termine com sucesso. Preocupa-se principalmente em definir e controlar o que está incluso no projeto e o que não está. PMBOK (2008), Santos (2006)
3	Iniciar o Termo de Abertura do Projeto (TAP)	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Esse documento registra as necessidades do projeto e também conecta o projeto aos trabalhos em andamento na organização. PMBOK (2008).
4	Buscar apoio de alguém influente (Patrocinadores)	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Os patrocinadores são importantes pois podem ajudar a contornar divergências e negociar prioridades durante o

			desenvolvimento dos trabalhos. Cledes (2001).
<b>5</b>	Decidir a forma do levantamento de requisitos		São informações coletadas que representarão as funcionalidades do projeto e estas atenderão as expectativas das partes interessadas. PMI (2008). Nos itens 5.1 e 5.2 são apresentadas duas opções.
<b>5.1</b>	<i>Source-Driven</i>	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Os requisitos serão identificados com base na observação dos dados fornecidos pelas fontes. Santos (2006).
<b>5.2</b>	Métodos tradicionais	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Podem ser usadas as técnicas como: entrevistas, discussão em grupo, <i>workshops</i> , técnicas de criação em grupo, técnicas de tomada de decisão em grupo, questionários, pesquisas observações e protótipos. PMI (2008)
<b>5.3</b>	Iniciar o levantamento de requisitos usando as técnicas escolhidas.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Isso resultará nos requisitos documentados PMI (2008).
<b>6</b>	Avaliar infraestrutura de TI disponível.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Verificar no apêndice B
<b>7</b>	Compile o que foi levantado.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	PMBOK (2008).
<b>8</b>	Definir o tipo de implementação de arquitetura.		Há dois tipos mais famosos de arquitetura a <i>top-down</i>

			<i>Inmon (1997 apud Almeida, 2006) e a bottom-up Kimball (1988, apud Almeida 2006&gt;</i>
<b>8.1</b>	<i>Top-Down</i>	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	É criado um Data Warehouse como o centro analítico da empresa. O Data Marts são criados a partir dele.
<b>8.2</b>	<i>Bottom-Up</i>	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Primeiro são formados os Data Marts que formarão o Data Warehouse.
<b>9</b>	Definir o modelo	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	O modelo mais utilizado é chamado de “Estrela”. Trata-se de uma grande tabela central denominada “Tabela Fato” e pequenas tabelas conectadas à mesma denominadas de “Dimensões” formando o modelo dimensional Kimball (2013)
<b>10</b>	Finalizar TAP	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	

## APÊNDICE B – Checklist da infraestrutura.

INFRAESTRUTURA			
Nº	Item	Feito?	Observação
1	Avaliar infraestrutura de TI disponível.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Toda a implementação é influenciada pela infraestrutura de TI. IBM, (apud Cledes, 2001)
2	Agir caso infra atual não suportar a implementação do DW.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	<p>Caso a arquitetura atual não suporte o projeto, existem duas soluções possíveis:</p> <p>1- Readaptar o projeto para que ele se encaixe ao infra disponível.</p> <p>2- Negociar, se for possível, a compra e adequação do que for preciso para o bom funcionamento do DW.</p>
3	Reservar junto a TI os hardwares acessos e softwares necessários para projeto.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	É importante notar que sendo um projeto inevitavelmente outras equipes serão envolvidas, como nesse caso TI.
4	Definir as ferramentas de desenvolvimento.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Existem várias ferramentas no mercado, <i>Free, Open Source</i> e Pagas. O time de TI mais uma vez, provavelmente, será envolvido para configuração de ambientes, bem como o

			financeiro caso se opte por ferramentas pagas. Santos (2006).
5	Documentar o modelo lógico da infraestrutura.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Documentar o modelo lógico da infra é de grande importância para referencias e consultas. Vieira (2009).

## APÊNDICE C – Checklist do backup e recuperação de desastres

BACKUP E RECUPERAÇÃO DE DESASTRE			
Nº	Item	Feito?	Observação
1	Iniciar planos de backup de recuperação de desastre.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Tanto o <i>Data Warehouse</i> quanto os dados armazenados nele compõem os ativos da organização, e como tais devem ser protegidos. No caso do <i>DW</i> dois pontos importantes quanto a sua proteção não devem ser negligenciados, que são: Backup e Recuperação de desastres. NBR ISO/IEC 27001. Meira (2003)
2	Agir caso a política de backup e recuperação de desastres for inexistente.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Solicitar junto a TI a criação das políticas e planos necessários para implementação das rotinas de backup e restauração do serviço.
3	Adicionar o projeto ao escopo de backup e recuperação de desastre.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Na maioria das vezes o departamento de TI das empresas já tem uma política voltada para esse assunto, basta apenas incluir o novo projeto no escopo.
4	Documentar a rotina de backup e o plano de recuperação de desastre.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	É importante reunir toda a documentação relativa a continuidade do negócio, isso

			facilitará o entendimento de todos quanto ao projeto como um todo.
--	--	--	--

## APÊNDICE D – Checklist do desenvolvimento

DESENVOLVIMENTO			
Nº	Item	Feito?	Observação
1	Desenvolvimento.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Após a etapa de planejamento, que se deteve em um nível mais abstrato, deve ser dado início ao desenvolvimento onde um nível maior de detalhamento será aplicado.
2	Definir as tabelas de fatos.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Tabela central de um determinado assunto em um Data Warehouse. Kimball (2013)
3	Definir a granularidade	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Nível de detalhes dos dados que serão salvos no DW. É um aspecto de extrema importância para o design de um DW. Inmon (2002).
4	Definir as dimensões.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Tabelas que rodeiam a tabela de fato, têm a função descritiva.
5	Criar metadados.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Um ponto que não deve ser negligenciado, tão pouco deixado por último. Os metadados serão o dicionário dos dados. Inmon (2002).
	Criar modelo físico.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	O modelo com um nível maior de detalhamento em relação ao banco de dados. Golfarelli (2001 apud Dill 2002).

## APÊNDICE E – Checklist do ETL

DESENVOLVIMENTO - ETL			
Nº	Item	Feito?	Observação
1	Iniciar a criação do projeto logico de ETL.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	A fase de ETL é a mais demorada no processo de BI. Para que seja fácil seguir, aplicar e explicar aos interessados o processo de ETL deve-se criar um projeto logico. Vieira (2009)
2	Criar os scripts de extração.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	A etapa de extração “colherá” os dados de diferentes fontes dentro da organização. É importante checar com cada área que fornecerá os dados quanto a permissão de acesso aos locais que os dados se encontram.
3	Criar scripts de transformação.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	A etapa de transformação tornará os dados prontos para serão salvos.
4	Criar scripts de carga	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	A etapa de carga irá persistir os dados transformados nas usas respectivas tabelas, Fatos e Dimensões.

## APÊNDICE F – Checklist dos Testes

DESENVOLVIMENTO – Testes			
Nº	Item	Feito?	Observação
1	Subir os ambientes de <i>Staging Area</i> e DW de teste	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	<p>Agora que o projeto lógico já foi criado, deve-se começar a iniciar os ambientes para teste, nesse momento o acesso às funcionalidades é restrito a apenas usuários chave.</p> <p><i>Staging Area</i> é um local intermediário onde os dados já extraídos via ETL são armazenados para depois serem transformados e inseridos no DW. Kimball (2004)</p>
2	Configurar os scripts de ETL na plataforma escolhida.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Os scripts de ETL de ser configurados para execução na plataforma escolhido no projeto.
3	Validar o acesso a todas a fontes de dados.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Nesse momento é importante validade se nenhum dos acessos às fontes de dados foram perdidos. Caso algum acesso seja perdido deve ser tratado o problema para que não impacte o andamento do projeto.

4	Validar se os dados foram extraídos e persistido na Staging Area.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	É importante verificar se os dados esperados foram trazidos. Caso negativo, corrigir o problema.
5	Iniciar transformação de carga no DW.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Os dados que estão na <i>Staging Area</i> devem ser transformados através dos scripts de <i>ETL</i> e carregados no <i>DW</i> .
6	Validar se todos os dados foram carregados no DW.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	É importante verificar se os dados esperados foram carregados conforme esperado. Caso negativo corrigir o problema.
7	Atualizar documentação do projeto.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Caso haja mudanças consideradas relevantes para o projeto é importante que a documentação sempre esteja atualizada.

## APÊNDICE G – Checklist da Visualização

DESENVOLVIMENTO – Visualização			
Nº	Item	Feito?	Observação
1	Configurar ambiente de consultas aos dados no ambiente de teste.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	É necessário um ambiente para poderem ser realizadas as consultas aos dados na fase atual, esse ambiente será disponibilizado para o usuário final posteriormente. Kimball (2004).
2	Criar índices para melhorar o desempenho das consultas.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	A criação de índices aumentará a velocidade com que os resultados retornaram das consultas. Golfarelli (1998)
3	Realizar teste no ambiente de consulta	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Este teste é necessário para verificar se as informações solicitadas no levantamento de requisitos estão sendo obtidas. Caso haja alguma discrepância nas informações exibidas deve-se atuar no problema a fim de alinhar o que é exibido com o que foi solicitado.

## APÊNDICE H – *Checklist* da Implementação

DESENVOLVIMENTO – Implementação			
Nº	Item	Feito?	Observação
1	Ao término da etapa de desenvolvimento disponibilizar o acesso ao ambiente.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	O momento é que os usuários finais terão acesso ao ambiente, seu uso deve ser incentivado para que seja possível avaliar como ele se comporta no uso diário.
2	Liber acesso apenas para pessoas autorizadas.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	É importante que o acesso ao ambiente seja protegido com login e senha, evitando assim que pessoas não autorizadas tenha acesso as informações. A necessidade de um ambiente protegido se intensifica se o ambiente também fica acessível pela internet.
3	Distribuir material de apoio.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Material que irá auxiliar ao uso do ambiente.

## APÊNDICE I – Checklist da Pós-Implementação

DESENVOLVIMENTO – Pós-Implementação			
Nº	Item	Feito?	Observação
1	Avaliar o desempenho e a qualidade das informações.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Este é o momento de avaliar como o serviço de BI está se comportando bem como a qualidade das informações obtidas, ver se essas informações continuam a estar alinhadas com o levantamento de requisitos. Clemes (2001).
2	Colher feedback.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Colha o máximo de feedback que puder dos usuários do sistema.
3	Agir sobre os feedbacks negativos.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Até os feedbacks negativos podem fornecer ótimas oportunidade para melhorar o ambiente, aja sobre os eles com prioridade.
4	Notas a necessidade de novas funcionalidades.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	No momento em que os usuários finais para a utilizar de fato o sistema o entendimento é aprimorado e novas ideias podem surgir.
5	Atualizar o material de apoio.	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO	Adeque o material de apoio caso haja alguma mudança no sistema ou o material está confuso para os usuários.