



Universidade de Brasília - UnB
Faculdade UnB Gama - FGA
Engenharia de Software

**Estudo da Eficácia e Eficiência do Uso de um
Ambiente de *Data Warehousing* para Aferição
da Qualidade Interna de Software: um Estudo
de Caso no TST**

Autor: Pedro da Cunha Tomioka
Orientador: Prof. Msc. Hilmer Rodrigues Neri

Brasília, DF
2014



Pedro da Cunha Tomioka

**Estudo da Eficácia e Eficiência do Uso de um Ambiente
de *Data Warehousing* para Aferição da Qualidade Interna
de Software: um Estudo de Caso no TST**

Monografia submetida ao curso de graduação
em Engenharia de Software da Universidade
de Brasília, como requisito parcial para ob-
tenção do Título de Bacharel em Engenharia
de Software.

Universidade de Brasília - UnB

Faculdade UnB Gama - FGA

Orientador: Prof. Msc. Hilmer Rodrigues Neri

Brasília, DF

2014

Pedro da Cunha Tomioka

Estudo da Eficácia e Eficiência do Uso de um Ambiente de *Data Warehousing* para Aferição da Qualidade Interna de Software: um Estudo de Caso no TST/
Pedro da Cunha Tomioka. – Brasília, DF, 2014-

23 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Msc. Hilmer Rodrigues Neri

Trabalho de Conclusão de Curso – Universidade de Brasília - UnB
Faculdade UnB Gama - FGA , 2014.

1. Métricas de Código-Fonte. 2. *Data Warehousing*. I. Prof. Msc. Hilmer Rodrigues Neri. II. Universidade de Brasília. III. Faculdade UnB Gama. IV. Estudo da Eficácia e Eficiência do Uso de um Ambiente de *Data Warehousing* para Aferição da Qualidade Interna de Software: um Estudo de Caso no TST

CDU 02:141:005.6

Pedro da Cunha Tomioka

Estudo da Eficácia e Eficiência do Uso de um Ambiente de *Data Warehousing* para Aferição da Qualidade Interna de Software: um Estudo de Caso no TST

Monografia submetida ao curso de graduação em Engenharia de Software da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em Engenharia de Software.

Trabalho aprovado. Brasília, DF, Ainda não se sabe:

Prof. Msc. Hilmer Rodrigues Neri
Orientador

Ainda não se sabe
Convidado 1

Ainda não se sabe
Convidado 2

Brasília, DF
2014

Este trabalho é dedicado à minha mãe, Elizabeth, e à minha irmã, Luisa.

Agradecimentos

Matheus Tristão, Nilton César Araruna, Guilherme Baufaker

Agradeço a meus amigos e companheiros de graduação: Pedro Henrique Potiguara, Aline Gonçalves, Hichemm Khalyd, Tiago Pereira, Thabata Granja, Guilherme Fay, Guilherme de Lima, Pedro Matias, Mayco Henrique, Carlos Filipe, Marcos Ronaldo, Rafael Ferreira, Fágner Rodrigues, Bruno Motta, Paulo Acés, Matheus Patrocínio, <equipe de mds>, Greg Ouyama, Eusyar <sobrenome>, João Pedro Carvalho, Artur Potiguara , Pedro Guilherme Moreira, Pedro Inazawa.

"The Truth is Out There"
(The X-Files)

Resumo

Palavras-chaves: Métricas de Código-Fonte. *Data Warehousing*. *Data Warehouse*

Abstract

Palavras-chaves: *Source Code Metrics. Data Warehousing. Data Warehouse*

Lista de ilustrações

Figura 1 – Legenda	17
--------------------	----

Lista de tabelas

Tabela 1 – Add caption	21
----------------------------------	----

Lista de abreviaturas e siglas

ACC	<i>Afferent Connections per Class</i>
ACCM	<i>Average Cyclomatic Complexity per Method</i>
AMLOC	<i>Average Method Lines of Code</i>
ANPM	<i>Average Number of Parameters per Method</i>
CBO	<i>Coupling Between Objects</i>
CSV	<i>Comma-Separated Values</i>
DER	Diagrama Entidade Relacionamento
DIT	<i>Depth of Inheritance Tree</i>
DW	<i>Data Warehouse</i>
ETL	<i>Extraction-Transformation-Load</i>
FTP	<i>File Transfer Protocol</i>
GQM	<i>Goal-Question-Metric</i>
IEC	<i>International Electrotechnical Commission</i>
IPHAN	Instituto do Patrimônio Histórico e Artístico Nacional
ISO	<i>International Organization for Standardization</i>
JSON	<i>JavaScript Object Notation</i>
LCOM4	<i>Lack of Cohesion in Methods</i>
LOC	<i>Lines of Code</i>
NPA	<i>Number of Public Attributes</i>
NOC	<i>Number of Children</i>
NOM	<i>Number of Methods</i>
OLAP	<i>On-Line Analytical Processing</i>
OLTP	<i>Online Transaction Processing</i>

RFC	<i>Response For a Class</i>
SCAM	<i>IEEE International Working Conference on Source Code Analysis and Manipulation</i>
SGBD	Sistema de Gerenciamento de Bancos de Dados
SICG	Sistema Integrado de Conhecimento e Gestão
XML	<i>Extensible Markup Language</i>
YAML	<i>YAML Ain't Markup Language</i>

Sumário

1	INTRODUÇÃO	14
1.1	Contexto	14
1.2	Problema	14
1.3	Questão de Pesquisa	14
1.4	Objetivos	14
1.5	Hipótese	14
1.6	Organização do Trabalho	14
2	MÉTRICAS DE SOFTWARE	15
3	<i>DATA WAREHOUSING</i>	16
4	PROJETO DE ESTUDO DE CASO	17
4.1	Definição	17
4.2	Background	19
4.3	Design	19
4.4	Seleção	19
4.5	Fonte e Método de Coleta de Dados	19
4.6	Validade	19
4.7	Trabalho de Campo	20
4.8	Análise	20
4.9	Considerações Finais do Capítulo	20
5	CONCLUSÃO	22
	Referências	23

1 Introdução

1.1 Contexto

1.2 Problema

1.3 Questão de Pesquisa

1.4 Objetivos

1.5 Hipótese

1.6 Organização do Trabalho

2 Métricas de Software

3 *Data Warehousing*

Os principais fatores para a adoção de um programa de métricas em organizações de desenvolvimento de software são regularidade da coleta de dados; a utilização de uma metodologia eficiente e transparente nessa coleta; o uso de ferramentas (não-intrusivas) para automatizar a coleta; o uso de mecanismos de comunicação de resultados adequados para todos os envolvidos; o uso de sofisticadas técnicas de análise de dados;

4 Projeto de Estudo de Caso

BREVE INTRODUCAO DO CAPITULO (YIN, 2001) 123

4.1 Definição

O estudo de caso representa uma maneira de se investigar um tópico empírico seguindo-se um conjunto de procedimentos pré-especificados Yin (2001). Um estudo de caso tem a vantagem de prover respostas que vão além das questões de “O que”

Consiste no estudo profundo e exaustivo de um ou poucos objetos, de maneira que permita seu amplo e detalhado conhecimento, tarefa praticamente impossível mediante outros delineamentos já considerados

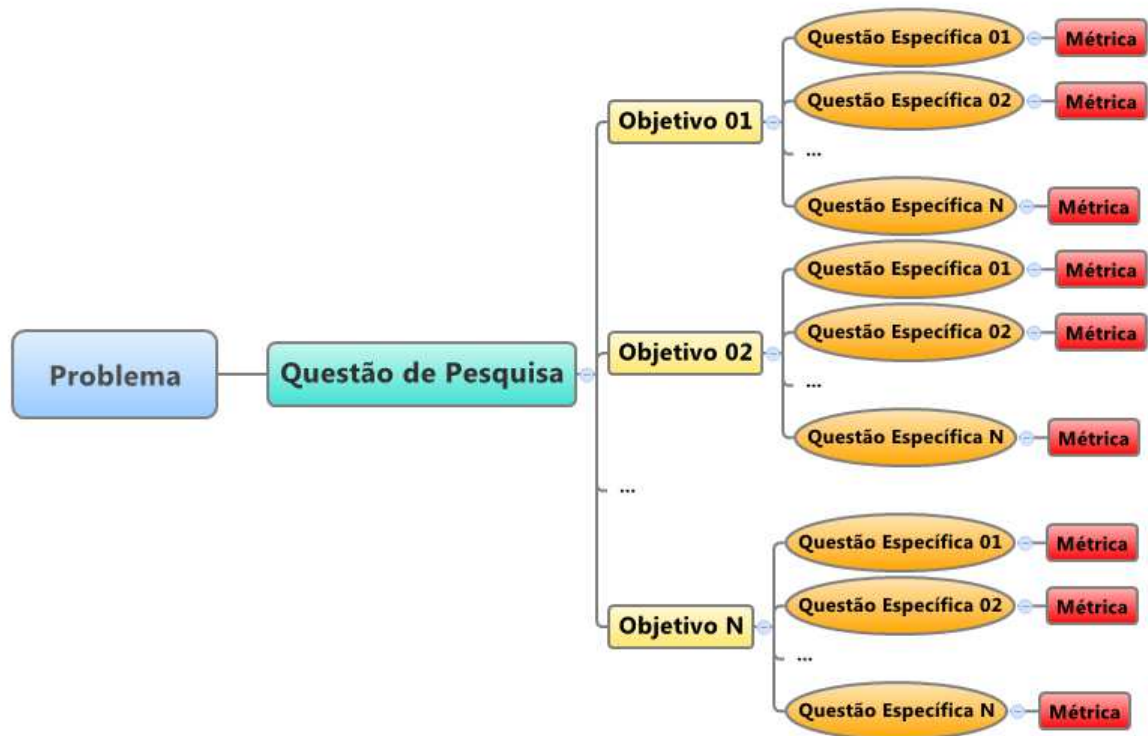


Figura 1 – Legenda

Objetivo 01 Avaliar a eficácia e eficiência da solução de *Data Warehousing* do ponto de vista da <equipe de qualidade> no contexto de desenvolvimento de software.

Questão 01: Quantas tomadas de decisão foram realizadas pela equipe baseado no uso da solução?

Fonte: Observação em campo

Métrica: número de decisões tomadas

Questão 02: Quantas tomadas de decisão foram realizadas pela equipe baseado no uso da solução em um <período de tempo>?

Fonte: Observação em campo (audio/video)

Métrica: número de decisões tomadas/tempo

Questão 03: Qual a avaliação da equipe de qualidade quanto a detecção de cenários de limpeza de código?

Fonte: Questionário com a equipe de qualidade

Métrica: muito bom, bom, regular, ruim, muito ruim.

Questão 04: Com que frequência a equipe de qualidade encontra falhas relacionados à utilização da ferramenta?

Fonte: Observação em campo (audio/video)

Métrica: quantidade de falhas

Questão 05: Com que frequência a equipe de qualidade encontra falhas relacionados à utilização da ferramenta por tempo?

Fonte: Registro de Observação em campo (audio/video)

Métrica: quantidade de falha / tempo

Questão 06: Qual a proporção de uso da ferramenta para a tomada de decisões?

Fonte: Questionário com a equipe de qualidade

Métrica: número de decisões tomadas / número de vezes que a solução foi usada

Interpretação do valor da métrica: —

Questão 07: Qual a quantidade de cenários que foram corrigidos após utilização da solução?

Fonte: código fonte

Métrica: números de cenários corrigidos por release / número de cenários encontrados por release

Questão 08: Qual o nível de satisfação do uso da solução em comparação à solução anterior?

Fonte: Equipe de Qualidade

Métrica: muito satisfeito, satisfeito, neutro, insatisfeito, muito insatisfeito.

Questão 09: Qual a taxa de oportunidade de melhoria de código da solução em uma <intervalo de tempo (sprint, release)>?

Fonte: Código-fonte

Métrica: A Taxa de Aproveitamento de Oportunidades de Melhoria de Código, em uma release, é calculada como:

$$T_r = \frac{\sum_{i=1}^n Ce_i}{\sum_{i=1}^n Cl_i}$$

onde Ce é o total de cenários de limpezas identificados e Cl é total de classes em uma release.

Interpretação do valor da métrica: —

4.2 Background

4.3 Design

4.4 Seleção

4.5 Fonte e Método de Coleta de Dados

4.6 Validade

- Validade do Constructo: Está presente na fase de coleta de dados onde deve ser evidenciado as múltiplas fontes de evidência e a coleta de um conjunto de métricas para que se possa saber exatamente o que medir e quais dados são relevantes para o estudo, de forma a responder as questões de pesquisa (YIN, 2001) buscou-se garantir a validade de construção ao se definir objetivos com evidências diferentes. Estas por sua vez estão diretamente relacionadas com os objetivos do estudo de caso e os objetivos do trabalho

- Validade interna: Para Yin (2001), o uso de várias fontes de dados e métodos de coleta permite a triangulação, uma técnica para confirmar se os resultados de diversas fontes e de diversos métodos convergem. Dessa forma é possível aumentar a validade interna do estudo e aumentar a força das conclusões. A triangulação de dados se deu pelo resultado da solução de DW que utiliza o código-fonte e foi explicada no capítulo 3, de base de documentos, de questionários e de entrevistas para coleta de dados. A triangulação de métodos ocorreu pelo uso de métodos de coleta quantitativos e qualitativos
- Validade externa: Por este ser um caso único a generalização do estudo de caso se dá de maneira pobre (YIN, 2001), assim é necessário a utilização do estudo em múltiplos casos para que se comprove a generalidade dos resultados. O fato deste trabalho ser o primeiro a utilizar a solução para o estudo de caso no órgão, faz com que não seja possível correlacionar os resultados obtidos a nenhum outro estudo.
- Confiabilidade: Sobre isso, Yin (2001) associa à repetibilidade, desde que seja usada a mesma fonte de dados. Nesse trabalho o protocolo de estudo de caso apresentado nessa seção garante a repetibilidade desse trabalho e consequentemente a validade relacionada a confiabilidade

4.7 Trabalho de Campo

4.8 Análise

- Categorização: Organização dos dados em duas categorias: Qualitativos e quantitativos. Os dados qualitativos referem-se aos questionários realizados. Os dados quantitativos, por sua vez, referem-se aos valores numéricos da solução de DW para monitoramento de métricas
- Exibição: Consiste na organização dos dados coletados para serem exibidos através de gráficos, tabelas e texto para poderem ser analisados.
- Verificação: Atestar padrões, tendências e aspectos específicos dos significados dos dados. Procurando assim gerar uma discussão e interpretação de cada dado exibido.
- Conclusão: Agrupamento dos resultados mais relevantes das discussões e interpretações dos dados anteriormente apresentados.

4.9 Considerações Finais do Capítulo

Tabela 1 – Add caption

coluna qualquer	valor	valor
A	1	3
B	2	4

5 Conclusão

Referências

YIN, R. K. *Estudo de caso: planejamento e metodos*. Porto Alegre: Bookman, 2001. ISBN 8573078529 9788573078527. Citado 3 vezes nas páginas [17](#), [19](#) e [20](#).