

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

**Métricas de apoio a estimativa de esforço para
o desenvolvimento de sistemas de *Business
Intelligence***

Autor: Luciano Hideaky de Macedo Endo
Orientadora: Msc. Fabiana Freitas Mendes

Brasília, DF
2014



Luciano Hideaky de Macedo Endo

Métricas de apoio a estimativa de esforço para o desenvolvimento de sistemas de *Business Intelligence*

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.

Universidade de Brasília - UnB

Faculdade UnB Gama - FGA

Orientador: Msc. Fabiana Freitas Mendes

Brasília, DF

2014

Luciano Hideaky de Macedo Endo

Métricas de apoio a estimativa de esforço para o desenvolvimento de sistemas de *Business Intelligence*/ Luciano Hideaky de Macedo Endo. – Brasília, DF, 2014-87 p. : il. (algumas color.) ; 30 cm.

Orientador: Msc. Fabiana Freitas Mendes

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

1. Métricas. 2. Business Intelligence. I. Msc. Fabiana Freitas Mendes. II. Universidade de Brasília. III. Faculdade UnB Gama. IV. Métricas de apoio a estimativa de esforço para o desenvolvimento de sistemas de *Business Intelligence*

CDU 02:141:005.6

Luciano Hideaky de Macedo Endo

Métricas de apoio a estimativa de esforço para o desenvolvimento de sistemas de *Business Intelligence*

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, 27 de novembro de 2014:

Msc. Fabiana Freitas Mendes
Orientadora

Dr. Wander Cleber M. P. da Silva
Professor Avaliador

Msc. George Marsicano Corrêa
Professor Avaliador

Brasília, DF
2014

"Só se pode alcançar um grande êxito quando nos mantemos fiéis a nós mesmos."
(Friedrich Nietzsche)

Resumo

Objetivo: Este trabalho se propõe a identificar métricas capazes de estimar o esforço necessário para o desenvolvimento de um sistema de Business Intelligence.

Metodologia: Foi realizado um estudo secundário com o objetivo de identificar métricas existentes no contexto de BI, e métricas para aferir a complexidade de processos de negócio. A partir disso, foi elaborada uma hipótese que relaciona complexidade de processos de negócio com esforço necessário para implementá-lo em um Sistema BI. Esta hipótese foi explorada em um estudo de caso.

Resultados: Foi demonstrada que a hipótese inicialmente identificada é verdadeira, apesar de serem necessários mais estudos para confirmar essa relação. Também foi proposto um método que explora a relação definida pela hipótese. Dessa forma, utilizando o método, é possível estimar o esforço necessário para implementar um Sistema BI. **Conclusões:** Este trabalho enfatiza a importância de mensurar Sistemas de BI e, representa uma fonte de informação para ambos acadêmicos e desenvolvedores de aplicações BI que estão interessados em aumentar a precisão da estimativa de esforço para o desenvolvimento de Sistemas BI. Entretanto, a pesquisa comprova uma hipótese por meio de um estudo de caso, não sendo este o melhor procedimento técnico para este fim. Desta forma, é necessário executar experimentos que mostrem mais adequadamente a relação. Além disso, o método proposto não foi validado em casos reais, sendo este, portanto, um trabalho futuro relacionado.

Palavras-chave: BI. Métricas. Desenvolvimento de Sistemas BI. Complexidade de Processos de Negócio. Medição de Esforço.

Abstract

Objective: This study aims to identify metrics able to estimate the effort necessary for the development of a business intelligence system.

Methodology: A secondary study was conducted in order to identify metrics existing in the context of BI, and metrics to measure the complexity of business processes. From this, a hypothesis that relates complexity of procedures was drawn up business with the effort required to implement it in a BI system. this hypothesis was explored in a case study.

Results: It was shown that the hypothesis is true initially identified, although of the need for more studies to confirm this relationship. It has also been proposed a method that explores the relationship defined by the hypothesis. Thus, using this method, You can estimate the effort required to implement a BI system. conclusions: This work emphasizes the importance of measuring BI systems, and is a source information for both academics and developers of BI applications that are interested effort to increase the estimation accuracy for the development of BI systems. However, research by proving a hypothesis of a study case, which is not the best technical procedure for this purpose. Thus, it is necessary perform experipments that show better the relationship. In addition, method was not validated in real cases, which is therefore a further work related.

Keywords: BI. metrics. BI system development. process complexity. effort measurement.

Lista de abreviaturas e siglas

APF	Análise por Ponto de Função
BI	Business Intelligence
COSMIC	Common Software Measurement Consortium
CPD	Centro de Processamento de Dados
DM	Data Mart
DW	Data Warehouse
ETL	Extract-Transform-Load
IEEE	Institute of Electrical and Eletronics Engineers
ODS	Operational Data Store
OLAP	Online Analytical Processing
PCU	Pontos por Caso de Uso
PDCA	Plan-Do-Control-Act
PF	Pontos por Função
PMBok	Process Management Book of Knowledge
PMI	Project Managment Institute
RUP	Rational Unified Process
TCC	Trabalho de Conclusão de Curso
TI	Tecnologia da Informação
UC	Casos de Uso
UML	Unified Modeling Language

Sumário

1	INTRODUÇÃO	17
1.1	Justificativa	17
1.2	Objetivos	18
1.3	Metodologia	18
1.4	Organização do Trabalho	19
2	REFERENCIAL TEÓRICO	21
2.1	Bussiness Intelligence	21
2.1.1	Evolução	22
2.1.2	Benefícios e Dificuldades	22
2.1.3	Dados, Informação e Conhecimento	25
2.1.4	Componentes do BI	26
2.1.4.1	Dados Operacionais	26
2.1.4.2	Operational Data Store - ODS	26
2.1.4.3	Ferramentas de ETL (Extração, Transformação e Carga)	26
2.1.4.4	Data Warehouse e Data Mart	27
2.1.4.5	Mineração de Dados	27
2.1.4.6	Visualização dos Resultados	27
2.1.5	Modelo de Desenvolvimento de BI	28
2.1.5.1	Justificativa	28
2.1.5.2	Planejamento	28
2.1.5.3	Análise de Negócio	29
2.1.5.4	Desenho do Sistema	29
2.1.5.5	Desenvolvimento	29
2.1.5.6	Implementação	30
2.2	Medição de Software	30
2.2.1	Análise de Pontos por Função (APF)	31
2.2.2	Pontos por Caso de Uso (PCU)	33
2.2.3	COSMIC	34
2.3	Medição de Processo de Negócio	34
2.3.1	<i>McCabe's Cyclomatic Complexity</i>	34
2.3.2	<i>Control-Flow Complexity</i>	35
2.3.3	<i>Cognitive Weight</i>	37
2.3.4	<i>Number of Activities in a Process</i>	37
2.3.5	<i>Halstead-based Process Complexity</i>	38

3	METODOLOGIA	41
3.1	Visão Geral	41
3.2	Segunda Etapa do Estudo Secundário	41
3.2.1	Planejamento do Estudo	42
3.2.1.1	Objetivo	42
3.2.1.2	Questão de Pesquisa	42
3.2.1.3	Crerérios de Seleção de Trabalhos	42
3.2.1.4	Procedimentos de Seleção	43
3.2.1.5	Snowballing	43
3.2.1.6	Extração dos Dados	44
3.2.2	Condução da Segunda Etapa do Estudo Secundário	44
3.3	Estudo de Caso	51
3.4	Processo do Estudo de Caso	51
3.4.1	Contexto	51
3.4.2	Design e Planejamento	52
3.4.2.1	Hipótese	52
3.4.2.2	Objeto de Estudo	52
3.4.2.2.1	Bigdata	52
3.4.2.2.2	Equipe	53
3.4.2.2.3	Casos Selecionados	53
3.4.2.3	Papéis	53
3.4.2.4	Considerações Éticas	54
3.4.3	Definição de Conceitos e Medidas	54
3.4.4	Coleta de Dados	54
3.4.5	Análise de Dados	55
3.4.6	Relatório	55
4	RESULTADOS	57
4.1	Relação entre complexidade e esforço	57
4.1.1	Hipótese	57
4.1.2	Artefatos e Outros Aspectos do Caso	57
4.1.3	Processo de coleta de dados	60
4.1.4	Síntese dos resultados	61
4.1.5	Lições Aprendidas	62
4.2	Método CFCBI	63
4.2.1	Fase de Definição	63
4.2.1.1	Determinar escopo do projeto	64
4.2.1.2	Coletar documentação	64
4.2.1.3	Definir cronograma	65
4.2.1.4	Definir fatores de influência do projeto	65

4.2.1.5	Analisar produtividade da equipe	65
4.2.2	Fase de Aplicação	65
4.2.2.1	Analisar e classificar atividades do processo	66
4.2.2.2	Calcular valor da divisão AND	66
4.2.2.3	Calcular valor da divisão XOR	66
4.2.2.4	Calcular valor da divisão OR	66
4.2.2.5	Calcular o valor CFC	67
4.2.3	Fase de Conclusão	67
4.2.3.1	Calcular esforço	67
4.2.3.2	Atualizar base histórica	67
4.2.3.3	Relatar resultados	68
5	CONSIDERAÇÕES FINAIS	69
5.1	Resultados	70
5.2	Limitações e Trabalhos Futuros	70
	Referências	73
	 APÊNDICES	 77
	APÊNDICE A – PRIMEIRA ETAPA DO ESTUDO SECUNDÁRIO	79
A.1	Planejamento do Estudo	79
A.1.1	Objetivo	79
A.1.2	Questão de Pesquisa	79
A.1.3	Critérios de Seleção de Trabalhos	80
A.1.4	Procedimentos de Seleção	80
A.1.5	Avaliação de Qualidade	80
A.1.6	Extração dos Dados	81
A.2	Condução da Primeira Etapa do Estudo Secundário	81
	 ANEXOS	 85
	ANEXO A – TERMO DE CESSÃO DE DIREITO DE PESQUISA	87

1 Introdução

No cenário contemporâneo, a competição entre as empresas de um mesmo ramo enseja das companhias que se antecipem aos concorrentes quando na tomada de decisões gerenciais para modificar, inovar ou melhorar processos. É nesse contexto que surgiu a definição de Business Intelligence (BI), ou simplesmente Inteligência de Negócio. Apesar do conceito de BI ser relativamente novo, de 1989, a sua aplicação já era feita desde a Idade Antiga, em que as pessoas precisavam coletar, armazenar e cruzar informações. Naquela época, os dados recolhidos estavam relacionados, principalmente, à sobrevivência humana (PRIMAK, 2008).

A diferença fundamental entre as aplicações do conceito de BI antes e hoje é a tecnologia desenvolvida nos últimos anos, que resultou em facilidade e eficiência na execução de tarefas que antes poderiam ser consideradas difíceis ou inexistentes. Com o advento dos instrumentos da Tecnologia da Informação (TI), tornou-se mais fácil trabalhar com dados, da coleta à análise, passando pelo seu armazenamento (PRIMAK, 2008).

Segundo Gangadharan (2004) a principal utilização do BI nos dias atuais é agregar, em uma mesma ferramenta, variadas informações produzidas e disponibilizadas pelos mais diversos setores de uma organização. Assim, a alta administração tem acesso facilitado e sistematizado a um conjunto de dados passíveis de variadas análises, que subsidiarão a tomada de decisões gerenciais.

O desenvolvimento de sistemas de BI no campo da Engenharia de Software requer um planejamento diferenciado no que diz respeito ao estabelecimento de mecanismos de medida de tamanho para ajudar na tomada de decisões iniciais. Isso quer dizer que a estimativa de tamanho e, conseqüentemente, de prazo e custo para desenvolvimento de um sistema de BI, não pode ser encarada da mesma forma que em um sistema de software funcional. (GANGADHARAN, 2004)

1.1 Justificativa

Com rápidos avanços em tecnologia, as empresas frequentemente procuram por novos meios de estabilizar o valor da sua posição. Sistemas de Business Intelligence bem desenvolvidos podem prover a habilidade de analisar as informações de negócio com o intuito de realizar a tomada de decisões através do resultado das atividades do negócio. (BARA et al., 2009)

Sistemas de BI têm o potencial de maximizar o uso de informação, aumentando a capacidade da empresa em estruturar uma grande quantidade de informação, criando

assim uma vantagem competitiva. (BARA et al., 2009)

Segundo Davenport (2006) "enquanto investimentos em sistemas de BI continuam acelerados, há uma completa falta de um método específico e rigoroso para medir tais sistemas". A falta dessas métricas pode acarretar na dificuldade de monitorar e controlar o desenvolvimento de novas aplicações de Business Intelligence.

Após a realização da pesquisa, os profissionais da Engenharia de Software terão uma perspectiva, ainda que não aprofundada, a respeito da utilização de métricas para avaliação de tamanho de um sistema de BI, com conclusões obtidas a partir do teste e da coleta de resultados.

Assim, este trabalho tem como questão motivadora: Como estimar o esforço para desenvolver uma aplicação de BI?

1.2 Objetivos

O objetivo geral deste trabalho é propor um método de estimativa de esforço para o desenvolvimento de Sistemas de Inteligência de Negócio.

Para alcançar o objetivo geral deste trabalho, foram definidos os objetivos específicos abaixo:

- Revisar a literatura relacionada a BI e assuntos correlatos;
- Investigar métricas existentes para aferir complexidade de modelos de processos de negócio;
- Relacionar complexidade de modelos de processos de negócio com esforço requerido para customização de sistemas BI;
- Definir forma de uso de métrica de aferição de complexidade de processos de negócio para ser utilizada no contexto de estimativa de esforço para desenvolvimento de sistemas BI.

1.3 Metodologia

A metodologia adotada para o desenvolvimento deste Trabalho de Conclusão de Curso (TCC) consiste, basicamente, em quatro macroetapas conforme discriminado a seguir.

A primeira etapa foi constituída do Referencial Teórico dos assuntos Medição de Software, Business Intelligence e Medição de Complexidade de Processo de Negócio. No campo da BI, foi tratada a parte de desenvolvimento de sistemas, conceituação, evolução e

arquitetura. No âmbito de medição de software, foi estudado a sua definição, importância e foram citados algumas metodologias de medição de complexidade.

Na segunda etapa, foi levantado um conjunto de métricas, baseado em pesquisa na literatura de trabalhos relacionados e levantamento de métricas já existentes e na adaptação delas para o contexto de Inteligência de Negócio.

Na terceira etapa, foi realizado um estudo de caso sobre a utilização das métricas desenvolvidas em sistemas reais de BI. Esta fase necessitou um esforço específico, que incluiu o planejamento, design de experimento e execução.

Na quarta e última etapa, foi proposto um método de aplicação da métrica levantada e parcialmente comprovada através do estudo de caso. Esta etapa finaliza o TCC respondendo a questão levantada na Introdução.

1.4 Organização do Trabalho

Este capítulo introdutório apresentou o contexto, a justificativa, a questão, o objetivo e a metodologia que norteou o desenvolvimento desta pesquisa.

A organização textual deste trabalho segue a estrutura abaixo:

- Capítulo 2 – Referencial Teórico: apresenta uma revisão de literatura, onde são tratados os temas fundamentais para que seja possível entender a que se propõe o trabalho. Dentre os temas tratados encontram-se: Medição de Software, Business Intelligence, Revisão Sistemática e Estudo de Caso.
- Capítulo 3 – Metodologia: apresenta passo-a-passo que foi realizado para se chegar aos resultados. O capítulo está constituído das duas etapas de estudo secundário, baseado em Revisão Sistemática, e no planejamento e design do estudo de caso.
- Capítulo 4 - Resultados: apresenta o resultado obtido através dos estudos secundários realizados, uma lista compilada das métricas utilizadas para mensurar Processos de Negócios. Apresenta, também, a comprovação da hipótese levantada no estudo de caso utilizando uma métrica selecionada após o estudo secundário. A seção apresenta, também, uma proposta de aplicação da métrica selecionada após o Estudo de Caso.
- Capítulo 5 – Considerações Finais: apresenta um resumo do que foi feito, dos resultados alcançados e das limitações e trabalhos futuros.

2 Referencial Teórico

Esse referencial teórico relaciona os temas importantes para fundamentar a proposta deste trabalho. Encontra-se dividido em seções, de forma que cada uma aborda assuntos importantes para o entendimento do que tratará o projeto, conforme explicitado a seguir.

Na Seção 2.1 são abordados os conceitos de Business Intelligence, a sua evolução, os benefícios de se utilizar um BI e as dificuldades de sua implementação. Diferencia, também, dados, informações e conhecimento; conceitua as ferramentas envolvidas no processo de BI e promove a integração entre elas.

Na Seção 2.2 são apresentados conceitos de medição de software. Esta conceituação é feita a partir da visão e da obra de diferentes autores, o que permite consolidar o conhecimento sobre o assunto. Também mostra a importância de se utilizar medição, a diferença entre alguns conceitos-chave da medição de software (medida, métrica e indicador) e apresenta, ainda, as principais métricas para medir o tamanho de sistemas: Análise por Ponto de Função, Pontos por Caso de Uso e Pontos COSMIC.

Finalmente a Seção 2.3 apresenta a definição de métricas utilizadas para mensurar a complexidade de processos de negócio. Tais métricas foram levantadas na execução da segunda etapa do estudo secundário do presente trabalho.

2.1 Business Intelligence

De forma direta, Inteligência de Negócio ou Business Intelligence (BI) é um conjunto de conceitos e metodologias que utiliza-se de dados extraídos da organização e serve de apoio à tomada de decisões. [Barbieri \(2001\)](#) define o conceito de BI como “[...]a utilização de variadas fontes de informação para se definir estratégias de competitividade nos negócios da empresa”. Há diversos outros autores que definem, de forma mais detalhada, o conceito de Business Intelligence, como [Angeloni e Reis \(2006\)](#):

"[...] o conceito de Business Intelligence com o entendimento de que é Inteligência de Negócios ou Inteligência Empresarial compõe-se de um conjunto de metodologias de gestão implementadas através de ferramentas de software, cuja função é proporcionar ganhos nos processos decisórios gerenciais e da alta administração nas organizações, baseada na capacidade analítica das ferramentas que integram em um só lugar todas as informações necessárias ao processo decisório. Reforça-se que o objetivo do Business Intelligence é transformar dados em conhecimento, que suporta o processo decisório com o objetivo de gerar vantagens competitivas."([ANGELONI; REIS, 2006](#))

A utilização do Business Intelligence vem crescendo consideravelmente na medida em que a aplicação dos seus conceitos e práticas possibilitam às organizações realizar uma grande quantidade de análises e projeções, apoiando, assim tomada de decisões gerenciais (PRIMAK, 2008).

2.1.1 Evolução

Apesar de o termo Business Intelligence ser relativamente novo, datado do ano de 1989 pela empresa Gartner, a sua aplicação já era feita na Antiguidade. Povos do Oriente Médio já colocavam em prática os princípios mais comuns de BI quando precisavam cruzar informações obtidas com base de observação nos fenômenos da natureza em prol de suas aldeias. Algumas aplicações iam da análise das marés até o posicionamento dos astros, resultando em informações utilizadas a favor da melhoria de vida dos habitantes das aldeias. (PRIMAK, 2008)

Serra (2002) diz que a história contemporânea do BI teve início na década de 1970, quando alguns produtos de BI foram fornecidos para os analistas de negócios. A problemática enfrentada com esses produtos estava relacionada às exigências de um esforço muito grande em programação, o alto custo de implantação e a demora em disponibilizar a informação.

Em termos de tecnologia, a evolução da TI foi um fator crucial para a utilização das ferramentas de BI: os computadores ficaram menores e mais potentes e os dados, agora, poderiam ser tratados de forma mais rápida e precisa - dados que, após a sua devida análise, se transformavam em valiosas informações gerenciais.

A Figura 1 mostra a evolução do BI

2.1.2 Benefícios e Dificuldades

As ferramentas de BI, dependendo de como as organizações as utilizam, podem trazer grandes benefícios. É importante frisar que a forma de implantação da ferramenta irá ditar o sucesso ou o insucesso da análise de dados.

Os benefícios obtidos por meio da implantação bem sucedida das ferramentas de BI, segundo Primak (2008), são:

- Redução de custos com softwares;
- Redução de custos com administração e suporte;
- Redução de custos na avaliação de projetos;
- Redução de custos com treinamentos aos colaboradores;

Período	O que aconteceu?
Idade Antiga	<ul style="list-style-type: none"> • Povos do Oriente Médio cruzavam informações da natureza, marés, etc. para analisarem a viabilidade do cultivo de determinadas espécies, períodos de pesca abundante, etc.
Século XVI	<ul style="list-style-type: none"> • Rainha Elizabeth I, determinou que a base da força inglesa fosse “informação e comércio” e ordenou então ao filósofo Francis Bacon que inventasse um sistema dinâmico de informação, o qual foi amplamente aplicado pelos ingleses.
Década de 60	<ul style="list-style-type: none"> • Evolução dos computadores que deixaram de ocupar salas enormes, na medida em que diminuíram de tamanho. • As empresas passaram a perceber os dados como uma possível e importante fonte geradora de informações decisórias e que renderiam eventuais lucros. • Época dos cartões perfurados. • Predominância das linguagens de programação COBOL, ASSEMBLY e FORTRAN.
Década de 70	<ul style="list-style-type: none"> • Evolução das formas de armazenamento e acesso a dados – DASD e SGBD (Sistema Gerenciador de Banco de Dados). • Desenvolvimento do EIS.
Década de 80	<ul style="list-style-type: none"> • Desenvolvimento e evolução das linguagens de programação CLIPPER e PASCAL. • Início da aplicação do termo Business Intelligence.
Década de 90	<ul style="list-style-type: none"> • Início do termo CPD nas empresas. • Desenvolvimento do Data Warehouse. • O setor corporativo passou a se interessar pelas soluções de BI.
Década de 2000	<ul style="list-style-type: none"> • Evolução dos conceitos de DSS (Decision Support System – sistema de suporte à decisão). Planilhas Eletrônicas, Geradores de Consultas e de Relatórios, Data Marts, Data Mining, Ferramentas OLAP, entre outras. • Estreitamento das conexões entre BI e ERP. • Surgimento do termo E-business.

Figura 1 – Evolução do BI (PRIMAK, 2008)

- ROI (Retorno sobre Investimento) mais rápido;
- Maior controle e menos dados incorretos;
- Maior segurança da informação;
- Alinhamento de informações estratégicas e operacionais;
- Facilidade de controle de acesso e definição de níveis de gerência;
- Melhor alinhamento dos usuários corporativos;
- Rapidez na informação para tomada de decisões estratégicas;
- Informação consistente em vários locais dispersos;
- Vantagem competitiva.

Apesar dos excelentes benefícios obtidos por meio da implantação das ferramentas de BI serem excelentes, sua implementação deve ser cautelosa e bem planejada. Primak (2008) cita as principais dificuldades de se implementar um BI:

- Dados operacionais estão dispersos e, muitas vezes, incoerentes com a organização;
- Deficiência dos sistemas operacionais utilizados pelas organizações, que não armazenam dados úteis para futura tomada de decisão;
- A organização não reconhece as necessidades de informação, somente identificando-as quando, muitas vezes, é tarde demais;
- A falta de conhecimento dos gestores pode deixar um projeto de BI sem utilidade prática;
- Necessidade de uma boa inter-relação entre a área de negócio e a equipe de tecnologia da informação;
- As ferramentas técnicas operacionais da atualidade são dispersas e ineficientes e necessitam de uma reconstrução para serem utilizadas para o BI;
- A obtenção de informações de diversas fontes externas é feita de uma maneira que a relação custo-benefício pode não ser favorável;
- Alguns projetos falham em decorrência de uma adoção equivocada de hardware e software;
- Muitos problemas podem ocorrer devido à falta de experiência e conhecimento do fornecedor da ferramenta de BI;
- O tratamento dos dados (ETL) e o armazenamento (DW/DM) são um processo que deve ser bem planejado, em decorrência de ser trabalhoso e complexo, necessitando de profissionais de alto gabarito para garantir o sucesso desta etapa;
- Simples erros na elaboração e desenvolvimento de um DW podem ser fatais e trazer resultados negativos ao projeto;
- Dificuldade de realizar o nivelamento entre o BI e a gestão do conhecimento da organização;
- O custo para implementação de um projeto de BI não é barato.

2.1.3 Dados, Informação e Conhecimento

O BI, na sua essência, trata apenas da tomada de decisão eficaz. No entanto, é de fundamental importância saber de onde partiu determinada decisão no âmbito da empresa. Nenhuma decisão é tomada por acaso; elas ensejam a existência de dados, que se transformam informações, e, junto com o conhecimento, levam à tomada de decisões. A Figura 2 reflete a relação Dados X Informação X Conhecimento X Decisão.

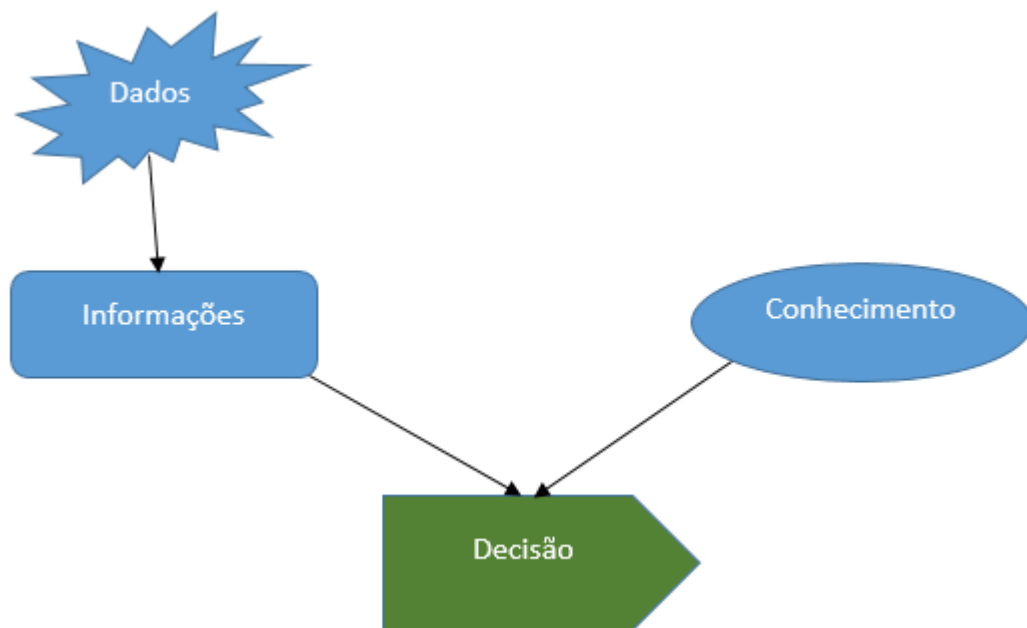


Figura 2 – Dados x Informação x Conhecimento x Decisão (ANTONELLI, 2009)

Dados, segundo Alter (1996), são "[...] fatos, imagens ou sons que podem ser pertinentes ou úteis para uma tarefa particular". Dados sozinhos, geralmente, não trazem nenhuma significância.

Informação, segundo Romney e Steinbart (2000), é um conjunto de "[...] dados que foram organizados e processados de forma que sejam significativos". É possível aferir que a informação é a organização dos dados de forma pré-determinada e pode ser aplicada em uma ou mais situações.

Conhecimento, segundo Alter (1996), é "[...] uma combinação de instintos, ideias, regras e procedimentos que guiam ações e decisões". Ou seja, conhecimento agregado as informações gera decisões.

O termo **decisão** pode ser definido com o processo pelo qual são escolhidas algumas ou várias ações a serem realizadas. Esta escolha deve ser pautada na maior quantidade de informação e conhecimento possível.

2.1.4 Componentes do BI

O BI é composto por ferramentas que realizam trabalhos de armazenamento, mineração e análise dos dados. A Figura 3 representa essas ferramentas e a interação que elas possuem umas com as outras.

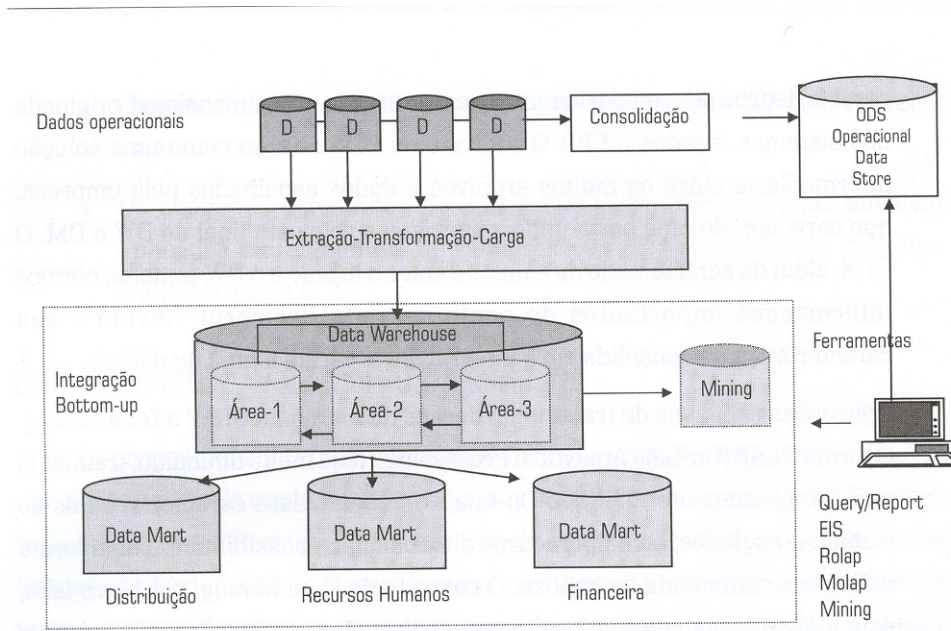


Figura 3 – Componentes de um ambiente de BI (BARBIERI, 2001)

As subseções seguintes irão detalhar cada um dos elementos presentes na Figura 3.

2.1.4.1 Dados Operacionais

São todos os dados relevantes para a empresa que deseja implementar o BI. São originados das operações realizadas na organização e armazenados no sistema de informação utilizado. É a matéria-prima do Business Intelligence, o que vai ser trabalhado e transformado em informação.

2.1.4.2 Operational Data Store - ODS

Segundo Primak (2008), *Operational Data Store* tem a função de realizar “[...] o armazenamento de dados operacionais de forma consolidada, porém não possui características dimensionais”. O ODS é, portanto, uma ferramenta especial de armazenamento de informações.

2.1.4.3 Ferramentas de ETL (Extração, Transformação e Carga)

É uma das ferramentas mais importantes e fundamentais para o processo de BI. As ferramentas de ETL são responsáveis pelo tratamento dos dados antes de eles serem

armazenados em um Data Warehouse (DW). Primak (2008) diz ainda que a etapa de ETL é uma das mais críticas de um DW, principalmente porque envolve a movimentação de grande volume de dados.

2.1.4.4 Data Warehouse e Data Mart

Data Warehouses (DW) e Data Marts (DM) são estruturas especiais de armazenamento de informações capazes de sustentar a camada de inteligência da empresa, com o intuito de serem utilizadas na tomada de decisão como elementos diferenciais e competitivos. De forma mais clara, DW e DM são dados organizados por temas de interesse e integrados por data, facilitando o gerenciamento de grandes quantidades de dados e modelando-os para ajudar na tomada de decisão.

A definição dada por Primak (2008) ajuda a compreender a principal diferença entre um Data Warehouse e um Data Mart de um banco de dados comum. Segundo ele, um banco de dados comum é uma “[...] coleção de dados operacionais armazenados e utilizados pelo sistema de aplicação de uma empresa específica”. Já DW e DM, segundo Primak (2008), é “[...] uma coleção de dados derivados dos dados operacionais para sistema de suporte a decisão. Esses dados derivados são, muitas vezes, referidos como dados gerenciais, informacionais e analíticos”. Os conceitos dados por Primak (2008) sugerem que DW/DM armazenam dados de forma mais eficiente, possibilitando uma consulta mais eficaz e rápida. Uma das tecnologias que possibilitam essa melhora na consulta é a ferramenta OLAP (Online Analytical Processing), que armazena os dados em formato de cubo multidimensional, facilitando a agregação rápida de dados e detalhamento de análises.

Vale salientar que as tecnologias presentes no Data Warehouse e no Data Mart são as mesmas; entretanto, há variações no escopo de criação. O DM é um DW departamental, um DW reduzido, e não fornece informação para toda a empresa, mas somente para o seu departamento.

2.1.4.5 Mineração de Dados

Para que o uso de DW e DM seja interessante, é necessário utilizar boas ferramentas de mineração de dados (Data Mining), que também é um fator decisivo para o processo de BI. Primak (2008) diz que Data Mining está mais relacionado ao processo de análise de inferência que com o processo de análise dimensional. A forma de fazer mineração de dados é buscar uma informação através de algoritmos que objetivam o reconhecimento de padrões escondidos nos dados.

2.1.4.6 Visualização dos Resultados

O resultado obtido por meio de todo o processo descrito acima precisa ser visualizado de forma amigável pelo usuário do BI e, para isso, é necessário utilizar ferramentas

que auxiliem nessa visualização.

2.1.5 Modelo de Desenvolvimento de BI

Segundo [BARA et al. \(2009\)](#), existem diversas diferenças entre o modelo de desenvolvimento de sistemas transacionais e sistemas de BI que dependem de características de sistemas de decisões. Entretanto, as mesmas técnicas e fases são utilizados para o desenvolvimento: estudo inicial, planejamento do projeto, análise, desenho, construção e implementação. A Figura 4 ilustra as fases supracitadas.



Figura 4 – Fases de desenvolvimento de sistemas BI ([BARA et al., 2009](#))

Além das fases de desenvolvimento, [BARA et al. \(2009\)](#) define os passos necessário para atingir o objetivo de cada fase definida.

2.1.5.1 Justificativa

A fase de justificativa é composta pelo passo de **Avaliação de Caso de Negócio**, onde as necessidades e oportunidades são identificadas, e, então a equipe propõe uma solução inicial justificando os custos e benefícios. Um relatório é gerado nessa fase.

2.1.5.2 Planejamento

A fase de planejamento é composta por dois passos. A **Avaliação da Infraestrutura da Organização**, onde é avaliada a capacidade de manter e apoiar um projeto de

desenvolvimento de sistemas BI em termos de: infraestrutura, componentes, equipamento, rede e equipamentos futuros.

O passo de **Planejamento de Projeto** consiste na planejamento dinâmico da tecnologia utilizada, organização, necessidade de negócio, recursos humanos e equipe de implementação.

2.1.5.3 Análise de Negócio

A fase de análise de negócio é composta por quatro passos. A **Definição da Necessidade de Negócio** consiste em entrevistas e reuniões com executivos e gerentes, a necessidade do negócio e os requisitos são identificados e definidos.

O passo de **Análise de Dados** envolve a identificação e desenho das fontes de dados, contém os modelos de dados detalhados com os atributos e referências entre dados. O modelo lógico é desenhado.

O passo de **Prototipação** consiste na construção e teste de uma versão inicial com objetivo de validar a necessidade de negócio.

O passo de **Análise de Metadados** consiste no desenho de metadados e no mapeamento das fontes de dados na estrutura dos metadados.

2.1.5.4 Desenho do Sistema

A fase de desenho do sistema é composta por três passos. O **Desenho dos Dados** consiste no detalhamento do modelo lógico. O modelo de dados para processar e armazenar é escolhido.

O passo de **Desenhar o Processo ETL** é o mais difícil de todo o ciclo e depende da qualidade da fonte de dados. É recomendado que o processo deva ser construído em uma organização que integra todos os módulos da organização.

O passo de **Desenhar o Repositório de Metadados** consiste no desenho em termos do modelo lógico do metadado: relacional, orientado à objeto ou multidimensional.

2.1.5.5 Desenvolvimento

A fase de desenvolvimento é composta por cinco passos. O **Desenvolvimento da ETL** consiste nas ferramentas de filtro, procedimento e operadores são usados para construir o processo de ETL. O filtro de dado e transformação depende da qualidade da fonte de dados.

O passo de **Desenvolvimento da Aplicação** consiste na construção de templates, interface e procedimentos, direitos e privilégios dos usuários são definidos.

O passo de **Mineração de Dados** consiste no teste de algoritmos, técnicas de mineração de dados, predição e organização dos métodos.

O passo de **Desenvolver o Repositório de Metadados** consiste no desenvolvimento do dicionário de metadados e na interface de acesso aos dados.

2.1.5.6 Implementação

A fase de implementação é composta por dois passos. A **Implementação** é a entrega do processo onde a equipe de desenvolvimento organiza treinamentos para gerentes, a documentação final e o suporte técnico são entregues.

O passo de **Teste do Sistema** consiste na construção do relatório final contendo as conclusões preliminares, a performance do sistema e algumas partes onde pode haver melhorias.

2.2 Medição de Software

A determinação da proporção entre a dimensão de um atributo do objeto e uma determinada unidade de medida se denomina medição. Segundo [Fenton \(1991\)](#), “[...] medição é o processo pelo qual números ou símbolos são anexados aos atributos de entidades no mundo real para defini-los de acordo com regras claramente estabelecidas”. A aplicação da medição em Engenharia de Software é de extrema importância para aumentar nossa percepção de entidades particulares e o resultado da medição auxilia na tomada de decisão em uma organização. A utilização da medição resulta em um maior entendimento dos atributos do sistema e, conseqüentemente, ajuda na avaliação da qualidade dos softwares que foram desenvolvidos através de um processo da organização. Cada métrica de processo coletada por meio da atividade de medição contribui para o aumento da qualidade do software. Segundo [Sommerville \(2007\)](#), há duas maneiras de utilizar as medições de software: para realizar previsões gerais sobre um sistema e para identificar componentes atípicos.

[Pressman \(2011\)](#) diz que, para ajudar engenheiros de software a ter uma visão total do projeto e da construção do software, é necessária a utilização de métricas de produto, focalizando os atributos específicos e mensuráveis dos artefatos da Engenharia de Software. Ainda de acordo com [Pressman \(2011\)](#), o engenheiro de software não pode se basear apenas em aspectos qualitativos, sendo necessários critérios objetivos para ajudá-lo no direcionamento do projeto de dados, da arquitetura, das interfaces e dos componentes. As métricas de produto proporcionam a base para conduzir e avaliar, objetivamente, os dados quantitativos nas etapas de análise, projeto e codificação.

No contexto de medição, existem termos que ajudam a consolidar o entendimento

na área: medida, métrica e indicador. [Pressman \(2011\)](#) define que “[...] medida proporciona uma indicação quantitativa da extensão, quantidade, capacidade ou tamanho de algum atributo de um produto ou processo”. Medida é, portanto, um único ponto de dado coletado, como, por exemplo, o número de linhas de código de um sistema bancário. Métrica, segundo o [IEEE \(1990\)](#), é definida como “[...] uma medida quantitativa do grau com o qual um sistema, componente ou processo possui determinado atributo”. [Sommerville \(2007\)](#), por sua vez, define métrica de software como “[...] qualquer tipo de medição que se referia a um sistema de software, processo ou documentação relacionada”. Isto posto, pode-se dizer que métrica é o relacionamento entre as medidas individuais coletadas, ilustrando, como exemplo, a média de linhas de códigos por funções do sistema.

Já os indicadores são obtidos por meio da coleta de medidas e/ou do desenvolvimentos de métricas por Engenheiros de Software. [Pressman \(2011\)](#) diz que “[...] um indicador é uma métrica ou combinação de métricas que proporcionam informações sobre o processo de software, em um projeto de software, ou no próprio produto”. Os indicadores, normalmente, são representados por meio de tabelas e gráficos e fornecem profundidade na visão do processo, do projeto e do produto final, para auxiliar no processo de tomada de decisão.

As métricas de produto, segundo [Sommerville \(2007\)](#), “[...] se dedicam às características do próprio software”. O principal desafio encontrado na medição de produto está relacionado à complexidade do software. Enquanto existirem diferentes visões sobre o que é complexidade e quais os atributos de um sistema que levam à complexidade, fica impossível estabelecer um valor único dessa métrica de qualidade. Apesar da dificuldade existente, é necessário medir e controlar a complexidade do software e, para isso, diversas medidas são coletadas de diferentes atributos internos. Todas as medidas e métricas provenientes dos atributos são utilizadas como indicadores dos modelos de requisito e produto.

Nas subseções abaixo serão abordados três metodologias de medição de tamanho de software, a análise de pontos por função, o pontos de caso de uso e a métrica COSMIC.

2.2.1 Análise de Pontos por Função (APF)

Existem diversas métricas de produto diferentes e uma das mais usadas é a métrica Ponto por Função (PF) utilizada para medir a funcionalidade do sistema. [Dekkers \(1998\)](#) reafirma que a métrica PF mede o tamanho do que o sistema faz, e não como ele é desenvolvido e implementado. A Análise de Pontos por Função (APF) é um método padrão para dimensionar o tamanho funcional de um software, sob a perspectiva do usuário. É uma métrica que independe de implementação física e linguagens de programação. Os objetivos da APF, definidos por [IFPUG \(2005\)](#) são:

- Medir o software, quantitativamente, por meio das funcionalidades solicitadas e recebidas pelo cliente;
- Medir o desenvolvimento e a manutenção do sistema, independentemente da implementação e da linguagem utilizada; e
- Realizar a medição em todos os projetos e organizações.

Com base em dados históricos, é possível empregar a PF para, segundo [Pressman \(2011\)](#), “[...] estimar o custo ou trabalho necessário para projetar, codificar e testar o software; prever o número de erros que serão encontrados na fase de testes; e prever o número de componentes e/ou linhas projetadas de código-fonte no sistema implementado”. Os principais benefícios obtidos com a utilização da APF, segundo [IFPUG \(2005\)](#), podem ser representados como:

- Uma ferramenta para determinar o tamanho de um pacote de software;
- Uma forma de estimar custos e recursos necessários para desenvolver e realizar a manutenção de um sistema; e
- Uma ferramenta que apoia a análise de produtividade e de qualidade por meio da medição das unidades de um produto de software.

Para realizar a contagem, é necessário que haja um projeto lógico e sejam seguidos os seguintes passos, definidos por [IFPUG \(2005\)](#) para realizar a análise:

1. Determinação do propósito da contagem
2. Determinação do tipo de contagem
3. Definição do escopo da contagem
4. Identificação da fronteira da aplicação
5. Identificação das funcionalidade do software
6. Cálculo dos PFs sem ajuste
7. Determinação da influência das características gerais da aplicação
8. Cálculo do fator de ajuste
9. Cálculo dos pontos por função ajustados

2.2.2 Pontos por Caso de Uso (PCU)

Outra métrica utilizada para analisar o tamanho de projeto do software orientado a objetos é a métrica Use Case Point (UCP, ou Pontos por Caso de Uso). [Sommerville \(2007\)](#) diz que Casos de Uso (UC, da sigla original em inglês) “[...] constituem uma técnica baseada em cenários para elicitación de requisitos”. Os UCs são representados pela notação UML (Unified Modeling Language) para descrição de modelos de sistema orientados a objetos. Os diagramas de Casos de Uso são compostos pelos atores envolvidos e pela sua interação com o sistema. A métrica UCP foi desenvolvida por Gustav Karner, da Object Systems, hoje IBM, e foi concebida por desenvolvedores que utilizavam UML e RUP (Rational Unified Process)([JONES, 2008](#)).

Segundo [Jones \(2008\)](#), a métrica Pontos por Caso de Uso se inicia com os fatores básicos da APF, mas adiciona muitos outros parâmetros, como a relevância dos atores do sistema. Outra diferença é que a UCP também inclui alguns tópicos subjetivos, como capacidade do analista líder, motivação e familiaridade com UML.

As vantagens da utilização da métrica UCP, segundo [Jones \(2008\)](#), são:

- Uso da UML, amplamente utilizada em desenvolvimento web;
- Suporte da IBM, base de projeto grande;
- Técnica simples; e
- Não necessita certificação.

As desvantagens, ainda segundo [Jones \(2008\)](#) são:

- Os Casos de Uso variam muito e a técnica UCP só usa três níveis de complexidade; e
- Há 21 fatores que influenciam o tamanho do software, mas que deveriam influenciar apenas a produtividade.

Os passos para realizar a contagem, segundo [Cohn \(2011\)](#), são:

1. Classificar e calcular os atores
2. Classificar e calcular os casos de uso
3. Calcular UCP não ajustado (soma dos atores com os casos de uso)
4. Calcular complexidade técnica

5. Calcular complexidade ambiental
6. Calcular UCP ajustado

2.2.3 COSMIC

A terceira métrica a ser tratada nessa revisão de literatura é a COSMIC (Common Software Measurement International Consortium). Segundo Jones (2008), é uma das métricas mais atuais de medição funcional, com sua primeira publicação em 1999, aproximadamente. Entretanto, as pesquisas relacionadas ao método COSMIC foram iniciadas anos antes. A métrica é aplicada em sistemas de negócios e em softwares em tempo real, não sendo aplicável, por outro lado, em softwares com uso de algoritmos matemáticos complexos.

As principais vantagens da utilização da métrica COSMIC, segundo Jones (2008), são:

- Possibilidade de realizar estimativas em fases iniciais de projeto;
- Identificação de camadas;
- Maior flexibilidade da ferramenta;
- Aplicabilidade mais simples e menos ambígua; e
- Possibilidade de aplicar a visão do usuário final e do desenvolvedor.

Já a principal desvantagem de se utilizar a métrica COSMIC é a ainda pequena existência de relatos sobre a experiência prática em sua utilização.

As metodologias supracitadas foram selecionadas para exemplificar a existência de métricas capazes de medir o tamanho do software a ser desenvolvido.

2.3 Medição de Processo de Negócio

Esta seção apresenta a definição de cinco métricas de processo de negócio encontradas através de uma abordagem de estudo secundário. Ao todo, foram identificadas 21 métricas. São elas: *McCabe's Cyclomatic Complexity*, *Control-Flow Complexity*, *Cognitive Weight*, *Number of Activities in a Process* e *Halstead-based Process Complexity*.

2.3.1 McCabe's Cyclomatic Complexity

Segundo Solichah et al. (2013), existem diversas métricas de complexidade de software publicadas na literatura. Uma das mais populares é a métrica *McCabe's Cyclomatic*

Complexity (MCC), que "[...] conta o número de caminhos comuns que podem gerar todos os caminhos possíveis através do gráfico de fluxo de controle". O gráfico de fluxo de controle é construído com o apoio de um programa de software, considerando apenas estruturas de controle (*if-else*, *loop* e *switch*).

McCabe derivou a medida de complexidade de software da Teoria dos Gráficos usando a definição de número ciclomático, que corresponde ao número de caminhos lineares independentes em um programa. Desde a sua criação, a métrica MCC tem sido amplamente utilizada e aplicada em mais de dez milhões de linha de código em aplicações comerciais e no Departamento de Defesa Americano. A base de conhecimento resultante da aplicação da métrica possibilitou que desenvolvedores de software fossem capazes de calibrar as medidas dos seus próprios programas e chegassem a um entendimento comum do que seria complexidade. (CARDOSO et al., 2006)

A métrica McCabe, segundo Cardoso et al. (2006), é um indicador da complexidade do fluxo de controle do módulo de um programa e se tornou um indicador confiável da complexidade em grandes projetos de software.

McCabe (1976) define a métrica como:

$$v = e + (n - 2)$$

onde e (*edge*) se refere ao vértice e n (*nodes*) se refere ao nó do gráfico de fluxo de controle. Os nós representam declarações e expressões computacionais e o vértice representa a transferência de controle entre nós.

Para cálculo da complexidade de Processos de Negócio, a adaptação é direta. Segundo Thammarak (2010), os nós são aplicados às atividades e os vértices são os fluxos entre as atividades.

2.3.2 Control-Flow Complexity

Cardoso (2008) propôs a métrica Control-Flow Complexity (CFC) para medir a complexidade do fluxo de controle de modelos de Processos de Negócio. Esta é uma métrica adaptada da métrica McCabe's Cyclomatic Complexity (MCC). Enquanto a métrica MCC atribui a mesma semântica para todos os nós de decisão, a métrica CFC define os vários nós como portadores de semânticas diferentes, como *AND-Split*, *XOR-Split* e *OR-Split*.

Para ajudar no entendimento da métrica CFC, Cardoso (2008) introduz alguns conceitos chaves que constituem a base para cálculo da métrica.

O **processo** é definido como uma coleção de atividades que necessita de uma ou mais entradas e cria uma saída que é o valor para o cliente (CARDOSO, 2008 apud DAVENPORT, 1993).

A **propriedade do processo** é uma característica ou um atributo do processo como a complexidade, manutenibilidade, custo, confiabilidade, entre outros. As propriedades do processo podem ser avaliadas e quantificadas utilizando modelos, métodos ou algoritmos (CARDOSO, 2008).

A **medida do processo** é uma atribuição empírica de um número no processo com o objetivo de caracterizar uma propriedade específica (CARDOSO, 2008).

A **métrica do processo** é qualquer tipo de medição relacionada ao processo. As métricas permitem a quantificação de atributos do processo (CARDOSO, 2008).

O **fan-out de uma atividade** é o número de transições saindo de uma atividade e é computado através da função $fan-out(a)$ (CARDOSO, 2008).

O **estado mental do fluxo de controle induzido** deve ser considerado quando o processo está sendo modelado. Uma divisão (*Split*) introduz a noção de estado mental em processos.

A definição de **XOR-Split**, segundo Cardoso (2008), consiste no número de estados mentais que são introduzidos com esse tipo de divisão exclusiva. Para o cálculo do $CFC_{XORsplit}(a)$, deve-se simplesmente somar a quantidade de fluxos saindo de uma atividade.

A definição de **OR-Split** também consiste no número de estados mentais que são introduzidos com esse tipo de divisão não exclusiva. Para o cálculo do $CFC_{ORsplit}(a)$, a fórmula é

$$2^n - 1$$

onde n é a quantidade de fluxos saindo de uma atividade. (CARDOSO, 2008)

Por último, Cardoso (2008) define a complexidade de **AND-Split** simplesmente como a saída única de uma atividade:

$$CFC_{ANDsplit}(a) = 1$$

Matematicamente, a métrica CFC é aditiva. Embora seja muito fácil calcular a complexidade absoluta de um processo simplesmente adicionando os valores de todos os *Splits*:

$$CFC_{abs} = CFC_{ORsplit}(a) + CFC_{XORsplit}(a) + CFC_{ANDsplit}(a)$$

a complexidade relativa é resultante da razão entre a complexidade absoluta pela quantidade de *Splits* existentes no processo, segundo Cardoso (2008):

$$CFC_{abs}/\Sigma_{Splits}$$

A métrica CFC foi escolhida para dar continuidade a pesquisa, ela é a métrica que teve o maior número de referências em trabalhos acadêmicos.

2.3.3 *Cognitive Weight*

Shao e Wang (2003) definiram *Cognitive Weight* (CW) ou Peso Cognitivo como "[...] uma métrica para medir o esforço necessário para compreender um pedaço de software". Baseados em estudos empíricos, Shao e Wang (2003) definiram CW para estruturas básicas de controle de software. Cada estrutura básica de controle possui um peso atrelado a ela (ex.: estrutura com *if-then* ou *if-else* tem peso 2). Para realizar o cálculo de complexidade de um componente de software, é necessário realizar o somatório dos CW, como mostra a fórmula a seguir:

$$\Sigma CW$$

Para utilizar a métrica CW no contexto de Processo de Negócio, é preciso considerar diferenças importantes entre a modelagem do Processo de Negócio e programas de software.

Segundo Gruhn e Laue (2006a), a complexidade de um software depende em três fatores: processamento interno, entradas e saídas. O tamanho cognitivo de um componente de software é uma função do Peso Cognitivo das estruturas de controle e das entradas e saídas. Entretanto, Processo de Negócio frequentemente mostra um controle de fluxo, mas omitem as entradas e saídas que possuem outra terminologia no Processo de Negócio. Por esse motivo, Gruhn e Laue (2006a) não levam em consideração as entradas e saídas para cálculo da complexidade do Processo de Negócio. A métrica *Cognitive Weight* utiliza as definições de Cardoso (2008) para as divisões *XOR*, *OR* e *AND*.

Gruhn e Laue (2006a) estipulam pesos para cada elemento de um processo de negócio, conforme mostra a Tabela 1. Para realizar o cálculo da complexidade do processo de negócio, basta efetuar o somatório da multiplicação dos elementos pelo peso.

2.3.4 *Number of Activities in a Process*

A métrica *Number of Activities in a Process* (NOA) é adaptada por Cardoso et al. (2006) de uma métrica já existente na Engenharia de Software, a *Lines of Code* (LOC). A LOC é uma das métricas mais antigas e fundamentais, baseada na contagem simples de linhas de código de um programa.

Apesar de ser bastante criticada como uma métrica de complexidade, a sua popularidade cresceu devido à sua simplicidade em ser utilizada. A base da métrica é que o tamanho do código do programa pode ser utilizado para prever algumas características, como a ocorrência de erros, a confiabilidade e a facilidade de manutenção. (CARDOSO et al., 2006)

Segundo Cardoso et al. (2006), se a atividade de um processo for enxergada como uma declaração de um programa de software, pode-se derivar uma métrica muito trivial que simplesmente conta o número de atividades (NOA) em um processo de negócio. A

Tabela 1 – Pesos Cognitivos para Elementos BPM com Adaptações (GRUHN; LAUE, 2006a)

Padrões de Workflow	Estrutura de controle BPM	W
Sequência	Passos consecutivos em um workflow	1
Escolha exclusiva	XOR-Split (escolher um de duas opções)	2
	XOR-Split (escolher um de \geq três opções)	3
Divisão paralela	AND-Split que ativa atividades paralelas	4
Múltipla escolha	OR-Split (escolher mais de uma opção)	7
Nenhum	Subtarefa	2
Padrão de instância múltipla	Múltiplas instâncias de uma atividade ocorrendo concorrentemente	6
Atividade cancelada	Escolher uma atividade que cancela a outra	1
Caso cancelado	Escolher uma atividade que cancela todos os elementos de um pedaço do modelo	2 ou 3

métrica pode ser definida como:

$$NOA = \Sigma_{Atividades}$$

Vale ressaltar que a métrica não leva em consideração a complexidade ou funcionalidade; ela conta apenas uma visão específica do tamanho, o comprimento do processo.

Além da métrica NOA, Cardoso et al. (2006) adaptou outras duas métricas: *Number of Activities and Control-Flow Elements in a Process* (NOAC) e *Number of Activities, Joins and Splits in a Process* (NOAJS), além de levar em consideração o comprimento do processo, consideram também os elementos de fluxo de controle e os *Joins* e *Splits*, respectivamente.

2.3.5 Halstead-based Process Complexity

Cardoso et al. (2006) também fez uma adaptação da métrica de software *Halstead-based Process Complexity* (HPC) para o contexto da complexidade do Processo de Negócio. As medidas HPC são as mais conhecidas e mais estudadas no que diz respeito à complexidade de software. Essas medidas foram desenvolvidas como meio de determinar de forma quantitativa a complexidade do software baseado na compreensão dos operadores (operações aritméticas) e dos operandos (variáveis e constantes).

A métrica HPC é composta de um conjunto de quatro medidas primitivas: n_1 (número de operadores únicos), n_2 (número de operandos únicos), N_1 (número total de

ocorrências de operadores) e $N2$ (número total de ocorrências de operandos), que podem ser derivadas do código-fonte. (CARDOSO et al., 2006)

Para adaptar a métrica HPC para medir a complexidade do Processo de Negócio, Cardoso et al. (2006) sugere mapear os elementos do Processo de Negócio no conjunto de medidas proposto por Halstead, que, segundo Kluza e Napela (2012), resulta em: $n1$ (número de atividades únicas, *splits* e *joins* e outros elementos de fluxo de controle), $n2$ (número de variáveis de dados únicos que são manipulados pelo processo e suas atividades) e $N1$ e $N2$ (número total de elementos de fluxo de controle e dados).

A fórmula para cálculo da métrica HPC pode ser dividida em três partes, que segundo Cardoso et al. (2006), são:

- Comprimento do Processo: $N = n1 * \log_2(n1) + n2 * \log_2(n2)$
- Volume do Processo: $V = (N1 + N2) * \log_2(n1 + n2)$
- Dificuldade do Processo: $D = (n1/2) * (N2/n2)$

Apesar de a métrica possuir algumas vantagens como não precisar ser realizada uma análise aprofundada da estrutura do processo e poder ajudar em prevenção de erros; não foi definido como as medidas primitivas devem ser adaptadas para o processo de negócio. (KLUZA; NAPELA, 2012)

A seguinte seção apresenta os resultados obtidos através da execução de um estudo de caso utilizando a métrica CFC selecionada, e, também, é constituída da validação da hipótese previamente levantada.

3 Metodologia

Para atender os objetivos propostos nesta pesquisa, fez-se necessário seguir um conjunto de passos e adotar alguns procedimentos metodológicos. Este capítulo tem como objetivo mostrar cada um desses itens.

3.1 Visão Geral

A metodologia de pesquisa realizada neste trabalho está representada, de forma macro, no fluxograma ilustrado na Figura 5.

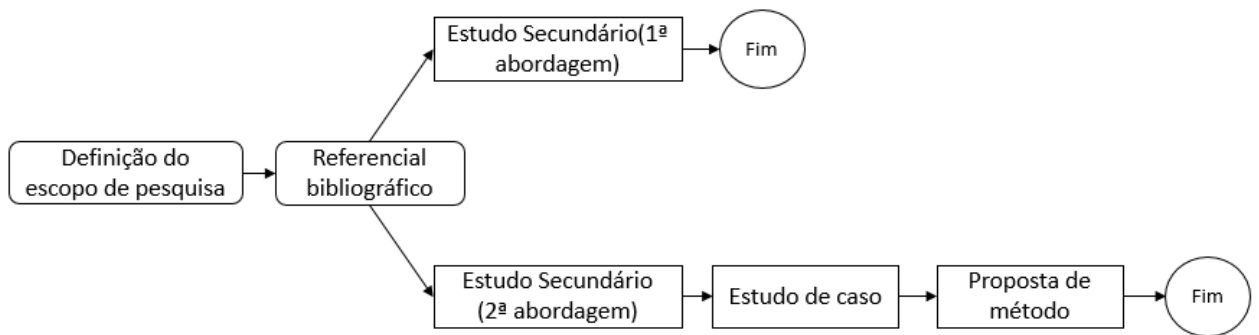


Figura 5 – Passos realizados no TCC

O estudo secundário foi primeiro passo da metodologia do trabalho. A primeira abordagem desse estudo não obteve sucesso, retornando trabalhos que não atendiam ao protocolo definido. Após a execução da primeira abordagem do estudo secundário e da obtenção dos resultados insatisfatórios, foi necessário mudar a abordagem do trabalho. O protocolo utilizado na primeira abordagem do estudo secundário se encontra no Apêndice A. Na segunda abordagem do estudo, o protocolo foi redesenhado e os resultados foram melhores, o que possibilitou a continuidade do trabalho.

O estudo de caso foi executado para comprovar a hipótese levantada previamente através do estudo secundário.

As seções subsequentes explicam com maior grandeza de detalhe como foi a execução dos estudos secundário e do estudo de caso.

3.2 Segunda Etapa do Estudo Secundário

Após a mudança de foco, uma nova revisão de literatura teve de ser realizada. Este capítulo consiste na descrição da segunda parte da execução do estudo.

3.2.1 Planejamento do Estudo

Os modelos para definição do protocolo deste estudo nessa segunda etapa são exatamente os mesmos utilizados na primeira etapa. Para que o enriquecimento do estudo, foi acrescido um procedimento de Snowballing, que, segundo WHOLIN (2014), se refere ao uso das referências do artigo, ou do artigos que citam o presente artigo.

3.2.1.1 Objetivo

O objetivo desta revisão é *analisar* estudos e relatos de iniciativas de medição de elementos do processo de negócio *com o propósito de* identificar e analisar métricas de processo de negócio *com relação à* complexidade *do ponto de vista* das organizações *no contexto de* monitoramento e controle da execução do processo.

3.2.1.2 Questão de Pesquisa

Quais métricas são utilizadas para mensurar a complexidade de um modelo de processo de negócio?

Outros aspectos relacionados ao escopo e à especificidade das questões de pesquisas são apresentadas a seguir:

- **População:** artigos publicados sobre métricas utilizadas para mensurar a complexidade de processos de negócios;
- **Intervenção:** medição da complexidade do processo de negócio no contexto do desenvolvimento de uma aplicação de BI;
- **Resultados:** elementos de influência na atividade de medição do processo de negócio;
- **Aplicação:** servir de base ou apoiar pesquisas envolvendo metodologia de medição da complexidade do processo de negócio; definição de métricas para medição da complexidade do processo de negócio; e criação de métricas para medição da complexidade de processos de negócio.

3.2.1.3 Critérios de Seleção de Trabalhos

Foram definidos os seguintes critérios de inclusão (CI) e exclusão (CE):

- **CI1.** Os artigos devem estar escritos em inglês.
- **CI2.** Os resultados devem estar disponíveis integralmente.
- **CI3.** Os artigos devem apresentar algum estudo sobre métricas de complexidade de elementos do processo de negócio.

- **CE1.** Os artigos que apresentam um estudo superficial sobre o tema.
- **CE2.** Artigos que estão duplicados.
- **CE3.** Artigos que não considerem o contexto da complexidade do processo.

3.2.1.4 Procedimentos de Seleção

A *string* de busca foi executada, assim como na primeira etapa, nas bases de pesquisas do IEEEExplore, da ACM, da Capes e Scopus e foram escolhidas pelo mesmo motivo. Todos os artigos retornados, assim como na primeira etapa, foram armazenados na ferramenta que auxilia e automatiza o processo de revisão sistemática (StArt). Cada vez que a *string* for executada, os dados obtidos serão documentados e depois contabilizados os registros repetidos. Os títulos, resumos e palavras-chaves foram lidos, avaliando-se os critérios de inclusão e exclusão definidos anteriormente. Os artigos que passaram pelos critérios de inclusão e exclusão seguem para o procedimento de Snowballing e a extração dos dados.

3.2.1.5 Snowballing

Para realizar snowballing, WHOLIN (2014) sugere-se que os passos existentes na Figura 6 sejam executados.

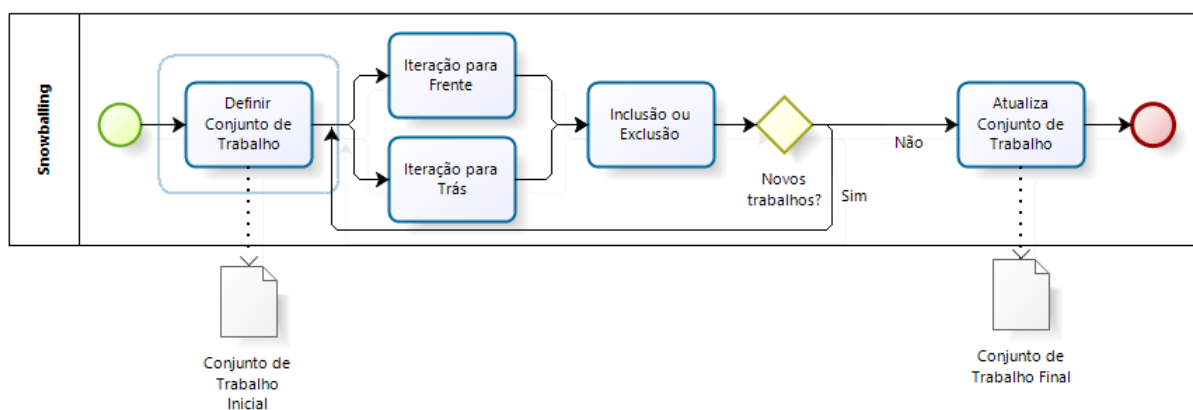


Figura 6 – Procedimentos para Snowballing com adaptações (WHOLIN, 2014)

O primeiro passo, segundo WHOLIN (2014), consiste na identificação de um **Conjunto de Trabalhos Inicial**, para que seja possível a execução de snowballing. Esse conjunto de trabalhos inicial é resultado do procedimento dos critérios de inclusão e exclusão e dos procedimentos de seleção definidos no planejamento deste estudo secundário.

WHOLIN (2014) diz que o segundo passo consiste nas **Iterações para Frente** e nas **Iterações para Trás**. Para que possam ser executadas, há a necessidade de já ter o conjunto de trabalhos bem definidos e analisados. As iterações para frente consistem

na análise dos trabalhos que citam o artigo sendo examinado. Já as iterações para trás consistem na utilização da lista de referências do trabalho sendo examinado. Devido ao prazo de tempo limitado, o presente trabalho realizou apenas as iterações para trás. Para realizar a Inclusão e Exclusão dos novos trabalho, foram utilizados os mesmos critérios de seleção definidos no protocolo.

O último passo consiste na **Extração dos Dados** e foi realizado de acordo com o protocolo.

3.2.1.6 Extração dos Dados

Para extrair informações dos artigos resultantes, foram selecionadas as seguintes informações para serem coletadas:

- Título;
- Autor;
- Data de publicação;
- Fonte de publicação;
- Listagem das métricas de medição de complexidade de Processos de Negócio.

3.2.2 Condução da Segunda Etapa do Estudo Secundário

A segunda etapa deste estudo foi executada em setembro de 2014. A definição da string foi inspirada na que foi utilizada na primeira etapa do estudo secundário, entretanto, foi mudada a temática, resultando na seguinte string: ("Business Process"OR "Processo de Negocio) AND (Metrics OR Metricas) AND (Complexity OR Complexidade). A string foi criada de maneira a ser genérica e retornar a maior quantidade possível de artigos; entretanto, isso não foi observado, pois poucos resultados foram retornados, e, por isso foi necessário aumentar esse montante através do snowballing.

Na base da IEEE, 21 resultados foram obtidos; na ACM, apenas dois; na base de dados Scopus, 55; e na Capes, onze. Houve 8 publicações em comum entre a base da Capes e da Scopus. Em conferências nacionais, nenhum artigo que atendesse a expressão foi retornado.

Após a execução da string, todos os trabalhos resultantes foram submetidos aos critérios de seleção de trabalho. A tabela resultante segue abaixo:

O resultado obtido após a etapa de critério de seleção de trabalhos foi de cinco artigos selecionados como estudos primários. O procedimento de iterações para trás do

Tabela 2 – Resultado da Segunda Etapa do Estudo Secundário

Base	String	Resultado	CI1	CI2	CI3	CE1	CE2	CE3
IEEE	"Business Process"AND Metrics AND Complexity	21	21	21	2	2	2	2
ACM	(Abstract:("Business Process") AND Abs- tract:(Metrics) AND Abstract:(Complexity))	2	0	0	0	0	0	0
Scopus	ABS("Business Process") AND ABS(Metrics) AND ABS(Complexity)	55	44	14	3	3	3	3
Capes	("Business Process"OR "Processo de Negocio") AND (Metrics OR Metri- cas) AND (Complexity OR Complexidade)	41	11	0	0	0	0	0

snowballing foi realizado logo em seguida e foram selecionados mais sete trabalhos, a partir da análise do título do artigo.

Após a execução de todos os procedimentos de seleção de trabalhos, a etapa seguinte consistiu na síntese dos resultados. Todos os doze artigos resultantes passaram pela extração de dados de acordo com o que foi definido na Seção [A.1.6](#).

O resultado da coleta de dados pode ser verificado na Tabela [4](#).

Após a análise dos doze artigos resultantes da segunda etapa do estudo, foram coletadas diversas métricas utilizadas para medir a complexidade de Processos de Negócio. A Tabela [3](#) apresenta as métricas coletadas de acordo com sua frequência de utilização nos trabalhos analisados.

Tabela 3 – Lista de métricas

Métricas	Frequência
<i>Control-Flow Complexity (CFC)</i>	10
<i>Number of activities in a process (NOA)</i>	4
<i>Halstead-based Process Complexity (HPC)</i>	4
<i>Cognitive Weight (CW)</i>	4
<i>McCabe's cyclomatic complexity (MCC)</i>	4
<i>Number of activities and control-flow elements (NOAC)</i>	3
<i>Number of activities, joins, and splits (NOAJS)</i>	3
<i>Extended Cardoso Metric (ECaM)</i>	3
<i>Extended Cyclomatic Metric (ECyM)</i>	3
<i>Coefficient of Network Complexity (CNC)</i>	2
<i>Maximum Nesting Depth (MaxND)</i>	2
<i>Mean Nesting Depth (MeaND)</i>	2
<i>Structuredness Metric (SM)</i>	2
<i>Total Number of Activities (TNA)</i>	1
<i>Number of Sequence Flows between Activities (NSFA)</i>	1
<i>Antonini's Business Process Measure</i>	1
<i>Cognitive Functional Size (CFS)</i>	1
<i>Cross-Connectivity</i>	1
<i>Number of Handles</i>	1
<i>Fan-in and Fan-out Metric</i>	1
<i>Event-driven Process Chain (EPC)</i>	1

Tabela 4 – Resultado da Coleta de Dados do Estudo Secundário

ID	Título	Autor	Local de Publicação	Métricas
1	<i>Proposal of Square Metrics for Measuring Business Process Model Complexity</i>	(KLUZA; NAPELA, 2012)	<i>Federated Conference on Computer Science and Information Systems 2012</i>	Number of activities in a process (NOA). Number of activities and control-flow elements (NOAC). Number of activities, joins, and splits (NOAJS). Control-Flow Complexity (CFC). Halstead-based Process Complexity (HPC). Coefficient of Network Complexity (CNC). Total Number of Activities (TNA). Number of Sequence Flows between Activities (NSFA). Maximum Nesting Depth (MaxND). Mean Nesting Depth (MeaND). Extended Cardoso Metric (ECaM). Extended Cyclomatic Metric (ECyM).
2	<i>Adopting the Cognitive Complexity Measure for Business Process Models</i>	(GRUHN; LAUE, 2006a)	<i>5th. IEEE Int. Conf. on Cognitive Informatics</i>	Cognitive Weight (CW).
3	<i>Exploration on Software Complexity Metrics for Business Process Model and Notation</i>	(SOLICHAH et al., 2013)	<i>International Conference on Advanced Computer Science and Information Systems 2013</i>	Control-Flox Complexity (CFC). Antonini's Business Process Measure. McCabe's cyclomatic complexity (MCC).

ID	Título	Autor	Local de Publicação	Métricas
4	<i>Process control-flow complexity metrics: An empirical validation</i>	(CARDOSO, 2006)	<i>Conference on Services Computing 2006</i>	Control-Flox Complexity (CFC).
5	<i>A Survey of Business Process Complexity Metrics</i>	(MUKETHA et al., 2010)	<i>Information Technology Journal 2010</i>	Control-Flox Complexity (CFC). Cognitive Functional Size (CFS). Extended Cardoso Metric (ECaM). Extended Cyclomatic Metric (ECyM). Cross-Connectivity. Structuredness Metric (SM).
6	<i>Evaluating the Process Control-flow Complexity Measure</i>	(CARDOSO, 2005)	<i>International Conference on Web Services 2005</i>	Control-Flox Complexity (CFC).
7	<i>Business Process Control-Flow Complexity: Metric, Evaluation, and Notation</i>	(CARDOSO, 2008)	<i>International Journal of Web Services Research 2008</i>	Control-Flox Complexity (CFC).

ID	Título	Autor	Local de Publicação	Métricas
8	<i>A Discourse on Complexity of Process Models</i>	(CARDOSO et al., 2006)	<i>BPM 2006 Workshops</i>	Number of activities in a process (NOA). Number of activities and control-flow elements (NOAC). Number of activities, joins, and splits (NOAJS). McCabe's cyclomatic complexity (MCC). Halstead-based Process Complexity (HPC). Control-Flox Complexity (CFC). Coefficient of Network Complexity (CNC).
9	<i>Complexity Metrics for Business Process Models</i>	(GRUHN; LAUE, 2006b)	IEEE	Number of activities in a process (NOA). Control-Flox Complexity (CFC). Maximum Nesting Depth (MaxND). Mean Nesting Depth (MeaND). Number of Handles. Cognitive Weight (CW).
10	<i>Survey Complexity Metrics for Reusable Business Process</i>	(THAMMARAK, 2010)	<i>1st National Conference on Applied Computer Technology and Information System 2010</i>	McCabe's cyclomatic complexity (MCC). Cognitive Weight (CW). Control-Flox Complexity (CFC). Halstead-based Process Complexity (HPC). Cognitive Weight (CW). Fan-in and Fan-out Metric.

ID	Título	Autor	Local de Publicação	Métricas
11	<i>Software Measures for Business Processes</i>	(ANTONINI et al., 2011)	<i>Advances in Database and Information Systems 2011</i>	Number of activities in a process (NOA). Number of activities and control-flow elements (NOAC). Number of activities, joins, and splits (NOAJS). Halstead-based Process COMplexity (HPC). McCabe's cyclomatic complexity (MCC). Control-Flox Complexity (CFC). Event-driven Process Chain (EPC).
12	<i>Complexity Metrics for Workflow Nets</i>	(LASSEN; AALST, 2009)	<i>Information and Software Technology 2009</i>	Extended Cardoso Metric (ECaM). Extended Cyclomatic Metric (ECyM). Structuredness Metric (SM).

3.3 Estudo de Caso

Esta Seção destina-se a apresentar os procedimentos necessários de um estudo de caso, criado para validar a hipótese levantada após a execução dos estudos secundários. O protocolo foi criado a partir do template definido por [Brereton et al. \(2008\)](#).

3.4 Processo do Estudo de Caso

Segundo os passos propostos por [Runeson et al. \(2012\)](#), foi modelado o processo que será seguido na execução do presente estudo de caso. A Figura 7 ilustra o processo a ser seguido.

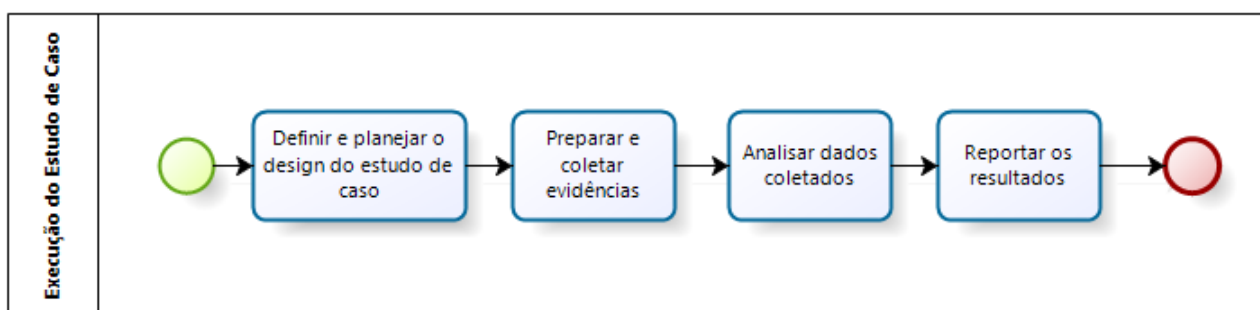


Figura 7 – Processo do Estudo de Caso

A atividade de **Definir e planejar o design do estudo de caso** tem como objetivo definir os objetivos do estudo de caso e planejar a sua execução.

A atividade de **Preparar e coletar evidências** tem como objetivo definir os procedimentos e os protocolos que serão utilizados na coleta de dados.

A atividade de **Analisar dados coletados** tem como objetivo a execução dos procedimentos da coleta de dados e a análise dos mesmos.

A atividade de **Relatar os resultados** tem como objetivo a divulgação do resultado e da conclusão do estudo de caso.

3.4.1 Contexto

Como resultado da condução da segunda etapa dos estudos secundários, foram encontradas algumas métricas capazes de calcular a complexidade de processos de negócio. Para que fosse possível responder a questão de pesquisa inicial proposta para este Trabalho de Conclusão de Curso, foi necessária a condução deste estudo de caso, com o objetivo de identificar um conjunto de métricas de complexidade para o desenvolvimento de sistemas de Inteligência de Negócio.

3.4.2 Design e Planejamento

Para o design e planejamento deste estudo de caso, foi utilizado o *checklist* desenvolvido por Host e Runeson (2007) junto ao template proposto por Brereton et al. (2008).

3.4.2.1 Hipótese

A métrica que obteve a maior quantidade de trabalhos referenciados, isto é, CFC; foi utilizada como base para a construção de uma hipótese a qual é a questão principal a ser respondida pelo estudo. A Figura 9 demonstra a relação que este estudo deseja mostrar.

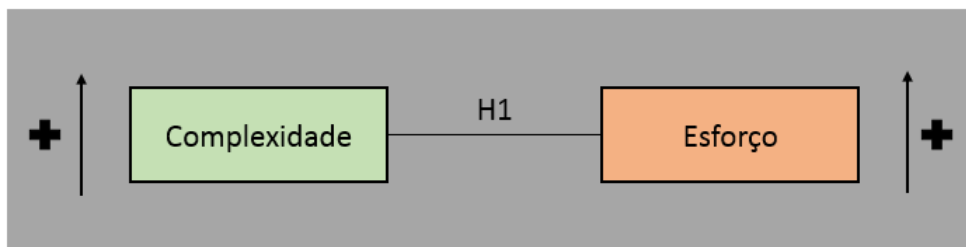


Figura 8 – Complexidade x Esforço

A hipótese pode ser resumidamente descrita como:

H1: Quanto maior a *complexidade* obtida através da aplicação da métrica em um Processo de Negócio maior também, será o *esforço* necessário para desenvolver uma aplicação de Business Intelligence.

3.4.2.2 Objeto de Estudo

Para definir o objeto de estudo do presente trabalho é necessário descrever o cenário onde o estudo de caso será executado que são a empresa e a equipe selecionados. A seguir, cada um deles é contextualizado.

3.4.2.2.1 Biggdata

Biggdata é uma empresa especializada em prover soluções de BI para a Administração Pública Federal. As soluções de BI estão sempre pautadas nas boas práticas de governança de TI, em alinhamento às conformidades da Governança Corporativa. A metodologia de desenvolvimento utilizada é uma adaptação de algumas práticas do PMBoK e outras provindas da metodologia ágil Scrum.

3.4.2.2.2 Equipe

A equipe de desenvolvimento é composta por quatro desenvolvedores juniores e um desenvolvedor sênior. O desenvolvedor sênior possui mais de 15 anos de experiência com BI, enquanto os desenvolvedores juniores estão alcançando o seu segundo ano na área. Toda a equipe possui formação ou estão formando na área de Tecnologia da Informação, e possuem treinamento na ferramenta de Data Discovery Qlikview.

3.4.2.2.3 Casos Selecionados

Para realização do estudo foram selecionados três casos de contextos distintos. Os casos são projetos já desenvolvidos pela empresa e pela equipe supracitada, e o critério de seleção desses casos se deu pela disponibilidade dos dados para análise.

Para manter o sigilo dos dados, não será explicitado nenhum nome e nenhuma informação advinda da empresa contratante.

O primeiro caso é um macroprocesso de licitação e contratos constituído por dois processos detalhados em quatro subprocessos. O segundo caso é um macroprocesso de auditoria constituído de dois processos. Finalmente, o terceiro caso é um macroprocesso de captação e administração financeira constituído de sete processos.

A Tabela 5 compila as principais informações referentes aos processos utilizados nesse Estudo de Caso.

Tabela 5 – Objetos de Estudo

Processo	Descrição	Observações
Processo de Aquisição de Materiais e Serviços	Consiste em um conjunto de procedimentos que auxiliam as áreas requisitantes na aquisição de materiais e serviços	O processo é composto por quatro subprocessos
Processo de Auditoria Operacional	Consiste em um conjunto de procedimentos que provê um diagnóstico inicial auxiliando a detecção e classificação de falhas	
Processo de Captação e Administração Financeira	Consiste no controle e gestão das contribuições, arrecadações, contábil e financeiro	O processo é composto por sete subprocessos

3.4.2.3 Papéis

A equipe responsável pela execução do Estudo de Caso é composta por duas pessoas:

- Luciano Endo, responsável pela condução e execução do Estudo de Caso, graduando em Engenharia de Software pela Universidade de Brasília, consultor de Business Intelligence na empresa Biggdata e interessado pela área de Business Intelligence; e
- Fabiana Freitas Mendes, responsável por auxiliar na condução e na definições de diretrizes de execução do Estudo de Caso. Possui graduação em Ciência da Computação pela Universidade Federal de Goiás (2008) e mestrado em Ciência da Computação pela Universidade Federal de Goiás (2010). Atualmente é professora da Universidade de Brasília, atuando principalmente temas de qualidade de software e melhoria de processos de software.

3.4.2.4 Considerações Éticas

Para manter a integridade dos dados da empresa contratante e assegurar que as informações da empresa provedora de serviços estejam de acordo com as suas normas, um termo de concessão de direitos de pesquisa foi elaborado e assinado pelos responsáveis pelo Estudo de Caso e pela Biggdata. O termo se encontra no Anexo A do presente trabalho.

3.4.3 Definição de Conceitos e Medidas

Foram mensurados, através da aplicação da métrica CFC, a complexidade dos processos de negócio para que se faça possível o relacionamento entre o valor quantitativo obtido, com o valor quantitativo do esforço despendido para a automatização dos processos em sistemas BI.

Alguns conceitos são importantes para entendimento do estudo de caso. São eles:

- CFC Seção 2.3.2.
- Processo Seção 2.3.2.
- Medida do Processo Seção 2.3.2.
- Métrica do Processo Seção 2.3.2.
- *Fan-out* da atividade Seção 2.3.2.

3.4.4 Coleta de Dados

O principal método para coleta de dados foi a aplicação da métrica CFC nos três processos de negócio. O resultado da aplicação dessa métrica resultou em um dado quantitativo relacionado à complexidade dos processos de negócio. O segundo conjunto de dados contém os dados relacionados ao esforço para desenvolvimento das três aplicações de BI provenientes dos processos de negócio.

Para aplicação da métrica foi utilizado a metodologia proposta por [Cardoso \(2008\)](#), que consiste no cálculo dos fatores que influenciam na complexidade do processo. Todos os dados coletados foram armazenados e analisados em planilha.

3.4.5 Análise de Dados

Devido à pequena quantidade de amostras utilizadas nesse estudo, a amostragem de dados não poderá ser feita através de métodos estatísticos, o método de amostragem utilizado se dará através de uma abordagem **não-probabilística** que segundo [Mattar \(2012\)](#) "é aquela em que a seleção dos elementos da população para compor a amostra depende ao menos em parte do julgamento do pesquisador ou do entrevistador no campo".

A análise dos dados foi realizada através da comparação e observação direta entre o valor quantitativo obtido com a métrica CFC e o esforço demandado para desenvolver as aplicações de Business Intelligence.

3.4.6 Relatório

Com a intuito de revelar as descobertas obtidas com a execução do estudo de caso, o presente trabalho tem como objetivo final conter as informações resultantes das atividades executadas. Essas informações podem beneficiar a comunidade acadêmica e empresas desenvolvedoras de sistemas BI. Em um primeiro momento, o relatório terá como público-alvo a comunidade acadêmica e pretende ajudar no entendimento da relação entre esforço e complexidade de processo de negócio e servir de guia para trabalhos futuros relacionados ao tema.

4 Resultados

Este capítulo se destina a apresentar o resultado do estudo de caso e apresenta, também, uma proposta de modelo de utilização da métrica CFC.

4.1 Relação entre complexidade e esforço

Esta Seção apresenta os resultados obtidos após a execução do estudo de caso. Neste capítulo serão apresentados os artefatos e instrumentos utilizados para essa execução, os resultados obtidos e as lições aprendidas após o término do estudo.

4.1.1 Hipótese

Utilizando a métrica que obteve a maior quantidade de resultados no procedimento de Revisão Sistemática, uma hipótese foi levantada e pode ser considerada a questão principal a ser respondida por um estudo de caso. A figura 9 demonstra a relação entre complexidade e esforço.

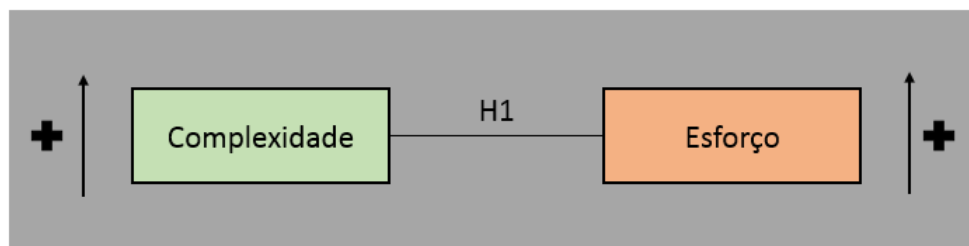


Figura 9 – Complexidade x Esforço

H1: Quanto maior a *complexidade* obtida através da aplicação da métrica em um Processo de Negócio maior também, será o *esforço* necessário para desenvolver uma aplicação de Business Intelligence.

4.1.2 Artefatos e Outros Aspectos do Caso

Para contagem dos valores requeridos pela métrica CFC, foram analisados os artefatos relacionados à descrição do processo de negócio da área. Sendo eles, o **Documento de Análise de Cenário** e a **Modelagem do Processo**.

O Documento de Análise de Cenário tem como objetivo descrever o cenário geral das necessidades e características gerais da área, para definição do escopo do projeto de BI da organização. Ele se concentra nas necessidades dos usuários, nas razões que levam

a essas necessidades e como elas são atendidas, garantindo consistência e conformidade com os requisitos de negócio a serem implementadas. Além disso, o documento descreve de forma objetiva os processos que são executados pela área.

A Modelagem do Processo tem como objetivo mostrar graficamente o processo executado pela área, o documento da modelagem do processo normalmente é feito pela própria organização. Entretanto, quando a área não possui esse artefato, pode ser necessário a modelagem por parte da empresa para questões de esclarecimento e entendimento do negócio.

Além dos dois artefatos, foi fundamental analisar alguns aspectos envolvidos no contexto de desenvolvimento da aplicação de BI, isto é, a quantidade de pessoas envolvidas no desenvolvimento da aplicação; a complexidade de entendimento do negócio; e a complexidade e a qualidade dos dados legados da organização.

A Tabela 6 compila as informações coletadas dos objetos de estudo.

Tabela 6 – Casos e aspectos envolvidos

Processo	Descrição	Observações	Equipe envolvida	Complexidade de entendimento	Dados
Processo de Aquisição de Materiais e Serviços	Consiste em um conjunto de procedimentos que auxiliam as áreas requisitantes na aquisição de materiais e serviços	O processo é composto por quatro subprocessos	2	Médio	Dados de complexidade média e de média qualidade
Processo de Auditoria Operacional	Consiste em um conjunto de procedimentos que provê um diagnóstico inicial auxiliando a detecção e classificação de falhas	O processo é composto por dois subprocessos	1	Alto	Dados de complexidade baixa e de média qualidade
Processo de Captação e Administração Financeira	Consiste no controle e gestão das contribuições, arrecadações, contábil e financeiro	O processo é composto por sete subprocessos	2	Alto	Dados de complexidade alta e de baixa qualidade

A Seção seguinte mostra o processo seguido para extração dos dados para composição da métrica CFC nos objetos de estudo.

4.1.3 Processo de coleta de dados

O processo ilustrado na Figura 10 reflete as atividades necessárias para coletar os dados do estudo para composição da métrica CFC.

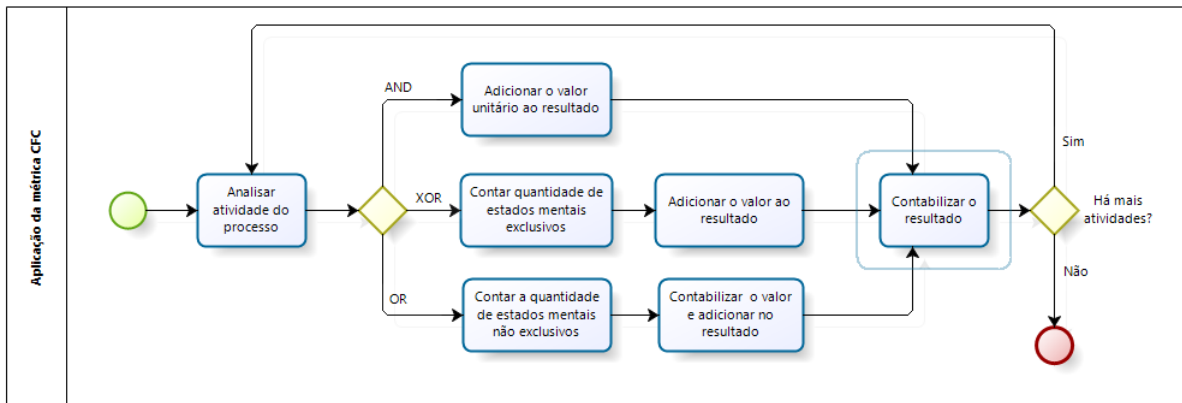


Figura 10 – Processo de utilização da métrica CFC

A atividade **Analisar atividade do processo** teve como objetivo classificar a atividade do processo de acordo com sua divisão (*Split*). A divisão da atividade pode ser classificada em: *AND-Split*, *XOR-Split* e *OR-Split*.

Caso a divisão da atividade fosse do tipo *AND-Split*, a atividade **Adicionar o valor unitário ao resultado** foi executada e consistiu na adição do valor **um** ao resultado da métrica.

Caso a divisão da atividade fosse do tipo *XOR-Split*, a atividade **Contar quantidade de estados mentais exclusivos** foi executada e consistiu na contagem de estados mentais exclusivos, ou seja, quando o estado mental da atividade é um **OU** outro.

A atividade **Adicionar valor ao resultado** consiste na adição da quantidade de estados mentais exclusivos ao resultado da métrica.

Caso a divisão da atividade seja do tipo *OR-Split*, a atividade **Contar quantidade de estados mentais não exclusivos** foi executada e consistiu na contagem de estados mentais não exclusivos, ou seja, quando o estado mental da atividade é um **E** outro(s).

A atividade **Contabilizar o valor e adicionar ao resultado** consiste no cálculo utilizando a fórmula $2^n - 1$, onde n é a quantidade de estados mentais não exclusivos possíveis. O valor resultante é adicionado ao resultado da métrica.

A atividade **Contabilizar o resultado** consiste na soma dos valores resultantes das atividades anteriores.

Caso exista mais atividades, o processo é reiniciado, caso contrário o processo é finalizado e o valor contabilizado é o valor CFC daquele processo.

4.1.4 Síntese dos resultados

Os dados obtidos após o processo de coleta foram armazenados na Tabela 7 abaixo.

Tabela 7 – Resultados

Processo	<i>AND-Split</i>	<i>XOR-Split</i>	<i>OR-Split</i>	CFC	Esforço (horas)	Produtividade (h/CFC)
Processo de Aquisição de Materiais e Serviços	43	13	2	59	244h	4.13
Processo de Auditoria Operacional	7	4	0	13	42h	3.23
Processo de Captação e Administração Financeira	41	12	3	61	274h	4.49

Nota-se em um primeiro momento uma certa relação entre os valores obtidos da métrica CFC e o esforço despendido para o desenvolvimento das aplicações de BI dos processos de negócio. Quantitativamente é possível visualizar o crescimento do esforço com o aumento da complexidade CFC medida, isso ajuda a comprovar a hipótese levantada, ainda que fracamente.

Vale ressaltar, que o contexto de aplicação da métrica CFC, neste trabalho, não levou em consideração os possíveis fatores externos e os aspectos de influência que podem interferir no resultado final. Há a necessidade de realizar um experimento em que seja possível analisar todas as variáveis envolvidas para uma validação mais forte da hipótese.

Um fator que pode interferir no resultado final, e por isso merece ser mencionado no presente trabalho, se refere à metodologia de desenvolvimento de aplicações de Business Intelligence. É a partir da metodologia de desenvolvimento que se faz possível determinar o escopo do projeto, além da definição dos indicadores envolvidos e o alinhamento da aplicação ao plano estratégico da organização. Todos os fatores previamente citados devem ser incluídos em um experimento para validação, de forma mais segura, da hipótese levantada.

Além do fator supracitado, vale ressaltar que cada ferramenta possui uma característica específica de desenvolvimento, o que deve acarretar, também, na alteração dos resultados do esforço necessário para automatização de um processo de negócio em uma aplicação BI.

Para auxiliar os desenvolvedores e interessados no desenvolvimento de sistemas BI é necessário construir uma base histórica contendo as informações relevantes, os fatores

externos e os aspectos de influência para que seja possível conduzir os desenvolvimentos de forma controlada. A base histórica irá prover parâmetros de comparação para acompanhamento e melhoria da precisão da estimativa.

Pode-se concluir que existem diversos fatores e aspectos que devem influenciar na utilização da métrica CFC, para reduzir o nível de influência é necessário levantar, tratar e registrar cada um desses fatores. O estudo de caso mostrou a validade da hipótese desconsiderando os fatores supracitados e por isso não pode ser considerado como absoluta, entretanto, o estudo abriu oportunidades para trabalhos e descobertas futuras.

4.1.5 Lições Aprendidas

Como resultado do estudo de caso, algumas lições foram aprendidas:

Limitações do Estudo. Devido à falta de tempo hábil para executar um experimento, o estudo de caso foi realizado com os recursos disponíveis. A seleção dos casos se deu de forma não-probabilística e por conveniência, foram selecionados os casos que possuíam as informações necessárias para a execução do estudo. Como não foi possível controlar diversas variáveis envolvidas no estudo, o mesmo ficou limitado àquilo que já estava pronto. O estudo foi realizado utilizando dados já existentes.

Contexto teórico. Em estudo exploratórios em que existem poucos trabalhos relacionados, o pesquisador deve examinar cuidadosamente sobre como organizar os dados e realizar a análise.

Protocolo de estudo de caso estável. A escrita e atualização do protocolo do estudo de caso pode ser desafiadora. Alguns eventos aconteceram durante o estudo e o método de análise teve que ser modificado devido à amostragem pequena e restrita.

Equilibrando interesses pessoais e gerais na estratégia de publicação. É importante estabelecer uma estratégia de publicação da pesquisa que leva em consideração a possibilidade da utilização da pesquisa no mercado e na possibilidade de continuidade do estudo de forma mais controlada. Em um primeiro momento, a publicação seria estritamente para a comunidade acadêmica. Entretanto, com a possibilidade de dar continuidade à pesquisa em um projeto de mestrado ou até mesmo a aplicação do estudo no mercado, a estratégia de publicação deve ser modificada para abranger essas possibilidades.

A partir da validação da hipótese, foi possível desenvolver um método para medir o esforço necessário para desenvolver um sistema BI, levando em consideração a complexidade do processo de negócio e a produtividade da equipe. Este método é detalhado na seção a seguir.

4.2 Método CFCBI

O método CFCBI é composto por três fases, cada fase é composta por atividades necessárias para a conclusão com êxito do cálculo do tamanho de aplicações de Business Intelligence. O método deve ser aplicado na etapa de Planejamento do Projeto de desenvolvimento de uma aplicação de BI. A Figura 11 ilustra as fases e as suas respectivas atividades.

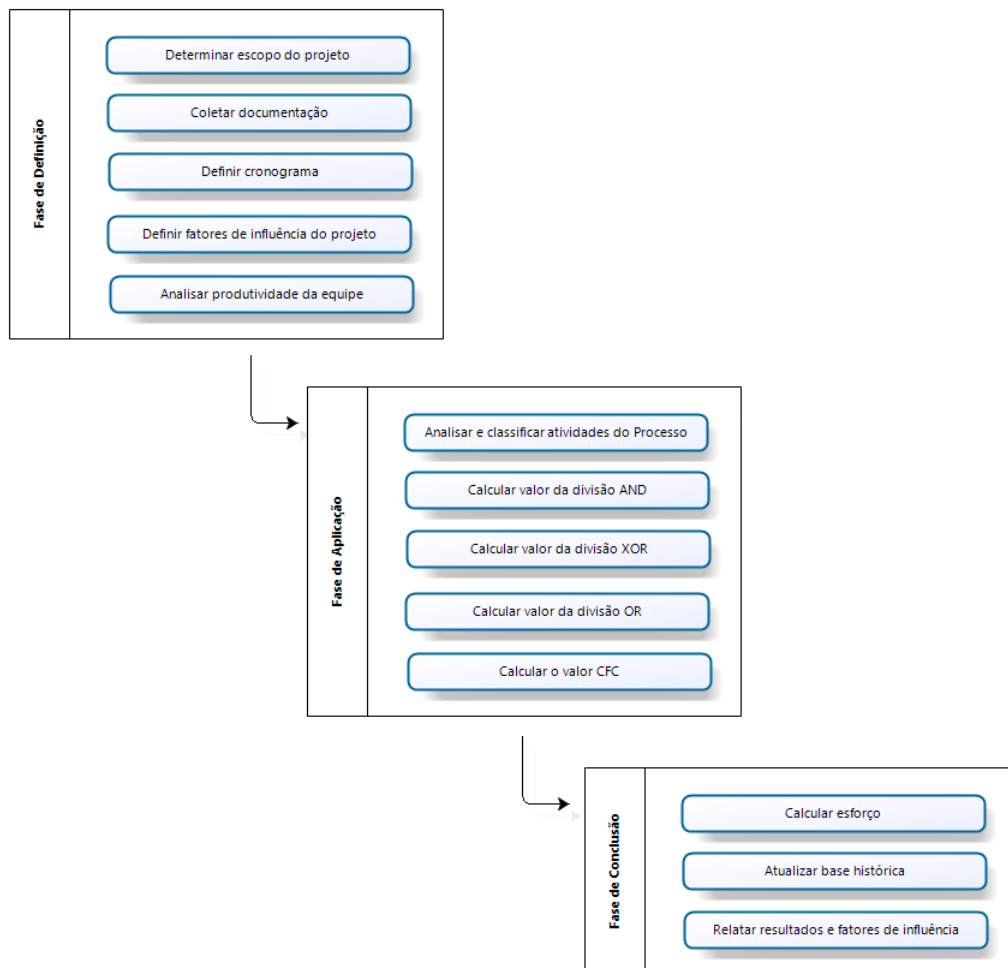


Figura 11 – Método CFCBI

As seções subsequentes descrevem com maior clareza cada fase e suas respectivas atividades.

4.2.1 Fase de Definição

A Fase de Definição abrange as tarefas necessárias de planejamento. Nesta fase é definido o cronograma contendo os recursos alocados, os riscos envolvidos e as estimativas de prazos. A fase tem como objetivo principal, definir o escopo do processo da aplicação de BI que será desenvolvida. A Figura 12 ilustra as atividades existentes nessa etapa.

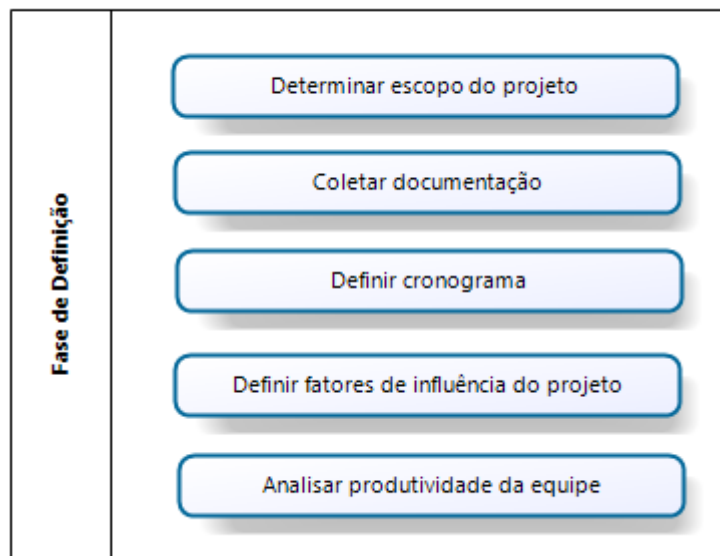


Figura 12 – Fase de Definição

As próximas subseções descrevem as atividades vistas na Figura 12.

4.2.1.1 Determinar escopo do projeto

A atividade de **Determinar o escopo do projeto** envolve capturar o contexto em que os processos que serão medidos estão incluídos dentro da organização. É importante ressaltar que uma aplicação de BI deve estar alinhada ao planejamento estratégico da organização e por isso é necessário o entendimento do contexto estratégico em que os processos estão incluídos.

Entrada: Planejamento Estratégico da Empresa

Saída: Escopo do Projeto

4.2.1.2 Coletar documentação

A atividade de **Coletar a documentação** é realizada após a definição do escopo, todos os documentos necessários para o entendimento do processo são coletados e preparados para análise futura. Os documentos de processos devem incluir a modelagem e descrição do processo e podem conter também o documento de negócio, a análise de cenário e outros.

Entrada: Escopo do Projeto e Documentos de Processo

Saída: Documentos de Processo Selecionados

4.2.1.3 Definir cronograma

A atividade de **Definir cronograma** é composta por algumas tarefas, são elas: listar as atividades que serão executadas e as datas da sua execução, definir os recursos envolvidos, estabelecer os riscos envolvidos em cada atividade e identificar e analisar o caminho crítico.

Entrada: Atividades e Recursos

Saída: Cronograma

4.2.1.4 Definir fatores de influência do projeto

A atividade de **Definir fatores de influência do projeto** abrange o levantamento e definição de fatores que podem alterar no resultado final da aplicação da métrica. Esses fatores podem estar ligados ao processo, aos dados que serão utilizados, a documentação fornecida ou fatores externos. Podem ser listados como fatores de influência: a complexidade de entendimento do processo de negócio, a qualidade dos dados utilizados, a complexidade dos dados, documentação confusa ou inexistente, fontes de dados externas à organização, entre outros.

Entrada: -

Saída: Fatores de influência

4.2.1.5 Analisar produtividade da equipe

A atividade de **Analisar produtividade da equipe** consiste no levantamento de CFCs produzidos por hora da equipe. A produtividade será necessária na hora de calcular o esforço necessário para desenvolver uma aplicação de BI. É necessário que a empresa desenvolvedora de sistemas de Business Intelligence possua uma base histórica contendo a produtividade da sua equipe de desenvolvedores.

Entrada: Base histórica de produtividade

Saída: Produtividade da equipes

4.2.2 Fase de Aplicação

A Fase de Aplicação abrange as tarefas necessárias para o cálculo da métrica CFC. Nesta fase são executadas as atividades relacionadas ao cálculo CFC proposto por [Cardoso et al. \(2006\)](#). A Figura 13 ilustra as atividades existentes nessa etapa.

As próximas subseções descrevem as atividades vistas na Figura 13.

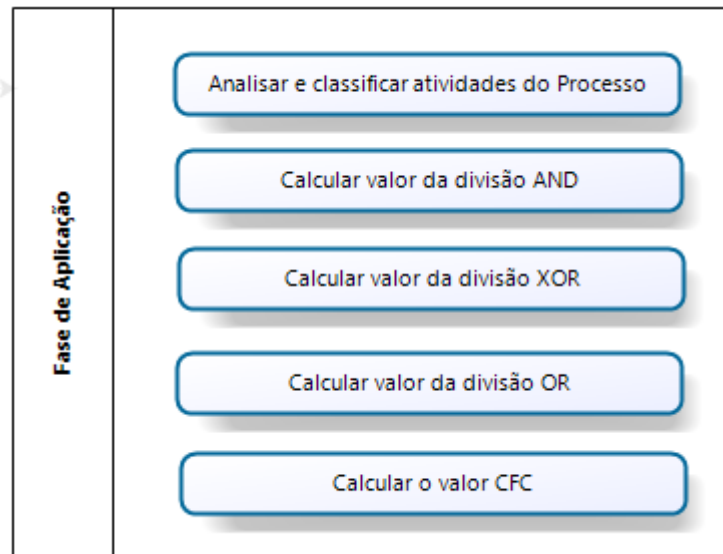


Figura 13 – Fase de Aplicação

4.2.2.1 Analisar e classificar atividades do processo

A atividade **Analisar e classificar atividades do processo** consiste na análise de cada atividade existente no processo e na sua classificação. A classificação da atividade depende da divisão resultante da sua execução, ela pode ser classificada em: AND, XOR e OR.

4.2.2.2 Calcular valor da divisão AND

A atividade **Calcular valor da divisão AND** é executada quando as atividades do processo analisado são classificadas como tal. Nessa etapa é atribuído o valor um para as atividades com divisão AND.

4.2.2.3 Calcular valor da divisão XOR

A atividade **Calcular valor da divisão XOR** é executada quando as atividades do processo analisado são classificadas como tal. Nessa etapa é atribuído o valor resultante da soma de quantidade de fluxos saindo da atividade.

4.2.2.4 Calcular valor da divisão OR

A atividade **Calcular valor da divisão OR** é executada quando as atividades do processo analisado são classificadas como tal. Nessa etapa é atribuído o valor resultante da fórmula $2^n - 1$, onde n é a quantidade de fluxo saindo da atividade.

4.2.2.5 Calcular o valor CFC

A atividade **Calcular o valor CFC** consiste na soma dos valores atribuídos anteriormente para cada atividade existente no processo sendo analisado. O valor da métrica CFC se dará no término do somatório dos valores resultantes.

4.2.3 Fase de Conclusão

A Fase de Conclusão abrange as tarefas necessárias para calcular o esforço necessário para desenvolver uma aplicação de BI, relatar o resultado e atualizar a base histórica. A Figura 14 ilustra as atividades existentes nessa etapa.

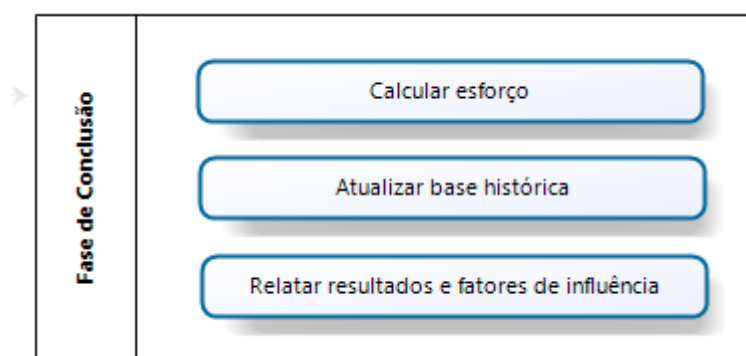


Figura 14 – Fase de Conclusão

As próximas subseções descrevem as atividades vistas na Figura 14.

4.2.3.1 Calcular esforço

A atividade de **Calcular esforço** consiste no cálculo do esforço necessário para desenvolver um sistema de Business Intelligence. O esforço (e) é resultante da relação da produtividade da equipe (p) com o valor resultante da complexidade CFC. A fórmula para cálculo do esforço é: $e = CFC/p$.

4.2.3.2 Atualizar base histórica

A atividade de **Atualizar base histórica** consiste na atualização da base histórica da empresa com as informações resultantes da aplicação do método CFCBI. As informações devem ser compostas pelo valor CFC, o esforço estimado, o esforço realizado, a produtividade utilizada, a produtividade real, as informações relacionados aos fatores de influência e os recursos humanos utilizados na aplicação do método.

4.2.3.3 Relatar resultados

A atividade de **Relatar resultados** consiste na compilação dos resultados em formato de relatório. Esse relatório deverá ser entregue à empresa após a execução do método. Deve conter as informações relacionadas ao processo, o cronograma de execução, as informações relevantes da aplicação do método, os fatores de influência e os resultados obtidos.

5 Considerações Finais

O objetivo geral desse trabalho foi descobrir alguma métrica capaz de mensurar o tamanho de sistemas de Business Intelligence. Para que o objetivo geral fosse atendido, foi levantada uma questão que deveria ser respondida ao término do trabalho. Nesse sentido, foi definida uma metodologia de pesquisa visando dar suporte a proposição do trabalho. A metodologia consistiu em quatro macroetapas: Referencial Bibliográfico, Revisão Sistemática, Estudo de Caso e Proposta de um Método.

Ao término da primeira abordagem do Estudo Secundário, nenhuma métrica relacionada ao tamanho de sistemas de Business Intelligence foi encontrada, impossibilitando a continuidade do trabalho. Em uma segunda abordagem, foram encontradas métricas de complexidade de processos de negócios e a partir delas foi possível estabelecer uma hipótese que poderia responder a questão levantada. A lista das cinco métricas mais citadas na literatura, resultante da segunda abordagem de Estudo Secundário, foi catalogada no presente trabalho.

A partir da lista de métricas previamente levantadas, foi selecionada a métrica com maior número de referências e desenhado um estudo de caso que tinha como objetivo provar, mesmo que de forma fraca, a relação da complexidade do processo de negócio com o esforço necessário para o desenvolvimento de uma aplicação de Business Intelligence. O estudo de caso foi executado e a hipótese foi fracamente comprovada, possibilitando assim que o trabalho pudesse ser continuado e que a questão levantada fosse respondida.

A última etapa do trabalho foi propor, a partir da métrica fracamente comprovada, um método de aplicação da mesma. Parte dos objetivos específicos já haviam sido alcançados mas a questão ainda não tinha sido respondida. Com a proposta do método de aplicação da métrica, o objetivo geral era alcançado e a questão, finalmente, respondida.

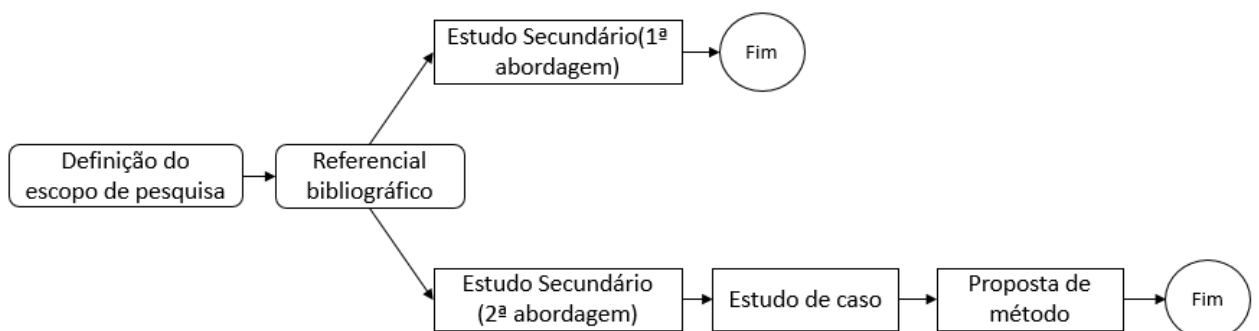


Figura 15 – Passos realizados no TCC

O fluxograma ilustrado na Figura 15 mostra, de forma macro, os passos realizados

no Trabalho de Conclusão de Curso.

5.1 Resultados

Com a execução das atividades do trabalho diversos resultados foram obtidos. Dentre os resultados mais significativos podemos citar uma lista de métricas utilizada na literatura para medir a complexidade de processos de negócio, uma hipótese validada, de forma fraca, através da aplicação de um estudo de caso e uma proposta de método de aplicação da métrica validada.

A lista de métricas é resultante da execução da segunda abordagem da revisão sistemática, ela é composta por um ranking de métricas referenciadas na literatura para mensurar a complexidade de processos de negócio. Para a execução do trabalho, foram selecionadas as cinco primeiras métricas do rank, elas foram descritas na Seção 4 do presente trabalho.

A hipótese levantada para execução do estudo do caso afirmava que quanto maior a complexidade do processo de negócio, maior, também, seria o esforço para desenvolver uma aplicação de BI. Durante o estudo de caso foi possível confirmar a relação afirmada pela hipótese.

O método de aplicação da métrica validada só foi proposto após o término do estudo de caso, e pode ser considerado o resultado de maior importância do trabalho. É através do método que a questão desta pesquisa, apresentada no Capítulo 1, é respondida e os objetivos do trabalho são alcançados. Além disso, o método deixa espaço para a realização de outras pesquisas.

5.2 Limitações e Trabalhos Futuros

Ao longo do desenvolvimento do TCC algumas limitações foram encontradas, tais limitações expandem o leque de oportunidades para o desenvolvimento de trabalhos futuros. Dentre as limitações encontradas podemos citar o uso de estudo de caso ao invés de experimento, os fatores de influência não controlados e a não utilização do método proposto.

A utilização do estudo de caso ao invés do experimento torna a validação da hipótese menos confiável, isso se dá devido à falta de controle das variáveis que envolvem o estudo. Como trabalho futuro, é possível executar um experimento em laboratório controlando pelo menos uma variável envolvida na pesquisa e mantendo as demais constantes.

Os fatores de influência não controlados podem alterar o resultado da aplicação da métrica CFC no estudo de caso. Devido ao curto espaço de tempo e a indisponibilidade

de projetos reais para a execução do estudo, tais fatores não tiveram suas influências estudadas. Como trabalho futuro, é possível estudar a influência desses fatores no resultado final da complexidade do processo de negócio.

O método proposto não foi validado e nem aplicado em nenhuma situação, ainda que ele esteja embasado em métodos já conhecidos na academia. Como trabalho futuro, é possível utilizar o método proposto em contextos reais de desenvolvimento de sistemas de BI e validar os resultados obtidos após essa execução.

Referências

- ALTER, S. *Information systems: a management perspective*. 3th. ed. [S.l.]: Addison-Wesley Educational Publishers Inc, 1996. Citado na página 25.
- ANGELONI, M. T.; REIS, E. S. Business intelligence como tecnologia de suporte a definicao de estrategias para melhoria da qualidade do ensino. *XXX Encontro Nacional de Pos-Graduacao em Administracao*, 2006. Citado na página 21.
- ANTONELLI, R. (Ed.). *Conhecendo o Business Intelligence*. 2009. Citado na página 25.
- ANTONINI, A. et al. Software measures for business processes. 2011. Citado na página 50.
- BARA, A. et al. A model for business intelligence systems development. *Informatica Economica*, 2009. Citado 3 vezes nas páginas 17, 18 e 28.
- BARBIERI, C. *Business Intelligence: modelagem e tecnologia*. Rio de Janeiro: Axcel Books, 2001. Citado 2 vezes nas páginas 21 e 26.
- BIOLCHINI, J.; AL. et. *Systematic review in software engineering*. [S.l.], 2005. Citado na página 79.
- BRERETON, P.; AL. et. Lessons from applying the systematic literature review process within the software engineering domain. *Journal of systems and software*, v. 80, n. 4, 2007. Citado na página 79.
- BRERETON, P. et al. Using a protocol template for case study planning. *European Association of Science Editor*, 2008. Citado 2 vezes nas páginas 51 e 52.
- CARDOSO, J. Evaluating the process control-flow complexity measure. *Proceedings of the IEEE International Conference on Web Services*, 2005. Citado na página 48.
- CARDOSO, J. Process control-flow complexity metric: An empirical validation. *IEEE International Conference on Services Computing*, 2006. Citado na página 48.
- CARDOSO, J. Business process control-flow complexity: Metric, evaluation, and validation. *International Journal of Web Services*, 2008. Citado 5 vezes nas páginas 35, 36, 37, 48 e 55.
- CARDOSO, J. et al. A discourse on complexity of process models. *BPM 2006 Workshops*, 2006. Citado 6 vezes nas páginas 35, 37, 38, 39, 49 e 65.
- COHN, M. *Estimating With Use Case Points*. [S.l.], 2011. Citado na página 33.
- DAVENPORT, T. Process innovation: Reengineering work through information technology. *Harvard Business School Press*, 1993. Citado na página 35.
- DAVENPORT, T. H. Competing on analytics. *Harvard Business Review*, 2006. Citado na página 18.

- DEKKERS, C. (Ed.). *Demystifying Function Points: Lets Understand Some Terminology*. 1998. Citado na página 31.
- DYBA, T.; DINGSOYR, T.; HANSSEN, G. K. Applying systematic reviews to diverse study types: An experience report. *Empirical Software Engineering and Measurement*, 2005. Citado na página 80.
- FENTON, N. E. *Software Metrics: A Rigorous Approach*. [S.l.]: Chapman and Hall, 1991. Citado na página 30.
- GANGADHARAN, G. R. Business intelligence systems: Design and implementation strategies. *2dh Int. Conf. Information Technology Interfaces IT1*, 2004. Citado na página 17.
- GRUHN, V.; LAUE, R. Adopting the cognitive complexity measure for business process models. *5th. IEEE Int. Conf. on Cognitive Informatics*, 2006. Citado 3 vezes nas páginas 37, 38 e 47.
- GRUHN, V.; LAUE, R. Complexity metrics for business process models. *5th. IEEE Int. Conf. on Cognitive Informatics*, 2006. Citado na página 49.
- HOST, M.; RUNESON, P. Checklist for software engineering case study research. *International Symposium on Empirical Software Engineering and Measurement*, 2007. Citado na página 52.
- IEEE, E. Ieee std 610.12-1990(r2002). *IEEE Standard Glossary of Software Engineering Terminology*, 1990. Citado na página 31.
- IFPUG. *Manual de Praticas de Contagem de Pontos de Funcao*. 4.2.1. ed. [S.l.], 2005. Acessado em:. Citado 2 vezes nas páginas 31 e 32.
- JONES, C. *Applied Software Measurement: Global Analysis of Productivity and Quality*. 3. ed. [S.l.]: McGraw-Hill Osborne Media, 2008. Hardcover. ISBN 0071502440. Citado 2 vezes nas páginas 33 e 34.
- KITCHENHAM, B.; BRERETON, P. A systematic review of systematic review process research in software engineering. *Information and Software Technology*, v. 55, n. 12, 2013. Citado na página 79.
- KLUZA, K.; NAPELA, G. Proposal of square metrics for measuring business process model complexity. *Proceedings of the Federated Conference on Computer Science Information Systems*, 2012. Citado 2 vezes nas páginas 39 e 47.
- LASSEN, K.; AALST, W. Complexity metrics for workflow nets. *Information and Software Technology*, 2009. Citado na página 50.
- MATTAR, F. *Pesquisa de Marketing*. [S.l.]: Elsevier Editora Ltda., 2012. Citado na página 55.
- MCCABE, T. A complexity measure. *IEEE Transactions on Software Engineering*, Vol. SE-2, 1976. Citado na página 35.
- MUKETHA, G. et al. A survey of business process complexity metrics. *Information Technology Journal*, 2010. Citado na página 48.

- PAPASTEFANATOS, G. et al. Metrics for the prediction of evolution impact in etl ecosystems: A case study. *Spring - Verlag*, 2012. Citado na página 82.
- PRESSMAN, R. S. *Engenharia de Software: Uma Abordagem Profissional*. 7. ed. Porto Alegre: McGraw-Hill, 2011. ISBN 978-85-63308-33-7. Citado 3 vezes nas páginas 30, 31 e 32.
- PRIMAK, F. V. *Decisoes com BI (Business Intelligence)*. [S.l.]: Fabio Vinicius Primak, 2008. ISBN 8573937149, 9788573937145. Citado 6 vezes nas páginas 17, 22, 23, 24, 26 e 27.
- ROMNEY, M. B.; STEINBART, P. J. *Accounting Information Systems*. 8th. ed. New Jersey: Prentice-Hall, Inc., 2000. Citado na página 25.
- RUNESON, P. et al. *Case Study Research in Software Engineering: Guidelines and Examples*. [S.l.]: John Wiley and Sons, 2012. Citado na página 51.
- SERRA, L. *A essencia do Business Intelligence*. 1st. ed. Sao Paulo: Berkely Brasil, 2002. Citado na página 22.
- SHAO, J.; WANG, Y. A new measure of software complexity based on cognitive weights. *IEEE Canadian Journal of Electrical and Computer Engineering*, 2003. Citado na página 37.
- SOLICHAH, I. et al. Exploration on software complexity metrics for business process model and notation. *ICACISIS*, 2013. Citado 2 vezes nas páginas 34 e 47.
- SOMMERVILLE, I. *Engenharia de Software*. 8. ed. Sao Paulo: Pearson Addison-Wesley, 2007. ISBN 978-0-13-703515-1. Citado 3 vezes nas páginas 30, 31 e 33.
- THAMMARAK, K. Survey complexity metrics for reusable business process. *Proceedings from 1st National Conference*, 2010. Citado 2 vezes nas páginas 35 e 49.
- WHOLIN, C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. *EASE'14*, 2014. Citado 2 vezes nas páginas 42 e 43.

Apêndices

APÊNDICE A – Primeira Etapa do Estudo Secundário

Conforme mostrado, um dos passos desta pesquisa constitui na condução de um estudo secundário baseado nos procedimentos definidos para a execução de uma Revisão Sistemática com o objetivo de encontrar e listar as principais métricas utilizadas para mensurar as aplicações de Business Intelligence.

A.1 Planejamento do Estudo

O protocolo do estudo secundário foi definido baseando-se nos modelos propostos por BIOLCHINI e al. (2005) e nas recomendações apresentadas no estudo de KITCHENHAM e BRERETON (2013). A StArt foi utilizada para apoiar o estudo. A StArt é uma ferramenta para auxiliar e automatizar o processo do estudo secundário.

A.1.1 Objetivo

O objetivo desta revisão foi *analisar* estudos e relatos de iniciativas de medição de sistemas de Business Intelligence no seu processo de desenvolvimento *com o propósito de* identificar e analisar métricas que auxiliem na estimativa de tempo e esforço necessários para a condução do projeto *com relação ao* dimensionamento do sistema na fase de planejamento, *do ponto de vista* da equipe de desenvolvimento, *no contexto de* fornecimento de soluções BI.

A.1.2 Questão de Pesquisa

BRERETON e al. (2007) recomenda que, após a definição do objetivo, sejam elaboradas as questões de pesquisa. A seguinte questão foi identificada para este estudo: Quais métricas são utilizadas para mensurar sistemas BI?

É necessário também, segundo BIOLCHINI e al. (2005), definir alguns outros aspectos relacionados ao escopo e à especificidade das questões de pesquisas:

- **População:** artigos publicados sobre medição de sistemas BI;
- **Intervenção:** medição de sistemas BI dentro do processo de desenvolvimento no contexto de fornecedoras de solução em BI;
- **Resultados:** elementos de influência na atividade de medição de sistema BI;

- **Aplicação:** servir de base ou apoiar pesquisas envolvendo: metodologia de medição de sistema BI; definição de métricas para medição de sistemas BI; e criação de métricas para medição de sistema BI.

A.1.3 Critérios de Seleção de Trabalhos

Foram definidos os seguintes critérios de inclusão (CI) e exclusão (CE):

- **CI1.** Os artigos devem estar escritos em inglês.
- **CI2.** Os resultados devem estar disponíveis integralmente.
- **CI3.** Os artigos devem apresentar algum estudo sobre métricas de sistema BI na visão do desenvolvedor.
- **CI4.** Os artigos devem apresentar estudos sobre frameworks, métodos ou boas práticas ou qualquer outro item aplicável em medição de sistema BI.
- **CE1.** Os artigos que apresentam um estudo superficial sobre o tema.
- **CE2.** Artigos que estão duplicados.
- **CE3.** Artigos que não considerem o contexto de desenvolvimento de sistemas BI.

A.1.4 Procedimentos de Seleção

A *string* de busca foi executada nas bases de pesquisas: IEEEXplore, ACM, Capes e Scopus. Dentre as bases supracitadas, as bases IEEEXplore e ACM são consideradas bases específicas da Engenharia de Software, enquanto as da Capes e Scopus englobam todas as áreas de conhecimento. Todos os artigos retornados foram armazenados na ferramenta que auxilia e automatiza o processo de revisão (StArt). A *string* construída foi executada tanto em bases cuja língua é o inglês quanto em bases cuja língua é o português. Cada vez que a *string* foi executada, os dados obtidos foram documentados e, depois, contabilizados os registros repetidos. Os títulos, resumos e palavras-chaves foram lidos, avaliando os critérios de inclusão e exclusão definidos anteriormente. Os artigos que passaram pelos critérios de inclusão e exclusão prosseguiram para a etapa de análise de qualidade e extração de resultados.

A.1.5 Avaliação de Qualidade

Para realizar a avaliação de qualidade dos estudos, foi utilizada uma lista de verificação apresentada por [DYBA, DINGSOYR e HANSEN \(2005\)](#), que consiste em cinco aspectos: se há algum interesse para o estudo; se as evidências são válidas; se as evidências

são importantes; se elas podem ser usadas na prática; e se as evidências no estudo em questão são consistentes com as evidências dos outros estudos disponíveis.

A.1.6 Extração dos Dados

Para extrair informações dos artigos resultantes, foram selecionadas os seguintes dados para serem coletados:

- Título;
- Autor;
- Data de publicação;
- Fonte de publicação;
- Resumo;
- Listagem das métricas de medição de sistema BI.

A.2 Condução da Primeira Etapa do Estudo Secundário

A pesquisa utilizando a primeira versão da string foi executada em agosto de 2014. A string foi a seguinte: ((`"BUSINESS INTELLIGENCE"`OR `"BI"`OR `"BI SYSTEMS"`OR `"Inteligência de Negócio"` OR `"ANALYTICS BUSINESS"` OR `"ANÁLISE DE NEGÓCIO"`) AND (`"METRICS"`OR `"SIZE METRICS"`OR `"MÉTRICAS"` OR `"ESTIMATIVA DE TAMANHO"` OR `"SIZE ESTIMATION"` OR `"SOFTWARE SIZING"`)).

A string foi criada de maneira a ser genérica e retornar a maior quantidade possível de artigos. Entretanto, poucos foram os artigos resultantes e menos ainda os que respondiam à questão identificada.

Na base da IEEE, nenhuma publicação foi obtida. Na ACM, também não houve resultados. Na base de dados Scopus, 91 publicações, e na Capes, doze. Houve 40 publicações em comum entre as bases da Capes e da Scopus. Em conferências nacionais, não foram encontrados artigos que atendessem a expressão. A falta de publicações mostra que a temática ainda possui poucos estudos relacionados, apesar da importância e do crescimento que em que o tema supracitado se encontra.

Após a execução da busca, todas as publicações existentes foram submetidas aos procedimentos de seleção, avaliação de qualidade e extração dos dados descritos no protocolo, resultando na Tabela 8.

Conforme pode ser visto na Tabela 8 os resultados não foram satisfatórios. Entretanto, um dos artigos analisados, apesar de não responder a pergunta apresenta um estudo de caso relacionado a métricas.

O artigo, intitulado *Metrics for the Prediction of Evolution Impact in ETL Ecosystems: A Case Study* de PASTEFANATOS et al. (2012), foi escrito em 2012 e tem como objetivo tentar minimizar os efeitos de manutenção da etapa de Extract-Transform-Load (ETL), através do gerenciamento de métricas de banco de dados. Durante a fase de design e de todo o ciclo de vida da ETL, o desenvolvedor da arquitetura precisa desenhar e melhorar o design da ETL de forma que a performance e a garantia de precisão sejam satisfatórias.

Dentro do contexto de BI, esse é um estudo que consegue medir parte do sistema com foco no banco de dados. Entretanto, esse estudo não se mostrou completo o suficiente para ser aproveitado para a revisão de literatura que o encontrou; ou seja, o artigo não responde à questão levantada no planejamento.


Após o término deste estudo secundário e dos resultados poucos satisfatórios, uma decisão importante foi tomada: o foco de pesquisa precisou ser repensado. Em vez de tentar encontrar estudos que definam e utilizam métricas de sistema BI já prontas, o foco tornou-se encontrar métricas para medir a complexidade de modelos de processo de negócio e, a partir disso, definir métricas derivadas que consigam mensurar o esforço necessário para desenvolver um sistema BI que apoie o processo em questão.

Tabela 8 – Resultado da Primeira Etapa do Estudo Secundário

Base	String	Resultado	CI1	CI2	CI3	CI4	CE1	CE2	CE3
IEEE	("BUSINESS INTELLIGENCE"OR "BI SYSTEMS"OR "Inteligência de Negócio" OR "ANALYTICS BUSINESS" OR "ANÁLISE DE NEGÓCIO") AND ("METRICS"OR "SIZE METRICS"OR "MÉTRICAS" OR "ESTIMATIVA DE TAMANHO" OR "SIZE ESTIMATION" OR "SOFTWARE SIZING"))	0	0	0	0	0	0	0	0
ACM	(Abstract:("BUSINESS INTELLIGENCE"OR "BI SYSTEMS"OR "Inteligência de Negócio" OR "ANALYTICS BUSINESS" OR "ANÁLISE DE NEGÓCIO") AND Abstract:("METRICS"OR "SIZE METRICS"OR "MÉTRICAS" OR "ESTIMATIVA DE TAMANHO" OR "SIZE ESTIMATION" OR "SOFTWARE SIZING"))	0	0	0	0	0	0	0	0
Scopus	(ABS("BUSINESS INTELLIGENCE"OR "BI SYSTEMS"OR "Inteligência de Negócio"OR "ANALYTICS BUSINESS"OR "ANÁLISE DE NEGÓCIO") AND ABS("METRICS"OR "SIZE METRICS"OR "MÉTRICAS"OR "ESTIMATIVA DE TAMANHO"OR "SIZE ESTIMATION"OR "SOFTWARE SIZING"))	91	50	4	2	1	0	0	0
Capes	(("BUSINESS INTELLIGENCE"OR "BI SYSTEMS"OR "Inteligência de Negócio"OR "ANALYTICS BUSINESS"OR "ANÁLISE DE NEGÓCIO") AND ("METRICS"OR "SIZE METRICS"OR "MÉTRICAS"OR "ESTIMATIVA DE TAMANHO"OR "SIZE ESTIMATION"OR "SOFTWARE SIZING"))	58	56	5	3	0	0	0	0

Anexos

ANEXO A – Termo de Cessão de Direito de Pesquisa



Brasília, ____ de _____ de 2014

Termo de Cessão de Direitos de Pesquisa

Pelo presente instrumento particular as partes: Biggdata, pessoa jurídica de direito privado, com sede no endereço SAUS Quadra 4, Bloco A - Sala 209 - Ed. Victória Office Tower, Cidade: Brasília - DF, Cep: 70070-938, inscrito sob o CNPJ/MF n.º 26.430.710/0001-45, abaixo assinado por seu representante local, adiante simplesmente denominada **CEDENTE**; e Luciano Hideaky de Macedo Endo CPF n.º 036.917.011-32 e Fabiana Freitas Mendes CPF n.º 957.658.831-68 ambos de nacionalidade brasileira adiante simplesmente denominados **CESSIONÁRIOS**.

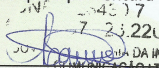
O **CEDENTE**, **CEDE** ao **CESSIONÁRIO** os direitos relacionados a marca Biggdata para uso exclusivo das informações da empresa no desenvolvimento do Trabalho de Conclusão de Curso.

O **CESSIONÁRIO** manifesta conhecer que:

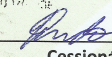
- a) a marca é de propriedade da empresa;
- b) é obrigatória a menção expressa da Biggdata em todo trabalho realizado com envolvimento parcial ou total dos dados da empresa.

Além disso, por intermédio do presente instrumento, o **CESSIONÁRIO** expressa comprometimento e concordância em:

- a) reconhecer que todos os documentos, dados e informações relativos à empresa Biggdata não podem ser divulgadas com outros intuitos;
- b) reconhecer que todos os documentos que envolvam terceiros não podem ser divulgados ou sequer mencionados no trabalho;
- c) manter sigilo (escrito e verbal) de todos os dados, informações e dos materiais obtidos e utilizados, aos quais porventura tenha acesso.



 SAUS Quadra 4 - Bloco A - Sala 209 - Ed. Victória Office Tower
 Laura Solano de Carvalho
 CNPJ nº 26.430.710/0001-45 e 8131 1324
 Brasília-DF



 Cessionário
 Luciano Hideaky de Macedo Endo

 Cessionário
 Fabiana Freitas Mendes