

**INSTITUTO  
FEDERAL**  
Paraíba

**Instituto Federal de Educação, Ciência e Tecnologia da Paraíba**

**Campus João Pessoa**

**Programa de Pós-Graduação em Tecnologia da Informação**

**SAVYO IGOR DA NÓBREGA SANTOS**

**UMA INVESTIGAÇÃO SOBRE O FENÔMENO DA  
ESTIMATIVA DE ESFORÇO EM PROJETOS DE BIG  
DATA NA ADMINISTRAÇÃO PÚBLICA FEDERAL**

**DISSERTAÇÃO DE MESTRADO**

**JOÃO PESSOA – PB**

**2022**

**Savyo Igor da Nóbrega Santos**

**Uma Investigação sobre o Fenômeno da Estimativa de Esforço  
em Projetos de Big Data na Administração Pública Federal**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Tecnologia da Informação pelo Programa de Pós-Graduação em Tecnologia da Informação do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba – IFPB.

Orientadora: Prof. Dra. Crishane Azevedo Freire

João Pessoa – PB

2022

Dados Internacionais de Catalogação na Publicação (CIP)  
Biblioteca Nilo Peçanha do IFPB, *campus* João Pessoa.

S237i Santos, Savyo Igor da Nóbrega.

Uma investigação sobre o fenômeno da estimativa de esforço em projetos de Big data na administração pública federal / Savyo Igor da Nóbrega Santos. – 2022.

70 f. : il.

Dissertação ( Mestrado em Tecnologia da Informação ) – Instituto Federal de Educação da Paraíba / Programa de Pós-Graduação em Tecnologia da Informação, 2022.

Orientação: Profa. D.ra Crishane Azevedo Freire.

1. Medição de software. 2. Estimativa de esforço. 3. Big Data. 4. Análise de pontos de função. 5. Administração pública federal. I. Título.

CDU 004.41(043)

Lucrecia Camilo de Lima  
Bibliotecária – CRB 15/132


**Savyo Igor da Nóbrega Santos**

**Uma Investigação sobre o Fenômeno da Estimativa de Esforço em Projetos de Big Data na Administração Pública Federal**

*Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Tecnologia da Informação pelo Programa de Pós-Graduação em Tecnologia da Informação do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba – IFPB.*

**Aprovado em 27 de maio de 2022.**


**BANCA EXAMINADORA:**

Documento assinado digitalmente  
 Juliana Dantas Ribeiro Viana de Medeiros  
Data: 30/06/2022 12:12:56-0300  
Verifique em <https://verificador.iti.br>

---

*Profa. Dra. Juliana Dantas Ribeiro Viana de Medeiros*

*Avaliador Interno*

Documento assinado digitalmente  
 Nadja da Nobrega Rodrigues  
Data: 29/06/2022 06:16:47-0300  
Verifique em <https://verificador.iti.br>

---

*Profa. Dra. Nadja da Nóbrega Rodrigues*

*Avaliador Externo*



---

*Profa. Dra. Crishane Azevedo Freire*

*Orientador*

Visto e permitida a impressão

João Pessoa

Prof. Dr. Francisco Petrônio A. de Medeiros

Coordenador PPPGTI

# AGRADECIMENTOS

Agradeço primeiramente a Deus, pelo dom da vida e por permitir que eu chegasse até aqui. A minha família, meu porto seguro, por todo o amor e apoio recebidos em todos os momentos. Aos meus amigos, pelas mensagens de carinho e encorajamento na caminhada. Sem vocês, eu também não teria chegado até aqui.

Agradeço aos professores do Programa de Pós-Graduação em Tecnologia da Informação, do Instituto Federal da Paraíba (IFPB), em especial a minha orientadora Dra. Crishane Azevedo Freire, por estarem sempre dispostos a ajudar e a tornarem esse caminho mais fácil de trilhar.

Aos colegas da empresa, que também me deram força e vivem as alegrias e lutas do trabalho comigo durante toda a semana.

A todos que, direta ou indiretamente, colaboraram para o sucesso deste trabalho.

## RESUMO

Os avanços promovidos pela adoção em larga escala das Tecnologias de Informação e Comunicação (TIC) em vários setores da sociedade possibilitaram a geração e armazenamento de uma grande quantidade de dados, que vêm aumentando com o passar dos anos em diversas fontes de informação. O termo *Big Data* foi criado para representar essa grande quantidade de dados, com variadas formas e dimensões, com o objetivo de gerar valor para o usuário de forma rápida. Devido ao seu alto poder de processamento e complexidade, um dos desafios do *Big Data* atualmente está relacionado à estimativa de esforço em softwares desenvolvidos nas empresas e organizações. No contexto da Administração Pública Federal, onde os órgãos estão submetidos a leis reguladoras, restringindo sua forma de estimativa, a métrica Análise de Pontos de Função (APF) é utilizada para medir o esforço nos contratos de projetos de software. No entanto, a métrica possui limitações e, por isso, não vem se mostrando adequada para este fim. Em projetos de *Big Data*, não apenas APF, mas também outros métodos de esforço não se mostram adequados para esse contexto. A partir do que foi exposto, este trabalho visa investigar o fenômeno da estimativa de esforço na literatura para propor um modelo mais adequado a projetos de Big Data. Para isso, foi realizado um estudo de caso numa empresa pública federal com o objetivo de analisar o seu processo de medição utilizando modelos com base em unidades de serviços técnicos (UST) e em requisitos não funcionais (SNAP), em complemento à APF. O resultado obtido no estudo mostrou a boa aplicação dos modelos SNAP e UST, com base na cobertura dos requisitos, granularidade das tarefas e tempo médio de medição. A avaliação do modelo de estimativa de esforço, obtido como resultado do estudo, foi realizada utilizando o algoritmo de classificação KNN (*K Nearest Neighbor*) e o resultado validou o que foi apresentado no trabalho.

**Palavras-chaves:** Medição de Software; Estimativa de Esforço; Big Data; Análise de Pontos de Função; Administração Pública Federal; SNAP; Unidade de Serviço Técnico.

# ABSTRACT

The advances promoted by the large-scale adoption of Information and Communication Technologies in various sectors of society have enabled the generation and storage of a large amount of data, which has been increasing over the years in various sources of information. The term Big Data was created to represent this large volume of data, with different shapes and dimensions, with the aim of generating value for the user quickly. One of the challenges of Big Data today is related to the estimation of effort in software developed in companies and organizations due to its high procession power and complexity. In the context of Federal Public Administration, where bodies are subject to regulatory laws, restricting their estimation methods, the Function Point Analysis (FPA) is used to measure effort in software project contracts. However, the metric has limitations and, therefore, has not been shown to be adequate for this purpose. In Big Data Projects, not only FPA, but also other methods of effort are not suitable for this context. Based on what has been exposed, the purpose of work is to investigate the phenomenon of effort estimation in the literature and in the industry, in order to propose a model that is more suitable for Big Data Projects. For this, a case study was carried out in a federal public company with the objective of analyzing its measurement process using models based on technical service units and based on non-functional requirements (SNAP) in addition to FPA. The results showed the good application of the models SNAP and based on technical service units based on requirements coverage, task granularity and average measurement time. The evaluation of the effort estimation model obtained as a result of the study was performed using the KNN (K Nearest Neighbor) classification algorithm and the result validated what was presented in the work.

**Keywords:** Software Measurement; Effort Estimation; Big Data; Function Point Analysis; Federal Public Administration; SNAP.

## LISTA DE FIGURAS

|   |    |
|---|----|
| Figura 1 - Etapas da pesquisa   | 20 |
| Figura 2 - Modelo de Análise de Pontos de Função                                      | 19 |
| Figura 3 - Modelo de contagem SNAP  | 20 |
| Figura 4 - Passos do estudo de caso realizado   | 31 |
| Figura 5 - Linha do tempo da coleta de dados  | 19 |
| Figura 6 - Tempo de experiência no processo de contagem em APF dos Grupos 2 e 3       | 40 |
| Figura 7 - Modelos de medição conhecidos além de APF pelos Grupos 2 e 3               | 19 |
| Figura 8 - Nível de conformidade de APF na medição das demandas dentro da organização | 41 |
| Figura 9 - Classificação do treinamento com o Grupo 1                                 | 19 |
| Figura 10 - Projetos e cenários do estudo de caso                                     | 45 |
| Figura 11 - Relação dos cenários com os parâmetros do estudo de caso                  | 47 |
| Figura 12 - Relação dos Vs do Big Data com os cenários do estudo de caso              | 47 |
| Figura 13 - Nível de conformidade do uso complementar de SNAP e UST a APF             | 50 |
| Figura 14 - Construção do modelo  | 53 |
| Figura 15 - Modelo após o estudo de caso  | 54 |
| Figura 16 - Gráfico da variação da acurácia na validação cruzada 5-fold               | 61 |



## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 1 – Questões de Pesquisa e etapas da pesquisa  | 15 |
| Tabela 2 – Trabalhos relacionados   | 30 |
| Tabela 3 – Script das entrevistas realizadas após alinhamento com Grupos 2 e 3              | 37 |
| Tabela 4 – Script do questionário realizado após o treinamento com o Grupo 1                | 37 |
| Tabela 5 – Script do questionário realizado após a execução dos cenários com o Grupo 1      | 38 |
| Tabela 6 – Script das entrevistas realizadas após a execução dos cenários com o Grupo 1     | 38 |
| Tabela 7 – Script das entrevistas realizadas após avaliação dos resultados com Grupos 2 e 3 | 38 |
| Tabela 8 – Relação das tarefas realizadas e níveis da coleta de dados                       | 38 |
| Tabela 9 – Cenários – Projeto A   | 44 |
| Tabela 10 – Cenários – Projeto B  | 44 |
| Tabela 11 – Cenários – Projeto C  | 45 |
| Tabela 12 – Valores de referência para pergunta 1 do questionário                           | 46 |
| Tabela 13 – Níveis dos parâmetros do estudo de caso   | 49 |
| Tabela 14 – Pseudocódigo do algoritmo de classificação KNN                                  | 56 |
| Tabela 15 – Resultados obtidos a partir do modelo   | 57 |
| Tabela 16 – Resultados das combinações possíveis de cenários, após análise dos Grupos 2 e 3 | 57 |
| Tabela 17 – Pseudocódigo do algoritmo de recomendação dos cenários feito pelos Grupos 2 e 3 | 58 |
| Tabela 18 – Exemplo do algoritmo de recomendação para a linha 3 da Tabela 17                | 58 |
| Tabela 19 – Exemplo do algoritmo de recomendação para a linha 21 da Tabela 17               | 58 |

# LISTA DE ABREVIATURAS E SIGLAS

|          |   |
|----------|---|
| AIE      | Arquivo de Interface Externa  |
| ALI      | Arquivo Lógico Interno  |
| ALR      | Arquivo Lógico Referenciado   |
| APF      | Análise de Pontos de Função   |
| BI       | <i>Business Intelligence</i>  |
| CE       | Consulta Externa  |
| CPM      | Manual de Práticas de Contagem  |
| DW       | <i>Data Warehouse</i>   |
| EE       | Entrada Externa   |
| HDFS     | <i>Hadoop Distributed File System</i>   |
| ID       | Item de Dado  |
| IFPB     | Instituto Federal da Paraíba  |
| IFPUG    | <i>International Function Point User Group</i>  |
| IRAMUTEQ | <i>Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires</i> |
| KNN      | <i>K-nearest neighbor</i>   |
| LGPD     | Lei Geral de Proteção de Dados  |
| MP       | Ministério do Planejamento  |
| PF       | Ponto de Função   |
| RL       | Registro Lógico   |
| RNF      | Requisito Não-Funcional   |
| SE       | Saída Externa   |
| SLTI     | Secretaria de Logística e Tecnologia da Informação  |
| SISP     | Sistema de Administração dos Recursos de Tecnologia da Informação                           |
| SNAP     | Processo de Avaliação de Software Não-Funcional   |
| TI       | Tecnologia da Informação  |
| TIC      | Tecnologia da Informação e Comunicação  |
| TCU      | Tribunal de Contas da União   |
| UML      | <i>Unified Modeling Language</i>  |
| UST      | Unidade de Serviço Técnico  |
| USTIBB   | Unidade de Medida para Serviços Técnicos  |

# SUMÁRIO

|   |    |
|---|----|
| <b>INTRODUÇÃO</b>   | 11 |
| Motivação e Definição do Problema   | 11 |
| Objetivos   | 13 |
| Objetivo geral  | 13 |
| Objetivos específicos   | 13 |
| Metodologia de Pesquisa   | 14 |
| Aplicabilidade e Abordagem  | 15 |
| Estrutura do Documento  | 16 |
| <b>FUNDAMENTAÇÃO TEÓRICA</b>  | 17 |
| Medição de Software   | 17 |
| Aspectos da Administração Pública Federal   | 23 |
| Big Data  | 24 |
| <b>TRABALHOS RELACIONADOS</b>   | 28 |
| Estimativa de esforço em lógica fuzzy e métricas de desempenho                                  | 28 |
| Análise de métricas na Administração Pública Federal  | 28 |
| Métricas de medição em projetos de Big Data   | 28 |
| Modelo híbrido de estimativa de esforço com Pontos de Casos de Uso                              | 29 |
| COCOMO II, APF e Casos de Uso em projetos de Big Data   | 29 |
| Métricas de software em sistemas de BI  | 29 |
| Análise e sumarização dos trabalhos   | 29 |
| <b>ESTUDO DE CASO</b>   | 31 |
| Planejamento  | 31 |
| Protocolo e Preparação dos Casos  | 32 |
| Coleta de Dados   | 35 |
| Resultados e Discussões   | 38 |
| Considerações   | 51 |
| <b>MODELO DE ESTIMATIVA DE ESFORÇO EM PROJETOS DE BIG DATA DA ADMINISTRAÇÃO PÚBLICA FEDERAL</b> | 53 |
| Construção  | 53 |
| Validação   | 55 |
| <b>CONCLUSÃO</b>  | 61 |
| Contribuições da Pesquisa   | 62 |
| Trabalhos Futuros   | 62 |

# 1. INTRODUÇÃO

Neste capítulo, é apresentada uma visão geral do problema que será investigado neste trabalho, os seus objetivos gerais e específicos, uma breve introdução da metodologia utilizada na pesquisa, isto é, o estudo de caso, e, por fim, a forma que este documento está estruturado.

## 1.1. Motivação e Definição do Problema

Os avanços promovidos pela adoção em larga escala das Tecnologias de Informação e Comunicação (TIC) em vários setores da sociedade possibilitaram a geração e armazenamento de dados sobre serviços, produtos, pessoas e diversas fontes de informação (SANTOS, SALVADOR e CRUZES, 2011). Com o passar dos anos, a quantidade desses dados aumentou consideravelmente (IBM, 2016), o que fez o mundo se tornar cada vez mais conectado, com as informações fluindo rapidamente, afetando, dessa maneira, muitas pessoas (ALAOUI, GAHI e MESSOUSSI, 2019). Além disso, com a chegada de novos dispositivos e tecnologias, como GPS e a Internet das Coisas (IoT), essa quantidade está aumentando cada vez mais (MACHADO, 2018) (REIS e SÁ, 2020).

O termo Big Data foi criado para representar essa grande quantidade de dados cujo crescimento é exponencial e cuja dimensão está além da habilidade das ferramentas tradicionais de processamento e análise de dados (TAURION, 2013).

A definição do Big Data surgiu no início dos anos 2000 (LANEY, 2001) e está baseada em 3Vs principais. São eles: (i) volume, (ii) variedade e (iii) velocidade. Isto é, um Big Data pode ser definido como um grande conjunto de dados, de variadas formas e dimensões, cujo objetivo está concentrado na geração de respostas rápidas, atendendo aos requisitos do usuário. Além dos 3Vs principais, outros Vs são utilizados na literatura para mostrar outras características do Big Data, como, por exemplo, a veracidade e o valor das informações que são geradas, além da forma em que elas são visualizadas ao final do processo (ALAOUI, GAHI e MESSOUSSI, 2019) (KHAN et al., 2018) (SAMOSIR et al., 2017).

Essa quantidade enorme de dados vem sendo trabalhada e processada dentro das empresas e organizações, que precisam se adaptar e desenvolver novas soluções para acomodar o processamento de Big Data. Das fases existentes no desenvolvimento de software, o planejamento é muito importante, pois é nele onde é medida, entre outras métricas, a estimativa de esforço dos projetos, que é necessária para que o projeto cumpra, entre outros fatores, com o orçamento de forma organizada (NATHANAEL, HENDRADJAYA e SUNINDYO, 2015).

No contexto da Administração Pública Federal, as empresas de Tecnologia da Informação (TI) desenvolvem soluções para o Governo Federal e precisam estar alinhadas com as tecnologias existentes. A forma de medir esforço dos projetos nessas empresas segue a recomendação da Instrução Normativa SLTI/MP N° 4, de 11 de setembro de 2014 (MPOG, 2014), e da Portaria SLTI/MP N° 31, de 29 de novembro de 2010 (MPOG, 2010), os quais utilizam o modelo de estimativa de esforço baseado em Análise de Pontos de Função (APF) em contratos de projetos de software.

APF é um modelo de estimativa de esforço e custo do desenvolvimento de software, cujo objetivo é dimensionar os requisitos funcionais existentes na solução a ser desenvolvida, com base no ponto de vista do usuário. A análise é feita baseada na documentação existente do projeto e não depende da tecnologia envolvida (HAI, NHUNG e THÁI, 2019).

Em outras palavras, a métrica Ponto de Função (PF) é utilizada como referência para medir e custear os contratos de serviços de desenvolvimento e manutenção de sistemas estabelecidos entre instituições públicas e empresas prestadoras desse tipo de serviço (SISP, 2018).

No entanto, diversos órgãos federais têm apresentado dificuldades na medição das soluções de Tecnologia da Informação (TI) utilizando PF. Os relatos indicam que a métrica utilizada não mensura corretamente o esforço condizente com o trabalho executado. Isso acaba gerando divergência e mostra um desequilíbrio econômico-financeiro para a prestadora dos serviços (ALMEIDA e FURTADO, 2019).

Com relação às soluções de Big Data, que lidam com grandes conjuntos de dados, de variadas formas, e que possuem etapas complexas de análise, transformação e visualização de dados (MACHADO, 2018), os métodos atuais de estimativa de esforço não parecem ser adequados para esse contexto específico (VENKATRAMAN R. e VENKATRAMAN S., 2019) (NATHANAEL, HENDRADJAYA e SUNINDYO, 2015).

Ainda no contexto da estimativa de esforço dentro das organizações federais, é interessante citar outros aspectos que norteiam as decisões a serem tomadas nessas empresas e que podem dificultar esse processo de estimativa. São as leis e instruções normativas que cercam a Administração Pública Federal, com relação a medição de softwares contratados pela União, que restringem a forma de estimar as soluções de TI (MPOG, 2010) (MPOG, 2014) (GOVERNO DIGITAL, 2020), gerando valores diferentes do esperado pela organização, por não conseguir estimar as soluções de alta complexidade (ALMEIDA, MONTEIRO e FURTADO, 2018).

A partir do que foi exposto, temos, nesse contexto, que o problema de pesquisa deste trabalho consiste na dificuldade em estimar o esforço de projetos de software que utilizam Big Data e que estão inseridos no âmbito da Administração Pública Federal.

## 1.2. Objetivos

Nesta seção, serão apresentados o objetivo geral, com as respectivas questões de pesquisa, e os objetivos específicos desta pesquisa.

### 1.2.1. Objetivo geral

O objetivo geral da pesquisa é investigar o fenômeno da estimativa de esforço na literatura para propor um modelo mais adequado a projetos de Big Data no contexto de uma empresa que pertence à Administração Pública Federal.

Levando em consideração a necessidade de utilizar diferentes ferramentas para averiguar o fenômeno citado no parágrafo anterior, as seguintes questões de pesquisa foram elaboradas para nortear todo o trabalho:

*QP1. O modelo atual de medição em Pontos de Função atende aos projetos de Big Data nas empresas públicas federais?*

*QP2. O uso de modelos de medição com base em unidades de serviços técnicos (UST) e em requisitos não funcionais (SNAP) complementa as medições dos projetos de Big Data nas organizações que fazem parte da Administração Pública Federal?*

*QP3. As características relacionadas ao volume, variedade e velocidade dos dados de um projeto de Big Data devem ser levadas em consideração na sua estimativa de esforço?*

### 1.2.2. Objetivos específicos

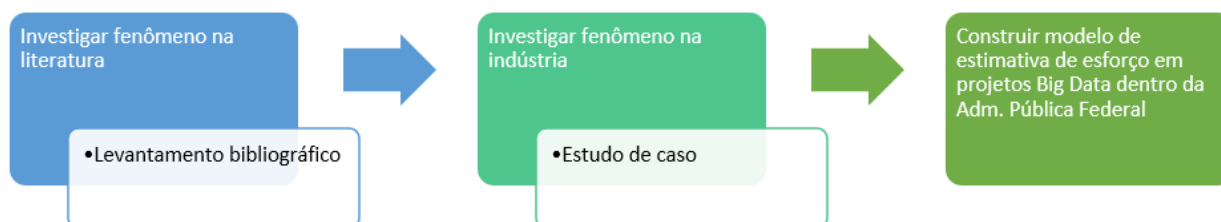
Para alcançar o objetivo geral, o trabalho prevê os seguintes objetivos específicos:

- Fazer uma revisão da literatura sobre o tema proposto na pesquisa;
- Descrever o processo atual de medição em projetos de Big Data dentro de uma empresa pública federal;
- Realizar um estudo de caso com a finalidade de averiguar a eficácia do processo atual e comparar otimizações a esse processo;
- Analisar os dados do estudo de caso realizado na empresa pública federal;
- Propor um modelo de medição em projetos de Big Data no âmbito das organizações que fazem parte da Administração Pública Federal.

### 1.3. Metodologia de Pesquisa

Com base nos objetivos gerais e específicos descritos na seção anterior, as etapas da pesquisa são mostradas na Figura 1.

**Figura 1 - Etapas da pesquisa.**



**Fonte: Desenvolvido pelo autor.**

A primeira etapa da pesquisa consistiu no levantamento bibliográfico sobre o estado da arte em que o trabalho está inserido, ou seja, no campo da medição e estimativas de esforço em projetos de Big Data no contexto da Administração Pública Federal.

A importância desse levantamento foi, além de investigar o estado da arte da pesquisa, identificar possíveis lacunas existentes na área. No entanto, foram encontrados poucos trabalhos relacionados com o tema proposto neste documento. Isso fez com que houvesse a necessidade de investigar o problema definido dentro de um contexto real, realizando uma análise mais aprofundada da questão de pesquisa, para que, com esse conhecimento, seja proposta uma abordagem que ajude a solucionar o problema.

Por essas razões, a metodologia que será utilizada para cumprir o objetivo da pesquisa consiste no desenvolvimento de um estudo de caso.

De acordo com Yin (2005), um estudo de caso é um tipo de pesquisa empírica que investiga um fenômeno contemporâneo dentro de seu contexto da vida real, especialmente quando as fronteiras entre o contexto e o fenômeno analisado não estão claramente definidos. Diferente de um experimento, em que as amostras são analisadas com mudanças nos fatores, o estudo de caso analisa a amostra somente do caso que está sendo estudado.

Para Rodrigues (2006), o estudo de caso consiste numa pesquisa aplicada, na qual ocorre uma análise mais aprofundada do problema, de maneira que seja permitido o seu amplo e detalhado conhecimento.

O estudo de caso que será realizado nesta pesquisa pode ser classificado como exploratório, uma vez que visa investigar um fenômeno de forma inicial para que sejam derivadas algumas hipóteses e, eventualmente, teorias a respeito do assunto.

Após a realização e como um dos resultados do estudo de caso, a última etapa da pesquisa consistiu na construção do modelo de estimativa de esforço para projetos de Big Data no contexto das empresas públicas federais. Os detalhes sobre o modelo desenvolvido no trabalho estão descritos no Capítulo 5.

As etapas da pesquisa descritas nesta seção possuem relação com todas as questões de pesquisa apresentadas na seção anterior, conforme pode ser visto na Tabela 1.

**Tabela 1 – Questões de Pesquisa e etapas da pesquisa.**

| Questão de Pesquisa | Etapas da pesquisa         |                |                      |
|---------------------|----------------------------|----------------|----------------------|
|                     | Levantamento Bibliográfico | Estudo de Caso | Construção do modelo |
| QP1                 | x                          | x              |                      |
| QP2                 |                            | x              | x                    |
| QP3                 |                            | x              | x                    |

Fonte: Desenvolvido pelo autor. Adaptação de Medeiros (2017).

#### 1.4. Aplicabilidade e Abordagem

A pesquisa desenvolvida neste trabalho consistiu numa pesquisa aplicada, que, segundo Rodrigues (2006), tem o propósito de produzir conhecimento que tenha aplicação prática e que cumpra a necessidade de solução de problemas concretos.

Além disso, esse tipo de estudo possui a característica de desenvolver ou testar teorias em campos práticos e tem o objetivo de ampliar o conhecimento de processos organizacionais e produzir importantes conclusões a respeito do problema analisado (FLICK, 2013).

No contexto deste trabalho, a pesquisa aplicada está relacionada com a análise do ambiente real de uma empresa pública federal, com relação ao processo de estimativa de esforço em projetos de Big Data, com o objetivo de propor alternativas ao modelo de medição existente na empresa.

O processo atual da organização que será analisada neste trabalho segue o que foi publicado no acórdão do TCU (BRASIL, 2015), no qual é mostrado que os modelos de medição utilizando APF não conseguem suprir as necessidades das empresas. No caso da organização supracitada, a utilização apenas do Ponto de Função faz com que a estimativa de esforço das demandas seja subestimada, isto é, não corresponde ao esforço real da equipe de desenvolvimento.

Com relação à abordagem adotada na pesquisa, ela foi qualitativa e quantitativa. A pesquisa qualitativa pode ser construída a partir de várias teorias ou abordagens metodológicas, com a possibilidade de utilizar vários métodos, como a observação, entrevistas, questionários e análise de documentos (GIL, 2008). A quantitativa, por sua vez, resume-se na utilização de métodos estatísticos para coletar e analisar os dados que foram obtidos para solucionar o problema levantado (FLICK, 2013). No contexto da pesquisa, a análise quantitativa está



relacionada com a coleta dos dados sobre estimativa de esforço em projetos de Big Data dentro da organização federal. Já a pesquisa qualitativa foi realizada por meio de questionários e entrevistas com funcionários da empresa, além da análise de demandas existentes no contexto citado (MCGRATH, PALMGREN e LILJEDAHL, 2019).

## **1.5. Estrutura do Documento**

Este documento compreende outros capítulos subsequentes a este, que estão organizados da seguinte forma:

- Os conceitos de medição de software, APF, SNAP, UST e Big Data são apresentados em detalhes no Capítulo 2, incluindo uma descrição de aspectos da Administração Pública Federal; e de trabalhos relacionados;
- No Capítulo 3, são apresentados os trabalhos relacionados;
- O protocolo e resultados do estudo de caso realizado são mostrados no Capítulo 4;
- No Capítulo 5, é descrito o modelo proposto de medir esforço de projetos de Big Data dentro do contexto da Administração Pública Federal, como resultado do estudo de caso realizado;
- Finalmente, as considerações finais da pesquisa, além de propostas para trabalhos futuros, são elencadas no Capítulo 6.

## **2. FUNDAMENTAÇÃO TEÓRICA**

Neste capítulo, são apresentados conceitos relevantes ao contexto da pesquisa e trabalhos que compõem parte do seu estado da arte. Ao final do capítulo, é feita uma análise comparativa, de forma resumida, das pesquisas relacionadas.

### **2.1. Medição de Software**

De acordo com Sommerville (2015), o processo de medição de software tem o objetivo de quantificar alguns atributos de um produto ou de um processo de software. Comparando essas informações, é possível tirar conclusões sobre a qualidade do software medido. Além disso, a medição permite mensurar se mudanças organizacionais (a adoção de novas ferramentas ou metodologias, por exemplo) estão sendo positivas ou não para os processos de desenvolvimento utilizados. Nesse caso, são feitas medições antes e depois da mudança a fim de verificar se ela foi positiva ou não para a organização.

A estimativa de esforço de desenvolvimento de um software, por sua vez, vem sendo utilizada desde os anos 50 nos projetos existentes (VERA; OCHOA; PEROVICH, 2018) e, com o aumento da complexidade das máquinas nos anos 80, outras técnicas foram surgindo para aperfeiçoar essa estimativa, como a SEER (JENSEN, 1983) e Price-S (PARK, 1988), que projetam os recursos dos requisitos em função do seu tempo de desenvolvimento. Atualmente, ela é um requisito importante de todas as empresas de desenvolvimento de software (KAUSHIK; TAYAL; YADAV, 2019). Em outras palavras, estimar o esforço é um meio de avaliar o total de custo necessário para executar um projeto ou produto.

Essas estimativas de esforço são comumente baseadas na predição do tamanho do sistema que será desenvolvido (ABRAHAO; INSFRAN, 2008). Para projetos de melhoria, o seu tamanho pode ser uma referência para o cálculo do esforço a ser feito pela equipe.

Portanto, estimar o esforço de um software é crucial no seu desenvolvimento, seja para projetos de sistemas novos ou melhorias de um já existente, pois é a partir dessa estimativa que é possível se ter uma ideia do tamanho e do tempo que precisará ser gasto no sistema que será desenvolvido ou evoluído e, assim, com esta projeção, o projeto será contratualmente regido (WANDERLEY, 2015). Dessa forma, uma alta estimativa de esforço pode indicar que o tamanho do software é grande. E uma baixa estimativa pode ser um indicativo que o projeto é pequeno.

Durante o processo de estimativa do esforço, é comum que se fale numa “curva de incerteza” existente no processo (LEDERER; PRASAD, 1989). De modo geral, nos estágios iniciais de desenvolvimento, a quantidade de incerteza na estimativa é alta e, com isso, é produzida uma medição com baixa qualidade. À medida que o projeto evolui, o valor estimado melhora, portanto sua incerteza diminui.

Existem dois tipos de métodos de estimativa de esforço: os métodos heurísticos, que se concentram em uma experiência humana (um estimador), e os métodos algorítmicos, que consistem em fórmulas matemáticas para calcular o total de esforço necessário. Além desses dois tipos, existem outros métodos propostos na literatura, tais como o método de regressão, analogia e *soft-computing* (AMIT; VERMA, 2013).

O método heurístico de estimativa se concentra na utilização de conhecimentos e experiência de especialistas. O chamado estimador e o especialista de determinado domínio estimam o esforço total necessário do desenvolvimento do software. Alguns métodos heurísticos utilizados na literatura são a técnica Delphi (baseada em consenso), o *Planning Poker* (estimativa de tarefas feita com base na sequência de Fibonacci) e o *Work Breakdown Structure - WBS* (dividir a carga de trabalho em tarefas simples para que facilite a estimativa) (AMIT; VERMA, 2013).

O método algorítmico consiste em formulações matemáticas para realizar a estimativa. É mais utilizado na literatura devido a sua consistência e por não depender do fator humano, que tende a ser subjetivo. Alguns dos métodos algorítmicos mais citados na literatura são a Análise de Pontos de Função (APF), o COCOMO II e a Estimativa Baseada em Casos de Uso (AMIT; VERMA, 2013) (VERA; OCHOA; PEROVICH, 2018) (HAI; NHUNG; THAI, 2019).

Os modelos de estimativa apresentados e descritos nas subseções a seguir estão relacionados com o contexto da pesquisa descrita neste documento, i.e., são métodos utilizados em projetos dentro da Administração Pública Federal.

### **2.1.1. Análise de Pontos de Função (APF)**

Análise de Pontos de Função (APF) (ALBRECHT, 1979) é um modelo de estimar o custo e esforço do desenvolvimento de software que quantifica as funcionalidades existentes no produto com base no ponto de vista do usuário, analisando seu projeto lógico e especificação funcional, independente da tecnologia envolvida.

Atualmente, a APF é mantida como padrão internacional através da norma ISO/IEC 20.9261 (ISO/IEC 20926, 2002) e pelo *International Function Point User Group* (IFPUG)<sup>1</sup>, responsável por definir suas regras de contagem, no documento Manual de Práticas de Contagem (CPM) (IFPUG, 2010). Além disso, existem roteiros de métricas elaborados por sistemas de cada país que visam complementar a estimativa de projetos que não estão contemplados no manual do IFPUG (SISP, 2015) (SISP, 2018). No Brasil, o órgão responsável por esses manuais complementares é o Sistema de Administração dos Recursos de Tecnologia da Informação (SISP), que foi criado em 2011 com o objetivo de organizar a operação, controle,

---

<sup>1</sup> <https://www.ifpug.org/>

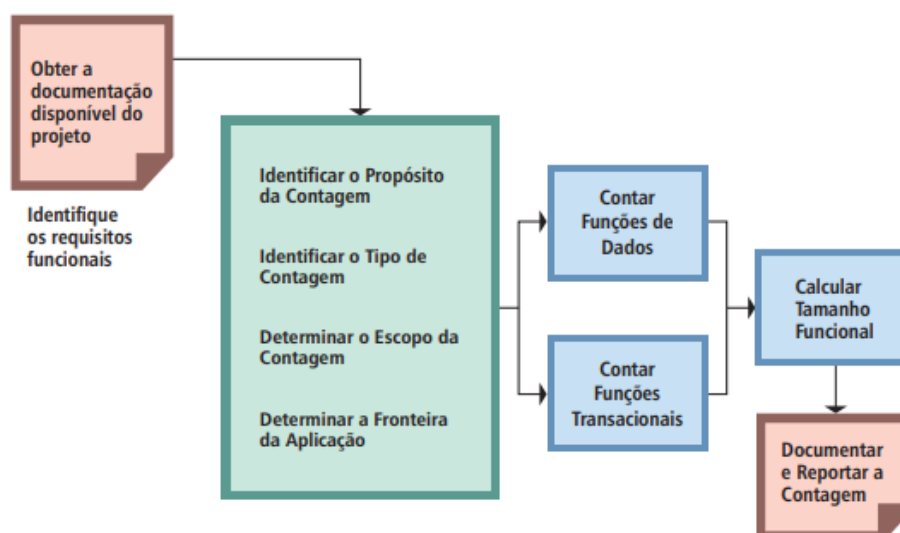
supervisão e coordenação dos recursos de tecnologia da informação da administração direta, autárquica e fundacional do Poder Executivo Federal (SISP, 2018).

Os manuais do SISP são compostos por itens de contagem provenientes de demandas de serviço que não estão presentes no CPM, e que fazem parte do dia a dia das organizações, como manutenções corretivas, adaptativas e cosméticas em requisitos não funcionais dos produtos (SISP, 2015).

No contexto específico de projetos de *Data Warehouse* (DW), um item importante que exerce grande influência no desenvolvimento das funcionalidades das soluções, e que é mensurado utilizando APF, de acordo com o manual do SISP, é a modelagem multidimensional dos dados, considerando as tabelas Fato e Dimensão definidas neste modelo, para que sejam atendidas as necessidades do usuário com relação aos relatórios e consultas e relatórios, à análise de grandes volumes de dados e à obtenção de informações estratégicas para a tomada de decisão (SISP, 2015).

O modelo de contagem por APF é mostrado na Figura 2.

**Figura 2 – Modelo de Análise de Pontos de Função.**



**Fonte: SISP (2018).**

Como pode ser visto na Figura 1, o processo de estimativa utilizando APF se inicia com a análise da documentação disponível do projeto a ser estimado, visando a identificação dos seus requisitos funcionais.

A próxima fase é a identificação do propósito da contagem, que fornece uma resposta para uma questão de negócio a ser resolvida, por exemplo: necessidade de dimensionar um novo projeto da empresa x.

Com base no propósito da contagem, podem ser definidos o tipo e o escopo dela. O tipo pode ser um projeto de melhoria, de desenvolvimento ou de uma aplicação instalada. O escopo compreende o conjunto de funcionalidades que serão contadas, por exemplo: apenas as funcionalidades de front-end.

O próximo passo é definir a fronteira da aplicação da contagem, que é feita de acordo com a visão do usuário e estabelece os limites das aplicações envolvidas no projeto que será contado (SISP, 2018).

Após a designação da fronteira da aplicação, ocorre a medição propriamente dita das funções de dados e transacionais. As funções de dados são divididas entre Arquivos Lógicos Internos (ALI) e Arquivos de Interface Externa (AIE), cada um com seus Registros Lógicos (RL) e Itens de Dados (ID). As funções de transação, por sua vez, constituem de processos elementares distintos, podendo ser classificadas como Entradas Externas (EE), Saídas Externas (SE) e Consultas Externas (CE), cada uma com seus Arquivos Lógicos Referenciados (ALR) e Itens de Dados (ID).

Em seguida a essas definições, pode-se calcular o tamanho funcional do projeto estimado e, com isso, documentar a contagem realizada.

Conforme dito anteriormente, APF pode ser utilizado tanto em projetos de desenvolvimento quanto de melhoria. No caso de projetos de melhoria, como uma demanda de manutenção evolutiva de um produto, por exemplo, é necessário que seja avaliado, também, o impacto que essa melhoria trará à solução existente, visto que ela pode se tratar de inclusões, alterações ou exclusões nas funcionalidades.

A medição em PF é feita, então, utilizando um fator de impacto em todas as funções de dados e de transação afetadas pela mudança.

### **2.1.2. SNAP**

O modelo SNAP (Processo de Avaliação de Software Não-Funcional) também foi definido pelo IFPUG e compreende a medição de requisitos não-funcionais (RNF) de um sistema, complementando o APF, que mede os requisitos funcionais das aplicações, com base na visão do usuário (IEEE, 2019).

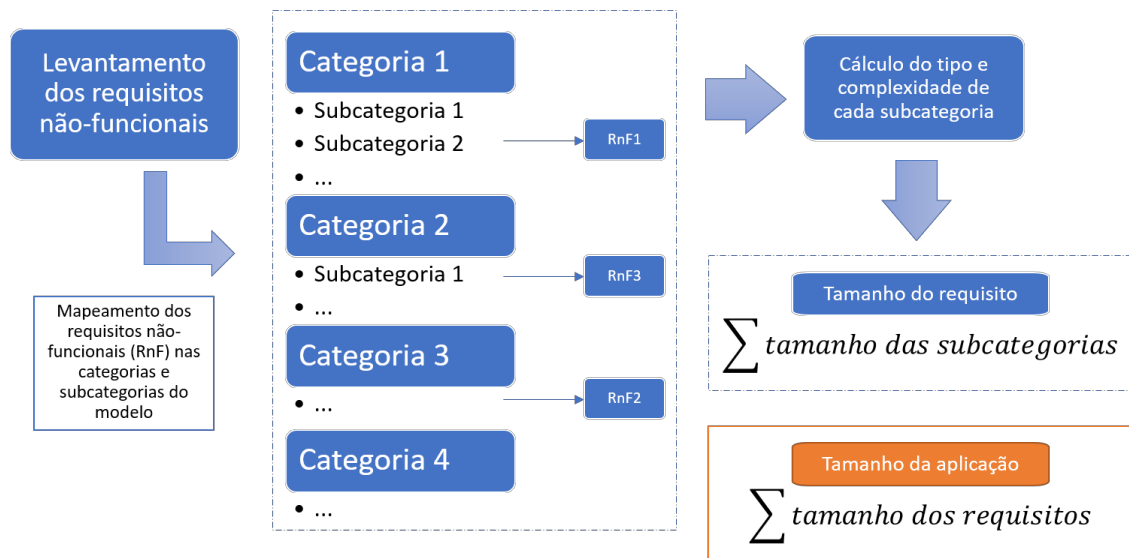
O SNAP é composto por quatro categorias e catorze subcategorias de itens de medição. Cada categoria tem a finalidade de classificar o RNF que será mapeado e, posteriormente, medido. Dentro de cada categoria, existem várias subcategorias, que possuem características comuns e são utilizadas para satisfazer o RNF.

As categorias e subcategorias do modelo SNAP são descritas a seguir:

1. Operações de Dados - descrevem como os dados são processados para satisfazer os RNFs da aplicação. Suas subcategorias são:

- a. Validações na Entrada de Dados
  - b. Operações Lógicas e Matemáticas
  - c. Formatação de Dados
  - d. Movimentações de Dados Internos
  - e. Entregar Valor Agregado aos Usuários por Configuração de Dados
2. Projeto de Interface - está relacionada com a experiência do usuário final, ou seja, com os métodos e artefatos que envolvem a interação do usuário com a aplicação. Suas subcategorias são:
    - a. Interfaces do Usuário
    - b. Métodos de Ajuda
    - c. Múltiplos Métodos de Entrada
    - d. Múltiplos Métodos de Saída
  3. Ambiente Técnico - refere-se aos aspectos do ambiente ao qual a aplicação está inserida, i.e., a tecnologia, alterações nos dados internos e outras configurações que não resultam em alterações nos requisitos funcionais do sistema. Suas subcategorias são:
    - a. Múltiplas Plataformas
    - b. Tecnologia de Banco de Dados
    - c. Processos Batch
  4. Arquitetura - além de avaliar a complexidade do desenvolvimento em módulos ou baseado em componentes, a categoria também está relacionada com técnicas de codificação usadas para construir ou melhorar a aplicação que está sendo medida. Suas subcategorias são:
    - a. Software Baseado em Componentes
    - b. Múltiplas Interfaces de Entrada / Saída

As etapas do modelo de contagem SNAP são mostradas, em linhas gerais, na Figura 3.

**Figura 3 – Modelo de contagem SNAP.**

Fonte: Desenvolvido pelo autor. Adaptação de IFPUG (2020).

O processo da contagem SNAP segue a mesma linha da contagem utilizando APF, sendo que o SNAP consiste em elencar os requisitos não-funcionais da aplicação e mapeá-los para as subcategorias existentes no framework. Cada subcategoria é medida de acordo com o seu tipo e complexidade, e o tamanho de cada requisito é a soma dos tamanhos de suas subcategorias. Esses tamanhos são, então, totalizados para fornecer a medida do tamanho não-funcional da aplicação (IFPUG, 2020).

### 2.1.3. Unidades de Serviço Técnico (UST)

O método de estimativa utilizando Unidades de Serviço Técnico (UST) começou a ser utilizado recentemente pelas organizações federais com o intuito de contratar os serviços de TI e remunerá-los por resultado (FUNDECC, 2017). Ele surgiu como uma forma complementar de mensurar esforço e custo de demandas de serviços que envolvem esforço humano, com variada complexidade e que não é possível ser medido com precisão utilizando outro modelo de estimativa, como o APF ou SNAP.

O cálculo de USTs segue o raciocínio definido pela CGU em 2011, em que uma UST equivale a uma hora de trabalho. Essa relação pode ser alterada de acordo com a complexidade da tarefa a ser realizada para desenvolver o serviço. Em outras palavras, uma hora de trabalho pode valer, de acordo com a complexidade da tarefa:

- 1 UST, se a complexidade for baixa;
- 1,5 USTs, se a complexidade for intermediária;
- 3,5 USTs, se a complexidade for mediana; e
- 6 USTs, se a complexidade for alta.

De forma genérica, o processo para a organização definir o seu catálogo de serviços, com suas respectivas USTs, contempla as seguintes tarefas (FINEP, 2017):

1. Identificar o módulo e o(s) cenário(s) envolvido(s) para desenvolver o serviço;
2. Atribuir uma complexidade (baixa, média ou alta) a cada cenário;
3. Identificar outros fatores que influenciam no desenvolvimento do serviço (infraestrutura, interface, regras de negócio, entre outros fatores) e suas complexidades;
4. Verificar a dependência de cada fator citado no item 3 com o cenário envolvido.

O número de USTs de cada cenário envolvido no desenvolvimento do serviço é o produto dos fatores relativos à complexidade do cenário com os demais que são influentes.

O modelo UST vem sendo utilizado em processos contratuais de soluções de TI da Administração Pública Federal. No entanto, até o final de 2019, a métrica era usada sem padronização, ou seja, cada organização mensurava seus produtos e serviços de acordo com o seu processo interno de medição. Após uma auditoria do TCU (BRASIL, 2020), foi constatado uma deficiência na estimativa de preços da UST, além do uso de parâmetros injustificados e impossibilidade de preços condizentes com o mercado. Com isso, foi elaborada, pela Secretaria de Governo Digital do Ministério da Economia, o conjunto de orientações para padronizar as novas contratações e renovações de contratos das empresas que utilizam essa métrica (GOVERNO DIGITAL, 2020).

## **2.2. Aspectos da Administração Pública Federal**

No contexto da Administração Pública Federal, o Ponto de Função é o modelo de medição utilizado como referência para mensurar o esforço e custo dos contratos de serviços de desenvolvimento e manutenção de soluções firmados entre empresas e autarquias públicas com organizações que prestam esse tipo de serviço. A recomendação pelo uso dessa métrica nesse tipo de contrato veio a partir da portaria SLTI/MP N° 31 (MPOG, 2010) e da Instrução Normativa SLTI/MP N° 4 (MPOG, 2014).

A portaria SLTI/MP N° 31, de 29 de novembro de 2010, reconhece o PF como uma boa prática utilizada na contratação dos serviços de desenvolvimento e manutenção de software, devido a sua objetividade na medição. Além disso, é recomendado o uso dos roteiros de contagem em PF aos integrantes do SISP, que foi criado no ano seguinte.

A Instrução Normativa SLTI/MP N° 4, de 11 de setembro de 2014, dispõe sobre a forma de contratação dos serviços de TI pelos órgãos que integram o SISP, recomendando novamente o uso de Pontos de Função como boa prática de medição e remuneração das soluções de software nas empresas e autarquias públicas, restringindo o uso da métrica de esforço Homem-Hora, utilizada até então pela União.



Em consonância com a portaria e instrução normativa citadas anteriormente, o Tribunal de Contas da União (TCU) publicou vários acórdãos recomendando, também, o uso dessa métrica nos contratos de serviços de TI (SISP, 2018). Os acórdãos também reforçam a necessidade de não vincular a métrica de tamanho funcional (PF) com a de esforço baseada em produtividade (Homem-Hora).

Com base no cenário posto, vários órgãos relataram dificuldades na adoção e uso exclusivamente do PF nas suas soluções de TI. Então, o TCU emitiu o Acórdão 2.362/2015 (BRASIL, 2015), em que houve a análise dos contratos de prestação de serviços dessas empresas e foi relatado que a adoção do PF não tem sido bem aproveitada nos órgãos auditados. Foi visto que, em vários casos, a APF não mediu o esforço e custo condizentes com o trabalho executado, gerando desequilíbrio econômico e financeiro para a prestadora dos serviços (ALMEIDA, MONTEIRO e FURTADO, 2018).

Em junho de 2020, a Secretaria de Governo Digital do Ministério da Economia, órgão responsável pelo SISP, emitiu uma série de orientações para novas contratações e renovações de contratos feitos utilizando a métrica Unidade de Serviço Técnico (UST), que já vinha sendo usada por alguns órgãos como alternativa ou em complemento à métrica Ponto de Função. Essas orientações têm o objetivo de auxiliar os órgãos e empresas contratantes na tomada de decisão (GOVERNO DIGITAL, 2020).

## **2.3. Big Data**

Nesta seção, serão definidos o conceito e as características de um Big Data, além de elucidar as etapas de um projeto de Big Data.

### **2.3.1. Características**

Em 2001, o analista Doug Laney evidenciou, em seu trabalho, que as condições e meios de negócios do início do século estavam levando os princípios tradicionais de gerenciamento de dados ao limite, dando origem a novos conceitos e abordagens. Foi mostrado, então, que a gerência dos dados deveria se preocupar com o seu volume, variedade e velocidade desses dados (LANEY, 2001).

A partir do problema levantado na pesquisa de Laney, foi elucidado o primeiro conceito de Big Data, baseado nos seus 3Vs principais. São eles:

1. Volume: está relacionado à grande quantidade de dados oriundos de variadas fontes diferentes;
2. Variedade: os dados são concebidos em diferentes tipos de formatos (de dados estruturados outros sem estrutura fixa, como documentos de texto, e-mails, áudios, vídeos, transações financeiras, entre outros tipos);

3. Velocidade: eles devem ser analisados e processados numa velocidade sem precedentes e devem ser abordados em tempo hábil.

Em outras palavras, um Big Data pode ser definido como um grande conjunto de dados, de diferentes formatos e dimensões, que tem o propósito de gerar respostas rápidas, atendendo aos requisitos do usuário.

Outras definições existentes na literatura (REIS e SÁ, 2020) (ALAOUI, GAHI e MESSOUSSI, 2019) (SAMOSIR et al., 2017) também relacionam o termo Big Data com a dificuldade de analisar e processar grandes conjuntos de dados, porém trazem novas características importantes e que levam as organizações a melhores decisões e estratégias de negócio.

Essas características utilizadas na literatura compõem outros Vs do Big Data. Alguns deles são mostrados a seguir.

1. Veracidade: relacionado com a confiabilidade dos dados. Eles precisam ser confiáveis para auxiliar na tomada de decisão;
2. Valor: diz sobre as informações e percepções que os dados fornecem ao final do processo do Big Data;
3. Variabilidade: diferente da característica “variedade”, aqui trata, geralmente, os dados com significado em constante evolução;
4. Visualização: também parte do final do processo do Big Data, está relacionado com a apresentação dos dados analisados para o usuário de forma legível e de fácil entendimento.

Essa grande quantidade de dados vem sendo trabalhada e processada dentro das empresas e organizações, que precisam se adaptar e desenvolver novas soluções para acomodar o seu processamento. Os projetos de Big Data, ou *Big Data Analytics*, conseguem gerar valor para o usuário, combinando as características de Big Data com técnicas de análises de dados, e necessitam de suporte para paralelizar o processamento e armazenar a enorme concentração de dados (SAMOSIR et al., 2017).

Com relação às tecnologias utilizadas num projeto de Big Data, as principais estão relacionadas com a infraestrutura, armazenamento e processamento dos dados, em bancos de dados não relacionais (NoSQL), além das que estão envolvidas com análise de dados, cujo principal framework existente é o Apache Hadoop (Pereira, 2016).

O Apache Hadoop é um framework open source, cuja principal função é o armazenamento e processamento em larga escala de Big Data. O seu sistema é compartilhado e distribuído, ou seja, feito em clusters de computadores. É composto por dois componentes principais (Taurion, 2013). São eles:

1. *Hadoop Distributed File System (HDFS)*: sistema de armazenamento distribuído. Os dados presentes no HDFS são divididos em pequenos blocos e são distribuídos por diversos servidores. Desta forma, o processamento subsequente é muito acelerado pois, em vez de uma pesquisa sequencial, os dados são pesquisados de forma simultânea, em paralelo.
2. *Hadoop MapReduce*: sistema de processamento e geração de grandes conjuntos de dados. O termo MapReduce representa as tarefas de (i) mapear os dados (Map), ou seja, acessar um conjunto de dados e convertê-los em elementos individuais, que são quebrados em pares chave/valor; e (ii) a tarefa Reduce, que combina os pares gerados na tarefa anterior e gera um novo e menor conjunto de dados, obtendo o resultado.

Outras tecnologias que são utilizadas num projeto de Big Data estão relacionadas com as etapas de visualização dos dados (Visual Analytics); linguagens de programação adequadas ao Big Data, como Python e Pig Hive; entre outras técnicas de aprendizado de máquina e processamento de linguagem natural (REIS e SÁ, 2020).

### 2.3.2. Projeto de Big Data

Os projetos existentes de Big Data estão relacionados com o conceito de *Big Data Analytics*, em que ocorre a utilização e manipulação desse grande conjunto de dados combinado com técnicas de análises de dados, com o objetivo de gerar valor (ELGENDY e ELRAGAL, 2014).

O processo que vai permitir a transformação desses dados em *insights* relevantes para as tomadas de decisão, dentro de um projeto de Big Data, inicia-se com a definição do *business case* do projeto. Nessa etapa, são elencados os desafios e objetivos do negócio a ser desenvolvido, além de questões como área de atuação e principais interessados na solução (TAURION, 2013).

Após a definição dos objetivos de negócio, as demais etapas que envolvem um projeto de Big Data são definidas a seguir:

1. **Coleta dos dados** (de diferentes fontes e formatos) para armazenamento e com base no objetivo específico do projeto;
2. **Pré-processamento dos dados**: uso de métodos estatísticos que avaliam desvios e, com base em alguns critérios, definem a relevância dos dados para a análise futura que será realizada;
3. **Mineração dos dados**: processamento realizado para identificar padrões e com a aplicação de técnicas de, por exemplo, *machine learning* e Inteligência Artificial;
4. **Análise de conteúdo** com os padrões identificados na fase anterior, com o objetivo de entender o cenário ao qual o projeto está inserido e, a partir disso, descrever, prever, prescrever ou diagnosticar informações relevantes;

5. **Visualização dos dados:** extrair os *insights* obtidos na fase 4 com o intuito de apresentar as informações ao usuário utilizando ferramentas visuais de fácil interpretação.

Além das etapas citadas anteriormente, é importante que os dados do projeto de Big Data estejam, ao final do processo, aptos a serem integrados de forma sistemática e com inteligência a outros projetos, uma vez que esse tipo de projeto deve ser adotado como uma estratégia de negócio, portanto é importante que os dados estejam em convergência (IBM, 2016).

### **3. TRABALHOS RELACIONADOS**

Os trabalhos relacionados a esta pesquisa são modelos e/ou abordagens escolhidos dentro do contexto da estimativa de esforço, e que foram aplicados em projetos de Big Data, da indústria, de *Business Intelligence* (BI) ou dentro do contexto da Administração Pública Federal.

A coleta dos trabalhos se deu a partir da busca realizada em bibliotecas digitais de relevância científica, considerando artigos completos e publicados nos últimos dez anos, em português e inglês, e que estivessem dentro do contexto da pesquisa.

#### **3.1. Estimativa de esforço em lógica fuzzy e métricas de desempenho**

O trabalho elaborado por Vijay (2018) propõe uma técnica híbrida de estimativa de esforço baseada em lógica fuzzy e métricas de desempenho (p.e. velocidade e latência) aplicadas dentro da APF, com o objetivo de tornar a sua estimativa mais precisa. A lógica fuzzy está inserida nesta pesquisa para diminuir a incerteza obtida em APF. As métricas de desempenho são utilizadas para calcular a precisão da estimativa. Os experimentos realizados para aferir a abordagem proposta utiliza dados de estimativas de esforço de projetos de BI. Os resultados indicaram que, para esse contexto, o uso de métricas de desempenho e lógica fuzzy melhorou a precisão da estimativa de esforço dos projetos.

#### **3.2. Análise de métricas na Administração Pública Federal**

O estudo desenvolvido por Almeida, Monteiro e Furtado (2018) teve o objetivo de identificar e analisar as métricas utilizadas em contratos de software realizados em empresas no contexto da Administração Pública Federal. Além disso, os autores buscaram uma maneira de mensurar serviços de software que não utilizassem apenas a Análise de Pontos de Função, amplamente utilizada nesse tipo de contratação. A pesquisa identificou que outras métricas de esforço estavam sendo utilizadas em algumas organizações federais: Unidade de Serviço Técnico (UST), que calcula o esforço com base num catálogo de serviços definido pela empresa, e Unidade de Medida para Serviços Técnicos (USTIBB), cujo esforço estimado se baseia na complexidade da tarefa, tempo (em horas), a produtividade e a qualificação da mão-de-obra existente.

#### **3.3. Métricas de medição em projetos de Big Data**

Samosir et al. (2017) propuseram métricas de medição que são aplicadas em projetos de Big Data Analytics, utilizando diagramas de classes UML para modelar e identificar métricas de medição estáticas (obtidas na fase de projeto), como o número de métodos numa classe, nível

de coesão do sistema, número de relações entre as classes e número de filhos; e dinâmicas (obtidas após a execução do projeto), como o número de objetos criados durante a execução e número de acoplamentos em tempo de execução. Elas serão usadas, de acordo com os autores, durante o desenvolvimento de um projeto de Big Data Analytics.

### **3.4. Modelo híbrido de estimativa de esforço com Pontos de Casos de Uso**

Azzeh e Nassif (2016) projetaram um modelo híbrido que utiliza máquinas de vetores de suporte e redes neurais para auxiliar os modelos que utilizam apenas Pontos de Casos de Uso e, às vezes, a produtividade para estimar o esforço de desenvolvimento do software. O modelo proposto apresentou bons resultados nos conjuntos de dados aplicados. Por isso, foi concluído pelos autores que o uso de modelos e técnicas de IA podem auxiliar os modelos de estimativas existentes.

### **3.5. COCOMO II, APF e Casos de Uso em projetos de Big Data**

Nathanael, Hendradjaya e Sunindyo (2015) elaboraram um estudo de estimativa de esforço de software de projetos de Big Data baseado em três modelos: COCOMO II, Análise de Pontos de Função e Estimativa Baseada em Casos de Uso. Os autores partiram do problema em se buscar um método e modelo que melhor se adequasse às necessidades de estimativa de esforço vindas de projetos de Big Data. Assim, criaram um modelo utilizando as três métricas citadas anteriormente e pode-se concluir que essa junção não foi a melhor combinação para realizar uma estimativa de esforço para o desenvolvimento de software de Big Data. No entanto, há muito espaço para aprimoramento para escolher as melhores opções, de uma maneira que possa acomodar as características especiais do software Big Data e o paradigma instável de um processo de desenvolvimento de software.

### **3.6. Métricas de software em sistemas de BI**

O trabalho de Endo (2014) tem o objetivo de identificar métricas capazes de estimar o esforço necessário para se desenvolver sistemas de BI. Foi realizado, no trabalho dele, 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.

### 3.7. Análise e sumarização dos trabalhos

Os trabalhos mostrados nesse capítulo contribuíram dentro de cada contexto envolvido e têm seus próprios méritos e deméritos. No entanto, não foi evidenciado, nesta pesquisa, uma técnica padrão que seja comumente aceita para estimativa de esforço, pois a depender do contexto utilizado na pesquisa, um método pode ser mais bem aplicado do que outro. Além disso, não foram encontrados, até o momento, estudos que mostrassem forma de estimar projetos de Big Data dentro do contexto apresentado para a administração pública federal.

Na Tabela 2 a seguir, um sumário dos trabalhos encontrados e que são relevantes para este documento é apresentado, com os seus autores, área de aplicação e técnica de medição utilizada nas pesquisas, além das etapas do projeto abordadas em cada contexto, os resultados e observações importantes.

**Tabela 2 – Trabalhos relacionados.**

| <b>Pesquisa</b>                                | <b>Área de aplicação</b>      | <b>Técnica de medição</b>   | <b>Etapas do projeto</b>                                     | <b>Resultados</b>   | <b>Observações</b>   |
|--|-------------------------------|---|--|---|--|
| <b>Vijay (2018)</b>                            | <i>Business Intelligence</i>  | Métricas de desempenho (velocidade, latência, entre outras), APF e lógica fuzzy           | Planejamento (estimativa de esforço)                         | O modelo proposto apresentou melhor desempenho que o modelo APF.  | Propõem um modelo híbrido de estimativa.   |
| <b>Almeida, Monteiro e Furtado (2018)</b>      | Administração Pública Federal | APF, Unidade de Serviço Técnico (UST) e Unidade de Medida para Serviços Técnicos (USTIBB) | Planejamento (estimativa de esforço e custo)                 | UST e USTIBB estão sendo utilizadas nas empresas federais para complementar a APF.  | A pesquisa trata-se de um survey com empresas públicas federais.                                       |
| <b>Samosir et al. (2017)</b>                   | Big Data Analytics            | Métricas estáticas e dinâmicas a partir de um diagrama de classes UML                     | Todas  | As métricas estáticas e dinâmicas são apresentadas e relacionadas com partes integrantes de um diagrama de classes UML.                       | A estimativa é feita apenas em cima do diagrama de classes UML.  |
| <b>Azeg e Nassif (2016)</b>                    | Indústria                     | Pontos de Casos de Uso  | Planejamento (estimativa de esforço)                         | O modelo proposto apresentou bons resultados nos conjuntos de dados aplicados.  | Propõem um modelo híbrido de estimativa.   |
| <b>Nathanael, Hendradjaya e Sunindy (2015)</b> | Big Data                      | COCOMO, APF e Estimativa baseada em Casos de Uso  | Definição do Business Case (estimativa de esforço)           | O modelo híbrido apresentado não obteve bons resultados na estimativa de esforço em projetos de Big Data.                                     | A junção dos modelos não serviu conforme esperado nesse contexto, mas pode ser útil em outro ambiente. |
| <b>Endo (2014)</b>                             | <i>Business Intelligence</i>  | APF, Pontos de Casos de Uso e COSMIC  | Planejamento (estimativa de tamanho, esforço e complexidade) | O autor conseguiu mostrar que quanto maior a complexidade do processo de negócio, maior seria o esforço para desenvolver uma aplicação de BI. | Foram realizados dois estudos na pesquisa, pois a 1ª abordagem não obteve bons resultados.             |

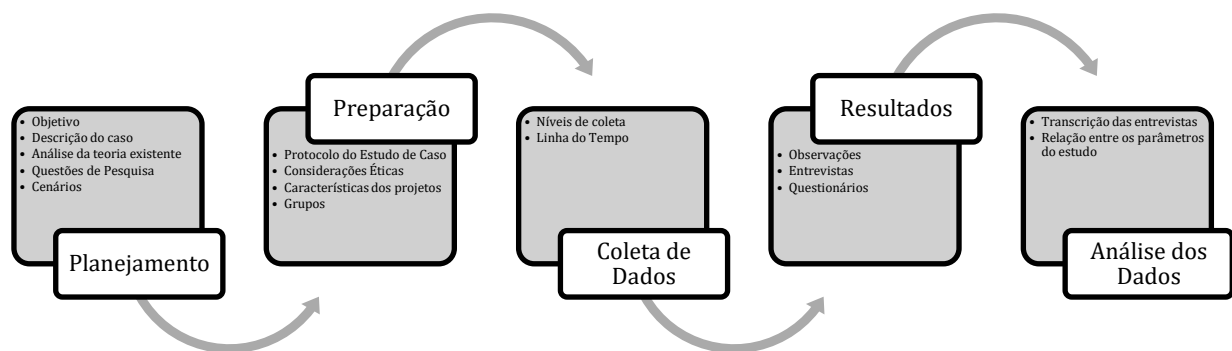
**Fonte: Desenvolvido pelo autor.**

## 4. ESTUDO DE CASO

Neste capítulo, serão descritas as informações sobre o estudo de caso realizado neste trabalho, além dos dados do protocolo e resultados da sua execução.

Os cinco passos que foram realizados no estudo de caso deste trabalho seguiram a abordagem proposta por Runeson e Höst (2009). As etapas são: planejamento, preparação dos casos, coleta dos dados, análise dos dados coletados e obtenção dos resultados do estudo. Elas são mostradas na Figura 4 e serão descritas nas seções a seguir.

**Figura 4 – Passos do estudo de caso realizado.**



**Fonte: Desenvolvido pelo autor.**

### 4.1. Planejamento

O planejamento do estudo de caso aborda definições sobre o objetivo (o que desejo alcançar no estudo?), descrição do caso (o que será estudado?), análise da teoria existente, questões de pesquisa (o que se deseja saber?), seleção dos casos (quais os cenários que serão estudados?), além dos métodos de coleta e análise.

O objetivo do estudo de caso realizado foi analisar o processo de estimativa de esforço em projetos de Big Data numa empresa que pertence à Administração Pública Federal. Essa organização, que foi o objeto de estudo, trata-se de uma empresa pública de Tecnologia da Informação, que desenvolve soluções para o governo e cidadãos brasileiros. Com relação a forma de estimar esforço dos seus contratos de software, ela segue a recomendação proposta pela portaria SLTI/MP N° 31 (MPOG, 2010) e pela Instrução Normativa SLTI/MP N° 4 (MPOG, 2014), isto é, utiliza como modelo de medição a APF.

É possível encontrar a descrição do caso estudado na seção 4.1 deste documento, assim como a análise da teoria existente, que está elucidada no Capítulo 2.



As questões de pesquisa deste estudo estão relacionadas com o objetivo geral da pesquisa, que se propõe a responder o problema levantado neste documento. Elas são mostradas a seguir:

*QP1. O modelo atual de medição em Pontos de Função atende aos projetos de Big Data nas empresas públicas federais?*

*QP2. O uso de modelos de medição com base em unidades de serviços técnicos (UST) e em requisitos não funcionais (SNAP) complementa as medições dos projetos de Big Data nas organizações que fazem parte da Administração Pública Federal?*

*QP3. As características relacionadas ao volume, variedade e velocidade dos dados de um projeto de Big Data devem ser levadas em consideração na sua estimativa de esforço?*

Para responder às questões de pesquisa elencadas no estudo de caso, foram selecionados os seguintes cenários:

*Cenário 1: Utilizar APF em conjunto com o modelo SNAP, que mede requisitos não-funcionais de um sistema, na medição de esforço dos projetos de Big Data;*

*Cenário 2: Utilizar APF em conjunto com o modelo UST, que utiliza unidades de serviços técnicos, na medição de esforço dos projetos de Big Data;*

*Cenário 3: Utilizar APF em conjunto com SNAP e UST na medição de esforço dos projetos de Big Data.*

O planejamento dos cenários levou em consideração o cruzamento de diferentes modelos de estimativa de esforço que pudessem ser utilizados como complemento a APF, uma vez que esta consegue quantificar os requisitos funcionais das soluções e os demais métodos se baseiam na medição de requisitos não-funcionais (SNAP) e em unidades de serviços técnicos (UST).

## **4.2. Protocolo e Preparação dos Casos**

A etapa de preparação dos casos está relacionada com o protocolo de estudo de caso realizado, que contém informações sobre os procedimentos utilizados e instrumentos de pesquisa e análise do estudo, e pode ser encontrado no Apêndice deste documento.

Além disso, como a metodologia desta pesquisa consiste em investigar o fenômeno da estimativa de esforço em projetos de Big Data dentro de uma empresa da Administração Pública Federal, são necessárias considerações éticas no que diz respeito ao tratamento de informações confidenciais da organização que será objeto de estudo deste trabalho.

#### 4.2.1. Considerações Éticas

As considerações éticas da pesquisa estão relacionadas ao uso e tratamento de dados sensíveis, além da regulamentação de questões de confidencialidade e publicação entre as partes envolvidas nesse processo, isto é, o pesquisador e a empresa estudada (RUNESON e HÖST, 2009).

No que diz respeito ao tratamento dos dados, os fatores éticos acertados estão conforme a Lei nº 13.709, de 14 de agosto de 2018 (BRASIL, 2018), conhecida como Lei Geral de Proteção de Dados (LGPD). A lei garante a realização do estudo caso seja garantida a anonimização dos dados, com sua utilização dentro da empresa e seguindo estritamente a finalidade proposta na pesquisa. Além disso, o pesquisador será o responsável pela segurança da informação, não permitindo, em nenhuma circunstância, a transferência dos dados a terceiros.

Do ponto de vista dos participantes do estudo de caso, eles receberam, no início do processo, um acordo de consentimento, com dados e procedimentos gerais da pesquisa, para a sua concordância e ciência da utilização dos dados.

Por essas questões éticas, após a realização da pesquisa, não será exibido nenhum dado sensível da organização neste documento. Os números obtidos na execução dos cenários e análise dos resultados serão novos dados calculados a partir dos originais, fornecidos pela empresa, de forma descaracterizada.

#### 4.2.2. Características dos projetos

Foram escolhidos três (3) projetos de Big Data para a execução do estudo de caso, que por questões de confidencialidade da organização, foram nomeados como **Projeto A**, **Projeto B** e **Projeto C**. Essa seleção se deu por conta de características em comum entre eles, para facilitar o processo de coleta e análise dos dados, de modo a se obter, posteriormente, o modelo de estimativa de esforço proposto neste trabalho.

A primeira característica dos projetos é que eles utilizam a solução de Big Data da empresa, portanto fazem parte do contexto da pesquisa.

Outro aspecto importante é que os projetos A, B e C estão concluídos, isto é, já passaram por todas as etapas de planejamento e desenvolvimento de software. Somente com os projetos concluídos é que foi possível a obtenção de seus indicadores de qualidade e métricas de esforço, que serão importantes para a coleta de dados do estudo.

Mais um atributo considerado é que os eles possuem a mesma natureza de demanda, ou seja, todos eles são demandas de manutenção adaptativa, que tratam de modificações de produtos de software para mantê-los funcionando de forma adequada em um ambiente que sofre mudanças (SISP, 2018). A portaria SLTI/MP Nº 31 (MPOG, 2010) explicita a importância de

medir projetos dessa natureza e considera relevante a definição de procedimentos complementares de medição nesse tipo de demanda.

Por fim, os projetos A, B e C possuem os mesmos passos de um projeto de Big Data, contendo etapas de coleta, pré-processamento, mineração e visualização dos dados, que foram descritas na subseção 2.3.1 deste documento.

#### **4.2.3. Grupos envolvidos no estudo de caso**

Com relação às pessoas envolvidas na pesquisa, elas foram separadas em três (3) grupos, com características distintas dentro da organização e papéis diferentes no estudo de caso.

##### **GRUPO 1 - Analistas de Conformidade**

Dentro da organização, o Grupo 1 atua no apoio técnico dos projetos, facilitando o processo ágil dentro da equipe, assim como no levantamento dos indicadores de cada time de desenvolvimento. Também é função do Grupo 1 analisar as estimativas iniciais em PF e validar as contagens finais de PF enviadas pelo time. No estudo de caso, esse grupo participou de um treinamento sobre os modelos de medição utilizados nos cenários (SNAP e UST), executou os cenários do estudo de caso e participou de questionários e entrevistas para avaliar sua experiência no estudo. É formado por 3 pessoas.

##### **GRUPO 2 - Líderes dos Times**

O Grupo 2 é responsável por gerenciar o time de desenvolvimento de software. Além disso, eles detêm conhecimento técnico e de negócio sobre o produto, que pode conter um ou mais projetos dentro dele. No estudo de caso, o Grupo 2 indicou e deu apoio aos participantes do Grupo 1, além de participar de entrevistas e questionários ao final do processo. É formado por 3 pessoas.

##### **GRUPO 3 - Gerentes de Capacidade e de Departamento**

Finalmente, as funções do Grupo 3 dentro da empresa envolvem toda a parte de gerenciamento dos líderes e outros gerentes de serviços, como o de conformidade. Eles também atuam no desenvolvimento e manutenção dos indicadores de qualidade da organização, detêm conhecimento sobre faturamento das demandas e outras informações financeiras, e realizam a gestão de demandas com o objetivo de não sobrecarregar os líderes dos times. No estudo de caso, o Grupo 3 não só indicou e deu apoio aos participantes dos Grupos 1 e 2, como também participou de questionários e entrevistas na última etapa do estudo de caso. É formado por 5 pessoas.

### 4.3. Coleta de Dados

A coleta dos dados deste estudo de caso utilizou diversas fontes de informação, com o objetivo de levar em consideração diferentes pontos de vista e situações do caso estudado.

No contexto da pesquisa, os dados foram coletados em três níveis (LETHBRIDGE, SIM e SINGER, 2005) e são descritos a seguir:

*Nível 1: Observações, entrevistas e questionários com recursos da organização;*

*Nível 2: Utilização de indicadores e métricas de esforço dos projetos obtidas nas ferramentas da empresa para auxiliar na execução dos cenários;*

*Nível 3: Análise de documentos de formalização e contagens de PF de projetos de Big Data.*

No primeiro nível da coleta de dados, foram aplicadas entrevistas estruturadas e questionários com os recursos da organização, que foram divididos em grupos e descritos na subseção 4.3.2 deste documento, além de observações feitas pelo pesquisador, que fez uso do método observação participante, proposto por Jorgensen (2015).

O método observação participante propõe que o pesquisador observe e reúna as informações necessárias para a sua pesquisa, enquanto participa ativamente da mesma em situações e ambientes cotidianos. Dessa forma, a coleta e observação permite a obtenção de outros dados que, muitas vezes, são inacessíveis a um observador externo não participante (JORGENSEN, 2015).

A pesquisa realizada considerou o método observação participante pois o seu interesse primário diz respeito aos significados, sentimentos e interações humanas vistos da perspectiva dos membros da organização objeto do estudo. Além disso, o fenômeno da estimativa de esforço em projetos de Big Data dentro do contexto da Administração Pública Federal é suficientemente limitado em escopo, tamanho e localização, e as questões de pesquisa investigadas são adequadas ao caso estudado.

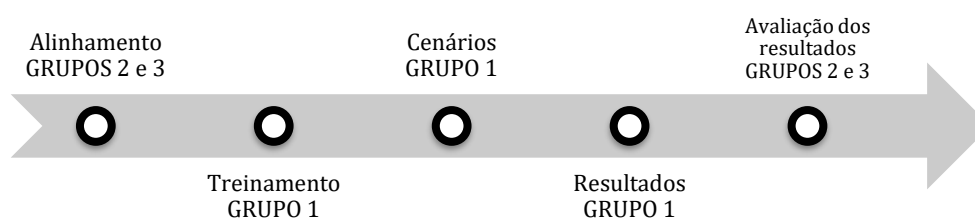
Sobre as entrevistas, elas foram realizadas nos três grupos participantes do estudo de caso ao final das etapas de alinhamento e de execução dos cenários, e seguiram a metodologia de pesquisa qualitativa proposta por McGrath, Palmgren e Liljedahl (2019). Os autores afirmam que esse tipo de entrevista permite que os pesquisadores explorem questões decorrentes das experiências dos entrevistados, então é possível compreender sobre a forma que diferentes fenômenos são vivenciados e percebidos. No presente trabalho, as entrevistas foram realizadas de forma semiestruturada, isto é, com um roteiro inicial de questões elaboradas pelo pesquisador e com a permissão para adicionar algumas questões que podem surgir nesse processo de acordo com as respostas do entrevistado.

Os questionários foram aplicados apenas com o Grupo 1 e seguiram a metodologia de McGrath, Palmgren e Liljedahl (2019). No entanto, diferente das entrevistas, elas possuem um formato estruturado, ou seja, com um roteiro pré-definido de questões que são aplicadas a todos os participantes da mesma forma, com o objetivo de obter respostas que seguem uma mesma linha.

Os demais níveis da coleta de dados, que contemplam a utilização de indicadores e métricas de esforço dos projetos de Big Data, além da análise de seus documentos de formalização e de suas contagens de PF, foram utilizados pelo Grupo 1 na execução dos cenários deste estudo, uma vez que eles contêm informações importantes e que serviram de auxílio nas estimativas de esforço dos projetos de Big Data utilizando SNAP e UST.

Com relação às tarefas realizadas durante a coleta de dados do estudo, elas são dispostas na linha do tempo da Figura 5 e descritas a seguir.

**Figura 5 – Linha do tempo da coleta de dados.**



**Fonte: Desenvolvido pelo autor.**

A primeira tarefa da coleta de dados consistiu no alinhamento com os Grupos 2 e 3 para ouvir relatos sobre o processo atual de medição, utilizando APF, além de explicar o objetivo e benefícios da pesquisa para a organização, e solicitar o apoio de recursos da empresa, que fizeram parte do Grupo 1. Além do alinhamento, foi assinado, nessa fase, o termo de confidencialidade entre as partes para uso e tratamento das informações que foram utilizadas no estudo de caso, detalhado na subseção 4.2.1. Ainda nessa reunião, foi solicitada, pelas partes, um treinamento com o Grupo 1 sobre os modelos de medição utilizados nos cenários do estudo realizado.

Um ponto de destaque durante o encontro inicial com os Grupos 2 e 3 foi o relato sobre a necessidade de aumentar a cobertura dos requisitos funcionais e não-funcionais, isto é, o percentual de funcionalidades e aspectos não-funcionais da solução contemplados pela medição realizada. Além disso, os grupos deram atenção ao nível de granularidade das tarefas criadas e utilizadas pelas equipes de desenvolvimento, isto é, ao seu nível de detalhamento. Por conta dessa necessidade, a cobertura dos requisitos e a granularidade das tarefas, além do tempo de

medição, foram utilizados como parâmetros dos cenários do estudo de caso. Dessa forma, foi possível analisar os diferentes cenários e comparar com o caso atual.

Durante essa fase inicial, o pesquisador observou as ações e comportamentos dos Grupos 2 e 3 e, ao final da reunião, realizou entrevistas com os participantes para avaliar os aspectos referentes ao modelo atual de medição da empresa e sobre as expectativas no uso complementar de outros modelos ao existente. As entrevistas seguiram o script mostrado na Tabela 3.

**Tabela 3 – Script das entrevistas realizadas após alinhamento com Grupos 2 e 3.**

- 
1. Tempo de experiência na métrica APF, em anos. \_\_\_\_\_
  2. Modelos de medição conhecidos. \_\_\_\_\_
  3. Vantagens e desvantagens de APF dentro da organização.
  4. Nível de conformidade de APF na medição de demandas dentro da organização.
  5. Expectativa para o uso complementar de outros modelos de medição ao existente.
- 

**Fonte: Desenvolvido pelo autor.**

A próxima tarefa foi o treinamento com o Grupo 1 sobre SNAP e UST, que fizeram parte dos cenários do estudo de caso. O treinamento ocorreu de forma virtual com os integrantes do Grupo 1, durou cinco dias, foi ministrado pelo pesquisador e foram passados os conceitos, métodos e aplicações de cada modelo de medição citado. Durante o treinamento, o pesquisador observou o comportamento dos participantes e, ao final do treinamento, elaborou um questionário para avaliar o conhecimento e impressões do Grupo 1. O questionário seguiu o script mostrado na Tabela 4.

**Tabela 4 – Script do questionário realizado após o treinamento com o Grupo 1.**

- 
1. Função/Cargo na organização. \_\_\_\_\_
  2. Tempo de experiência na função/cargo. \_\_\_\_\_
  3. Pontos positivos e negativos do treinamento.
  4. Classificação de pontos do treinamento.
  5. Avaliação geral do treinamento.
- 

**Fonte: Desenvolvido pelo autor.**

Após o treinamento, os cenários foram executados pelo Grupo 1. Essa fase durou cerca de 15 dias, distribuídos em 5 semanas. Conforme dito anteriormente, nessa fase, o Grupo 1 utilizou indicadores e métricas de esforço dos projetos de Big Data como insumos, além de analisarem suas contagens de Pontos de Função e documentos de formalização. Durante a execução dos cenários, o pesquisador realizou observações sobre o andamento da fase e, ao final da execução, após a coleta dos resultados, ele elaborou entrevistas e questionários com os participantes. O questionário e as entrevistas seguiram os scripts mostrados nas Tabelas 5 e 6, respectivamente.

**Tabela 5 – Script do questionário realizado após a execução dos cenários com o Grupo 1.**

1. Classificação dos cenários do estudo.
2. Grau de importância de cada cenário com o V do Big Data.

**Fonte: Desenvolvido pelo autor.**

**Tabela 6 – Script das entrevistas realizadas após a execução dos cenários com o Grupo 1.**

1. Pontos positivos e negativos da etapa de execução dos cenários.
2. Pontos positivos e negativos do uso complementar de outros modelos de medição a APF dentro da organização.
3. Recomendação para outros órgãos utilizarem modelos complementares a APF.

**Fonte: Desenvolvido pelo autor.**

Logo após a obtenção dos resultados, e finalizando a coleta de dados, foi marcado um encontro com os Grupos 2 e 3 para avaliar e consolidar os dados obtidos no estudo de caso, além de discutir a implicação dos resultados nos indicadores de qualidade da empresa. Nessa etapa, o pesquisador aplicou entrevistas com os participantes dos dois grupos, a fim de analisar a experiência deles no estudo realizado. As entrevistas seguiram o script mostrado na Tabela 7 a seguir.

**Tabela 7 – Script das entrevistas realizadas após avaliação dos resultados com Grupos 2 e 3.**

1. Vantagens e desvantagens do uso complementar de SNAP e UST a APF dentro da organização.
2. Nível de conformidade do uso complementar de outros modelos de medição a APF dentro da organização.
3. Impacto dos resultados nos indicadores de qualidade da empresa.
4. Recomendação para outros órgãos utilizarem modelos de medição de forma complementar a APF.

**Fonte: Desenvolvido pelo autor.**

Na Tabela 8 a seguir, são apresentadas as tarefas executadas e sua relação com os níveis da coleta de dados deste estudo.

**Tabela 8 – Relação das tarefas realizadas e níveis da coleta de dados.**

|  | Observações | Entrevistas | Questionários |
|--|-------------|-------------|---------------|
| <b>Alinhamento GRUPOS 2 e 3</b>              | x           | x           |               |
| <b>Treinamento GRUPO 1</b>                   | x           |             | x             |
| <b>Cenários GRUPO 1</b>                      | x           |             |               |
| <b>Resultados GRUPO 1</b>                    |             | x           | x             |
| <b>Avaliação dos resultados GRUPOS 2 e 3</b> |             | x           |               |

**Fonte: Desenvolvido pelo autor.**

#### 4.4. Resultados e Discussões

Nesta seção, serão apresentados os resultados e discussões do estudo de caso desta pesquisa, separados por fase da coleta de dados e respeitando os três níveis de obtenção dos dados relatados na seção 4.3.

Para auxiliar na consolidação dos dados, utilizou-se a ferramenta Google Formulários e os softwares RStudio e o IRAMUTEQ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*), ambos gratuitos, *open source* e ancorados no R, linguagem de programação para gráficos e cálculos estatísticos. O IRAMUTEQ permite fazer análises estatísticas sobre corpus textuais e tabelas de indivíduos por palavras (CAMARGO e JUSTO, 2013), e o RStudio é um ambiente de desenvolvimento integrado para R.

#### 4.4.1. Alinhamento GRUPOS 2 e 3

Na etapa inicial de coleta de dados, que consistiu na reunião de alinhamento com os Grupos 2 e 3, foram feitas observações e entrevistas pelo pesquisador, relatadas a seguir.

##### 4.4.1.1. Observações

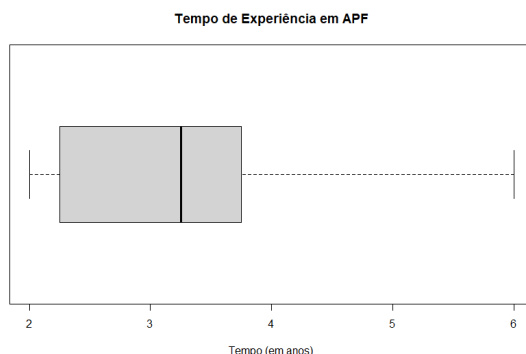
- Os participantes dos Grupos 2 e 3 estiveram dispostos e se mostraram animados para apoiar o estudo de caso;
- A ideia de realizar um treinamento com o Grupo 1 sobre os conceitos dos modelos de medição utilizados no estudo surgiu pela necessidade dos participantes do Grupo 1 não terem ainda o conhecimento sobre SNAP e UST para realizar as medições. Tal necessidade foi importante para a pesquisa, pois nivelou o aprendizado do grupo, ou seja, todos os participantes executaram os cenários do estudo de caso partindo do mesmo ponto de partida;
- O Grupo 3, formado por gerentes de capacidade e de departamento da empresa, destacou a importância da pesquisa para a boa manutenção dos indicadores de qualidade e gestão das demandas da organização, isto é, eles reforçaram o objetivo geral da pesquisa, descrito na seção 1.2 deste documento;
- Os participantes do Grupo 2, formado pelos líderes dos times de desenvolvimento, ficaram à disposição para apoiar e gerenciar situações que viessem a atrasar o andamento da pesquisa, como possíveis ausências dos participantes do Grupo 1, seja por razões da própria empresa ou por motivos de força maior.

##### 4.4.1.2. Entrevistas

As entrevistas foram realizadas ao final do alinhamento com os Grupos 2 e 3 e seguiram o script mostrado na Tabela 3, na seção 4.3.

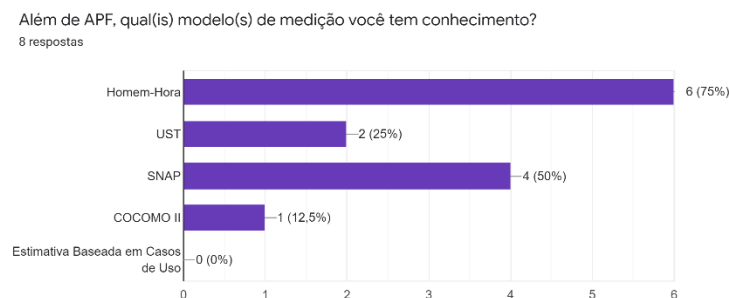
Com relação ao tempo de experiência no processo de contagem em APF, os dados são mostrados na Figura 6.



**Figura 6 – Tempo de experiência no processo de contagem em APF dos Grupos 2 e 3.**

**Fonte: Desenvolvido pelo autor.**

Como pode ser visto na Figura 6, o tempo de experiência no processo de contagem em APF médio dos Grupos 2 e 3 está entre três e quatro anos, e quase todos os dados concentram-se entre dois e quatro anos. Isso mostra que os participantes dos Grupos 2 e 3, que são líderes e gerentes, possuem boa experiência em APF.

**Figura 7 – Modelos de medição conhecidos além de APF pelos Grupos 2 e 3.**

**Fonte: Desenvolvido pelo autor.**

No que diz respeito aos modelos de medição conhecidos por eles, além de APF, é mostrado no gráfico da Figura 7 que o modelo Homem-Hora é conhecido por 75% dos entrevistados, seguido pelo SNAP (50%) e UST (25%). No entanto, conforme descrito na seção 2.2, o modelo Homem-Hora não vem sendo recomendado pelo TCU para ser utilizado nas empresas pertencentes a Administração Pública Federal (MPOG, 2014). Por isso, confirmou-se a utilização do SNAP e UST no estudo de caso, tanto na execução dos cenários, como no treinamento realizado com o Grupo 1.

A próxima pergunta da entrevista foi sobre as vantagens e desvantagens do uso de APF dentro da organização. As vantagens levantadas pelos Grupos 2 e 3 foram, em linhas gerais:

- Independe de linguagem de programação e solução tecnológica;

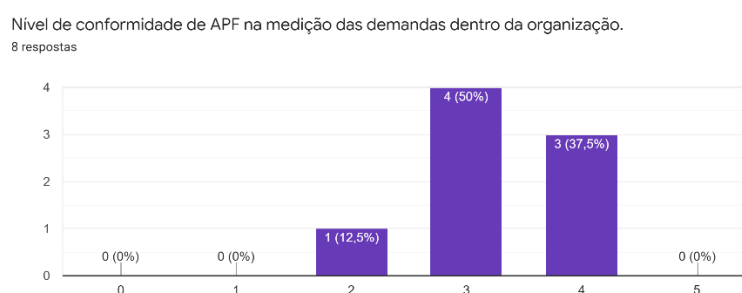
- Garante boa cobertura dos requisitos funcionais da solução;
- Métrica muito utilizada no mercado e aceita pelos órgãos de controle;
- Possui órgãos de regulamentação;
- Contribui para boa gestão de contratos e faturamento;
- Segura na conformidade de pagamentos.

As desvantagens obtidas nas entrevistas com os participantes dos Grupos 2 e 3 foram, também de forma genérica:

- Não representa o esforço efetivo de desenvolvimento da solução, por não contemplar requisitos não funcionais;
- Pode ter diferença entre o valor estimado e o valor final cobrado;
- Não mede a complexidade das demandas;
- Não tem relação com as tarefas desenvolvidas pelas equipes de desenvolvimento;
- Dificuldade de estimar traços de arquitetura e outras integrações existentes.

O nível de conformidade de APF na medição de demandas dentro da organização, numa escala de 0 a 5, onde 0 representa que APF não está em conformidade com o que é medido dentro da empresa, e 5 simboliza que APF está em total conformidade com as medições realizadas, é mostrado na Figura 8.

**Figura 8 – Nível de conformidade de APF na medição das demandas dentro da organização.**



**Fonte: Desenvolvido pelo autor.**

É mostrado na Figura 8 que metade dos entrevistados acreditam que, atualmente, o uso de APF dentro da organização está numa conformidade mediana na medição das suas demandas.

Finalizando as entrevistas, a expectativa dos Grupos 2 e 3 para o uso complementar de SNAP e UST ao modelo existente na empresa era alta, uma vez que os grupos acreditaram que a utilização dos outros modelos não vai impactar muito no tempo de medição das demandas e ajudará o andamento das atividades das equipes.

#### 4.4.2. Treinamento GRUPO 1

Na etapa do treinamento realizado com o Grupo 1, o pesquisador levantou observações e aplicou questionários com os participantes. Os resultados são mostrados nas subseções a seguir.

##### 4.4.2.1. Observações

- Assim como ocorreu na etapa anterior, os participantes do Grupo 1 se mostraram animados e dispostos para realizar o treinamento;
- Durante os cinco dias de treinamento, o pesquisador guiou e apresentou os conteúdos, mas o aprendizado foi colaborativo, isto é, os integrantes do Grupo 1 trabalharam juntos em prol do mesmo objetivo e foram responsáveis pelo aprendizado uns dos outros (GOKHALE, 1995);
- No início de cada dia de treinamento, ocorriam discussões sobre dúvidas que os participantes traziam a respeito de SNAP e UST. Essa etapa enriqueceu a atividade, pois motivou tanto o pesquisador como o Grupo 1 a buscar mais conceitos e relacioná-los com o cenário atual de medição da empresa, no que diz respeito a sua estimativa de esforço em projetos de Big Data;
- Conforme dito na subseção 2.1.3, é necessário um catálogo de serviços para que ocorra a medição em UST. Por isso, dado que não era possível criar um catálogo para a organização em tempo suficiente, surgiu a ideia de utilizar o catálogo de serviços de outra empresa pública federal, que utiliza UST para complementar sua estimativa de esforço. A criação de um catálogo adaptado às necessidades da empresa ficou como trabalho futuro;
- Os cinco dias de treinamento foram diluídos em duas semanas de reuniões, devido a fatores de força maior da empresa, o que não prejudicou o andamento da pesquisa;
- Ainda no período de treinamento, uma pessoa do Grupo 1 participou de uma reunião virtual sobre o futuro do UST e compartilhou um resumo do que foi discutido com o restante do grupo.

##### 4.4.2.2. Questionários

Os questionários foram aplicados de forma virtual, ao final do treinamento, por meio da ferramenta Google Formulários, e seguiram o script mostrado na Tabela 4, mostrado na seção 4.3.

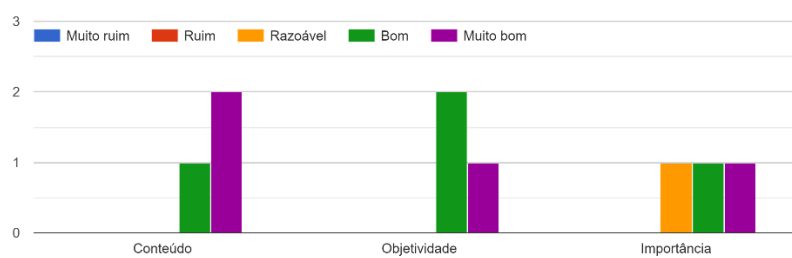
Os três participantes que compõem o Grupo 1 são analistas de conformidade/métricas na organização e possuem, no mínimo, 1 ano de experiência no cargo.

Os pontos positivos do treinamento, levantados por eles, se referiram ao conteúdo objetivo e em conformidade com a necessidade atual da empresa, além do aprendizado colaborativo, destacado pelo pesquisador na subseção anterior. Já o ponto negativo foi o curto espaço de tempo do treinamento, que não permitiu aprofundar em mais discussões, além de não ser possível criar um catálogo de serviços, também observado pelo pesquisador na subseção 4.4.2.1.

De maneira geral, o treinamento foi classificado como “esclarecedor” e “importante” pelo Grupo 1 e com relação ao conteúdo, objetividade e importância, as seguintes classificações foram feitas e mostradas na Figura 9.

**Figura 9 – Classificação do treinamento com o Grupo 1.**

Numa escala de 0 a 5, sendo 0 muito ruim e 5 muito bom, como você classifica os seguintes pontos do treinamento?



**Fonte: Desenvolvido pelo autor.**

#### 4.4.3. Cenários/Resultados GRUPO 1

Na próxima etapa do estudo de caso, os participantes foram alocados pelo pesquisador para realizar a medição de três projetos de Big Data da organização, utilizando os modelos SNAP e UST, e organizados nos cenários a seguir, anteriormente descritos na seção 4.1:

- **Cenário 1: APF + SNAP;**
- **Cenário 2: APF + UST;**
- **Cenário 3: APF + SNAP + UST.**

A execução de todos os cenários foi feita com base nos seguintes parâmetros, mencionados anteriormente na seção 4.3:

- **Cobertura dos requisitos:** proporção de funcionalidades (requisitos funcionais) e aspectos não-funcionais da solução (requisitos não-funcionais) contemplados pela medição realizada.

$$Cobertura = \left( \frac{RF' + RNF'}{RF + RNF} \right) \times 100, \text{ onde:}$$

RF' é a quantidade de requisitos funcionais contemplados pela medição;  
 RNF' é a quantidade de requisitos não-funcionais contemplados pela medição;  
 RF é a quantidade total de requisitos funcionais da solução;  
 RNF é a quantidade total de requisitos não-funcionais da solução;  
 $RF' \in RF$ ;  $RNF' \in RNF$ .

- **Granularidade:** percentual do nível de detalhamento das tarefas realizadas pela equipe de desenvolvimento na construção da solução.

$$\text{Granularidade} = \left( \frac{T'}{T} \right) \times 100, \text{ onde:}$$

T' é a quantidade de tarefas da equipe contempladas pela medição;

T é a quantidade total de tarefas da equipe;

$T' \in T$ .

- **Tempo de medição:** corresponde ao intervalo em horas para uma pessoa realizar a estimativa de esforço do projeto de Big Data utilizando o modelo de medição do cenário.

$$\text{Tempo de medição} = \sum_{i=1}^n h \times i, \text{ onde:}$$

n é a quantidade de dias da estimativa de esforço do projeto de Big Data;

h é a quantidade de horas trabalhadas por dia.

É importante destacar que, para efeitos de comparação, será utilizado o cenário base, ou **Cenário 0**, como sendo o que representa o modelo de medição atual da empresa, isto é, utilizando apenas APF.

O resultado da execução dos cenários, dividido por projeto, é mostrado nas Tabelas 9, 10 e 11 a seguir.

**Tabela 9 – Cenários – Projeto A.**

|                             | <b>Cenário 0</b> | <b>Cenário 1</b>    | <b>Cenário 2</b>   | <b>Cenário 3</b>          |
|-----------------------------|------------------|---------------------|--------------------|---------------------------|
|                             | <b>(APF)</b>     | <b>(APF + SNAP)</b> | <b>(APF + UST)</b> | <b>(APF + SNAP + UST)</b> |
| <b>Cobertura (%)</b>        | 52,3             | 85,7                | 66,6               | 85,7                      |
| <b>Granularidade (%)</b>    | 32               | 42,48               | 56,86              | 62,09                     |
| <b>Tempo de medição (h)</b> | 14               | 34                  | 18                 | 38                        |

**Fonte: Desenvolvido pelo autor.**

**Tabela 10 – Cenários – Projeto B.**

|                             | <b>Cenário 0</b> | <b>Cenário 1</b>    | <b>Cenário 2</b>   | <b>Cenário 3</b>          |
|-----------------------------|------------------|---------------------|--------------------|---------------------------|
|                             | <b>(APF)</b>     | <b>(APF + SNAP)</b> | <b>(APF + UST)</b> | <b>(APF + SNAP + UST)</b> |
| <b>Cobertura (%)</b>        | 21,8             | 53,1                | 46,8               | 59,3                      |
| <b>Granularidade (%)</b>    | 57,5             | 63,75               | 80                 | 86,25                     |
| <b>Tempo de medição (h)</b> | 22               | 38                  | 26                 | 42                        |

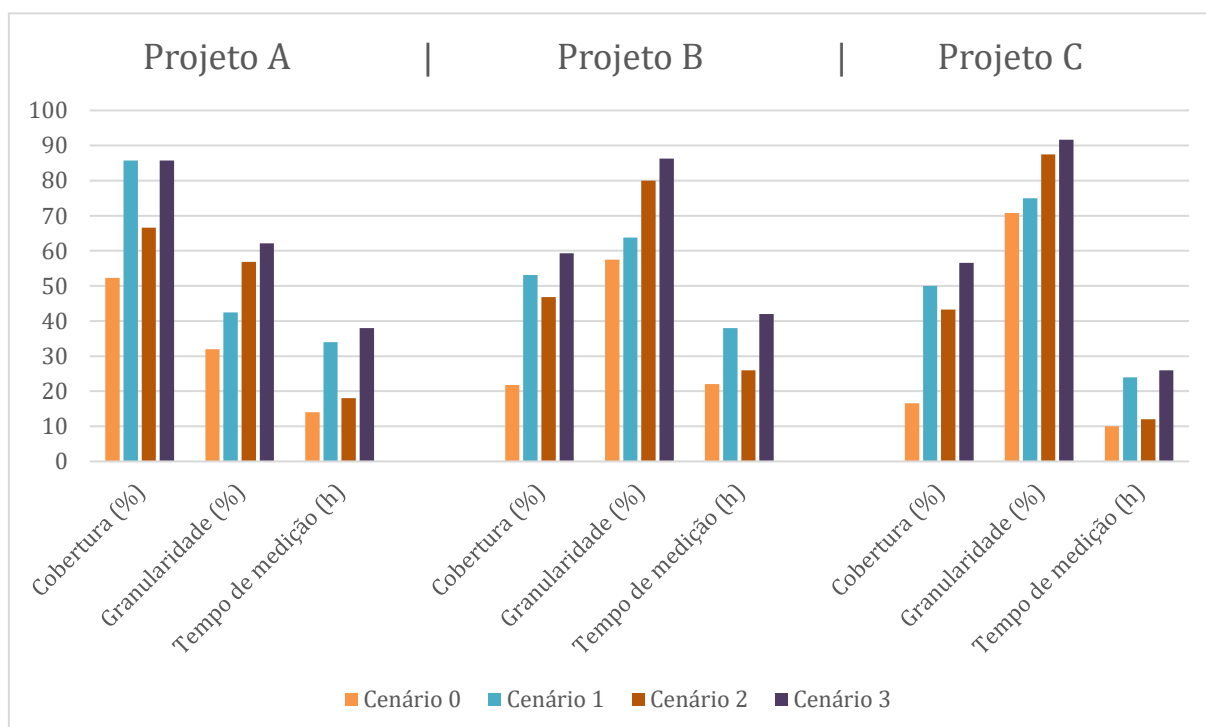
**Fonte: Desenvolvido pelo autor.**

**Tabela 11 – Cenários – Projeto C.**

|                             | <b>Cenário 0</b><br>(APF) | <b>Cenário 1</b><br>(APF + SNAP) | <b>Cenário 2</b><br>(APF + UST) | <b>Cenário 3</b><br>(APF + SNAP + UST) |
|-----------------------------|---------------------------|----------------------------------|---------------------------------|--|
| <b>Cobertura (%)</b>        | 16,6                      | 50                               | 43,3                            | 56,6                                   |
| <b>Granularidade (%)</b>    | 70,8                      | 75                               | 87,5                            | 91,67                                  |
| <b>Tempo de medição (h)</b> | 10                        | 24                               | 12                              | 26                                     |

**Fonte: Desenvolvido pelo autor.**

Conforme é visto nas tabelas acima, os valores da cobertura e granularidade aumentam à medida que outros modelos complementam a estimativa de esforço, nos três projetos do estudo de caso. O tempo de medição, por consequência, também sofre um aumento, mas ele varia de acordo com o cenário aplicado. Essa variação nos parâmetros também pode ser vista na Figura 10 a seguir.

**Figura 10 – Projetos e cenários do estudo de caso.**

**Fonte: Desenvolvido pelo autor.**

Percebe-se, pela Figura 10, que os três projetos possuíam o mesmo comportamento ao passo que mudavam os cenários do estudo. Comparando com o cenário base (cenário 0), a cobertura aumentou com o cenário 2, seguindo o aumento com o cenário 1 e, posteriormente, com o cenário 3. A granularidade aumentou na ordem dos cenários 1, 2 e 3. Já o tempo de

medição seguiu o mesmo comportamento da cobertura, aumentando na seguinte ordem de cenários: 2, 1 e 3.

Após a execução dos cenários, o pesquisador levantou observações e aplicou questionários e entrevistas com o Grupo 1. Os resultados são mostrados nas subseções a seguir.

#### 4.4.3.1. Observações

- A execução dos cenários foi diluída em cinco semanas, devido a fatores de força maior da empresa, isto é, demandas urgentes, o que não prejudicou o andamento da pesquisa;
- Durante o período de execução dos cenários, os participantes tiravam dúvidas entre si e revisavam alguns conteúdos vistos no treinamento, para relembrar alguns conceitos do modelo de medição SNAP e empregá-los no cenário 1;
- Ainda sobre o cenário 1, houve comentários sobre o tempo gasto na medição, a complexidade na sua execução e a relação existente com a cobertura de requisitos, que aumentou em todos os projetos;
- Na execução dos cenários 2 e 3, a interação com as equipes de desenvolvimento facilitou o trabalho, o que fez com o que o tempo de medição fosse menor que o esperado;
- Houve relatos sobre a qualidade da documentação das demandas, que foi insumo para a execução dos cenários.

#### 4.4.3.2. Questionários

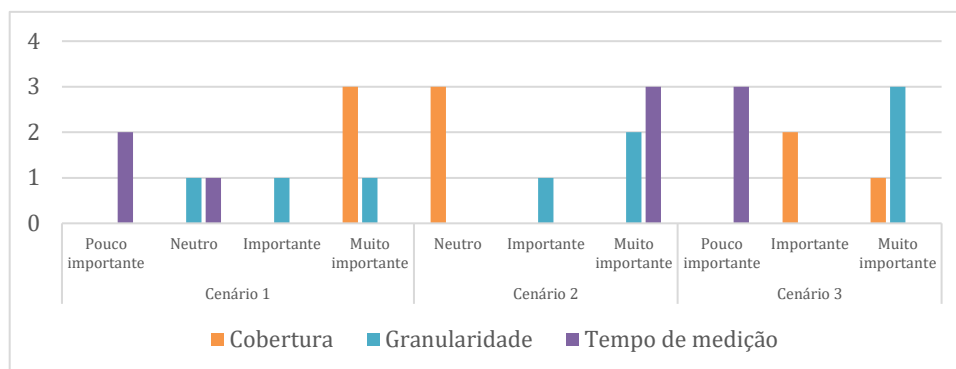
Os questionários foram aplicados de forma virtual, ao final da execução dos cenários, por meio da ferramenta Google Formulários, e seguiram o script mostrado na Tabela 5, mostrado na seção 4.3.

De acordo com os participantes, a relação de cada cenário executado com os parâmetros do estudo de caso é mostrada na Figura 11. Os valores de referência são mostrados na Tabela 12.

**Tabela 12 – Valores de referência para pergunta 1 do questionário.**

|   |                  |
|---|------------------|
| 1 | Não é importante |
| 2 | Pouco importante |
| 3 | Neutro           |
| 4 | Importante       |
| 5 | Muito importante |

**Fonte: Desenvolvido pelo autor.**

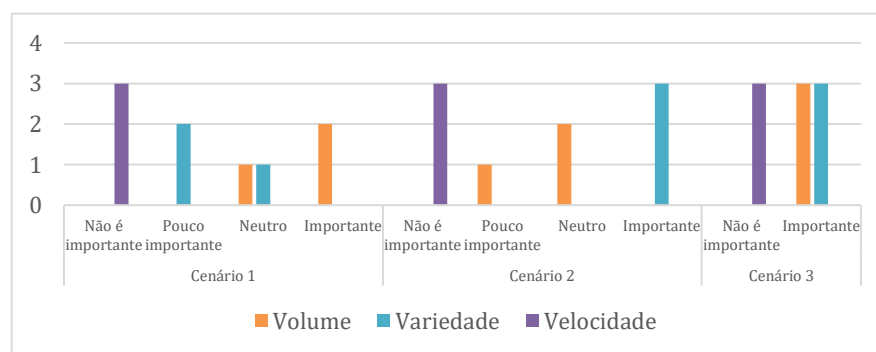
**Figura 11 – Relação dos cenários com os parâmetros do estudo de caso.**

**Fonte: Desenvolvido pelo autor.**

Conforme é visto na Figura 11, a cobertura dos requisitos foi considerada, pelos participantes, como “muito importante” no cenário 1, “neutro” no cenário 2, e “importante” e “muito importante” no cenário 3. A granularidade das tarefas obteve classificação de “neutro” a “muito importante” no cenário 1, “importante” e “muito importante” no cenário 2 e “muito importante” no cenário 3. E o tempo de medição foi visto como “pouco importante” e “neutro” no cenário 1, “muito importante” no cenário 2 e “pouco importante” no cenário 3.

Em linhas gerais, percebeu-se que a cobertura dos requisitos se destacou nos cenários 1 e 3, em que o SNAP esteve presente, a granularidade das tarefas teve relevância nos cenários 2 e 3, onde foi utilizado UST, e o tempo de medição foi enfatizado apenas no cenário 2 (APF + UST).

Outro questionamento feito ao Grupo 1 foi sobre a relação das características de Big Data (os Vs), vistos na seção 2.3, com cada cenário do estudo de caso. Os valores de referência são os mesmos da Tabela 12 e o resultado pode ser visto na Figura 12.

**Figura 12 – Relação dos Vs do Big Data com os cenários do estudo de caso.**

**Fonte: Desenvolvido pelo autor.**

O volume do Big Data é visto, na Figura 12, como “neutro” e “importante” no cenário 1, “pouco importante” e “neutro” no cenário 2 e “importante” no cenário 3. Já a variedade do



Big Data foi considerada, pelo Grupo 1, “pouco importante” ou “neutro” no cenário 1 e “importante” nos cenários 2 e 3. E a velocidade do Big Data obteve classificação de “não é importante” nos três cenários do estudo.

Portanto, de acordo com os dados do questionário, o volume foi evidenciado nos cenários em que o SNAP aparece (1 e 3), a variedade se destacou nos cenários onde o UST foi utilizado (2 e 3) e a velocidade não obteve relação com nenhum dos cenários.

#### 4.4.3.3. Entrevistas

As entrevistas também ocorreram de forma virtual com o Grupo 1, ao final dessa etapa, e seguiram o script mostrado na Tabela 6, na seção 4.3.

Os pontos positivos e negativos dessa etapa foram indagados ao Grupo 1. Os pontos positivos levantados por eles estão relacionados ao entrosamento do grupo com as equipes de desenvolvimento, o que colaborou para o bom andamento da execução dos cenários, principalmente do cenário 2, além da interação entre os participantes do Grupo 1, que já tinham experiência de trabalhar juntos, facilitando, assim, a curva de aprendizado deles. Dois pontos negativos dessa etapa foram o tempo para executar o cenário 1 e a qualidade de alguns insumos utilizados na execução.

Com relação ao uso complementar de outros modelos de medição a APF dentro da organização, os pontos positivos apresentados pelo Grupo 1 foram os seguintes:

- Representa, de forma mais adequada, a estimativa de esforço das soluções;
- Possibilidade de estimar o esforço em atividades de arquitetura, homologação e testes;
- Dependendo do modelo escolhido, o tempo de medição não sofre muita alteração;
- A cobertura dos requisitos e a granularidade das tarefas têm um impacto positivo e podem ajudar no faturamento e gerenciamento dos indicadores de qualidade da empresa.

Os pontos negativos que foram citados se referem ao tempo inicial para se construir um catálogo de serviços da organização, no caso do UST, além da curva de aprendizado com o modelo de medição SNAP.

A última pergunta foi sobre a recomendação para outros órgãos utilizarem modelos complementares de medição a APF. Todos os participantes do Grupo 1 recomendaram a utilização, desde que seja respeitado os pontos negativos levantados por eles na pergunta anterior.

#### 4.4.4. Avaliação dos Resultados GRUPOS 2 e 3

Na última etapa da coleta de dados, o pesquisador se reuniu com os Grupos 2 e 3, apresentou e avaliou os resultados dos cenários executados no estudo de caso.

Após a consolidação dos resultados, foi discutida a definição de níveis para os parâmetros do estudo de caso, para que os valores possam ser tratados dentro de uma faixa de valor, e não pelo valor absoluto. Essa definição foi feita de acordo com a experiência dos integrantes do Grupo 3, que detém conhecimento sobre gerenciamento e faturamento das demandas e parte das suas funções está voltada para a manutenção dos indicadores de qualidade da empresa. Por isso, os níveis definidos para os parâmetros do estudo estão mostrados na Tabela 13 a seguir.

**Tabela 13 – Níveis dos parâmetros do estudo de caso.**

|                         | <b>Baixo</b> | <b>Médio</b> | <b>Alto</b>  |
|-------------------------|--------------|--------------|--------------|
| <b>Cobertura</b>        | 0 a 50%      | 51 a 70%     | 71 a 100%    |
| <b>Granularidade</b>    | 0 a 50%      | 51 a 70%     | 71 a 100%    |
| <b>Tempo de medição</b> | 0 a 20h      | 21 a 30h     | Acima de 31h |

**Fonte: Desenvolvido pelo autor.**

Além dos níveis, foi estabelecida, pelos Grupos 2 e 3, uma ordem decrescente de preferência dos parâmetros e cenários do estudo de caso, também com base na experiência dos participantes dessa fase no gerenciamento das demandas que chegam na empresa.

Portanto, de acordo com os Grupos 2 e 3, a ordem de preferência dos parâmetros é o tempo de medição, seguido da cobertura e, por último, a granularidade. Com relação aos cenários, o que possui maior visibilidade e de interesse para a empresa é o seu próprio cenário base (cenário 0), seguido dos cenários 2, 1 e 3.

##### 4.4.4.1. Entrevistas

Após a avaliação dos resultados, o pesquisador realizou entrevistas com os participantes dos Grupos 2 e 3, que seguiram o script mostrado na Tabela 7, na seção 4.3.

Com relação às vantagens do uso complementar de SNAP e UST a APF dentro da organização, foi levantado o seguinte pelos Grupos 2 e 3:

- Representa, com mais efetividade, o esforço necessário de desenvolvimento das soluções que chegam na empresa;
- Estimula as equipes a se organizarem e a planejarem melhor suas tarefas;
- Consegue cobrir mais requisitos das soluções;

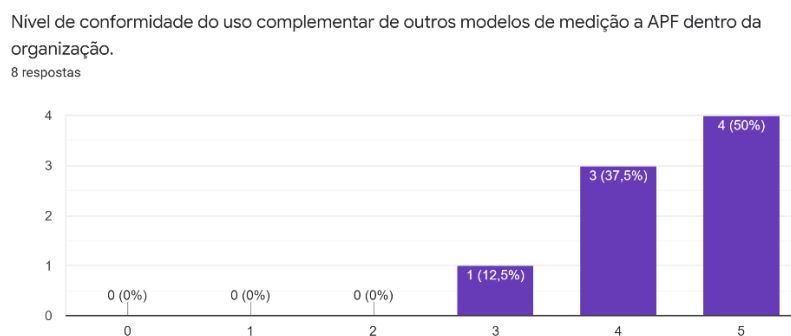
- É visto, de forma satisfatória, pelos órgãos de controle;
- As métricas são conhecidas, então o uso pode ser escalável.

As desvantagens apresentadas pelos entrevistados foram:

- A curva de aprendizado de um dos modelos precisa ser levada em consideração no planejamento das demandas;
- A relação entre a complexidade do Big Data e o detalhamento da solução realizado pela equipe pode aumentar o tempo de medição, considerando que a equipe precisará de tempo para entender e relacionar os requisitos não funcionais com os itens dos modelos de medição;
- O tempo de medição, quando se complementa APF com SNAP, aumentou de forma considerável.

O nível de conformidade do uso complementar de outros modelos de medição a APF na medição de demandas dentro da organização, numa escala de 0 a 5, onde 0 representa que o uso não está em conformidade e 5 simboliza que está em total conformidade com o que é praticado dentro da empresa, é mostrado na Figura 13.

**Figura 13 – Nível de conformidade do uso complementar de SNAP e UST a APF.**



**Fonte: Desenvolvido pelo autor.**

Conforme pode ser visto na Figura 13, a utilização de SNAP e UST em complemento a APF gerou um nível de conformidade positivo entre os membros dos Grupos 2 e 3, o que reforça a importância do uso dentro da empresa na estimativa de esforço de demandas de Big Data.

A próxima pergunta da entrevista foi sobre o impacto dos resultados do estudo de caso nos indicadores de qualidade da empresa. De maneira geral, foi relatado que o impacto ocorrerá de forma direta e indireta. Por exemplo, indicadores que envolvam o parâmetro “tempo” sofrerão impacto direto, já que utilizar UST e SNAP em complemento a APF aumentará o tempo de medição e, conseqüentemente, o tempo total de atendimento à demanda. No entanto, o efeito em um indicador de “uso das ferramentas da empresa” não ocorrerá de imediato, mas

com o passar do tempo, onde as equipes estarão mais preocupadas com o nível de detalhamento das tarefas de desenvolvimento, para que sejam feitas estimativas de esforço das demandas mais adequadas.

Finalmente, os Grupos 2 e 3 foram questionados sobre a recomendação para outros órgãos utilizarem modelos de medição de forma complementar a APF. Assim como ocorreu com o Grupo 1, todos os participantes recomendam a utilização de SNAP e UST, em complemento a APF, no entanto era necessário avaliar, de forma estratégica, a implicação desse uso no tempo de medição que a empresa utiliza.

#### **4.5. Considerações**

A investigação do fenômeno da estimativa de esforço em projetos de Big Data no contexto de uma empresa pública federal foi concluída após etapas de coleta de dados e com a exibição dos resultados na seção anterior.

Foi mostrado que, apesar da APF independer de solução tecnológica, ser bastante usada nos contratos e possuir baixo tempo de medição, ela necessita de um complemento na sua estimativa de esforço, principalmente quando se trata de projetos de alta complexidade, volume e variedade, como são os projetos de Big Data.

A utilização complementar de SNAP conseguiu aumentar a cobertura de requisitos, mas, de acordo com os participantes, a curva de aprendizado do modelo e seu tempo de medição podem não ser tão interessantes no contexto da estimativa de esforço.

No caso do UST, o tempo de medição foi baixo e o modelo conseguiu bons índices de cobertura e granularidade. No entanto, um ponto que merece destaque é o grande esforço inicial para se construir um catálogo de serviços da organização.

No último cenário, a combinação de SNAP e UST, em complemento a APF, traz consigo os pontos positivos e negativos de cada modelo, então ela fica sujeita a uma análise estratégica pelos responsáveis da organização.

Nesse contexto, o uso complementar de outros modelos de medição consegue representar a estimativa de esforço de forma mais adequada, uma vez que o SNAP e UST conseguem medir outros aspectos que a APF não mede, como, por exemplo, atividades de arquitetura e de testes não funcionais. Esse uso complementar também pode estimular as equipes a planejarem melhor suas tarefas de desenvolvimento e a discutirem mais sobre os documentos de requisitos com os responsáveis pela solução.

Falando sobre os parâmetros do estudo, percebeu-se que a cobertura e a granularidade possuem relação com a qualidade da documentação que é enviada e trabalhada pela equipe. Em outras palavras, uma boa documentação pode ser sinal de bons índices desses parâmetros.

A documentação também é importante para a compreensão e detalhamento dos requisitos do projeto de Big Data, por ser algo complexo e que necessita dessa explicação.

Dito isso, a partir do estudo de caso realizado, onde a análise ocorreu dentro de um contexto específico e com dados reais, foi proposto um modelo para representar uma forma de se estimar o esforço em projetos de Big Data dentro da organização, que será explicado no capítulo a seguir.

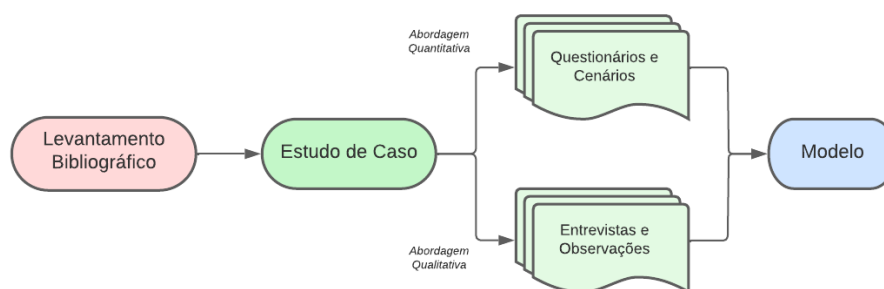
## 5. MODELO DE ESTIMATIVA DE ESFORÇO EM PROJETOS DE BIG DATA NA ADMINISTRAÇÃO PÚBLICA FEDERAL

O modelo de estimativa de esforço aplicado a projetos de Big Data no contexto da Administração Pública Federal, que foi tratado neste trabalho, e foi resultado do estudo de caso, é mostrado neste capítulo.

### 5.1. Construção

A construção do modelo ocorreu a partir da sequência de passos mostrados na Figura 14.

**Figura 14 – Construção do modelo.**



**Fonte: Desenvolvido pelo autor.**

É visto na Figura 14 que o modelo foi construído a partir do estudo de caso, que seguiu a abordagem qualitativa, com os dados obtidos das entrevistas e observações, e quantitativa, com as informações dos questionários e dos cenários. Sobre os cenários, foram utilizados os parâmetros (cobertura, granularidade e tempo de medição) com os níveis definidos pelos Grupos 2 e 3 na etapa de avaliação do estudo.

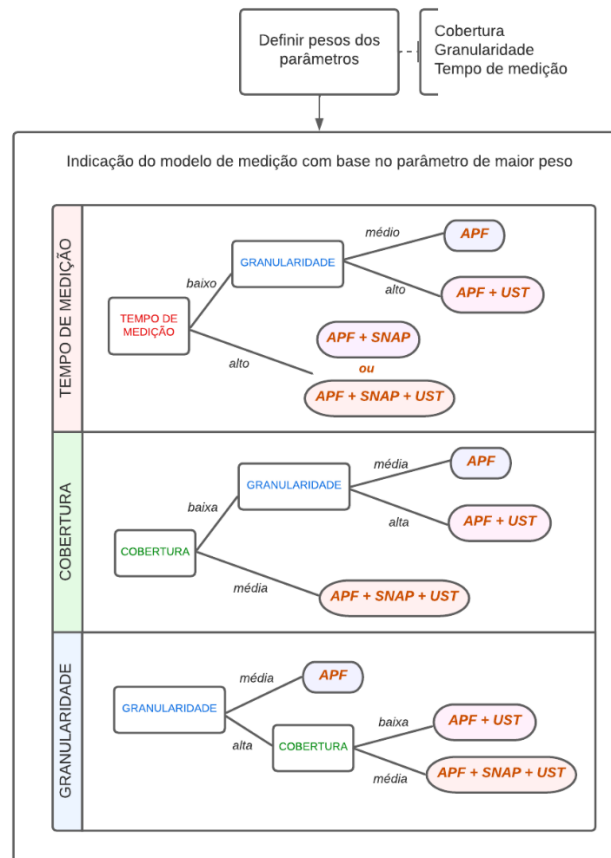
As informações de cobertura, granularidade e tempo de medição dos três projetos nos quatro cenários foram consolidadas em tabelas e, a partir disso, verificou-se a predominância dos níveis dos parâmetros em cada cenário, para que fosse criada uma representação do que foi obtido no estudo de caso. Então, observou-se que:

- No Cenário 0, o nível da cobertura foi predominantemente baixo, o da granularidade foi médio e o tempo de medição foi baixo;
- No Cenário 1, não houve predominância de níveis para cobertura e granularidade, só para o tempo de medição, que foi alto;
- No Cenário 2, o nível da cobertura foi predominantemente baixo, o da granularidade foi alto e o tempo de medição foi baixo;

- No Cenário 3, o nível da cobertura foi predominantemente médio, o da granularidade foi alto e o tempo de medição foi alto.

Assim, foi criado o modelo, descrito na Figura 15 e disposto em três raias, cada uma contendo um parâmetro do estudo como principal e relacionando ele com os outros parâmetros.

**Figura 15 – Modelo após o estudo de caso.**



**Fonte: Desenvolvido pelo autor.**

De acordo com o que foi observado nas entrevistas com os participantes do estudo, é importante que a tomada de decisão da empresa, com relação a sua estimativa de esforço, seja feita com base nas prioridades que ela possui no momento.

Por isso, a primeira etapa do modelo consiste em definir os pesos dos parâmetros, para que seja escolhido o mais importante, i.e., o de maior peso, para ser a raiz da árvore que indicará o(s) modelo(s) de medição adequado a situação.

Considerando o parâmetro “tempo de medição” como o de maior peso, é mostrado no modelo que ele pode assumir os valores “baixo” e “alto”. Caso seja alto, o modelo indica a utilização de dois cenários: APF + SNAP e APF + SNAP + UST. Caso o valor seja baixo, é necessário avaliar o nível de granularidade desejado, que pode ter valores “médio” e “alto”. Se o valor for médio, o modelo sugere apenas o cenário atual da empresa, isto é, utilizando APF.

No entanto, se o valor da granularidade for alto, a recomendação vai para o cenário em que UST complementa APF.

Na segunda raia, quando a “cobertura” tem maior peso, os valores assumidos pelo parâmetro são “baixo” ou “médio”. Quando for médio, o modelo sugere que se utilize SNAP e UST em complemento a APF para estimar o esforço do projeto de Big Data. Na hipótese de a cobertura ser baixa, é preciso avaliar a granularidade, que pode ser “média” ou “alta”. Caso a granularidade seja média, é indicado que a estimativa seja feita apenas com APF; porém, se ela for alta, é sugerido pelo modelo que se use APF e UST na medição.

A última raia contém o parâmetro “granularidade” como raiz da árvore, isto é, que possui o maior peso. Ele pode assumir os valores “médio” e “alto”. A recomendação será pelo cenário apenas com APF se a granularidade for média. Se ela for alta, avalia-se, também, a cobertura, que pode ser “baixa” ou “média”. O esforço da demanda será medido com APF e UST se a cobertura for baixa, e será APF + SNAP + UST se o parâmetro tiver valor “médio”.

É importante destacar que o modelo não consegue sugerir um ou mais modelos de medição caso a combinação dos parâmetros e seus níveis não esteja descrita nele. Exemplificando, ocasiões em que se deseja ter cobertura alta, granularidade baixa ou tempo de medição médio não são cobertas pelo modelo, ou seja, ele não consegue gerar a recomendação.

Portanto, o modelo pode ser aplicado como forma de recomendação à organização do(s) modelo(s) de medição indicado(s) para a ocasião de uma nova demanda de projeto de Big Data, com características semelhantes às que foram tratadas no estudo de caso, e que esteja alinhada à escolha estratégica da (i) ordem decrescente dos parâmetros utilizados no estudo, sendo o primeiro parâmetro o mais importante e o terceiro o menos importante, e (ii) dos seus níveis (baixo, médio e alto).

Por exemplo, considerando que a escolha estratégica da empresa, na ocasião, seja a ordem “tempo de medição = baixo” e “granularidade = alta”, o modelo recomendaria os modelos de medição APF e UST como indicados para medir o esforço da demanda.

Em outras palavras, o modelo pode ajudar a resolver um problema de classificação de uma demanda de projeto de Big Data que chega na empresa. Esse tipo de problema, na área de aprendizagem de máquina, consiste na predição da categoria de um dado não observado com base em outros dados de entrada já observados (ZHANG et al., 2017).

## **5.2. Validação**

A validação do modelo é uma etapa importante na pesquisa, pois é nela em que é possível avaliar a capacidade de predição dele.

Para realizar a validação do modelo proposto neste documento, foi considerado pelo pesquisador que existia um problema de classificação a ser resolvido, dado que a situação



consiste na chegada de um projeto de Big Data, com suas características, na empresa, e que necessitava de uma estimativa de esforço.

Para ajudar a resolver esse problema, e considerando o conjunto de dados da pesquisa, foi utilizado o algoritmo de classificação KNN (*K-nearest neighbor*) (FUKUNAGA e NARENDRA, 1975) (AGARWAL et al., 2021). Esse algoritmo é utilizado na área de aprendizagem de máquina e se baseia na classificação de um novo dado com base na distância para os  $k$  vizinhos mais próximos, que já foram classificados anteriormente (conjunto de treinamento). O pseudocódigo do KNN é descrito na Tabela 14.

**Tabela 14 – Pseudocódigo do algoritmo de classificação KNN.**

---

```
1 início:
2     Preparar conjunto de dados de treinamento e teste;
3     Informar o valor de  $k$ .
4 para cada nova amostra faça
5     Calcular distância para todas as amostras;
6     Determinar o conjunto das  $k$  distâncias mais próximas;
7     Verificar o rótulo com mais representantes no conjunto dos  $k$  vizinhos;
8     Escolher o rótulo.
9 fim para
10 retornar: conjunto de rótulos de classificação
```

---

**Fonte: Desenvolvido pelo autor e adaptado de Fukunaga e Narendra (1975).**

Como é mostrado na Tabela 14, a entrada do algoritmo consiste nos conjuntos de treinamento e teste, além do valor de  $k$ . Com essas definições, ele realiza, para cada dado que for classificado, os procedimentos de (i) cálculo da distância dele para todas as amostras, (ii) determinação do conjunto das  $k$  distâncias mais próximas dele e (iii) escolha da classe/conjunto de dados a qual o novo dado será classificado, que será a com mais representantes dentro do conjunto definido na etapa anterior.

No contexto da pesquisa, as quatro combinações dos modelos de medição, que foram os cenários do estudo de caso, serão as classes de classificação, e o dado a ser classificado será um projeto de Big Data com suas características baseadas nos parâmetros do estudo (cobertura, granularidade e tempo de medição). A ideia é recomendar uma das combinações dos modelos de medição para o projeto de Big Data utilizando o KNN.

Entretanto, para que seja realizada a validação do modelo proposto, é necessário um número representativo de amostras, para que diminua a ocorrência de erros nos resultados. Como pode ser observado no modelo proposto na Figura 15, o número de possíveis resultados é baixo, conforme é mostrado na Tabela 15. Para facilitar a leitura dos dados, optou-se em utilizar a nomenclatura dos cenários ao contrário do(s) nome(s) do(s) modelo(s) de medição recomendado, ou seja:

- Cenário 0 = APF;
- Cenário 1 = APF + SNAP;
- Cenário 2 = APF + UST;
- Cenário 3 = APF + SNAP + UST.

**Tabela 15 – Resultados obtidos a partir do modelo.**

| #  | Tempo de medição | Cobertura | Granularidade | Cenário |
|----|------------------|-----------|---------------|---------|
| 1  | Baixo            | -         | Média         | 0       |
| 2  | Baixo            | -         | Alta          | 2       |
| 3  | Alto             | -         | -             | 1       |
| 4  | Alto             | -         | -             | 3       |
| 5  | -                | Baixa     | Média         | 0       |
| 6  | -                | Baixa     | Alta          | 2       |
| 7  | -                | Média     | -             | 3       |
| 8  | -                | -         | Média         | 0       |
| 9  | -                | Baixa     | Alta          | 2       |
| 10 | -                | Média     | Alta          | 3       |

**Fonte: Desenvolvido pelo autor**

Por conta disso, foi necessário o pesquisador se reunir novamente com os Grupos 2 e 3 para realizar um refinamento nos dados obtidos e, conseqüentemente, no modelo. Nesse encontro, os participantes sugeriram uma expansão na lista de possíveis resultados baseada no modelo construído e na ordem decrescente de preferência de todos os parâmetros e cenários, feita por eles na última etapa do estudo de caso, e descrita na subseção 4.4.4. Dessa forma, a lista com os resultados para todas as combinações possíveis de cenários é mostrada na Tabela 16.

**Tabela 16 – Resultados das combinações possíveis de cenários, após análise dos Grupos 2 e 3.**

| #  | Tempo de medição | Cobertura | Granularidade | Cenário |
|----|------------------|-----------|---------------|---------|
| 1  | Baixo            | Baixa     | Baixa         | 0       |
| 2  | Baixo            | Baixa     | Média         | 0       |
| 3  | Baixo            | Baixa     | Alta          | 3       |
| 4  | Baixo            | Média     | Baixa         | 0       |
| 5  | Baixo            | Média     | Média         | 0       |
| 6  | Baixo            | Média     | Alta          | 3       |
| 7  | Baixo            | Alta      | Baixa         | 2       |
| 8  | Baixo            | Alta      | Média         | 1       |
| 9  | Baixo            | Alta      | Alta          | 2       |
| 10 | Médio            | Baixa     | Baixa         | 1       |
| 11 | Médio            | Baixa     | Média         | 0       |
| 12 | Médio            | Baixa     | Alta          | 0       |
| 13 | Médio            | Baixa     | Alta          | 3       |
| 14 | Médio            | Média     | Baixa         | 2       |
| 15 | Médio            | Média     | Média         | 0       |
| 16 | Médio            | Média     | Alta          | 1       |
| 17 | Médio            | Alta      | Baixa         | 2       |
| 18 | Médio            | Alta      | Média         | 3       |
| 19 | Médio            | Alta      | Alta          | 3       |
| 20 | Alto             | Baixa     | Baixa         | 0       |
| 21 | Alto             | Baixa     | Média         | 3       |
| 22 | Alto             | Baixa     | Alta          | 1       |
| 23 | Alto             | Média     | Baixa         | 0       |
| 24 | Alto             | Média     | Média         | 3       |
| 25 | Alto             | Média     | Alta          | 3       |
| 26 | Alto             | Alta      | Baixa         | 2       |
| 27 | Alto             | Alta      | Média         | 2       |
| 28 | Alto             | Alta      | Alta          | 3       |

**Fonte: Desenvolvido pelo autor**

A recomendação dos cenários feita pelos Grupos 2 e 3 seguiu o pseudocódigo do algoritmo descrito na Tabela 17.

**Tabela 17 – Pseudocódigo do algoritmo de recomendação dos cenários feito pelos Grupos 2 e 3.**

---

```

1   entrada:
2       Valores dos parâmetros.
3   início:
4       Colocar os parâmetros e cenários em ordem decrescente de preferência;
5       N = número de parâmetros.
6   enquanto não houver cenário recomendado e N for maior que zero faça
7       para cada cenário faça
8           Verificar no modelo a ocorrência dos N parâmetros no cenário;
9           se as ocorrências dos N parâmetros se confirmarem no cenário então
10              recomende o cenário
11          fim se
12      fim para
13      N = N - 1;
14  fim enquanto
15  se não houver cenário recomendado e N < 1 então
16      não recomende nenhum cenário
17  fim se
18  retornar: cenário recomendado

```

---

**Fonte: Desenvolvido pelo autor**

Para exemplificar o uso do algoritmo mostrado na Tabela 17 e o povoamento dos resultados visto na Tabela 16, serão consideradas a **linha 3**, com *tempo de medição baixo*, *cobertura baixa* e *granularidade alta*, e **linha 21**, com *tempo de medição alto*, *cobertura baixa* e *granularidade alta*.

A ordem decrescente de preferência dos parâmetros foi definida pelo Grupos 2 e 3 como sendo tempo de medição, cobertura e granularidade, e a ordem dos cenários foi o cenário 0, 2, 1 e 3 (ambas descritas na subseção 4.4.4. do documento).

**Tabela 18 – Exemplo do algoritmo de recomendação para a linha 3 da Tabela 17.**

---

**N = 3**

Loop 1: Não há cenário recomendado e N é maior que zero  
Parâmetros: tempo de medição, cobertura e granularidade  
Para o cenário 0, somente 2 parâmetros estão no modelo;  
Para o cenário 2, os três parâmetros estão no modelo;  
Recomenda o cenário 2.

**Retorno: Cenário 2**

---

**Fonte: Desenvolvido pelo autor**

**Tabela 19 – Exemplo do algoritmo de recomendação para a linha 21 da Tabela 17.**

---

**N = 3**

Loop 1: Não há cenário recomendado e N é maior que zero (N = 3)  
Parâmetros: tempo de medição, cobertura e granularidade  
Para o cenário 0, somente 1 parâmetro está no modelo;  
Para o cenário 2, somente 2 parâmetros estão no modelo;  
Para o cenário 1, somente 1 parâmetro está no modelo;  
Para o cenário 3, somente 1 parâmetro está no modelo.  
N = 2 (tempo de medição e cobertura)

Loop 2: Não há cenário recomendado e N é maior que zero (N = 2)  
Parâmetros: tempo de medição e cobertura  
Para o cenário 0, somente 1 parâmetro está no modelo;

---

---

```

Para o cenário 2, somente 1 parâmetro está no modelo;
Para o cenário 1, somente 1 parâmetro está no modelo;
Para o cenário 3, somente 1 parâmetro está no modelo.
N = 1 (tempo de medição)

```

Loop 3: Não há cenário recomendado e N é maior que zero (N = 1)

Parâmetro: tempo de medição

```

Para o cenário 0, o parâmetro não está no modelo;
Para o cenário 2, o parâmetro não está no modelo;
Para o cenário 1, o parâmetro está no modelo;
Recomenda o cenário 1.

```

**Retorno: Cenário 1**

---

### Fonte: Desenvolvido pelo autor

Os 28 resultados presentes na Tabela 16 ainda não representavam um número ideal da amostra para os participantes dos Grupos 2 e 3. Para obter um número mais significativo para a validação, foram realizados ensaios com replicação no número dos resultados até chegar a um valor numérico adequado para o processo. Então, chegou-se ao número que representa a quantidade de demandas que a organização tinha recebido no último ano até o momento do encontro. Não foi pensado no total de projetos de Big Data da empresa, pois esse valor ainda é baixo. Para que não houvesse repetição das linhas da tabela, uma vez que houve replicação dos dados, os valores dos parâmetros, que antes se referiam aos seus níveis, foram substituídos por números inteiros e aleatórios dentro da faixa de cada nível. Esses números estão na Tabela 13 e foram descritos na subseção 4.4.4. do documento.

Uma vez que os dados foram definidos, o processo de validação pode ser iniciado, de forma cruzada e do tipo *k-fold*. A validação cruzada efetua a divisão da base de dados em conjunto de treinamento e conjunto de testes para avaliar a sua predição. No método *k-fold*, a base é dividida em *k* subconjuntos de mesmo tamanho (ou o mais parecido possível) e o processo de validação ocorre em *k* iterações. Em cada uma delas, o conjunto de testes será uma das *k* partes da base e o restante forma o conjunto de treinamento. Então, ao final de todas as iterações, toda a base de dados é utilizada como treinamento e teste (BURMAN, 1989).

Com relação ao valor do *k*, é visto na literatura que o viés da validação cruzada *k-fold* diminui quando o *k* aumenta. Entretanto, um valor muito alto do *k* pode aumentar o custo computacional do processo, além de aumentar a variância com um conjunto pequeno de testes (CUNHA, 2019). Por isso, foi escolhido o valor de  $K = 5$ , que se adequou ao processo e ao tamanho da base de dados (MARCOT e HANEA, 2021). Portanto, o conjunto de treinamento ficará com 80% da base de dados e o conjunto de testes com 20% em cada uma das repetições.

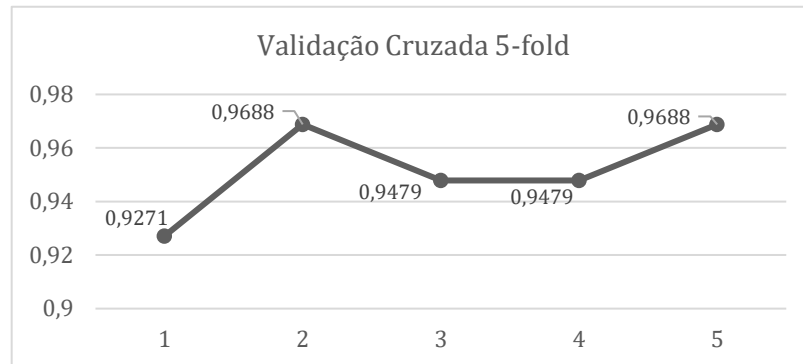
A métrica de avaliação escolhida para auxiliar na validação desse processo foi a acurácia, que indica uma performance geral do modelo, isto é, dentre todas as classificações realizadas, quantas o modelo classificou corretamente. Ela é calculada da seguinte maneira:

$$Acurácia = \frac{VP+VN}{VP+VN+FP+FN}, \text{ onde:}$$

VP é a quantidade de vezes que o algoritmo indica o cenário de forma correta (verdadeiros positivos);  
VN é a quantidade de vezes que o algoritmo não indica o cenário de forma correta (verdadeiros negativos);  
FP é a quantidade de vezes que o algoritmo indica um cenário, mas de forma incorreta (falso positivo);  
FN é a quantidade de vezes que o algoritmo não indica um cenário, mas de forma incorreta (falso negativo).

O valor médio da acurácia obtido ao final das  $K = 5$  iterações foi de 95,21%, conforme pode ser visto na Figura 16.

**Figura 16 – Gráfico da variação da acurácia na validação cruzada 5-fold.**



**Fonte: Desenvolvido pelo autor**

Como pode ser visto na Figura 16, os valores da acurácia variaram entre 0,9271 (92,71%), na primeira iteração, a 0,9688 (96,88%), na segunda iteração. O valor médio (95,21%) representa, no contexto da pesquisa, que a cada cem projetos de Big Data que chegam na organização, o modelo de estimativa de esforço proposto neste trabalho consegue prever corretamente o cenário para mais de 95 projetos. De acordo com a literatura, esse valor corresponde a um nível de alta precisão e baixo risco, trazendo mais segurança para a tomada de decisão e, dessa forma, validando o modelo proposto (SOUZA, 2020).

Foi observado, então, ao final dessa etapa, a importância dos Grupos 2 e 3 no refinamento do modelo proposto, tornando-o mais robusto e, agora, validado. Além disso, os resultados dessa validação podem indicar que o modelo pode ser instanciado em outros contextos de demandas e projetos, uma vez que ele permite a ponderação dos parâmetros e dos cenários de medição, que podem ser diferentes em outras circunstâncias.

## 6. CONCLUSÃO

A chegada de novas tecnologias, e suas soluções de TIC, vem trazendo mudanças importantes dentro das organizações de TI. No que se refere ao Big Data, com grande volume, variedade e velocidade, o desafio está relacionado ao seu processamento e, no contexto deste trabalho, à maneira de se estimar esforço dentro de uma empresa pertencente a Administração Pública Federal, regida por leis e normativos que limitam suas formas de medição.

Diante desse problema, o objetivo deste trabalho foi analisar o fenômeno descrito no parágrafo anterior na literatura, por meio de um estudo de caso, para sugerir um modelo de estimativa de esforço mais adequado a projetos de Big Data no contexto de uma empresa pública federal.

As questões de pesquisa levantadas nas seções 1.2.2. e 4.1. foram respondidas ao longo do documento, mas concluiu-se que:

- O modelo atual de medição em APF atende em partes os projetos de Big Data nas organizações da Administração Pública Federal. A partir do levantamento bibliográfico e estudo de caso realizado, foram encontradas lacunas nesse atendimento, que se referem, principalmente, à cobertura dos requisitos;
- A utilização de modelos de medição com base em requisitos não funcionais e em unidades de serviços técnicos complementa a medição em APF dos projetos de Big Data nessas empresas. Foi visto, no estudo de caso e na construção do modelo, que o uso complementar dos SNAP e UST ocorreu de forma satisfatória e positiva, com base nos parâmetros de cobertura dos requisitos, granularidade das tarefas e tempo de medição;
- Nem todas as características relacionadas ao Big Data (os Vs) tiveram impacto direto na sua estimativa de esforço. Também no estudo de caso e no modelo, foi mostrado que os Vs “volume” e “variedade” são associados em alguns cenários, porém isso não aconteceu com o V “velocidade”.

A fim de evitar que o andamento das resoluções das questões de pesquisa fosse prejudicado no decorrer do trabalho, foram identificadas algumas ameaças à validade da pesquisa, que foram minimizadas e gerenciadas pelo pesquisador. A mais importante delas se refere ao método observação participante utilizado na coleta de dados do estudo de caso. Como o pesquisador trabalha na empresa que foi caso do seu estudo, o risco de que houvesse um viés de sua parte era grande. No entanto, o pesquisador realizou um planejamento detalhado e, com o apoio do protocolo do estudo de caso, não se desviou do seu papel de observar e reunir as informações necessárias para a pesquisa.

Outras ameaças identificadas dizem respeito a aspectos de confiabilidade dos dados ou que também poderiam atrasar o tempo ou mudar o foco da pesquisa, como a disponibilidade de tempo dos entrevistados, o receio deles falarem o que pensam sobre a organização e a resolução das questões de pesquisa seguir por outro caminho. Para minimizar o impacto dessas ameaças,

o planejamento do estudo foi apresentado e validado em conjunto com todos os *stakeholders* entrevistados e observados, de forma incremental, e as sugestões de melhorias foram coletadas e analisadas pelo pesquisador de modo a não prejudicar o andamento do trabalho.

Nas próximas seções, serão levantadas as contribuições e sugestões de trabalhos futuros identificados no decorrer da pesquisa pelo pesquisador e demais envolvidos.

## 6.1. Contribuições da Pesquisa

A maior parte das contribuições da pesquisa, por ser aplicada a um contexto específico, estão voltadas para o ambiente da organização que foi caso do estudo realizado.

Como resultado do estudo de caso, o novo modelo de medição aplicado a projetos de Big Data dentro da Administração Pública Federal foi implementado e é passível de uso dentro da organização. No entanto, a mudança trazida com esse modelo seria a nível de processo dentro da empresa, para que todos os empregados sigam as etapas propostas e realizem as estimativas de esforço de uma forma complementar à maneira que se estimava antes, isto é, utilizando apenas APF.

Além disso, as lacunas identificadas na pesquisa durante a fase de levantamento bibliográfico e estudo de caso auxiliaram na construção do modelo. Ele foi elaborado a partir de dados reais, obtidos no estudo, e foi validado após refinamentos feitos pelos participantes da organização.

De maneira geral, a pesquisa contribuiu, no nível de literatura, com a possibilidade de investigação do fenômeno da estimativa de esforço aplicado a dois contextos específicos, que são o de projetos de Big Data e o da Administração Pública Federal, que possuem suas características bem definidas.

## 6.2. Trabalhos Futuros

Com relação aos trabalhos futuros, foram destacadas algumas melhorias na validação e construção do modelo, como também aspectos de generalização e replicação da pesquisa para outros contextos. Essas melhorias contam, então, como possibilidades de continuidade da pesquisa e são listadas a seguir:

- Instanciar o modelo de forma interna (dentro da empresa), fazendo uso dele em outros contextos, como projetos de Inteligência Artificial ou de Ingestão de Dados; ou de forma externa, aplicando em outras organizações dentro da Administração Pública Federal e/ou que trabalhem com projetos de Big Data;

- Ainda dentro da empresa, verificar o impacto que a estimativa de esforço realizada com o novo modelo de medição teve na parte (i) da estimativa de custo e orçamento e (ii) no gerenciamento de qualidade da organização;
- Para o modelo, pensar em outras formas de refinamento e validação. Por exemplo, utilizando dados de projetos reais;
- Outro ponto sobre o modelo é verificar o seu impacto com outras características dos projetos de Big Data consideradas na literatura, e descritas na seção 2.3. do documento, como a veracidade, valor, variabilidade e visualização dos dados.

Em resumo, conclui-se que o trabalho elaborado neste documento não se configura como uma finalização do estudo. As sugestões de trabalhos futuros podem contribuir, de maneira significativa, para o contexto da estimativa de esforço, dos projetos de Big Data e da Administração Pública Federal. Além disso, o avanço permanente do uso das TIC pode impulsionar a criação de novas tecnologias e as organizações precisam estar preparadas para o seu planejamento e atendimento, para que, dessa forma, possam estimar o esforço de suas soluções de maneira eficiente.



## REFERÊNCIAS BIBLIOGRÁFICAS

- AGARWAL, A.; SHARMA, P.; ALSHEHRI, M.; MOHAMED, A. A.; ALFARRAJ, O. Classification model for accuracy and intrusion detection using machine learning approach. In: PeerJ Computer Science. 2021.
- ALAOUI, I. E.; GAHI, Y.; MESSOUSSI, R. Big Data Quality Metrics for Sentiment Analysis Approaches. Proceedings of the 2019 International Conference on Big Data Engineering. Hong Kong. 2019.
- ALMEIDA, W. H. C; FURTADO, F. Análise sobre métricas nos contratos de fábricas de software no âmbito da Administração Pública. 1ª edição. Rio de Janeiro. Albatroz. 2019.
- ALMEIDA, W. H. C.; MONTEIRO, L. A.; FURTADO, F. Análise sobre métricas de software nos contratos de terceirização de serviços de desenvolvimento de sistemas (fábricas de software) no âmbito da administração pública federal. Anais Eletrônicos do Encontro Unificado de Computação do Piauí. 2018.
- AMIT, S.; VERMA, B. Software Development Effort Estimation: A Review. Proceedings of the International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, 2013.
- BRASIL. Tribunal de Contas da União. ACÓRDÃO Nº2362/2015. Plenário. 2015.
- BRASIL. Presidência da República. Secretaria Geral. Subchefia para Assuntos Jurídicos. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados. 2018.
- BRASIL. Tribunal de Contas da União. ACÓRDÃO Nº1508/2020. Plenário. 2020.
- BURMAN, P. A comparative study of ordinary cross-validation, v-fold cross-validation and the learning-testind methods. In: Biometrika, 76: 503-514. 1989.
- CAMARGO, B. V. e JUSTO, A. M. IRAMUTEQ: um software gratuito para análise de dados textuais. In: Temas psicol. vol.21, n.2, p. 513-518. 2013.
- CUNHA, J. P. Z. Um estudo comparativo das técnicas de validação cruzada aplicadas a modelos mistos. 59p. Dissertação (Mestrado em Estatística) - Instituto de Matemática e Estatística da Universidade de São Paulo. 2019.
- ELGENDY, N.; ELRAGAL, A. Big Data Analytics: A Literature Review Paper. In: Industrial Conference on Data Mining, p. 214–227. Springer, 2014.

---

FINEP. Guia de Métricas de Software FINEP. Versão 1.3. Ministério da Ciência, Tecnologia, Inovações e Comunicações. 2017.

FLICK, U. Introdução à Metodologia de Pesquisa: um guia para iniciantes. – 1. ed. – São Paulo: Penso Editora, 2013.

FUKUNAGA, K.; NARENDRA, P. M. A branch and bound algorithm for computing k-nearest neighbors. In: IEEE Transactions on Computers, v. 100, n. 7, p. 750–753. 1975.

FUNDECC. Fundação de Desenvolvimento Científico e Cultural. Processo de Contratação de Serviços. Anexo 1, Termo de Referência. 2017.

GOKHALE, A. A. Collaborative Learning enhances critical thinking. In: Journal of Technology Education, p. 22-30, Fall, 1995.

GOVERNO DIGITAL, 2020. Orientações para Novas Contratações e Renovações de Contratos Baseados em UST. Disponível em: <https://www.gov.br/governodigital/pt-br/contratacoes/orientacoes-para-novas-contratacoes-e-renovacoes-de-contratados-baseados-em-ust>. Último acesso em: 25/08/2020.

HAI, V. V.; NHUNG, H. L. T. K.; THÁI, H. H. A Review of Software Effort Estimation by Using Functional Points Analysis. Proceedings of the Computational Methods in Systems and Software. P. 408-422. 2019.

IBM. 10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations. 2016.

IEEE, 2019. IEEE Trial-Use Standard for Software Non-Functional Sizing Measurements. Disponível em: <https://standards.ieee.org/standard/2430-2019.html>. Último acesso em: 14/10/2020.

IFPUG. Manual de Práticas de Contagem de Pontos de Função. Versão 2.3.1. Traduzido de: IFPUG Counting Practices Manual V. 4.3. 2010.

IFPUG. Processo de Avaliação Não Funcional de Software (SNAP) - Manual de Práticas de Avaliação. Versão 2.4. Traduzido de: IFPUG Assessment Process Manual V. 2.4. 2020.

JENSEN, R. An Improved Macrolevel Software Development Resource Estimation Model. In Judge, G., W. Griffiths, and C. Hill, Learning and Practicing Econometrics, Wiley, New York, 1983.

---

JORGENSEN, D. L. Participant Observation. *Emerging Trends in the Social and Behavioral Sciences*, 1–15. doi:10.1002/9781118900772.etrds0247. 2015.

KHAN, N.; ALSAQER, M.; SHAH, H.; BADSHA, G.; ABBASI, A. A.; SALEHIAN, S. The 10 Vs, Issues and Challenges of Big Data. *Proceedings of the 2018 International Conference on Big Data and Education*. New York, NY. 2018.

LETHBRIDGE, T. C.; SIM, S. E.; SINGER, J. Studying software engineers: data collection techniques for software field studies. *Empir Softw Eng* 10(3):311–341. 2005.

MACHADO, F. N. R. *Big Data: O Futuro dos Dados e Aplicações*. São Paulo, Érica, Saraiva Educação. 1ª edição, 2018.

MARCOT, B. G.; HANEA, A. M. What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? In: *Computational Statistics*, 36:2009-2031. 2020.

MCGRATH, C.; PALMGREN, P. J.; LILJEDAHN, M. Twelve Tips for Conducting Qualitative Research Interviews. *Medical Teacher*, 41(9), 1002–1006. 2019.

MEDEIROS, J. D. R. V. An approach to support the requirements specification in agile software development. 2017. 180p. Tese (Doutorado em Ciência da Computação) - Universidade Federal de Pernambuco, Recife, 2017.

MPOG. Portaria SLTI/MP N° 31, de 29 de novembro de 2010. p. 2010.

MPOG. Instrução Normativa N° 4, de 11 de setembro de 2014. p. 2014.

NATHANAEL, E. H.; HENDRADJAYA, B.; SUNINDYO, W. D. Study of Algorithmic Method and Model for Effort Estimation in Big Data Software Development Case Study Geodatabase. In: *The 5th International Conference on Electrical Engineering and Informatics 2015*. Bali, Indonesia. 2015.

PARK, R. The Central Equations of the PRICE Software Cost Model. Park R., 4th COCOMO Users', Group Meeting, 1988.

PEREIRA, Vanessa Alves da Silva. *Big data: um estudo em gestão empresarial*. FACC, UFRJ, Rio de Janeiro. 2016.

REIS, L. C. R.; SÁ, M. I. F. Big Data: Um novo campo de atuação para bibliotecários. *Revista de Ciências e Tecnologias de Informação e Comunicação*, n. 41, 2020.

---

RUNESON, P.; HÖST, M. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software*. 14, 2, 131-164. 2009.

SAMOSIR, R. S.; HENDRIC, H. L.; GAOL, F. L.; ABDURACHMAN, E.; SOEWITO, B. Measurement Metric Proposed For Big Data Analytics System. *Proceedings of the 2017 International Conference on Computer Science and Artificial Intelligence*. New York, NY. 2017.

SANTOS, M. L.; SALVADOR, L. N.; CRUZES, D. FindYourHelp: an expert finder module on Virtual Learning Environments. *Informática na Educação: teoria & prática*, Porto Alegre, v. 14, n. 2, p. 95-112, jul./dez. 2011. Citado na página x.

SISP. Guia de Contagem de Pontos de Função do SISP para Projetos Data Warehouse: versão 1.0 / Ministério do Planejamento, Orçamento e Gestão. Secretaria de Logística e Tecnologia da Informação. Brasília: MP, 2015.

SISP. Roteiro de Métricas de Software do SISP: versão 2.3 / Ministério do Planejamento, Desenvolvimento e Gestão. Secretaria de Tecnologia da Informação e Comunicação - Setic. – Brasília: MP, 2018.

SOUZA, J. A. C. Acurácia das métricas de validação da classificação de imagens. *Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Minas Gerais, Escola de Engenharia*. 2020.

TAURION, C. *Big Data*. Rio de Janeiro: Brasport, 2013.

VENKATRAMAN, R.; VENKATRAMAN, S. Big Data Infrastructure, Data Visualisation and Challenges. In: *International Conference on Big Data and Internet of Things*. Melbourne, VIC, Australia. August 22–24, 2019.

VIJAY, J. F. Enrichment of accurate software effort estimation using fuzzy-based function point analysis in business data analytics. In: *Emerging Intelligent Algorithms for Edge-of-Things Computing*. Neural Computing and Applications, p. 1633–1639. 2018.

ZHANG, S.; LI, X.; ZONG, M.; ZHU, X.; CHENG, D. Learning k for kNN classification. In: *ACM Trans. Intell. Syst. Technol.* 8. 2017.

## **APÊNDICE**

## APÊNDICE – PROTOCOLO DO ESTUDO DE CASO

### **Objetivo:**

Analisar o processo de estimativa de esforço em projetos Big Data numa empresa que pertence à Administração Pública Federal.

### **Objeto:**

Trata-se de uma empresa pública de Tecnologia da Informação, que desenvolve soluções para o governo e cidadãos brasileiros há mais de 40 anos. Com relação a forma de estimar o esforço dos seus contratos de software, ela segue a recomendação proposta pela portaria SLTI/MP N° 31 (MPOG, 2010) e pela Instrução Normativa SLTI/MP N° 4 (MPOG, 2014), isto é, utiliza como modelo de medição a Análise de Pontos de Função (APF).

### **Descrição do caso:**

O processo atual de medição do objeto de estudo utiliza APF, que não consegue suprir suas necessidades, porque esse modelo de medição se preocupa apenas com a medição dos requisitos funcionais e que possuem valor para o usuário. Além disso, projetos de Big Data costumam ter vários requisitos não funcionais, que não são medidos utilizando APF.

### **Importância:**

A organização tem a oportunidade de ter seu processo de medição de esforço analisado e otimizado, fazendo com que ela aumente o seu faturamento e melhore as suas estimativas.

### **Questões de Pesquisa:**

*QP1. O modelo atual de medição em Pontos de Função está sendo bem utilizado em projetos de Big Data nas empresas públicas federais?*

*QP2. O uso complementar de modelos de medição com base em unidades de serviços técnicos (UST) e em requisitos não funcionais (SNAP) supre as necessidades das organizações que fazem parte da Administração Pública Federal?*

*QP3. As características relacionadas ao volume, variedade e velocidade dos dados de um projeto de Big Data devem ser levadas em consideração na sua estimativa de esforço?*

### **Cenários:**

1. Utilizar APF, que mede requisitos funcionais das soluções, em conjunto com o modelo SNAP, que mede requisitos não-funcionais de um sistema;
2. Complementar com UST a medição de esforço e custo dos projetos de Big Data, que atualmente é a partir de PF;
3. Combinar as três técnicas (APF, SNAP e UST) para estimar o esforço de projetos de Big Data dentro da organização.

### **Formas de contato com a organização:**

- Reunião com gestores e responsáveis, com assinatura de termo de confidencialidade entre as partes e indicação dos recursos participantes;
- E-mail de interesse aos possíveis participantes do estudo.

**Agendamento e duração:**

- Acerto das datas e durações de treinamentos e execução do estudo de caso com gestores e responsáveis;
- Envio do e-mail com 2 semanas de antecedência ao possível participante, solicitando retorno com, no máximo, 1 semana de antecedência à data combinada.

**Equipamentos:**

Computador e ferramentas de análise qualitativa dos dados das entrevistas, questionários e da execução dos cenários.

**Instrumentos de Pesquisa:**

- Questionários com os participantes após treinamento e execução dos cenários do estudo de caso;
- Entrevistas com gestores e responsáveis após a realização do estudo de caso;
- Observação dos participantes;
- Registro dos requisitos funcionais e não funcionais dos projetos;
- Catálogo de Serviços Misto para contagens UST;
- Medições dos projetos utilizando APF, SNAP e UST;

**Ameaças a Validade da Pesquisa:**

- Ameaças de validade de conclusão
  - Estão relacionadas à habilidade de chegar a uma conclusão correta a respeito dos relacionamentos entre o tratamento e o resultado do estudo.
- Ameaças de validade interna
  - Verificam se o relacionamento observado entre o tratamento e o resultado do estudo é casual e não é influenciado por outro fator.
- Ameaças de validade externa
  - Definem condições que limitam a habilidade de generalizar os resultados de um estudo para outros contextos.
- Ameaças de validade de construção
  - Consideram os relacionamentos entre o que foi dito no planejamento e o que foi observado no estudo feito.