**CENTRO UNIVERSITÁRIO DE JOÃO PESSOA – UNIPÊ**

**PRÓ-REITORIA ACADÊMICA – PROAC**

**CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

NITAI CHARAN ÁLVARES PEREIRA

CONSULTAS UTILIZANDO LINGUAGEM NATURAL NO SERVIDOR DE BUSCA DISTRIBUÍDO ELASTICSEARCH

JOÃO PESSOA – PB

2019

NITAI CHARAN ÁLVARES PEREIRA

CONSULTAS UTILIZANDO LINGUAGEM NATURAL NO SERVIDOR DE BUSCA DISTRIBUÍDO ELASTICSEARCH

Monografia apresentada ao Curso de Bacharelado em Ciência da Computação do Centro Universitário de João Pessoa - UNIPÊ, como pré-requisito para a obtenção do grau de Bacharel em Ciência da Computação, sob orientação do Prof. MS.c. Fábio Falcão de França.

JOÃO PESSOA – PB

2019

NITAI CHARAN ÁLVARES PEREIRA

CONSULTAS UTILIZANDO LINGUAGEM NATURAL NO SERVIDOR DE BUSCA DISTRIBUÍDO ELASTICSEARCH

Monografia apresentada ao Curso de Bacharelado em Ciência da Computação do Centro Universitário de João Pessoa – UNIPÊ, como pré-requisito para a obtenção do grau de Bacharel em Ciência da Computação, apreciada pela Banca Examinadora composta pelos seguintes membros.:

Aprovada em \_\_\_\_\_\_/\_\_\_\_\_\_\_/2018.

**BANCA EXAMINADORA**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Prof. (título ex.: Dr./Ms./Msc./Esp.) Nome do Orientador (a) (UNIPÊ)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Prof. (título ex.: Dr./Ms./Esp.) Nome do professor Examinador (a) (UNIPÊ)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Prof. (título ex.: Dr./Ms./Esp.) Nome do professor Examinador (a) (UNIPÊ)

**AGRADECIMENTOS**

Gostaria de agradecer minha avó Carminha (in memorian) e primo Luiz Felipe Abreu (in memorian).

Gostaria de agradecer aos meu pais Mathura Pati e Subala Das por sempre prezar e batalhar pela educação dos seu filhos, sem medir esforços para que todos pudesse se graduar e obter a formação acadêmica necessária.

Meus agradecimentos também a minhas irmã Jahnavi Caran por me apoiar, dar suporte e aconselhar na minha trajetória acadêmica. Meus sinceros agradecimentos e minha eterna gratidão, irmã.

Dedico este trabalho ao meu professor orientador MSc. Fábio Falcão de França por me acolher e aconselhar passando seus conhecimentos sem pestanejos e me aceitando como último orientando antes de sua jornada em busca do título de doutorado. Dedico também ao professor MSc. Hugo Vieira Lucena de Souza por todos seus conselhos, ensinamentos e correções que foram possível para conclusão deste título. Também dedico este trabalho e ao professor DSc. Erick France Meira de Souza por sua amizade, suporte e preciosos conselhos dados.

A todos os professores e pessoas não citadas diretamente neste texto, mas que participaram e contribuíram de alguma forma com a minha formação, meu muito obrigado.

**RESUMO**

<Copo Resumo>

**Palavras-Chave:** <Palavra chave>; <Palavra Chave>.

**ABSTRACT**

<Abstract boddy>

**Keywords:** <Keyword>; <Keyword>.

**LISTA DE ILUSTRAÇÕES**

Figura 1 – <Nome figura> .................................................................. <Página>

**LISTA DE QUADROS**

Quadro 1 – <Nome quadro> ............................ <Página>

**LISTA DE ABREVIATURAS E SIGLAS**

**<ABREVIATURAS/SIGLAS> –** <Nome>

**SUMÁRIO**

[**1 INTRODUÇÃO**](#_vswvcrkcs613) **12**

[1.1 JUSTIFICATIVA](#_hut1ovtmlp4f) 13

[1.2 OBJETIVO GERAL](#_1t3h5sf) 14

[1.3 OBJETIVOS ESPECÍFICOS](#_4d34og8) 14

[1.4 METODOLOGIA](#_2s8eyo1) 14

[ORGANIZAÇÃO DO TRABALHO](#_y128yvwdm9bv) 15

# 1 INTRODUÇÃO

A barreira do aprendizado para a interação entre homem e sistemas computacionais até os dias atuais ainda é um obstáculo a ser conquistado devido a necessidade de tempo e interesse por parte dos usuário. (BARBOSA; SILVA, 2010) afirmam que usuários que se dispõem a aprender novos sistemas interativos com características únicas e distintas precisam dispor de tempo e interesse para posteriormente ser capazes de usufruir e usar as funcionalidades deste sistema.

Consequentemente, pessoas que não possuem hábito de utilização da internet e sem muito conhecimento técnico em manusear softwares de acesso podem deixar de participar de eventos, oportunidades e movimentos que utilizam como meio a internet. Pessoas com estas características podem se ausentar da participação no controle sobre a Administração Pública atuando na verificação, acompanhamento e fiscalização da regularidade de gastos públicos devido a se depararem com barreiras até mesmo no momento do recolhimento dos dados públicos que utilizam como canal mediador a internet.

No que se refere a fiscalização, com a promulgação da Constituição de 1988, o Brasil se classificou se como um Estado Democŕatico de Direito onde uma de suas características, para este tipo de estado, é a participação popular no controle sobre a Administração Pública. Neste sentido, o cidadão tem o poder de acompanhar e fiscalizar a regularidade dos atos governamentais juntamente com os órgãos institucionais legalmente criados para esta finalidade (ARRUDA; TALES, 2010).

Tendo em vista que em 2000 foi promulgada a Lei de Responsabilidade Fiscal, que entre muito outros pontos, no artigo 48, definiu que as prestações de contas e outros instrumentos de transparência da gestão fiscal pública devem ter ampla divulgação, inclusive em meios eletrônicos de acesso público. Porém, todo benefício gerado a partir das regras impostas por esta lei, podem não ser devidamente aproveitadas por cidadãos que não possuem conhecimento prévio na utilização da internet.

Surgindo assim a pergunta central que este trabalho se propõe a intervir: como facilitar o acesso às informações públicas com o propósito que pessoas, sem conhecimento técnico em sistemas computacionais, possam auxiliar no processo de fiscalização de gastos governamentais, ajudando assim a diminuir os índices de corrupção no país?

## 1.1 JUSTIFICATIVA

A partir da promulgação da Lei de Acesso à Informação, Lei nº 12.527/2011, foi garantido o direito constitucional de acesso às informações públicas, possibilitando que qualquer pessoa física ou jurídica, sem a necessidade de apresentar motivos, receba informações públicas de órgãos e entidades. Desta forma, surgem novos mecanismos os quais possibilitam a participação do cidadão na fiscalização de gastos públicos e combate à corrupção.

A melhoria do acesso à informação pública e a criação de regras que permitem a disseminação de informações produzidas pelo governo reduzem os abusos que podem ser cometidos (STIGLITZ, 2002). Porém, tais dados estão mais disponíveis em sistemas computacionais mediante a internet, utilizando-se muitas vezes de ferramentas carentes de parametrização ou disponibilização dos dados em grandes planilhas, dificultando assim o entendimento por parte de não-especialistas em informática.

Na tentativa de melhorar a divulgação dos dados públicos, o Ministério da Transparência em conjunto com a Controladoria-Geral da União, em 2004, criou o site *Portal da Transparência do Governo Federal* que possibilita o acesso livre, no qual o cidadão pode encontrar informações sobre como o dinheiro público é utilizado, além de se informar sobre assuntos relacionados à gestão pública do Brasil. Ainda assim, o problema de acesso por pessoas sem muitos conhecimentos técnicos sobre internet e como utilizá-la continua a existir.

Como forma de intervenção, este trabalho sugere a utilização de linguagem natural para busca e manipulação dos dados em bases de dados disponibilizadas pelo Governo Federal. “O processamento de linguagem natural permite ao computador compreender e reagir a declarações e comandos de voz realizados em uma linguagem natural” (STAIRS; REYNOLDS , 2006, p. 508). Para tal, este trabalho propõe a utilização do serviço de API de reconhecimento vocal *Language Understanding Intelligent Service* (LUIS) que aplica inteligência de aprendizado personalizado de máquina a um texto de linguagem natural e extrai informações relevantes a futuras aplicações.

Para isso, este trabalho também propõe a utilização da ferramenta *Elasticsearch* que, entre muitas funcionalidades, possui recursos e tecnologias que permitem realizar consultas através de índices em grandes volumes de dados em tempo real (GORMLEY; TONG, 2015).

## 1.2 OBJETIVO GERAL

Desenvolver uma solução de software que utiliza *Graphical User Interface* (GUI) para aumentar a usabilidade a pessoas leigas. Recebe como entrada uma solicitação de consulta a dados públicos, em linguagem natural, e é respondia com ou sem os dados solicitados.

## 1.3 OBJETIVOS ESPECÍFICOS

* Realizar um estudo acerca das tecnologias utilizadas para a criação de um sistema de interação descrevendo suas características e possibilidades possíveis de utilização para se atingir o objetivo geral.
* Gerar protótipos de baixa fidelidade para os testes de usabilidade com os requisitos impostas na concepção do sistema.
* Avaliar a efetividade da aplicação com a realização de um estudo de caso uso solicitando a atores escolhidos de forma aleatória tentando buscar dados sobre gastos públicos sem e com a solução de software posposto.

## 1.4 METODOLOGIA

“Método é o conjunto das atividades sistemáticas e racionais que, com maior segurança e economia, permite alcançar o objetivo de produzir conhecimentos válidos e verdadeiros, traçando o caminho a ser seguido, detectando erros e auxiliando as decisões” (LAKATOS; MARCONE, p. 79, 2019). As autoras descrevem também que a ciência caracteriza-se pela utilização de métodos científicos e que estes não são de uso exclusivo pela ciência. Porém, não é possível a ciência estar apartado do emprego de métodos científicos (LAKATOS; MARCONE, 2019).

Com o objetivo de facilitar análise e avaliações futuras, este trabalho é dividido sobre a características dos tipos de pesquisas que foram aplicada do processo de construção deste trabalho.

Este trabalho se classifica como pesquisa exploratória pois entre outros fatores é tentado proporcionar maior familiaridade com as ferramentas utilizadas pela solução de software proposta. Realizando para isso, levantamento bibliográficos com o propósito de explicar e expor possibilidades proporcionadas por estas ferramentas. Como (GIL, p. 2018) explana que o levantamento bibliográfico utilizada pelas pesquisas exploratórias é uma das maneiras utilizadas para a coleta de dados relevantes.

Este trabalho também se classifica como uma pesquisa aplicada devido ao propósito ser uma possível abordagem ao problema identificado da dificuldade de acesso às informações públicas em que pessoas, sem conhecimento técnico em sistemas computacionais, possam posteriormente auxiliar no processo de fiscalização de gastos governamentais. (GIL, p. 25 2018) “pesquisa aplicada, abrange estudos elaborados com a finalidade de resolver problemas identificados no âmbito das sociedades em que os pesquisadores vivem”.

Com o intuito a avalizaçẽos futuras sobre a qualidade dos resultados mostrados neste trabalho, foi adotado a análise e interpretação dos dados expostos de forma qualitativa. Devido a descrição dos resultados serem em formas verbais e não em termos numéricos como proposto pelas pesquisas quantitativa. (GIL, p. 39 2018) “Nas pesquisas quantitativas os resultados são apresentadas em termos numéricos e, nas qualitativas, mediante descrições verbais”.

Os procedimentos adotados na análise, interpretação e coleta dos dados exposto neste trabalho são realizados de forma bibliográfica e experimental. A forma bibliográfica se dá pela referenciação de dados através de citações de materiais já publicados. Enquanto que a experimental se dá pela adoção de testes de caso de uso utilizado para demonstrar a eficácia do software proposto para intervenção ao problema.

ORGANIZAÇÃO DO TRABALHO

Após esse capítulo introdutório, o conteúdo deste trabalho organiza-se da seguinte forma:

No capítulo 2 irá conter um estudo a respeito da *Application Programming Interface* (API) da ferramenta utilizada *Language Understanding Intelligent Service* (LUIS).

O capítulo 3 tratará sobre o estudo da ferramenta Elasticsearch.

O capítulo 4 apresentará a metodologia baseada no design thinking, utilizada para criação do projeto, as fases e artefatos gerados que apoiam no desenvolvimento da aplicação produzida, assim como os resultados obtidos na validação.

**REFERÊNCIAS**

STIGLITZ, J. Transparency in government. In: WORLD BANK INSTITUTE. The right to tell: the role of mass media in economic development. Washington, 2002. Part.1, p. 27-44. http://books.google.com.br/books?id=56punueI7G0C&pg=PA22&hl=pt- BR&source=gbs\_toc\_r&cad=4#v=onepage&q&f=false

Acessado em: 5 mar. 2019

STAIR, R. M.; REYNOLDS, G. W. **Princípios de Sistemas de Informação**: 3. ed. São Paulo: Cengage Learning, 2015.

BARBOSA, S. D. .J; SILVA, B. S. **Interação Humano-Computador**: 11. ed. São Paulo: Elsevier Editora, 2010.

<https://books.google.com.br/books?hl=pt-BR&lr=&id=qk0skwr_cewC&oi=fnd&pg=PP2&dq=Intera%C3%A7%C3%A3o+Humano+Computador&ots=RNC25Ir2ks&sig=okHnPQAEoB93HYeN9NsmtyyTNB4#v=onepage&q=aprendizado&f=false>

Acessado em: 10 mar. 2019

ARRUDA, Ângela Furtado. TALES, José Silva. **A importância do controle social na fiscalização dos gastos públicos**: Volume 1, número 1, 2010. Disponível em: <http://institutoateneu.com.br/ojs/index.php/RRCF/article/view/3>

Acessado em: 27 mar. 2019

GORMLEY, Clinton. TONG, Zachary. Elasticsearch: **The Definitive Guide: A Distributed Real-Time Search and Analytics Engine**: 1. ed. Estados Unidos da América: O’Reilly Media, 2015. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=d19aBgAAQBAJ&oi=fnd&pg=PR3&dq=Elasticsearch&ots=NzhXsPzAgC&sig=vwBO86w9cdtmsJJpTeVjMxQT0tU#v=onepage&q=Elasticsearch%20real-time&f=false>

Acessado em: 27 mar. 2019

LAKATOS, Eva Maria; MARCONE, Marina de Andrade: **Fundamentos da metodologia científica**: 8. ed. São Paulo: Atlas S.A.: 2019.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**: 6 ed. São Paulo: Atlas, 2018.