

**CENTRO UNIVERSITÁRIO DE JOÃO PESSOA - UNIPÊ
PRÓ-REITORIA ACADÊMICA – PROAC
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

NITAI CHARAN ÁLVARES PEREIRA

**CONSULTAS UTILIZANDO LINGUAGEM NATURAL NO SERVIDOR DE BUSCA
DISTRIBUÍDO ELASTICSEARCH**

JOÃO PESSOA - PB

2019

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 da França.

JOÃO PESSOA - PB

2019

Nitai Charan Álvares Pereira

Consultas Utilizando Linguagem Natural no Servidor de Busca Distribuído Elastic-search/ Nitai Charan Álvares Pereira. – João Pessoa - PB, 2019-
27p. : il. (algumas color.) ; 30 cm.

Orientador: Fábio Falcão da França

Trabalho de Conclusão de Curso - TCC – Centro Universitário de João Pessoa - UNIPÊ, 2019.

1. Palavra-chave1. 2. Palavra-chave2. 2. Palavra-chave3. I. Orientador. II. Universidade xxx. III. Faculdade de xxx. IV. Título

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 da França.

Trabalho aprovado. João Pessoa - PB, 24 de novembro de 2012:

Fábio Falcão da França
Orientador

Professor
Convidado 1

Professor
Convidado 2

João Pessoa - PB
2019

Dedico este trabalho ao meu professor orientador MSc. Fábio Falcão da 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 Erick France Meira de Souza, da UFPB, por seus suporte, amizade e preciosos conselhos dados.

Agradecimentos

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

Gostaria de agradecer aos meu pais Mathura Pati Devi Das 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 minha irmã Jahnavi Caran por me apoiar, dar suporte e aconselhar na minha trajetória acadêmica. Meus sinceros agradecimentos e minha eterna gratidão, irmã.

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.

“Este trabalho é dedicado à criança a qual sempre falou que seria um cientista”

Resumo

Políticas públicas de transparência e fiscalização são essenciais para que o cidadão disponha de controle na fiscalização de gastos públicos. Porém, tais cidadãos certamente podem utilizar a *internet* como canal mediador entre as entidade portadora dos dados para fiscalização pública e o cidadão que atua como fiscal neste cenário. A *internet* acaba excluindo, assim, pessoas leigas em elementos com teor tecnológico.

Em vista disso, este trabalho propõe a criação de uma aplicação que auxilie na busca dos dados dados públicos persistidos em uma base de dados. Para facilitar a utilização, será utilizada a linguagem natural possibilitando a pessoas leigas em assuntos de conhecimentos tecnológicos, possam utilizar a aplicação e poder ajudar na fiscalização dos gastos governamentais.

Para o processamento da linguagem natural, será utilizado a suite de *software* Language Understanding Intelligent Service (LUIS) é uma suite de software pertencente a um dos Serviços Cognitivos da Microsoft que utiliza um ramo da Inteligência Artificial (IA) chamado Processamento de Linguagem Natural (PLN) (MAYO, 2017). Essa suite, possibilita que aplicações possam adquirir certo grau de inteligência. Formando assim, seu objetivo central

Palavras-chave: latex. abntex. editoração de texto.

1. Tecnologia no problema
2. Apresentar limitação das aplicações existentes
3. Mostrar que não há aplicações direcionadas a este propósito
4. Mostrar aplicação (solução) que ajudará no problema apresentado
5. Falar sobre a metodologia
6. Falar sobre o desenvolvimento da aplicação seguindo a metodologia

Construção do Resumo

- Indicativo sobre o trabalho e não sobre o conteúdo do trabalho
- Entre 150 e 500 palavras
- Não usar abreviações, símbolos, formulas, equações ou diagramas a não ser que seja estritamente indispensáveis
- Não usar citações
- Não faça mais de um paragrafo
- Não crie tópicos
- Sem frases negativas
- Após o título do trabalho e nome dos autores
- Primeira frase indica o tema do trabalho
- Informar os objetivos geral e específicos, metodologia, resultados
- 3 a 5 palavras chaves que represente a essência do trabalho separadas por pontos

Partes do Resumo

- Introdução
Objetivos e relevância do trabalho do trabalho.
- Metodologia
Elementos essenciais do método de pesquisa
- Resultados
Principais resultados ou de maior destaque que pode ser representativo da pesquisa
- Discussão
- Conclusão
Contribuição e limites do trabalho

Normas ABNT resumo

- Frases afirmativas, curtas, escritas em voz ativa e na terceira pessoa do singular
- Texto claro, conciso, seguir uma ordem lógica e sem completamente fiel ao trabalho

- Deve mencionar o tema, objetivo, métodos, resultados e conclusões
- Deve indicar, no fim, palavras chaves

Lista de ilustrações

Figura 1 – JSON - Resposta da chamada a API. Comando: curl http://localhost:9200/?pretty	18
Figura 2 – Diagrama de Atividade	22
Figura 3 – Fluxo de funcionamento da aplicação	25

Lista de abreviaturas e siglas

API	Application Programming Interface
LUIS	Language Understanding Intelligent Service
JSON	JavaScript Object Notation
IA	Inteligência Artificial
PLN	Processamento de Linguagem Natural
REST	Representational State Transfer

Sumário

1	INTRODUÇÃO	13
1.1	Justificativa	13
1.2	Objetivos	14
1.2.1	Objetivo Geral	14
1.2.2	Objetivos Específicos	14
1.3	Metodologia	15
1.4	Organização do Trabalho	16
2	CONJUNTOS DE TECNOLOGIAS UTILIZADAS	17
2.1	Estudo a respeito do Elasticsearch	17
2.2	Inteligência Artificial	18
2.2.1	Processamento de Linguagem Natural (PLN)	18
2.3	Estudo a respeito da Language Understanding Intelligent Service	19
3	ARQUITETURA DA SOLUÇÃO	21
3.1	Diagrama de Atividades	21
3.2	Fluxo de funcionamento da aplicação	22
	REFERÊNCIAS	26

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 à 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 Democrático 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; TELES, 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.

Surge assim a pergunta central na qual 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? Este é o problema que será trado no decorrer deste trabalho.

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 (ISLAM; DJANKOV, 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” (STAIR; REYNOLDS, 2015, 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 (GIL, 2010).

1.2 Objetivos

1.2.1 Objetivo Geral

Desenvolver uma ferramenta que utiliza processamento de linguagem natural, interface gráfica e busca em índices a fim de facilitar o acesso aos dados governamentais públicos disponíveis.

1.2.2 Objetivos Específicos

Realizar um estudo acerca dos conjuntos de tecnologias utilizadas para a criação de um sistema de interação descrevendo suas características e possibilidades possíveis de utilização para se atingir o objetivo geral.

Replicar uma base de dados públicos governamentais a fim de indexar as informações localmente;

Gerar o treinamento do algoritmo do LUIS, produto da Microsoft, a fim de entender e diferenciar ações diversas a serem realizadas.

Gerar versão de teste, utilizando o texto de entrada no algoritmo do LUIS, e gerando consultas parametrizadas e customizadas para retornar índices da base de dados replicada.

Validar o treinamento, a interface de entrada e os objetos retornados na busca indexada, a fim de alinhar-se com o objetivo geral deste trabalho.

1.3 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; MARCONI, 2019, p. 79). 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; MARCONI, 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, 2010) 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, 2010, p. 25) “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 avaliações 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, 2010, p. 39) “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.

1.4 Organização do Trabalho

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

- Capítulo <n> - DESENVOLVIMENTO DA SOLUÇÃO: apresentará a ferramenta proposta, assim como arquitetura de software, diagramas, mockups de telas e demais artefatos que contribuíram com o trabalho;
- Capítulo <n> - <TÍTULO>: apresentará quais teorias e respectivos autores mais contribuíram para a realização do estudo e as bases teóricas para a realização deste trabalho;
- Capítulo <n> - <TÍTULO>: apresentará os conceitos e as pesquisas que fundamentam a realização deste trabalho;
- Capítulo 2 - ESTUDO A RESPEITO DA API ELASTICSEARCH: apresentará um estudo detalhado a respeito da ferramenta utilizada para buscar, gravar e extrair informações de forma rápida e indexada;
- Capítulo 3 - ESTUDO A RESPEITO DA LANGUAGE UNDERSTANDING INTELLIGENT SERVICE (LUIS): apresentará um estudo detalhado a respeito da ferramenta de reconhecimento textual;
- Capítulo 4 - CONCLUSÃO: finaliza o trabalho apresentando os resultados obtidos, formas de utilizar o modelo para entender o desempenho dos alunos, discutindo os possíveis trabalhos futuros.

2 Conjuntos de tecnologias utilizadas

Neste capítulo é mostrada uma proposta de abordagem ao problema referido no início da introdução, através de uma aplicação a qual tem como objetivo facilitar o acesso a dados públicos através de linguagem natural.

2.1 Estudo a respeito do Elasticsearch

Para busca e persistência dos dados será utilizada a tecnologia *Elasticsearch*, que se caracteriza por ser um conjunto de mecanismo de busca *open-source* que realiza buscas e analisa dados em tempo real (GIL, 2010). Através disso, são criadas soluções próprias que utilizam as facilidades trazidas pelo conjunto para atingir seus objetivos.

Tendo como principal objetivo facilitar a busca por texto e a rapidez no acesso aos dados e gravação, o *Elasticsearch* possui a capacidade de funcionamento escalável, sendo possível a utilização em apenas um servidor (*standalone*) ou de forma distribuída em centenas de servidores (GORMLEY; TONG, 2015), agilizando assim a entrega ao usuário ou aplicação que solicitar os dados persistidos.

Para solicitar os dados persistidos é disponibilizada uma interface *API RESTful* que possibilita a obtenção e visualização dos dados por aplicações *web client* (GORMLEY; TONG, 2015). Esta funcionalidade disponibiliza a comunicação entre o *Elasticsearch* e outras aplicações independentemente de como foram feitas ou da linguagem utilizada na construção das aplicações *web clients*. Gera-se, assim, independência e maior usabilidade.

Possuindo como base o *Apach Lucene*, o qual se qualifica por ser atualmente a livreria de motor de busca mais avançada, performática e completa existente, o *Elasticsearch* é construído na linguagem Java para poder integrar e usufruir dessa livreria. Ao contrário, é necessário um grande volume de conhecimento para compreender como são realizadas as facilidades trazidas pelo *Lucene* devido a sua complexidade (GORMLEY; TONG, 2015).

Um dos maiores benefícios trazidos pela utilização do *Apach Lucene* é a ação de indexação. Utilizando isso, o *Elasticsearch* realiza busca, ordena, filtra e persiste dados fornecidos como objetos e documentos realizando a indexação em todas as suas ações. Consequentemente, esta tecnologia também é caracterizada como orientada a documento (GORMLEY; TONG, 2015).

Um documento é um dado constituído por campos os quais pode se repetir várias vezes. Todos os campos possuem um tipo como texto, numérico, data etc., ou tipos mais complexos como subdocumentos ou *arrays*. Para cada documento persistido é salvo um pequeno título, data de publicação e um *link* de acesso ao conteúdo. O ato de se persistir um documento é chamado

de indexação e o nome do dado persistido é chamado de índice (KUC; ROGOZINSKI, 2013).

Após a criação desta estrutura, é possível realizar pesquisas através de temas. Agiliza-se, assim, a obtenção dos dados desejados. A disponibilização dos dados solicitados ocorre através de uma simples *API RESTful* mediante a transformação e transporte dos dados no formato *JavaScript Object Notation*(JSON) como exemplificado na Figura 1.

Figura 1 – JSON - Resposta da chamada a API. Comando: curl http://localhost:9200/?pretty

```
{
  "name" : "XsyRmrG",
  "cluster_name" : "elasticsearch",
  "cluster_uuid" : "4sw9BMjIS_Sx_iqYfWfIRg",
  "version" : {
    "number" : "6.6.0",
    "build_flavor" : "oss",
    "build_type" : "zip",
    "build_hash" : "a9861f4",
    "build_date" : "2019-01-24T11:27:09.439740Z",
    "build_snapshot" : false,
    "lucene_version" : "7.6.0",
    "minimum_wire_compatibility_version" : "5.6.0",
    "minimum_index_compatibility_version" : "5.0.0"
  },
  "tagline" : "You Know, for Search"
}
```

Fonte: Autor

2.2 Inteligência Artificial

Inteligência Artificial é uma das ramificações da ciências computacionais que se dedica ao estudo de tentar simular a inteligência humana em máquinas, desenvolver a inteligência artificial de máquinas e softwares que podem raciocinar, aprender, reunir conhecimento, comunicar, manipular e perceber objetos. Inteligência é comumente considerado como a capacidade de resolver problemas complexos realizando previamente coleta de informações e conhecimentos (PANNU, 2015).

2.2.1 Processamento de Linguagem Natural (PLN)

Processamento de Linguagem Natural (PLN) é uma subárea da Inteligência Artificial (IA) que estuda as limitações, problemas e compreensão da linguagem humana para resoluções de

problemas solúveis através de uma inteligência artificial. Seu objetivo principal é possibilitar máquinas entenderem e interpretar a linguagem humana. Tal processo converte dados em texto em formatos que máquinas entendam como o numérico e binário (KULKARNI; SHIVANANDA, 2019).

Este processo é baseado em quatro etapas para extrair informações pertinentes do texto alvo. Primeiro o texto é realizado um processo de criação de *tokens* das palavras do texto. Nesta etapa o texto é quebrado em unidades, sem se preocupar no significado ou classificação gramatical das palavras. Formando assim uma lista de *tokens* que será usada em todos os outros processos com seus diferentes objetivos (REESE, 2015).

A segunda etapa é dedicada a detecção de sentenças utilizando unidades criadas na primeira etapa. A combinação de palavras em uma frase ou uma sentença pode ter diferentes significados mudando o relacionamento entre outras palavras ou outras sentenças. Pode ser dado como exemplo que, frequente é distinguido a função gramatical entre as palavras como substantivo e verbos (REESE, 2015).

A terceira é o processo de classificar cada palavra e documentos que consiste na rotulação destas informações encontradas no texto. No processo de rotulação, o rótulo pode ou não existir antes desta etapa. Caso exista, esta etapa é chamada de classificação. Antagônico, é chamado de agrupamento (REESE, 2015).

A quarta etapa é o processo de extração de relacionamentos identificado no texto. Esta etapa é útil para uma série de tarefas, incluindo a construção de bases de conhecimento, análise de tendências, coleta de inteligência e realização de pesquisas de produtos (REESE, 2015).

2.3 Estudo a respeito da Language Understanding Intelligent Service

Language Understanding Intelligent Service (LUIS) é uma suite de *software* pertencente a um dos Serviços Cognitivos da Microsoft que utiliza um ramo da Inteligência Artificial (IA) chamado Processamento de Linguagem Natural (PLN) (MAYO, 2017). Essa suite, possibilita que aplicações possam adquirir certo grau de inteligência. Formando assim, seu objetivo central da utilização.

O objetivo de utilização desta ferramenta é possibilitar a tradução de linguagem natural em texto plano para que programas de computadores possam entender e interagir utilizando tais dados. Para isso, a API extrai de uma sentença do tipo texto em linguagem natural as intenções e entidades, retornando após objetos no formato JavaScript Object Notation (JSON) (MAYO, 2017).

Previamente, é necessário a criação de modelos de linguagens para extração das entidades e intenções. Qualquer objeto, seja ele abstrato ou concreto, podem ser caracterizados como

entidades. Exemplos como lugares, tempo e números podem ser definidos como entidades alvo. Consequentemente possibilitando o reconhecimento de comandos dos usuários em entradas como frases ou sentenças em linguagem natural (LARSEN, 2017).

Porém, a LUIS não é uma aplicação cliente que deve conversar e solicitar dados ao cliente. LUIS possibilita para que outras aplicações como *chatbots*, *websites*, aplicativos móveis ou aplicações *desktops* possam se comunicar mediante a *internet* via API. Servido como um tradutor à aplicações se comunicarem e entenderem os desejos dos usuários.

3 Arquitetura da Solução

Este capítulo dedica-se a explicar como a aplicação proposta para a intervenção ao problema, anteriormente descrito, executa suas ações e descreve o fluxo de ação entre atividades e interação com outras aplicações. Para isso, é mostrado os fluxos e atividades de suas ações em figuras e descrito detalhadamente o que pode ser extraído das imagens.

3.1 Diagrama de Atividades

O intuito da utilização do Diagrama de Atividades é dar ênfase ao fluxo de controle na execução de um comportamento realizado por um sistema, mostrando o fluxo entre atividades em um sistemas. As atividades podem ser referidas como um fluxo sequencial ou ramificação de atividades que interagem entre si e outros objetos para realizar ou sofrer ações. Descrevendo assim, uma modelagens a função do sistema ([BOOCH; JACOBSON; JACOBSON, 2012](#)).

Ao iniciar aplicação é apresentada a tela inicial contendo um campo para entrada de dados para a pesquisa e botões de escolha. Este campo é utilizado para a entrada da sentença em linguagem natural para busca do usuário que esta utilizando a aplicação. Após a inserção da sentença, é necessário a escolha da próxima ação do usuário. A ação de pesquisar e a de limpar o campo da entrada.

Ao clicar no botão de limpar, é reapresentado a tela inicial contendo novamente os botões de escolha, juntamente com o campo para a entrada para a busca. Caso o botão de pesquisar for acionado, é iniciado o processo de busca dos dados solicitado ao usuário e retornado ao mesmo caso forem encontrados registros na base de dados onde-se encontra estes dados persistidos.

Todo este fluxo e atividades são desmostrado na [Figura 2](#) com o intuito de facilitar o esclarecimento e melhorar o entendimento a respeito da aplicação proposta.

Figura 2 – Diagrama de Atividade



Fonte: Autor

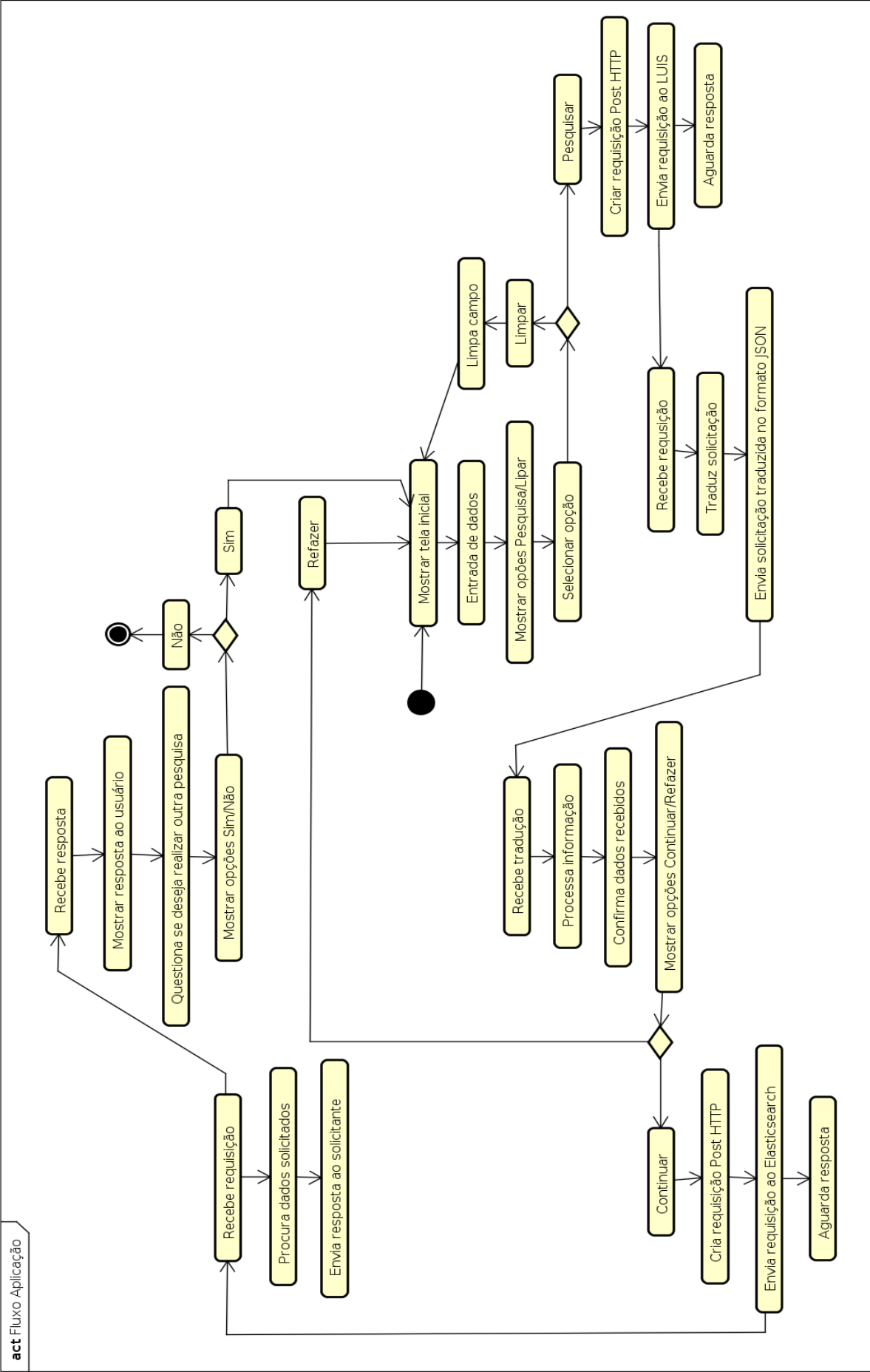
3.2 Fluxo de funcionamento da aplicação

Esta sessão dedica-se a explicar mais detalhadamente todo o fluxo da aplicação proposta. Usando como base a [Figura 3](#), é descrito cada atividade contido no formato: - <Nome atividade>: Detalhamento.

- **Mostrar tela inicial** - A aplicação renderiza tela inicial contendo o campo para entrada da busca e os botões de escolha Limpar e Pesquisar.
- **Entrada de dados** - É aguardado a entrada da sentença em linguagem natural solicitado pelo usuário.
- **Selecionar opção** - O usuário deve escolhe entre as opções Limpar campo e Pesquisar.
- **Limpar** - Caso o usuário escolha a opção Limpar, é limpada o campo e novamente é renderizado a tela inicial para entrada de nova busca.
- **Pesquisar** - Caso o usuário escolha a opção Pesquisar, a aplicação cria uma requisição HTTP do tipo Post contendo em seu corpo a sentença do usuário, para o endereço da API LUIS.
- **Envia requisição ao LUIS** - Após, o envio é realizado e a aplicação aguarda o retorno da solicitação.
- **Recebe requisição** - A API LUIS recebe a requisição e processa a sentença do usuário.
- **Envia solicitação traduzida no formato JSON** - Após o processamento, a API responde a solicitação traduzindo as entidades e intenções identificadas na sentença.
- **Recebe tradução** - Recebe a tradução da sentença no formato JSON.
- **Confirma dados recebidos** - A aplicação solicita confirmação das intenção e os atores indenticado pelo LUIS estão corretas.
- **Mostrar opções Continuar e Refazer** - Para confirmação, o usuário deve escolher as opções Continuar ou Refazer para dar continuidade no fluxo.
- **Refazer** - Caso escolha Refazer, é novamente mostrada a tela inicial dando a possibilidade que o usuário refaça a sentença, para posteriormente refazer novamente tradução.
- **Continuar** - Se a opção continuar for escolhida, a aplicação dará continuidade ao fluxo para busca dos dados.
- **Envia requisição ao Elasticsearch** - Posteriormente, a aplicação encaminha a requisição de busca dos dados na base da API Elasticsearch e aguarda o retorno da requisição.
- **Recebe requisição** - A API Elasticsearch recebe a requisição da aplicação e realiza o a procura pelos dados solicitados e retorna uma resposta caso os dados são encontrados ou não.
- **Recebe a resposta** - A aplicação recebe a resposta contendo os dados ou não no formato JSON.

- Mostra resposta ao usuário - Aplicação mostra os dados ao usuário e questiona se deseja fazer uma nova busca.
- Mostra opção Sim/Não - É mostrada as opções Sim e Não em forma de botões interativos o qual o usuário pode escolher continuar a realizar buscas por dados ou finalizar aplicação.
- Sim - Caso opção Sim for escolhida, a aplicação mostra a tela inicial possibilitando ao usuário realizar novas buscas.
- Não - Caso a opção escolhida for Não, a aplicação é encerrada.

Figura 3 – Fluxo de funcionamento da aplicação



Referências

- ARRUDA, Â. F.; TELES, J. S. A importância do controle social na fiscalização dos gastos públicos. *Revista Razão Contábil & Finanças*, v. 1, n. 1, 2010. Disponível em: <http://bit.ly/2H9qNph>. Acesso em: 27 mar. 2019. Citado na página 13.
- BARBOSA, S.; SILVA, B. *Interação humano-computador*. 11. ed. Elsevier Brasil, 2010. Disponível em: <http://bit.ly/2MdmcrQ>. Acesso em: 10 mar. 2019. Citado na página 13.
- BOOCH, G.; JACOBSON, I.; JACOBSON, R. *UML: Guia do usuário*. 12. ed. Rio de Janeiro, Brasil: Elsevier Editora Ltda., 2012. ISSN 978-85-352-8565-9. ISBN 978-85-352-1784-1. Disponível em: https://books.google.com.br/books?id=uGhaDwAAQBAJ&hl=pt-BR&source=gbs_navlinks_s. Acesso em: 2 jun. 2019. Citado na página 21.
- GIL, A. C. *Como elaborar projetos de pesquisa*. 6. ed. [S.l.]: Atlas, 2010. Citado 3 vezes nas páginas 14, 15 e 17.
- GORMLEY, C.; TONG, Z. *Elasticsearch: The Definitive Guide: A distributed real-time search and analytics engine*. 1. ed. O'Reilly Media, Inc., 2015. Disponível em: <http://bit.ly/2ExpWxW>. Acesso em: 27 mar. 2019. Citado na página 17.
- ISLAM, R.; DJANKOV, S. *The right to tell: The role of mass media in economic development*. 1. ed. The World Bank, 2002. Disponível em: <http://bit.ly/2K40dkj>. Acesso em: 5 mar. 2019. Citado na página 14.
- KUC, R.; ROGOZINSKI, M. *Elasticsearch server: A practical guide to building fast, scalable, and flexible search solutions with clear and easyto understand examples*. 2. ed. Packt Publishing Ltd, 2013. Disponível em: <http://bit.ly/2HyJ54G>. Acesso em: 23 mai. 2019. Citado na página 18.
- KULKARNI, A.; SHIVANANDA, A. *Natural Language Processing Recipes: Unlocking text data with machine learning and deep learning using python*. 1. ed. Apress, 2019. Disponível em: <http://bit.ly/2M6DfeZ>. Acesso em: 15 mai. 2019. Citado na página 19.
- LAKATOS, E. M.; MARCONI, M. d. A. *Fundamentos de metodologia científica*. 8. ed. [S.l.]: Editora Atlas S.A., 2019. Citado na página 15.
- LARSEN, L. H. *Learning Microsoft Cognitive Services*. 1. ed. Birmingham, Reino Unido: Packt Publishing, 2017. 137 p. (Série do livro, 15). Bibliografia: p. 131–132. ISBN 9781786460592. Disponível em: <https://books.google.com.br/books?id=eLkrDwAAQBAJ>. Citado na página 20.
- MAYO, J. *Programming the Microsoft Bot Framework: A multiplatform approach to building chatbots*. 1. ed. Microsoft Press, 2017. (Developer Reference Series). Disponível em: <http://bit.ly/2LDlyUi>. Acesso em: 5 mai. 2019. Citado na página 19.
- PANNU, A. Artificial intelligence and its application in different areas. *Artificial Intelligence*, v. 4, n. 10, p. 79–84, 2015. Disponível em: <http://bit.ly/2vXAp1i>. Acesso em: 13 mai. 2019. Citado na página 18.

REESE, R. M. *Natural language processing with Java*: Explore various approaches to organize and extract useful text from unstructured data using java. 1. ed. Packt Publishing Ltd, 2015. Disponível em: <<http://bit.ly/2VMSslf>>. Acesso em: 18 mai. 2019. Citado na página 19.

STAIR, R. M.; REYNOLDS, G. W. *Princípios de sistemas de informação*: Uma abordagem gerencial. 1. ed. [S.l.]: Cengage Learning, 2015. Citado na página 14.