

**UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI**

**Programa de Pós-graduação em Educação**

**Jésyka Milleny Azevedo Gonçalves**

**UMA PROPOSTA DE *DESIGN* DE INTERAÇÃO PARA UMA FERRAMENTA DE  
COLETA E MINERAÇÃO DE TEXTOS EM REDES SOCIAIS *ONLINE* PARA FINS  
DE PESQUISA CIENTÍFICA**

**Diamantina  
2019**



**Jésyka Milleny Azevedo Gonçalves**

**UMA PROPOSTA DE *DESIGN* DE INTERAÇÃO PARA UMA FERRAMENTA DE  
COLETA E MINERAÇÃO DE TEXTOS EM REDES SOCIAIS *ONLINE* PARA FINS  
DE PESQUISA CIENTÍFICA**

Dissertação apresentada ao Programa de Pós-Graduação em Educação da Universidade Federal dos Vales do Jequitinhonha e Mucuri, como requisito para obtenção do título de Mestre.

Orientadora: Profa. Dra. Maria Lúcia Bento Villela  
Coorientador: Prof. Dr. Marcus Vinicius Carvalho  
Guelpeli

**Diamantina  
2019**

Elaborado com os dados fornecidos pelo(a) autor(a).

G643p

Gonçalves, Jesyka Milleny de Azevedo.

Uma proposta de Design de Interação para uma ferramenta de coleta e mineração de textos em redes sociais online para fins de pesquisa científica / Jesyka Milleny de Azevedo Gonçalves, 2020.

147. : il.

Orientadora: Maria Lúcia Bento Villela

Dissertação (Mestrado – Programa de Pós-Graduação em Educação) - Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, 2020.

1. Mineração de textos. 2. Redes sociais. 3. Framework Oráculo. 4. Design da interface. I. Villela, Maria Lúcia Bento. II. Título. III. Universidade Federal dos Vales do Jequitinhonha e Mucuri.

**CDD 005.74**

JESYKA MILLENY AZEVEDO GONÇALVES

**Uma proposta de Design de Interação para uma ferramenta de coleta e mineração de textos em redes sociais online para fins de pesquisa científica**

Dissertação apresentada ao MESTRADO EM EDUCAÇÃO, nível de MESTRADO como parte dos requisitos para a obtenção do título de MESTRA EM EDUCAÇÃO.


Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Maria Lúcia Bento Villela

Coorientador: Prof. Dr. Marcus Vinicius Carvalho Guelpeli

Data da aprovação: 10/12/2019

  
Prof<sup>a</sup>. Dr<sup>a</sup>. MARIA LUCIA BENTO VILLELA - UFVJM

  
Prof<sup>a</sup>. Dr<sup>a</sup>. CAROLINE QUEIROZ SANTOS - UFVJM

  
Prof<sup>a</sup>. MSc.<sup>a</sup> GLÍVIA ANGÉLICA RODRIGUES BARBOSA - CEFET

DIAMANTINA





MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI  
DIAMANTINA - MINAS GERAIS  
PRO-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO

**ATESTADO DE DEFESA POR VIDEOCONFERÊNCIA**

Atesto para os devidos fins que no dia 10 de dezembro de 2020, às 13h, nas dependências da UFVJM – em Diamantina - MG, foi realizada a defesa de dissertação da discente Jésyka Milleny Azevedo Gonçalves, com o trabalho intitulado *“Uma Proposta de Design de Interação para uma Ferramenta de Coleta e Mineração de Textos em Redes Sociais Online para fins de Pesquisa Científica”*, no Programa de Pós-graduação em Educação – PPGEd.

Na qualidade de presidente da banca, atesto a Prof.<sup>a</sup> Msc. Glívia Angélica Rodrigues Barbosa, docente do Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), participou através de videoconferência.

Em virtude da participação remota do membro da banca acima indicado, eu, Prof.<sup>a</sup> Dr.<sup>a</sup> Maria Lúcia Bento Villela, enquanto servidor público, no gozo de fé pública, assino no lugar desse na Ata de Defesa e na Folha de Aprovação da referida defesa.

Por ser verdade, dou fé e assino o presente atestado.

Diamantina, 10 de dezembro de 2019.

  
Prof.<sup>a</sup> Dr.<sup>a</sup> Maria Lúcia Bento Villela

---





## **AGRADECIMENTOS**

A presente pesquisa não poderia chegar a bom porto sem as bênçãos de Deus, e a Ele agradeço primeiramente.

Agradeço ao meu marido, Wesley, por ser meu porto seguro e maior incentivador. Obrigada por todo o apoio, amor e compreensão essenciais para eu conseguir realizar mais este sonho.

Ao meu filho, Álvaro, por ser a minha motivação diária. Desculpe filho pelos momentos de ausência, tudo foi feito pensando em você!

Aos meus pais, Lu e Kincas, por serem o meu exemplo e por terem me ensinado o que tenho de mais valioso: a educação e a honestidade. Aos meus irmãos, Bruna e Hugo, por todo amor, incentivo, apoio e amizade. Agradeço também aos meus sogros, Anete e Isaias, por todo carinho comigo e ajuda com o meu pequeno. Vocês foram essenciais nesta jornada!

A minha orientadora, Maria Lúcia. Você não faz ideia da admiração e gratidão que tenho. Com seus ensinamentos e conselhos aprendi e amadureci tanto! Você é um exemplo de pessoa e profissional, possui um coração enorme. Desejo que tudo que fez e faz por mim volte a você e sua família em dobro, transformados em bênçãos.

A Lidy, Luanna e Hercules toda a parceria, ajuda, apoio e palavras de encorajamento e carinho. Amizades para a vida...Vocês são incríveis!

Aos participantes da pesquisa e aos membros das quatro bancas examinadoras, deste percurso, Carol, Glívia, Marcus e Geruza, pelas valiosas sugestões e contribuições.

Aos grupo MTPLNAM pela infraestrutura necessária, em especial ao Marcus por todo apoio, incentivo e orientações, não somente durante o mestrado, mas desde a graduação que estive comigo.

Enfim, a todas as pessoas que, direta ou indiretamente, contribuíram para a realização deste trabalho.



## RESUMO

As redes sociais *online* têm se tornando cada vez mais populares e, associado a este crescimento, o volume de informação gerado em formato digital tem aumentado de forma significativa. Independentemente do tipo de rede social, é possível aproveitá-la como fonte de informação para a construção de conhecimento científico, nas mais diferentes áreas. Porém realizar tais atividades manualmente é absolutamente inviável, sendo necessária a utilização das técnicas de Mineração de Textos para analisar e extrair informações úteis dessa extensa base de dados. Em geral, há boas opções de ferramentas no mercado, contudo, se considerarmos a questão do ponto de vista acadêmico, seja por falta de habilidade para operar a interface dos programas existentes, ou desconhecimento de sua existência, a utilização de ferramentas experimentais ainda é relativamente baixa na pesquisa acadêmica nacional da área de mídias sociais. Neste contexto, esta pesquisa teve o objetivo de desenvolver o design da interação do *Framework* Oráculo, uma ferramenta de apoio ao processo de coleta e mineração de textos, voltada para o apoio de usuários que coletam e mineram dados em redes sociais *online*, para fins de pesquisa científica. A pesquisa foi desenvolvida inicialmente com a identificação das necessidades e definição dos requisitos dos pesquisadores de instituições de ensino em relação à coleta e análise de dados em redes sociais *online* e, em seguida, foi feito o design do *framework*, seguindo a abordagem do Design Centrado na Comunicação, seguido da construção do protótipo funcional de alta fidelidade. Uma avaliação preliminar do protótipo foi realizada junto aos seus potenciais usuários, e os resultados apontaram a facilidade de uso e a aceitabilidade da sua interface, proporcionando a satisfação dos usuários, que foi refletida na vontade de utilizar o Oráculo em suas pesquisas futuras. Assim, acredita-se que, com o trabalho, foi possível contribuir com as áreas de IHC e mineração de textos, ao trazer considerações sobre aspectos relevantes para o design e avaliação de ferramentas de coleta e mineração de dados provenientes de RSOs, sob o ponto de vista da interação humano-computador.

**Palavras-chave:** Mineração de Textos, Redes Sociais, *Framework* Oráculo, *Design* da Interface



## **ABSTRACT**

Online social networks are becoming increasingly popular and, associated with this growth, the volume of information generated in digital format has increased significantly. Regardless of the type of social network, it can be used as a source of information for the construction of scientific knowledge in the most different areas. However, performing such activities manually is absolutely not feasible, requiring the use of Text Mining techniques to analyze and extract useful information from this extensive database. In general, there are good tool options on the market, however, if we consider the issue from the academic point of view, either because of the lack of ability to operate the interface of existing programs or lack of knowledge about them, the use of experimental tools is still relatively low. low in national academic research in the area of social media. In this context, this research aimed to develop the interaction design of the Oracle Framework, a tool to support the process of collecting and mining texts, aimed at supporting users who collect and mine data on online social networks, for the purpose of scientific research. The research was initially developed with the identification of the needs and definition of the requirements of the researchers of educational institutions in relation to the collection and analysis of data in online social networks. Then, the framework design was made, following the Centered Design approach. Communication, followed by the construction of the high fidelity functional prototype. A preliminary evaluation of the prototype was carried out with its potential users, and the results showed the ease of use and the acceptability of its interface, providing users' satisfaction, which was reflected in the desire to use the Oracle in their future research. Thus, it is believed that, with the work, it was possible to contribute to the areas of HCI and text mining, by bringing considerations on relevant aspects for the design and evaluation of data collection and mining tools from RSOs, under the point point of view of human-computer interaction.

**Keywords:** Text Mining, Social Networking, Oracle Framework, Interface Design



## LISTA DE ILUSTRAÇÕES

Figura 1- Processo de Mineração de Textos.....	29
Figura 2- Metacomunicação <i>designer</i> -usuário e comunicação usuário-sistema .....	33
Figura 3- Modelo do Framework Oráculo.....	35
Figura 4- Estrutura Metodológica.....	41
Figura 5- Passos para a Identificação das Necessidades e Definição dos Requisitos .....	46
Figura 6- Necessidades dos pesquisadores de dados para ferramentas de coletas de dados de mídias sociais.....	59
Figura 7- Persona 1: Pesquisadores com conhecimentos na área de computação.....	61
Figura 8- Persona 2: Pesquisadores que não possuem conhecimento na área de computação	62
Figura 9- Cenários de Análise dos pesquisadores que possuem conhecimento na área de computação (Persona 1).....	62
Figura 10- Cenários de Análise dos pesquisadores que possuem conhecimento na área de computação (Persona 2).....	63
Figura 11- Etapas do Design Centrado na Comunicação .....	66
Figura 12- Atividades de Design.....	67
Figura 13- Elementos da MoLIC.....	71
Figura 14- Modelagem dos Pontos de Abertura e Encerramento e acesso ubíquo na MoLIC.	72
<i>Figura 15- Processo do sistema e comunicação síncrona e consecutiva .....</i>	<i>72</i>
Figura 16- Metamensagem elaborada para o Oráculo.....	74
Figura 17- MoLIC – Diagrama de interação do framework Oráculo.....	77
Figura 18- Tela de Opções.....	80
Figura 19- Tela de Tour pelo Oráculo .....	80
Figura 20- Tela de Realizar Nova Coleta – Passo 1 .....	81
Figura 21- Tela de Realizar Nova Coleta – Passo 1 – Ajuda com Operadores.....	82
Figura 22- Tela de Realizar Nova Coleta – Passo 2 .....	83
Figura 23- Tela de Realizar Nova Coleta – Passo 3 .....	84
Figura 24- Tela Realizar Nova Coleta – Passo 4.....	85
Figura 25- Tela de Realizar Nova Coleta – Passo 4 – Confirmação Idioma e Localização.....	85
Figura 26- Tela de Processamento da Coleta – Alerta Sobreposição de Nome .....	86
Figura 27- Tela de Apresentação da Coleta.....	87
Figura 28- Tela de Configurar Filtro de Bots .....	88

Figura 29- Tela de Minhas Coletas.....	89
Figura 30- Tela de Configurar Mineração e Visualização – Passo 1 .....	90
Figura 31- Tela de Configurar Mineração e Visualização – Passo 2 .....	90
Figura 32- Tela de Configurar Mineração e Visualização – Passo 2 – Visão Geral .....	91
Figura 33- Tela de Configurar Mineração e Visualização – Passo 3 .....	92
Figura 34- Tela de Visualização da Coleta Minerada – Número Total de Termos.....	92
Figura 35- Tela de Visualização da Coleta Minerada – Tabela de Frequência de Palavras.....	93
Figura 36- Tela de Visualização da Coleta Minerada – Gráfico de Frequência de Termos.....	93
Figura 37- Tela de Visualização da Coleta Minerada – Nuvem de Palavras .....	94
Figura 38- Tela de Opções de Mineração Adicionais – Passo 1 .....	95
Figura 39- Tela de Opções de Mineração Adicionais – Passo 1 – Alerta Formato Arquivo ...	96
Figura 40- Tela de Opções de Mineração Adicionais – Passo 2 .....	97
Figura 41- Tela de Opções do Oráculo.....	99
Figura 42- Minhas Coletas .....	99
Figura 43- Nova Coleta (Passo 1).....	100
Figura 44- Conheça o Oráculo .....	100
Figura 45- Passos para realização da Avaliação.....	102
Figura 46- Síntese do modelo da pesquisa .....	119
Gráfico 1- Satisfação .....	108
Gráfico 2- Aprendizagem .....	109
Gráfico 3- Ajuda.....	110
Gráfico 4- Controle sobre a interface .....	112
Gráfico 5- Facilidade de Uso Percebida .....	114
Gráfico 6- Intenção Comportamental de Uso.....	115



## **LISTA DE TABELAS**

Tabela 1: Perfil dos participantes do estudo .....	51
Tabela 2: Elementos da Molic e suas formas de mapeamento .....	73
Tabela 3: Sugetões de melhorias .....	117

## LISTA DE ABREVEATURAS E SIGLAS

API - Application Programming Interface

CSV - Comma-separated values

DCC - Design Centrado na Comunicação

DECOM - Departamento de Computação

DMI - Digital Methods Initiative

EngSem - Engenharia Semiótica

FUP - Facilidade de Uso Percebida

Gexf - Graph Exchange XML Format

GNU - General Public License

HFC - *Human Factors Group*

ICU - Intenção Comportamental de Uso

IHC - Ciclo Básico de Design

Labic - Laboratório de estudos sobre Imagem e Cibercultura

MEDS - Método de Explicitação do Discurso Subjacente

MoLIC - *Modeling Language for Interaction as Conversation*

MT - Mineração de Textos

MTPLNAM - Mineração de Textos, Processamento de Linguagem Natural e Aprendizado de Máquina

QE1 – Questão 01

QE2 – Questão 02

RI - Recuperação da Informação

RSO - redes sociais *online*

SUMI - Software Usability Measurement Inventory

TAM - *Technology Acceptance Model*

TCAT - Twitter Capture and Analysis Toolset

TCLE - Termo de Consentimento e Livre Esclarecimento

UFDtna - Universidade de Federal de Diamantina –

UFES - Universidade Federal do Espírito Santo

UFVJM - Universidade Federal dos Vales do Jequitinhonha e Mucuri

WIMP - *Windows, Icons, Menus, and Pointers*

Ytk - YourTwrapperKeeper

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>20</b>
1.1. Objetivos .....	21
1.2. Justificativa .....	22
1.3. Problema .....	23
1.4. Contribuições .....	23
1.5. Estrutura da dissertação .....	24
<b>2. FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>26</b>
2.1. Redes Sociais <i>online</i> como fonte de informação para pesquisas científicas .....	26
2.2. Mineração de Textos em Redes Sociais <i>Online</i> .....	28
2.3. <i>Design</i> de Sistemas Interativos .....	30
2.3.1. Engenharia Semiótica.....	32
<b>3. FRAMEWORK ORÁCULO.....</b>	<b>34</b>
<b>4. TRABALHOS RELACIONADOS .....</b>	<b>37</b>
4.1. Mineração de Textos em Redes Sociais <i>Online</i> sob a perspectiva dos usuários .....	37
4.2. Ferramentas de Coleta e Mineração de Textos em Redes Sociais <i>Online</i> .....	39
<b>5. METODOLOGIA .....</b>	<b>41</b>
5.1. Levantamento Bibliográfico .....	42
5.2. Identificar as Necessidades e Definir os Requisitos .....	42
5.3. <i>Design</i> .....	43
5.4. Construção da Versão Interativa .....	43
5.5. Avaliação .....	44
<b>6. IDENTIFICAÇÃO DAS NECESSIDADES E DEFINIÇÃO DOS REQUISITOS...45</b>	
6.1. Considerações Metodológicas .....	45
6.1.1. Coleta de Dados .....	46
6.1.2. Análise e Interpretação dos Dados Coletados .....	47
6.1.3. Apresentação dos Resultados .....	48
6.1.3.1. Personas .....	48
6.1.3.1.1. Cenários de Análise .....	49
6.2. Execução .....	50
6.3. Resultados .....	51
6.4. Discussão .....	57
<b>7. DESIGN DO FRAMEWORK ORÁCULO .....</b>	<b>64</b>
7.1. Considerações Metodológicas .....	65
7.1.1. Design Centrado na Comunicação (DCC) .....	65
7.1.1.1. <i>Design</i> Conceitual .....	67
7.1.1.1.1. Construção da mensagem de metacomunicação de alto nível.....	68
7.1.1.1.2. Modelagem da Interação como Conversa .....	68
7.1.1.1.2.1. Definição da Estrutura da Conversa .....	69

7.1.1.1.2.2.	Detalhamento dos diálogos e falas.....	71
7.1.1.2.	<i>Design</i> Físico .....	72
7.2.	Artefatos de <i>Design</i> .....	74
7.2.1.	<i>Design</i> Conceitual .....	74
7.2.1.1.	Construção da mensagem de metacomunicação de alto nível.....	74
7.2.1.2.	Modelagem da Interação Como Conversa.....	75
7.2.2.	<i>Design</i> Físico.....	79
<b>8.</b>	<b>AVALIAÇÃO .....</b>	<b>101</b>
8.1.	Considerações Metodológicas .....	101
8.1.1.	Software Usability Measurement Inventory (SUMI).....	102
8.1.2.	Technology Acceptance Model (TAM) .....	103
8.2.	Execução .....	104
8.3.	Resultados e Discussão .....	106
8.3.1.	A satisfação do usuário durante a interação e a qualidade do design proposto ...	106
8.3.1.1.	Satisfação (S).....	107
8.3.1.2.	Aprendizagem (A).....	109
8.3.1.3.	Ajuda (D).....	110
8.3.1.4.	Controle sobre a interface (C) .....	111
8.3.2.	A aceitação do usuário durante a interação .....	113
8.3.2.1.	Facilidade de uso percebida (FUP).....	113
8.3.2.2.	Intenção Comportamental de Uso (ICU).....	115
8.3.3.	As impressões dos usuários e propostas de melhorias das telas .....	117
<b>9.</b>	<b>SÍNTESE DO MODELO DE DESIGN DA INTERAÇÃO DO FRAMEWORK ORÁCULO.....</b>	<b>119</b>
<b>10.</b>	<b>CONSIDERAÇÕES FINAIS.....</b>	<b>121</b>
10.1.	Contribuições .....	122
10.2.	Trabalhos Futuros.....	123
<b>11.</b>	<b>REFERÊNCIAS .....</b>	<b>124</b>
	<b>APÊNDICE A - Roteiro da Entrevista e Questionário .....</b>	<b>129</b>
	<b>APÊNDICE B - Questionários Pré Teste e TCLE.....</b>	<b>133</b>
	<b>APÊNDICE C - Tarefas, Questionários e Entrevista Pós Teste .....</b>	<b>136</b>

## 1. INTRODUÇÃO

Os grandes avanços das tecnologias e a popularização da internet têm estimulado a proliferação das redes sociais *online* (RSO). Conectadas à rede, as pessoas passaram a substituir por estas as formas tradicionais de socialização. Para Recuero (2014, p.6), “as redes sociais *online* representam um novo e complexo universo de fenômenos comunicativos, sociais e discursivos”. Neste ambiente virtual, as pessoas, de forma praticamente livre, interagem, produzem, disseminam e localizam informações, ficando a cargo dos internautas julgarem o que é ou não de seus interesses (COSTA, 2018).

Embora o conceito de redes sociais venha sendo delineado, desde a revolução industrial, como tema de preocupação científica, atualmente ele tem sugerido enfoques e vertentes diversos, atraindo pesquisadores de diferentes áreas, como ciência da computação, ciência da informação, jornalismo, psicologia, dentre outras (SILVA, 2015). Diante da possibilidade de, a partir do perfil dos usuários, mapear seus gostos, atos, ideias, conexões, e tantas outras informações, os sites de redes sociais tornam-se um grande repositório, mediador de uma multiplicidade de interesses e olhares sobre as mensagens em circulação (CASTRO; SPINOLA, 2015). Este potencial de produção de conteúdo tornou estas redes atraentes, tanto ao mundo do marketing quanto, mais atualmente, para o apoio das mais diversas pesquisas científicas no meio acadêmico (COSTA, 2018).

Grandes quantidades de conteúdo são compartilhadas e milhões de usuários interagem por meio de elos digitais (BENEVENUTO, 2010). Contudo, o uso desses dados, em seu formato bruto e desestruturado como são produzidos, é inviável para a descoberta de conhecimento. Dessa forma, é necessário realizar análises para extrair informações importantes que podem ser transformadas em conhecimento útil. Para auxiliar aos interessados neste processo, são utilizadas técnicas de Mineração de Textos (MT), um processo da Descoberta de Conhecimento que é considerado uma evolução da área de Recuperação da Informação (RI) (FERNEDA, 2003) e com aplicação cada vez mais corrente (SILVA *et al.*, 2015).

Neste contexto, o grupo de pesquisa em Mineração de Textos, Processamento de Linguagem Natural e Aprendizado de Máquina – MTPLNAM, do Departamento de Computação - DECOM da Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, tem empenhado esforços para o desenvolvimento do *Framework* Oráculo, uma ferramenta de coleta e mineração de textos em redes sociais *online*, apresentado no Capítulo

3. Este *framework* reúne um conjunto de funcionalidades que permitem ao usuário coletar determinados termos em postagens realizadas no *Twitter* e, após a finalização da coleta, realizar diferentes tipos de análises e pesquisas, suprimindo algumas limitações de outras ferramentas desenvolvidas para este propósito.

No entanto, no que tange especificamente à interação humano-computador, a literatura retrata pouca atenção por parte dos desenvolvedores de ferramentas no que diz respeito a suas interfaces, o que as tornam muitas vezes complexas e de difícil utilização (GOMES; CENDÓN, 2015). Nesse sentido, Silva (2015) mostra um panorama de afastamento entre os pesquisadores acadêmicos e as ferramentas de coleta. Este fato é confirmado por Santos (2014), que mostra que a utilização de ferramentas experimentais ainda é relativamente baixa na pesquisa acadêmica nacional da área de mídias sociais, seja por falta de habilidade dos usuários para operar os programas por estes possuírem interfaces complexas, ou mesmo desconhecimento de sua existência.

Assim, o presente trabalho vem atuar justamente neste ponto, no desenvolvimento da camada da interface do usuário do Oráculo, ao propor o *design* da interação deste *framework*, com foco em sua qualidade de uso. Para isso, a pesquisa executou, como passos metodológicos, as quatro atividades do Ciclo Básico de Design de IHC, proposto por Preece, Sharp e Rogers (PREECE *et al.*, 2002; SHARP *et al.*, 2007), que inclui: (1) identificação das necessidades e definição dos requisitos; (2) design, ao propor o design conceitual e físico; (3) construção da versão interativa, através de um protótipo funcional; e, por fim, a (4) avaliação da interação proposta.

Busca-se assim, apoiar os pesquisadores que utilizam dados provenientes de redes sociais *online* em suas pesquisas, ao lhes oferecer uma ferramenta interativa de coleta de dados que lhes proporcione uma experiência de uso positiva no alcance de seus objetivos.

### **1.1. Objetivos**

O objetivo deste trabalho é desenvolver o *design* da interação do *Framework* Oráculo, uma ferramenta de coleta e análise de dados em redes sociais *online*, voltada para usuários que realizam mineração de textos como parte de suas pesquisas científicas.

Os seguintes objetivos específicos deverão ser atingidos para o alcance do objeto geral:

- Identificar as dificuldades e necessidades dos usuários/pesquisadores e os requisitos de interação humano-computador para a ferramenta;
- Propor o *design* conceitual e o *design* físico da ferramenta;
- Disponibilizar um protótipo interativo da ferramenta a ser desenvolvida;
- Avaliar o protótipo proposto quanto à sua qualidade de uso, bem como a satisfação e aceitação do usuário em relação à mesma.

## 1.2. Justificativa

As redes sociais *online* têm produzido um volume cada vez maior de dados não estruturados (SILVA, 2017). Por permitirem que as pessoas se expressem e se comuniquem de forma rápida e livre sobre os mais variados assuntos, muitos pesquisadores, de diferentes áreas de conhecimento, estão extraindo e analisando dados das mídias sociais como fonte de informações para suas pesquisas científicas.

Para se obter conhecimento útil a partir das informações geradas pelas redes sociais, técnicas apuradas de mineração de textos (MT) precisam ser aplicadas, e para isso são disponibilizadas diferentes ferramentas que auxiliam os usuários em distintas etapas deste processo (POLONI; TOMAÉL, 2014). Porém existe, até o momento, pouca atenção por parte dos desenvolvedores destas ferramentas no que diz respeito a suas interfaces, o que as tornam muitas vezes complexas e de difícil utilização (GOMES; CENDÓN, 2015).

Silva (2015) mostra um panorama de afastamento entre os pesquisadores acadêmicos e as ferramentas de coleta. Tal fato é confirmado por Santos (2014), que mostra que a utilização de ferramentas experimentais ainda é relativamente baixa na pesquisa acadêmica nacional da área de mídias sociais, seja por falta de habilidade dos usuários para operar os programas por estes possuírem interfaces complexas, ou mesmo desconhecimento de sua existência.

A integração entre diferentes tipos de softwares e procedimentos de coleta e análise de dados tendem a enriquecer o processo de mineração em mídias sociais, por um lado, e as possibilidades de pesquisa acadêmica, por outro (COSTA, 2018). Assim, para que se consiga um uso eficiente, o entendimento de como desenvolver ferramentas que atendam as necessidades dos pesquisadores em mídias sociais, independente do seu conhecimento ou não sobre as diversas tecnologias, é um importante foco por parte dos *designers* e desenvolvedores (BROOKS, 2015).

Neste contexto, a presente proposta se justifica devido a escassez de estudos que relatam a preocupação de se planejar e desenvolver interfaces para ferramentas de coleta de dados e mineração de textos em redes sociais *online*, com foco em sua qualidade de uso.

### 1.3. Problema

Um dos primeiros passos do planejamento de uma pesquisa que utilize dados provindos das mídias sociais é escolher os aplicativos a serem utilizados, pois coletar e analisar dados manualmente é inviável, devido ao volume de tráfego da internet.

Em geral, há boas opções de ferramentas no mercado, como NodeXL<sup>1</sup>, Netlytic<sup>2</sup>, Uberlink<sup>3</sup> e outras. Contudo, se considerarmos a questão do ponto de vista acadêmico, alguns fatores devem ser levados em conta. Primeiramente, deve-se focar em ferramentas gratuitas, dado o pouco (ou nenhum) orçamento disponível para as pesquisas científicas em redes sociais. Outra questão é que, seja por falta de habilidade para operar a interface dos programas ou desconhecimento de sua existência, a utilização de ferramentas experimentais ainda é relativamente baixa na pesquisa acadêmica nacional da área de mídias sociais. (SILVA 2015)

A partir deste contexto, ressalta-se a seguinte questão norteadora:

Como proporcionar uma experiência de uso positiva aos pesquisadores, no processo de mineração de textos, ao utilizarem uma ferramenta de coleta e análise de dados em mídias sociais?

### 1.4. Contribuições

Com o *design* da interação do *Framework* Oráculo, uma ferramenta interativa de coleta e análise de dados, espera-se contribuir com pesquisadores de instituições de ensino ao:

- proporcionar qualidade de uso ao framework, proporcionando uma experiência de uso positiva no processo de mineração de textos.

---

<sup>1</sup> <https://nodexl.com/>

<sup>2</sup> <https://netlytic.org/index.php>

<sup>3</sup> <https://uberlink.com/>



Na área de pesquisa de IHC, o presente trabalho contribui ao propor um processo de design de interação, com atividades específicas para o design de uma ferramenta de mineração de textos em redes sociais *online*.

### 1.5. Estrutura da dissertação

Esta dissertação foi estruturada em onze capítulos, ordenados para o melhor entendimento desta pesquisa, a saber :

- O capítulo 2 – Fundamentação Teórica - apresenta teorias e conceitos que fundamentam o desenvolvimento desta pesquisa. São apresentadas seções sobre: as Redes Sociais *Online*, o seu uso como fonte de informações para pesquisas científicas através do processo de coleta de mineração de textos; as abordagens de Design de Sistemas Interativos; e por fim, o *Framework* Oráculo.
- O capítulo 3 – Framework Oráculo - apresenta do Framework Oráculo, ao qual esta pesquisa esta inserida.
- O capítulo 4 – Trabalhos Relacionados – apresenta alguns trabalhos relacionados à mineração de textos em redes sociais *online* sob o ponto de vista do usuário, bem como ferramentas que dão apoio a este processo, foram analisados.
- O capítulo 5 - Metodologia – apresenta os passos metodológicos que compreendem do presente trabalho, que compreendem: o Levantamento Bibliográfico referente aos temas envolvidos; a Identificação das Necessidades e Definição dos Requisitos dos usuários foco de estudo; o Desenvolvimento do Design para o *Framework* Oráculo; a Construção da Versão Interativa através de um Protótipo Funcional de Alta Fidelidade; e a Avaliação da solução proposta.
- O capítulo 6 – Identificação das Necessidades e Definição dos Requisitos – apresenta e discute os principais resultados obtidos com a coleta de dados realizada com pesquisadores, com o objetivo de identificar as necessidades e os requisitos dos usuários para o framework Oráculo.
- O capítulo 7 – Design do Framework Oráculo – apresenta o projeto da intervenção, ou seja, o design do *framework* Oráculo, bem como o

desenvolvimento da modelagem conceitual e física para a ferramenta, as discussões sobre os resultados obtidos.

- O capítulo 8 – Construção da Versão Interativa – apresenta o desenvolvimento da versão interativa do *framework* Oráculo, através da construção de um protótipo funcional de alta fidelidade.
- O capítulo 9 - Avaliação – apresenta a avaliação preliminar do design do *framework* Oráculo, que foi realizada a partir do protótipo funcional de alta fidelidade, bem como os resultados e discussões obtidos a partir dela.
- O capítulo 10 – Síntese do Modelo de Design da Interação – apresenta o modelo proposto pela pesquisa
- O capítulo 11 – Considerações Finais – apresenta as considerações finais sobre a pesquisa realizada, bem como suas contribuições e sugestões de trabalhos futuros a serem desenvolvidos.

## 2. FUNDAMENTAÇÃO TEÓRICA

Nesta seção serão apresentadas as principais abordagens que fundamentam esta pesquisa: redes sociais *online* como fonte de informação para pesquisas científicas, mineração de textos, design de sistemas interativos e o *Framework* Oráculo.

### 2.1. Redes Sociais *online* como fonte de informação para pesquisas científicas

Vivemos em uma sociedade acostumada com constantes inovações tecnológicas, onde o processo de criação e desenvolvimento de novas tecnologias se mostra cada vez mais apurado e dinâmico, acelerando o ritmo da evolução da comunicação e da interação virtual entre os indivíduos (KAYNNE *et al.*, 2015). Conectadas à internet, as pessoas passaram a substituir as formas tradicionais de socialização pela construída através das redes sociais *online*, onde mais de dois terços da população *online* global visita ou participa deste novo ambiente de interação (COSTA, 2018).

Para Carvalho e Tsunoda (2018), as redes sociais *online* (RSO) são conjuntos de pessoas ou grupos de pessoas com algum padrão de contato ou interação entre si através da internet. Por permitirem uma interação mais rápida e fácil entre seus usuários, as RSO têm experimentado um considerável crescimento, se tornando cada vez mais integradas à vida das pessoas (PERRIN, 2015).

Existem várias redes sociais *online* disponíveis na *Web*, que variam de acordo com seus objetivos primários, como por exemplo, as redes de profissionais (ex., LinkedIn), redes de amigos (ex., MySpace e Facebook), e redes para o compartilhamento de conteúdos específicos, como mensagens curtas (ex., Twitter), diários e blogs (ex., LiveJournal), fotos (ex., Flickr), e vídeos (ex., YouTube). Uma lista atualizada e exaustiva de RSO's, com 250 sítios *Web*, pode ser encontrada no site da Social Media List<sup>4</sup>.

As RSO acabaram revolucionando a forma como as pessoas se comunicam e expressam suas opiniões sobre os mais diversos aspectos. Estas se tornaram não somente uma nova forma de interação social, mas também um instrumento eficiente de “propagação de ideologias, culturas e identidades” (CASTRO; SPINOLA, 2015, p. 171).

---

<sup>4</sup> Disponível em: <https://socialmedialist.org/redes-sociais.html>. Acesso em setembro/2019

Diante desta realidade, surge a possibilidade de mapear gostos, atos, ideias e conexões de milhares de pessoas, investigar padrões entre essas múltiplas redes, principalmente através das interações que são mediadas por essas mídias (CASTRO; SPINOLA, 2015). Este potencial de produção de conteúdo, unido ao conjunto de informações secundárias que podem ser obtidas sobre o perfil dos usuários, tornou estas redes atraentes tanto ao mundo do marketing quanto, mais atualmente, para o apoio de pesquisas científicas no meio acadêmico (COSTA, 2018).

Assim, as RSO estão se tornando um tema central de pesquisas em diversas áreas, utilizando-sediferentes abordagens. Alguns exemplos de pesquisas são as de:

- Martins (2015), que realizou uma pesquisa bibliográfica objetivando contribuir com o campo educacional, através da reflexão sobre as redes sociais como lugar de aprendizagem. Como resultados, a autora percebeu a necessidade de reflexão do paradigma moderno de educação, onde o processo de aprendizagem efetivo só é possível a partir do encontro presencial físico, em salas de aula, mas enriquecido com a utilização das redes sociais, que se apresentam hoje como um dos principais espaços de expressão e informação da sociedade;
- Ahmed, Scheepers, Stokdale (2014) analisaram publicações entre 2009 e 2013, na área de Sistemas de Informação e observaram que há um aumento contínuo no número de artigos de pesquisa de mídia social publicados nos últimos anos, e descobriram que a maioria dos artigos era orientada por dados e não orientada por teoria. Neste contexto, estes sugeriram que a pesquisa em mídias sociais vem surgindo ao longo de um período de tempo e há inúmeras oportunidades para os pesquisadores realizarem estudos rigorosos nessa área;
- Silva (2015) analisaram artigos que realizaram coleta de dados em sites de redes sociais no período de 2005 a 2015. Nestes, o autor analisou variáveis relacionadas a quantidade de dados, ferramentas de coleta, plataformas analisadas, e apresentou os resultados apresentando questões referentes aos desafios e potencialidades destas modalidades de pesquisa;
- Jordan (2019) examinou os vínculos entre a rede *online* dos acadêmicos e os papéis institucionais, ao utilizar o *Twitter* na identidade acadêmica formal. O autor identificou várias estratégias subjacentes ao uso do site pelos acadêmicos, como contornar restrições institucionais, ampliar o espaço acadêmico, encontrar um nicho, promoção e impacto e liberdade acadêmica;

- Bista (2015) apresentou as perspectivas dos estudantes de pós-graduação em educação usando o *Twitter* como uma ferramenta pedagógica de estudos. Os resultados indicaram que os estudantes tiveram uma experiência positiva de aprendizado do uso do *Twitter* em cursos de pós graduação em educação;

- Kakush, Évora e Dora (2016) investigaram o uso de redes sociais na educação em enfermagem. Seus estudos encontraram poucas experiências utilizando as redes sociais nesse contexto, sendo que as contribuições evidenciaram vários benefícios e também dificuldades encontradas, tornando-se subsídios para o aprimoramento e reavaliação da utilização das redes sociais no processo de ensino e aprendizagem.

Apesar da grande importância das redes sociais, como fonte de informações para as mais variadas pesquisas científicas, Costa (2018) ressalta que também são muitos os desafios a este uso científico das redes. Isso se deve ao fato de que cada uma delas possui políticas específicas de funcionamento e nenhuma foi pensada propriamente para este fim, exigindo técnicas específicas de extração de conhecimento em suas bases de dados. Ainda assim, o uso de dados provenientes das mídias sociais para a pesquisa acadêmica pode ser bastante relevante, dado que estas consistem em uma fonte extremamente rica de dados de diferentes áreas e considerando os mais diferentes contextos (ARAÚJO *et al.*, 2019).

## 2.2. Mineração de Textos em Redes Sociais Online

Com o crescimento do número de usuários e do conteúdo gerado por eles, as RSO's se tornaram uma rica fonte de informação para empresas, governos, pesquisadores e pessoas em geral. Entender o alcance e as vantagens dessa forma de comunicação, e utilizar todas estas informações para gerar conhecimentos, pode trazer grandes benefícios e até mesmo ganhos financeiros a todos os interessados (MARTINS, RF; PEREIRA; BENEVENUTO, 2015).

O site *Internet Live Stats*<sup>5</sup> apresenta informações acerca da utilização da *internet* em todo o mundo em tempo real. Segundo este, o *Facebook* possui mais de 2,3 bilhões de usuários ativos, e gera a maior quantidade de dados dentre as redes sociais, e o *Twitter*, segunda maior rede social do mundo, possui mais de 350 milhões de usuários ativos.

---

<sup>5</sup> Dados de Internet Live Stats, em setembro de 2019. Disponível em: <https://www.internetlivestats.com>

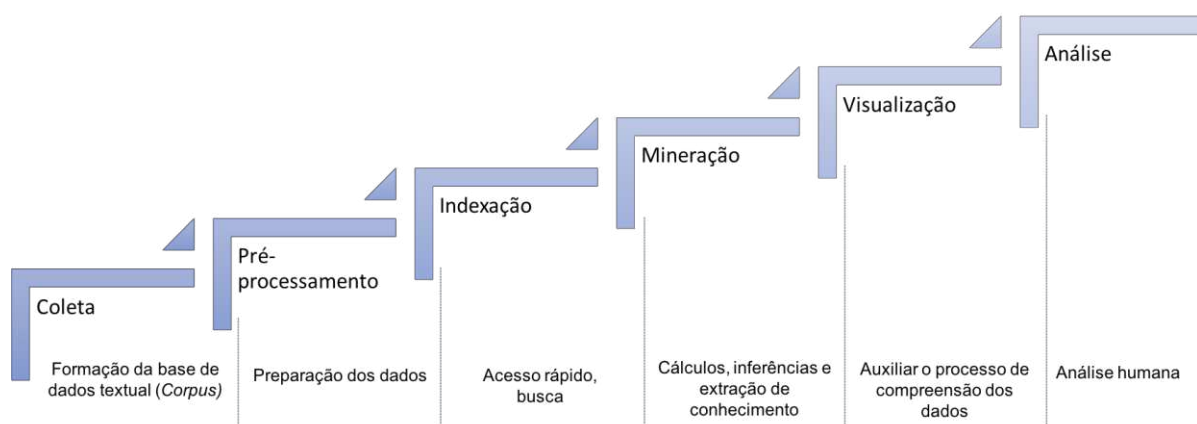
Recuero (2014) expõe que, deste gigante volume de informações disponível, cerca de 80% estão em formato de texto e não estruturado. Assim, a obtenção de conhecimento se tornou uma tarefa árdua, sendo cada vez mais necessários mecanismos eficientes e eficazes para extração e mineração de conhecimentos úteis da *web* (MARTINS, RF; PEREIRA; BENEVENUTO, 2015). Poloni e Tomaél (2014) completam ao afirmar que a análise do comportamento de indivíduos, e sua trajetória em uma RSO, depende de técnicas computacionais apuradas.

Neste contexto, destacam-se as abordagens de Mineração de Textos (MT), um processo de descoberta de conhecimento considerado uma evolução da área de Recuperação da Informação (RI) (FERNEDA, 2003), com aplicação cada vez mais corrente (SILVA, LEANDRO AUGUSTO *et al.*, 2015). Este processo consiste em um conjunto de atividades que abrangem o uso de técnicas, baseadas em modelos, capazes de encontrar padrões, sumarizar dados ou realizar previsões, com o objetivo de extrair novos conhecimentos com base em grandes volumes de dados (TAN, 1999).

A MT se preocupa com qualidade da busca e do processamento da informação, tendo impacto direto na qualidade do conhecimento obtido (CASTRO; FERRARI, 2016). Ela consiste em definir os dados a serem coletados, como a coleta será realizada e em como estes serão processados e apresentados; passos considerados base para a obtenção de dados, informações e conhecimentos que auxiliam a análise e a tomada de decisão (SANTOS; SILVEIRA, 2018).

Assim, a mineração de textos em RSO's abrange as etapas de coleta, pré-processamento, indexação, mineração, visualização e análise, como pode ser observado na Figura 1.

Figura 1- Processo de Mineração de Textos



Fonte: ARANHA E PASSOS, 2006, p.4. Adaptado.

Neste processo, a Coleta é a etapa inicial e tem como objetivo formar uma base de dados textual, conhecida na literatura como *Corpus*. Basicamente, nesta etapa são coletados os dados que serão analisados. A etapa de pré-processamento tem por finalidade melhorar a qualidade dos dados coletados e organizá-los, promovendo alguma formatação e representação da massa textual obtida na etapa anterior. Na etapa de indexação é realizado o processo que organiza todos os termos adquiridos facilitando o seu acesso e recuperação, garantindo rapidez e agilidade ao processo de localização dos dados. Na etapa de mineração são aplicadas técnicas para a extração do conhecimento, por meio de cálculos, inferências e algoritmos cujo objetivo é a descoberta de padrões e comportamentos importantes a serem observados (ARANHA; PASSOS, 2006). Na etapa de visualização, os dados são representados em uma linguagem de mais fácil compreensão, como gráficos e conceitos visuais (SANTOS; SILVEIRA; MANSSOUR, 2016). Por fim, ocorre a etapa de análise e interpretação dos dados pelas pessoas interessadas, através da procura por padrões e/ou análise das visualizações obtidas anteriormente, a fim de apoiá-las em suas tomadas de decisão (ARANHA; PASSOS, 2006).

### **2.3. Design de Sistemas Interativos**

Segundo Preece *et al.* (2013, p. 317), design é “atividade prática e criativa, com o objetivo de desenvolver um produto que ajude seus usuários a atingirem seus objetivos”. De acordo com Oxford English Dictionary, o termo é “um plano ou esquema concebido na mente, com o intuito de ser posteriormente executado”.

Ao projetar um sistema interativo, uma das preocupações do *designer* deve ser com a qualidade de uso relacionada à interação do usuário com o sistema (CARVALHO, 2012). Os critérios de qualidade de uso evidenciam certas características da interação e da interface que as tornam adequadas aos resultados esperados do uso do sistema (BARBOSA; SILVA, 2010).

Alguns dos critérios de qualidade de uso, citadas por estes autores são: usabilidade, experiência do usuário, acessibilidade e comunicabilidade. A usabilidade é um conceito que está relacionado com a facilidade de aprendizado e uso da interface, bem como a satisfação do usuário em decorrência deste uso (NIELSEN, 1993). A experiência do usuário engloba os sentimentos e as emoções dos usuários relacionadas ao uso dos sistemas computacionais interativos (SHARP *et al.*, 2007). O critério de acessibilidade está relacionado

à remoção das barreiras que impedem mais usuários de serem capazes de acessar a interface do sistema e interagirem com ele (BARBOSA; SILVA, 2010). A comunicabilidade, abordada pela Engenharia Semiótica, descrita na seção a seguir, diz respeito à capacidade da interface de comunicar ao usuário as intenções do *designer* e os princípios de interação definidos durante o processo de *design* (DE SOUZA, 2005).

Diante deste contexto, o design de interação é baseado na compreensão dos objetivos, desejos e necessidades dos usuários e tem como função principal simplificar a interação do usuário com o sistema.

Dessa forma, este se relaciona com a qualidade de uso das soluções computacionais, aperfeiçoando as maneiras como as pessoas trabalham, se comunicam e interagem (FERREIRA, 2015a), ou seja, reduzindo os aspectos negativos (por exemplo: frustração ou aborrecimento) e melhorando os aspectos positivos (por exemplo: divertimento e compromisso) da experiência do usuário.

O design é um processo que envolve as seguintes atividades básicas: análise da situação atual (identificação do problema), análise de uma intervenção e a avaliação dessa intervenção projetada ou já aplicada à situação atual. Cada processo de design detalha essas atividades básicas de uma forma particular, definindo “como executar cada atividade; a sequência que elas devem ser executadas; quais atividades podem se repetir e o porquê; e os artefatos consumidos e produzidos em cada uma delas” (BARBOSA; SILVA, 2010, p. 100).

A execução das atividades de forma iterativa, permitindo refinamentos sucessivos entre as três atividades básicas, é uma característica dos processos de *design*. Isso permite que o *designer* tenha maior flexibilidade, obtendo boas oportunidades de aprendizado, tanto sobre o problema a ser resolvido, quanto sobre a solução a ser apresentada para este (BARBOSA; SILVA, 2010).

Os usuários devem participar ativamente das decisões tomadas no processo de design, tendo em vista que são eles quem irão se beneficiar do produto desse processo. Veras (2017) afirma que, quanto mais cedo o usuário for envolvido no processo de *design*, mais cedo o *designer* irá aprender sobre suas necessidades, impactando positivamente na solução apresentada e identificando e corrigindo eventuais problemas de maneira muito mais ágil.

Muitos dos métodos, modelos e técnicas utilizados em IHC se baseiam em teorias, como a Psicologia Cognitiva Aplicada (CARD; MORAN; NEWELL, 1983), Engenharia Cognitiva (NORMAN; DRAPER, 1986) e Engenharia Semiótica (DE SOUZA; CLARISSE S., 2005). Na subseção a seguir, descrevemos esta última, na qual este trabalho se baseia.



### 2.3.1. Engenharia Semiótica

Dentre as abordagens teóricas que fundamentam os métodos, modelos e técnicas utilizados em IHC, a Engenharia Semiótica (EngSem) (DE SOUZA, 2005) é a mais recente. Esta é a base para o design centrado na comunicação, que consiste na abordagem de design utilizada no presente trabalho e descrita na seção 7.1.1.

A EngSem entende a interface de um sistema interativo como sendo uma comunicação do *designer* do sistema para os seus usuários. Esta mensagem do *designer*, transmitida por meio da interface, informa aos usuários a quem o sistema se destina, que problemas ele é capaz de resolver e como interagir com ele para resolvê-los. Assim, a interface de um sistema é tida como um artefato de metacomunicação, uma vez que, a comunicação entre *designer*-usuário se dá por meio da interação usuário-sistema (BARBOSA; SILVA, 2010; PRATES; BARBOSA, 2007).

Cabe ressaltar que esta comunicação é indireta, já que a interface age como preposto do *designer*, e unidirecional, pois, o usuário não pode se comunicar de volta com o *designer*, em tempo de interação (PRATES; BARBOSA, 2007; DE SOUZA, 2005). O conteúdo dessa mensagem de metacomunicação, ou metamensagem, pode ser parafraseado no seguinte modelo (DE SOUZA, 2005, p. 25):

Este é o meu entendimento, como *designer*, sobre quem você, usuário, é, do que aprendi que você quer ou precisa fazer, de que maneira prefere fazer, e por quê. Este, portanto, é o sistema que consequentemente projetei para você, e esta é a forma como você pode ou deve utilizá-lo para alcançar uma gama de objetivos que se encaixam nesta visão.

Dessa forma, como pode ser observado na Figura 2, os processos de comunicação são realizados em dois níveis distintos: a comunicação direta entre usuário-sistema e a comunicação do *designer* para o usuário, através da interface, chamada de metacomunicação (PRATES; BARBOSA, 2007).

Figura 2- Metacomunicação *designer*-usuário e comunicação usuário-sistema

Fonte: BARBOSA; SILVA, 2010, p. 78

Assim, podemos dizer que a interface atua como representante do *designer*, ou seja, ela repassa a intenção do *designer* ao usuário da aplicação. Logo, a interface é vista como o preposto do *designer*, tendo em vista que ela representa o *designer*, no momento da interação (DE SOUZA, 2005).

A comunicação proporcionada através dessa mensagem é realizada por meio dos signos dispostos na interface do sistema. Signo é tudo aquilo que significa algo para alguém (Peirce, 1992 e 1998). A Engenharia Semiótica classifica os signos de uma interface em estáticos, dinâmicos e metalinguísticos (DE SOUZA *et al.*, 2006 e 2010; DE SOUZA; LEITÃO, 2009). Signos estáticos são aqueles que expressam o estado do sistema, e podem ser percebidos ao se olhar para uma tela do sistema (ex. o texto de um botão). Já os signos dinâmicos expressam o comportamento do sistema e surgem da interação do usuário com ele (ex. ação disparada ao se clicar em um botão). Por fim, signos metalinguísticos explicam outros signos da interface de forma direta (ex. sistema de ajuda).

Nota-se que a EngSem entende o papel do *designer* como sendo tão importante quanto o do usuário no contexto de um sistema interativo. Através da metacomunicação, o processo de *design* foca em “como o *designer* vai apresentar a solução ao usuário” (BARBOSA; SILVA, 2010, p. 78). Portanto, através da metacomunicação, o *designer* procura mostrar ao usuário as possibilidades de interação presentes na interface (PEREIRA JR., 2016).

A qualidade de uso tratada pela EngSem é a comunicabilidade, definida como “a qualidade de sistemas interativos de se comunicarem com os usuários, de forma eficiente e eficaz, as intenções de projeto e princípios de interação” (DE SOUZA, 2005). Barbosa e Silva (2010) afirmam que, quando o usuário não é capaz de entender a comunicação pretendida pelo *designer*, pode-se dizer que ocorrem rupturas de comunicação. Estas podem dificultar ou

até mesmo impossibilitar a metacomunicação ou o uso do sistema, comprometendo então a comunicabilidade deste.

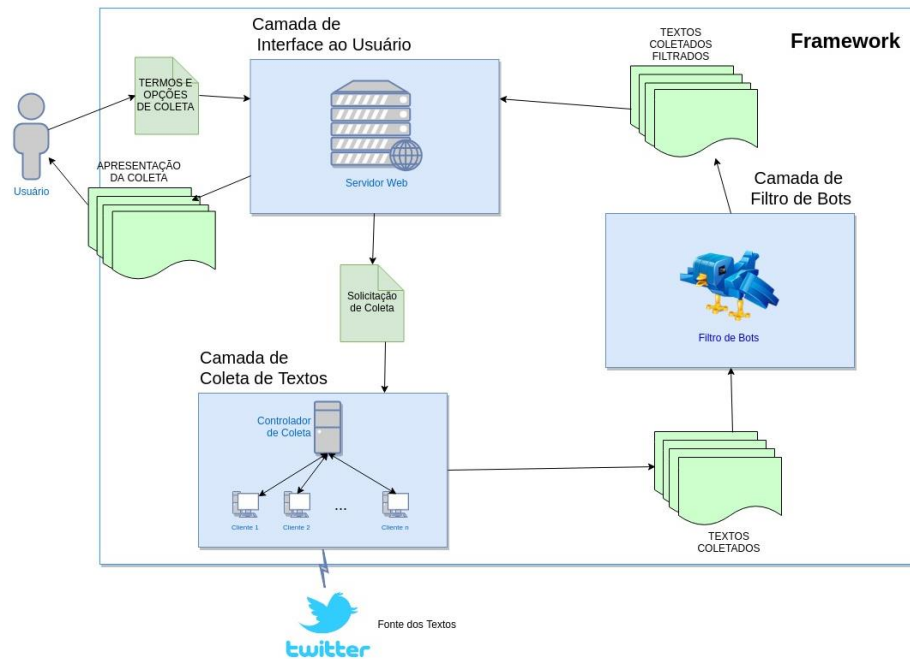
### 3. FRAMEWORK ORÁCULO

O Grupo de Pesquisa MTPLNM - Mineração de Textos, Processamento de Linguagem Natural e Aprendizado de Máquina, do Departamento de Computação – DECOM, da Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, está desenvolvendo um *Framework* de coleta de textos em redes sociais *online*, denominado Oráculo, e esta pesquisa de integra a uma das etapas de seu desenvolvimento.

Por meio do Oráculo, objetiva-se reunir, em uma única ferramenta, um conjunto de funcionalidades que permitam ao usuário agendar coletas de determinados termos em postagens, de forma que, após a finalização da coleta, os dados coletados possam passar por mineração de textos e serem utilizados para diferentes tipos de análises. Este visa ainda suprir as necessidades de ferramentas atuais, com a pretensão de elevar a qualidade do processo de coleta e também dos dados coletados (OLIVEIRA, 2019).

Conforme este autor, o *framework* Oráculo está sendo desenvolvido em três camadas, como ilustra a Figura 3.

Figura 3- Modelo do Framework Oráculo



Fonte: OLIVEIRA (2019, p. 32)

A Camada de Coleta de Textos é a responsável pela recuperação de informação junto ao *Twitter*, rede social escolhida para ser aplicado inicialmente as funcionalidades do Oráculo. Esta aplica diferentes técnicas e algoritmos de coleta de textos, utilizando-se de métodos de sistemas distribuídos, proporcionando uma coleta mais eficiente e eficaz, ao contornar limitações comuns a outras ferramentas com a mesma finalidade. Nesta camada ainda estão presentes funcionalidades desenvolvidas baseadas em técnicas de mineração de textos sobre os dados coletados (OLIVEIRA, 2019).

A Camada de Filtro de Bots é a camada intermediária, responsável por realizar a filtragem nos dados coletados. Tal camada, a partir de um determinado conjunto de regras, busca identificar quais as postagens são oriundas de bots e quais não são, possibilitando isolar as primeiras de forma que o usuário final possa desconsiderá-las de suas pesquisas, caso necessário. A ideia é que isso possa levar ao aumento da legitimidade das análises geradas a partir de coleta de dados de *tweets* (OLIVEIRA, 2019).

A Camada da Interface ao Usuário, desenvolvida neste trabalho, realiza a comunicação entre Usuário e o Oráculo por meio de uma plataforma *web* interativa. Esta será desenvolvida a partir de propostas de *Design* de IHC (ou da interação), com o objetivo de proporcionar ao usuário qualidade de uso, focando no modo e na forma que se dará a comunicação entre o usuário e o sistema.

Oliveira (2019) afirma que, com o desenvolvimento do *Framework* Oráculo, ao integrar as três camadas, espera-se obter ganhos de se utilizar sistemas distribuídos para realizar coletas de dados, visando a mineração de textos em redes sociais *online* por meio de uma ferramenta implantada numa plataforma *web* interativa. Acredita-se ainda que, os ganhos em desempenho, capacidade de armazenamento e qualidade nos dados coletados sejam satisfatórios.

O Oráculo ficará disponível para que outros pesquisadores, internos e externos à UFVJM, possam utilizá-lo. Este será parametrizável no momento de realização da coleta, a com a proposta de oferecer uma alta qualidade de uso, permitindo que usuários que não tenham conhecimento avançado em tecnologia, possam utilizar a ferramenta para realizar suas pesquisas e extrair o conhecimento almejado, de forma mais efetiva e satisfatória do que é feito atualmente.

## 4. TRABALHOS RELACIONADOS

Para maior conhecimento acerca do tema da pesquisa e análise dos estudos já realizados, alguns trabalhos relacionados à mineração de textos em redes sociais *online* sob o ponto de vista do usuário, bem como ferramentas que dão apoio a este processo, foram analisados.

### 4.1. Mineração de Textos em Redes Sociais *Online* sob a perspectiva dos usuários

A mineração de textos em redes sociais *online* é uma área de pesquisa que vem crescendo consideravelmente nos últimos anos. Várias pesquisas têm focado em desenvolver e melhorar técnicas computacionais que apoiam as diferentes atividades que compõem a mineração de textos (BARION e LAGO, 2015). Porém, ainda são escassos os trabalhos que exploram este processo em si, seja como um todo ou com foco em alguma de suas atividades, considerando a experiência e as necessidades dos usuários (SILVA, e STABILE, 2016).

Nesse sentido, Brooks (2015) investigou, por meio de um estudo etnográfico, como os pesquisadores trabalham com dados coletados de mídias sociais, buscando assim entender melhor o processo de análise. Com base no conhecimento adquirido com o estudo etnográfico, o autor discute as implicações de se utilizar uma nova tecnologia para ampliar a extensão da análise. Além disso, são discutidas duas propostas de *design* para apoiar esta atividade: uma considera a exploração visual dos dados e outra faz uso de colaboração na análise de dados *online*.

Ainda buscando entender as necessidades dos pesquisadores de dados de mídias sociais, Santos (2016), focaram especificamente em visualização, ao investigar se analistas de dados conhecem e utilizam a visualização e o quanto a consideram importante. Por meio de entrevistas semi-estruturadas, os autores coletaram informações que os permitiram descobrir que os analistas de dados de mídias sociais conhecem o conceito de visualização de dados e acreditam que o seu uso pode melhorar significativamente o processo de análise. Em um segundo trabalho, Santos e Silveira (2018), também, por meio de um estudo qualitativo exploratório, buscaram compreender como a visualização pode apoiar analistas de dados de mídias sociais a entenderem os dados e tomarem decisões. Os resultados revelam que os analistas de mídias sociais estão cientes das ferramentas de visualização existentes e destacam suas necessidades em relação a visualização de dados, a fim de fornecer uma base para o design de novas ferramentas para esta finalidade.

Chen *et al.* (2018) também conduziram um estudo formativo, baseado em entrevistas com cientistas sociais, para entender suas necessidades, práticas e restrições no trabalho com dados de mídias sociais. Com base nos resultados obtidos com esse estudo, os autores desenvolveram uma ferramenta de análise visual exploratória para dados de mídia social, que oferece suporte à construção e visualização de categorias, a fim de ajudar os pesquisadores a descobrirem e refinarem categorias importantes de conteúdo, além de comparar e analisar categorias com visualizações e resumos estatísticos.

Foram encontrados também estudos que visam descrever ou compreender etapas do processo de mineração de textos em mídias sociais, considerando que os pesquisadores de dados que executam essas etapas podem ser de diferentes áreas de conhecimento. Nesse sentido, Chen *et al.* (2018) descrevem as características, os elementos e o processo cronológico de análise de dados de mídia social sob a perspectiva de um acadêmico da área de comunicação de massa. O objetivo foi apresentar o processo cronológico em que um pesquisador lida com dados de mídia social na forma de estudos de caso.

Marques e Vidigal (2018) verificaram, através de entrevistas semiestruturadas com profissionais de diferentes empresas, como as redes sociais são utilizadas na fase de coleta de dados, com foco em inteligência competitiva como ferramenta fundamental para a empresa destacar-se frente a seus concorrentes. Os resultados demonstraram que as empresas estão atentas às redes sociais como fontes de informação, e destacam a necessidade de sua verificação, considerando-as muito importante e viável para a triagem das informações mercadológicas.

Brooker *et al.* (2016) propuseram um conjunto de metodologias para apoiar a análise de dados de mídias sociais. Para implementar estas metodologias, os autores apresentam uma ferramenta de *software* que ilustra uma abordagem analítica visual dos dados, ao fazer uso de visualizações de dados provenientes do *Twitter* como instrumento para formação e busca de hipóteses por parte dos analistas.

Embora os trabalhos acima explorem as atividades e necessidades de pesquisadores de dados de mídias sociais, todos eles focam em etapas específicas do processo de mineração de textos, como a coleta (MARQUES e VIDIGAL, 2018), a visualização (BROOKER *et al.*, 2016; CHEN *et al.*, 2017; SANTOS e SILVEIRA, 2018; SANTOS *et al.*, 2016) e a análise (BROOKER *et al.*, 2016; BROOKS, 2015; CHEN *et al.*, 2017; CHEN *et al.*, 2018).

No entanto, não encontramos pesquisas que realizem um estudo exploratório com pesquisadores de dados, no sentido de abordar suas experiências e necessidades, considerando

todas as atividades do processo de mineração de textos em RSOs de forma integrada. Além disso, não foram encontrados trabalhos que tivessem como foco o design da interação destas ferramentas, tomando como ponto de partida a coleta de dados, que é onde o processo tem início, e considerando também as demais atividades como sequência natural desse processo.

#### **4.2. Ferramentas de Coleta e Mineração de Textos em Redes Sociais *Online***

Aplicativos acadêmicos de código aberto e/ou scripts de qualidade para mineração nas mídias sociais vêm sendo desenvolvidos por pesquisadores. Contudo, ao contrário das ferramentas plenas, adequadas ao uso mercadológico e a projetos com maior orçamento, os aplicativos experimentais são alternativas eficientes para pesquisas acadêmicas e possuem a característica de exigir mais conhecimento e tempo dos pesquisadores, uma vez que não há alternativa gratuita disponível que realize todas as etapas de mineração em mídias sociais de forma rápida e intuitiva (SILVA, T; STABILE, 2016).

No meio acadêmico, em geral, cada aplicação tem uma função específica e cumpre uma tarefa determinada. Fato é que não faltam boas opções de ferramentas para mineração de dados das mídias sociais, mas deve-se atentar ao fato de que cada projeto possui características específicas que irão orientar a escolha. Algumas das principais ferramentas conhecidas pelo meio acadêmico, são YourTwrapperKeeper<sup>6</sup> e a DMI-TCAT<sup>7</sup>.

O YourTwrapperKeeper – yTK (BRUNS e LIANG, 2012) é uma ferramenta de código aberto, proposta pelo Laboratório de estudos sobre Imagem e Cibercultura (Labic)<sup>8</sup> do Departamento de Comunicação Social da Universidade Federal do Espírito Santo (UFES) e é licenciada através da GNU - General Public License (Licença Pública Geral GNU)<sup>9</sup>. Conforme sua proposta, o yTK está sujeito às limitações impostas pelas redes sociais, no caso o Twitter, por utilizar a sua API para realizar a coleta dos tweets. Desenvolvida em PHP e MySQL, essa ferramenta exporta os tweets coletados em formato Comma-separated values - CSV, permitindo, portanto, que os dados sejam acessados através de editores de textos e planilhas eletrônicas.

---

<sup>6</sup> <https://github.com/540co/yourTwrapperKeeper>

<sup>7</sup> <https://wiki.digitalmethods.net/Dmi/ToolDmiTcat>

<sup>8</sup> <https://www.labic.net/labic/>

<sup>9</sup> <https://github.com/540co/yourTwrapperKeeper>



A segunda ferramenta analisada é o Twitter Capture and Analysis Toolset (TCAT), que foi escolhida pelo grupo de pesquisa MTPLNAM, como ferramenta de referência para comparação do desenvolvimento e desempenho do framework Oráculo, por se tratar de uma ferramenta disponível e bem aceita pela comunidade de pesquisa desde 2014.

A TCAT foi desenvolvida e disponibilizada gratuitamente pelo grupo de pesquisa Digital Methods Initiative (DMI)<sup>10</sup>, que congrega em sua maioria pesquisadores da Universidade de Amsterdam. Seu objetivo principal é coletar e analisar *tweets* com finalidade acadêmica, inicialmente, em ciências humanas e sociais (BORRA *et al.*, 2014) . O DMI – TCAT fornece coleta e mineração de *posts* e comentários de perfis ou páginas institucionais, capturando dados realtime e retroativos, utilizando a API do Twitter, e ficando a mercê de suas limitações. O formato de saída é geralmente um CSV (arquivo de texto separado por vírgulas), suportado por uma gama de outros softwares que integram ao processo de mineração de textos, como planilhas eletrônicas e editores de textos.

Além dos textos presentes nos *tweets*, assim como o Oráculo, o DMI-TCAT consegue recuperar os metadados correspondentes, como menções, respostas, localização e urls, além de dados relacionados aos usuários dos tweets coletados, como, seguidores e tweets publicados. Tal ferramenta foi desenvolvida utilizando a linguagem de programação PHP e o banco de dados MySQL, sendo executado em um servidor *web* local. Ela possui interface básica, para algumas de suas funcionalidades voltadas para coletas *realtime*, sendo que suas funcionalidades relacionadas a coletas retroativas são disponibilizadas através de *scripts*.

Ambas as ferramentas acima utilizam a API do Twitter para realizar as coletas, tendo que se adequar às suas limitações, o que não ocorre com o Oráculo, uma vez que propõe a utilização de crawler para a realização das coletas. Além disso, elas disponibilizam suas funcionalidades relativas à coleta e mineração através de *scripts*, o que pode impactar de forma negativa a interação do usuário com tais ferramentas, diferentemente do Oráculo, que através desta pesquisa, vem disponibilizar ao usuário uma interface que se preocupa com a qualidade de uso da ferramenta.

---

<sup>10</sup> <https://wiki.digitalmethods.net/Dmi/WinterSchool2020>

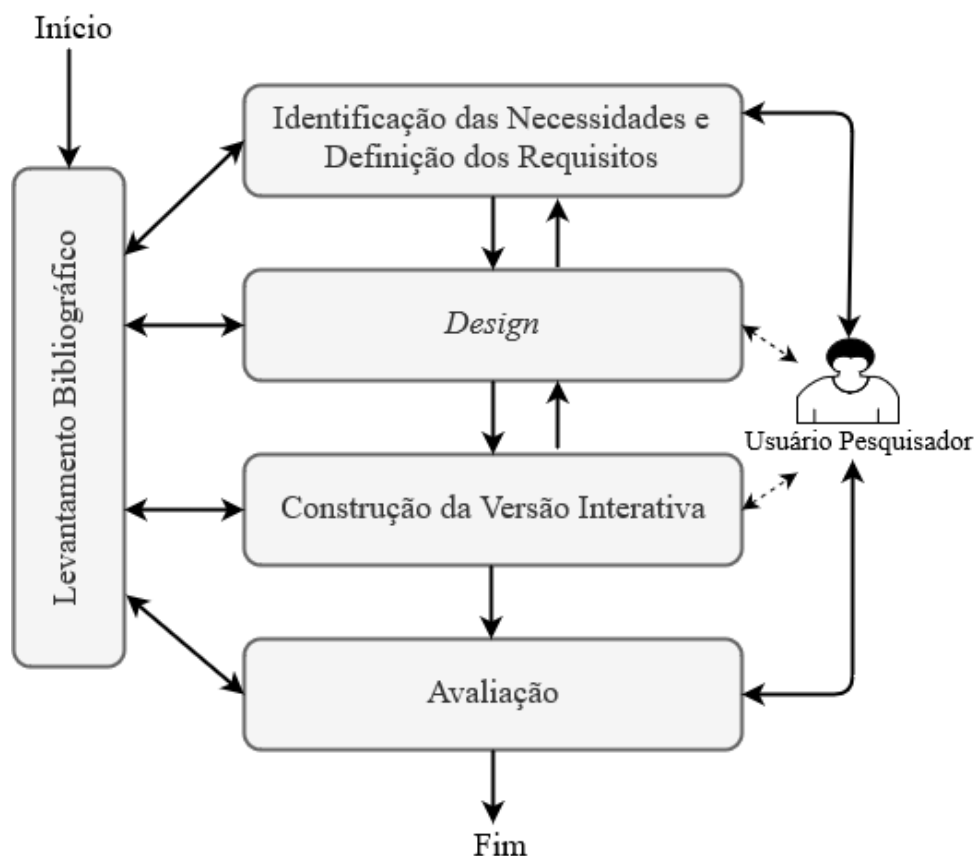
## 5. METODOLOGIA

A presente pesquisa é classificada, quanto à sua natureza e abordagem, como aplicada e qualitativa, respectivamente. Aplicada, por gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos, e qualitativa, por preocupar-se com aspectos da realidade que não podem ser quantificados (GIL, 2002).

Quanto aos seus objetivos, esta é classificada como uma pesquisa exploratória, por buscar proporcionar maior familiaridade com o fenômeno que está sendo investigado, a fim de descobrir novas explicações para o mesmo (GIL, 2002).

As etapas metodológicas do presente trabalho, descritas neste capítulo, são mostradas na Figura 4.

Figura 4- Estrutura Metodológica



Fonte: próprio autor

As etapas do design do *Framework* Oráculo seguiram o modelo simples de processo de design de IHC (SHARP; ROGERS; PREECE, 2007), que consiste em um conjunto de

atividades interligadas e iterativas, com os resultados de uma atividade fornecendo subsídios para a atividade subsequente.

Esse processo destaca a importância de envolver os usuários<sup>11</sup> em suas atividades, com a finalidade de aumentar as chances de o produto final ser bem aceito. Além disso, um ponto de destaque é a iteração entre as atividades, formando ciclos de design. Em cada ciclo, as atividades são realizadas no sentido de se construir uma versão interativa que, ao ser avaliada, permitirá que os resultados de tal avaliação sejam considerados em um novo ciclo de design, a fim de obter uma versão melhorada em relação à desenvolvida no ciclo anterior. Com isso, em cada ciclo, a versão do design vai gradativamente se aproximando do produto final. A seguir, serão descritas, de forma geral, cada uma das atividades que compõem a estrutura metodológica do presente trabalho. A descrição em detalhes de cada uma dessas atividades, incluindo os métodos utilizados e os resultados obtidos por elas, será feita posteriormente, nos capítulos de 6 a 9.

### **5.1. Levantamento Bibliográfico**

A atividade de “Levantamento Bibliográfico” consistiu na catalogação e leitura de artigos, dissertações, teses e livros sobre mineração de textos em redes sociais *online* e ferramentas voltadas para o usuário para esta finalidade, para melhor compreensão acerca dos temas envolvidos na pesquisa. Esta atividade foi realizada ao longo de toda a pesquisa, simultaneamente às demais atividades que compõem a estrutura metodológica.

### **5.2. Identificar as Necessidades e Definir os Requisitos**

Esta atividade envolveu uma pesquisa da situação atual, coletando as preferências e necessidades dos pesquisadores que coletam e analisam dados de redes sociais *online*. Buscou-se entender o espaço do problema e como a ferramenta proposta irá apoiar os usuários em suas atividades, além das restrições e condições que devem ser respeitadas (SHARP; ROGERS; PREECE, 2007). Para isso, foi realizado um estudo exploratório aprofundado, através da condução de entrevistas semiestruturadas, aplicando o Método de Explicitação do Discurso Subjacente (MEDS) (NICOLACI-DA-COSTA; LEITÃO; ROMÃO-DIAS, 2004). Através da execução etapas sugeridas pelo MEDS, os dados foram coletados, analisados e

---

<sup>11</sup> Projeto aprovado no Comitê de Ética em Pesquisa da UFVJM, sob o CAAE: 94091118.3.0000.5108

apresentados através de Personas (BARBOSA e SILVA, 2010) e Cenários de Análise (ROSSON e CARROL, 2002), sendo possível, assim, caracterizar as necessidades e os requisitos dos usuários.

### 5.3. Design

As atividades de design começaram de forma iterativa, após os requisitos terem sido estabelecidos (ROGERS; SHARP; PREECE, 2013). Esta atividade consiste no projeto da ferramenta computacional interativa, visando apoiar melhor os usuários no alcance de seus objetivos (BARBOSA; SILVA, 2010).

Com o objetivo de elaborar uma solução de *design* que comunique bem os princípios e intenções do *designer* aos usuários, foram utilizadas as abordagens do Design Centrado na Comunicação - DCC (DE SOUZA, 2005) para a elaboração de duas etapas principais: a primeira diz respeito a criação do modelo conceitual para o produto, ou seja, o Design Conceitual, que incluiu a construção da mensagem de metacomunicação de alto nível e a modelagem da interação do sistema, através da elaboração do Diagrama MoLIC (*Modeling Language for Interaction as Conversation*) (PAULA, 2003; BARBOSA e PAULA, 2003). A segunda etapa consistiu em definir a interface com a qual o usuário manterá contato direto, ou seja, a criação do *Design Físico*, a partir da escolha dos estilos de interação para o sistema e a representação da interface através de protótipos de baixa fidelidade<sup>12</sup>.

### 5.4. Construção da Versão Interativa

Nesta etapa, a partir do design físico criado anteriormente, foi desenvolvido um protótipo funcional de alta fidelidade<sup>13</sup> do *framework* Oráculo, para posterior avaliação da qualidade do seu *design*. Foi utilizado, nesta etapa, o Figma<sup>14</sup>, que consiste em uma ferramenta de prototipação gratuita, que vem sendo bastante utilizada e recomendada por

---

<sup>12</sup> Os protótipos de baixa fidelidade são utilizados em fases iniciais e exploratórias de um projeto para validar um conceito e decidir se uma ideia tem ou não valor funcional. Estes devem ser rápidos, rudimentares e possuírem baixo nível de detalhamento.

<sup>13</sup> Os protótipos de alta fidelidade se aproximam ao máximo dos aspectos visuais e funcionais do produto final, incluindo o conteúdo, fluxo de navegação e interações. São muito utilizados para testes e validação com usuários, ou para vender uma ideia, pois permitem que estes vejam o produto em funcionamento.

<sup>14</sup> Disponível em: <https://www.figma.com>

especialistas da área, como afirmam alguns sites de conteúdos sobre a experiência de usuários (UX), como: Medium<sup>15</sup>, Dicas UX<sup>16</sup>, Página do especialista Gabriel Silvestre<sup>17</sup>, o Coursera<sup>18</sup>, dentre outros.

## 5.5. Avaliação

Nesta etapa, foi realizada uma avaliação preliminar do *design* do produto proposto, a fim de determinar a qualidade da sua interação. Buscamos, com a avaliação, verificar a qualidade do *design* do *framework* Oráculo, sob o ponto de vista do usuário. Para isso, foram realizadas avaliações em ambiente controlado, envolvendo potenciais usuários do Oráculo, ou seja, pesquisadores que trabalham com análise de dados provenientes de redes sociais *online*. Em seguida, foram realizadas investigações sobre a experiência de uso, bem como a utilidade e facilidade de uso percebidas pelo usuários, fazendo uso dos questionários SUMI (Software Usability Measurement Inventory)<sup>19</sup> (KIRALOWSKI; CORBETT, 1993) e TAM (Technology Acceptance Model)<sup>20</sup> (DAVIS, 1989). Os dados coletados com a avaliação foram analisados e permitiram avaliar a qualidade do design proposto.

É interessante ressaltar que, devido aos prazos para conclusão desta pesquisa de mestrado, optou-se por se utilizar métodos de avaliação que nos exigisse menor tempo, não sendo possível executar as avaliações voltadas para a comunicabilidade do design proposto, sendo esta uma sugestão de trabalhos futuros.

---

<sup>15</sup> Disponível em: <https://medium.com/@kakaufb/%C3%A9-hora-de-aceitar-o-figma-na-sua-vida-7522954fe15b> e <https://medium.com/ui-lab-school/figma-5-motivos-para-ficar-de-olho-d8cfe3af07a1>

<sup>16</sup> Disponível em: <https://www.dicasux.com.br/figma/>

<sup>17</sup> Disponível em: <https://gabriel-silvestri.com.br/o-que-e-ux-design/>

<sup>18</sup> Disponível em: <https://pt.coursera.org/lecture/ux-ui-design-de-interface/ferramentas-de-prototipacao-rapida-E521Z>

<sup>19</sup> Disponível em: <http://sumi.ucc.ie>

<sup>20</sup> Disponível em: <https://pubsonline.informs.org/doi/abs/10.1287/mnsc.35.8.982>

## 6. IDENTIFICAÇÃO DAS NECESSIDADES E DEFINIÇÃO DOS REQUISITOS

Este capítulo apresenta e discute os principais resultados obtidos com a coleta de dados realizada com pesquisadores que trabalham com coleta e análise de dados de redes sociais *online*, com o objetivo de identificar as necessidades e os requisitos dos usuários para para uma ferramenta de coleta e mineração de textos, como o *framework* Oráculo.

Assim, o estudo realizado nesta etapa foi guiado pela seguinte questão de pesquisa: “Como é a experiência dos pesquisadores de dados no processo de mineração de textos provenientes das mídias sociais, tomando como ponto de partida a coleta de dados? ”. A fim de direcionar a discussão dos resultados, tal questão de pesquisa foi “quebrada” nas seguintes questões específicas:

- QE1: “Quais são as dificuldades enfrentadas pelos pesquisadores de dados durante a mineração de textos em RSOs, tomando como ponto de partida a coleta de dados?”
- QE2: “Quais são as necessidades dos pesquisadores de dados no processo de mineração de textos em RSOs, tomando como ponto de partida a coleta de dados?”

A seguir, na seção 6.1, apresentamos as considerações metodológicas da identificação das necessidade e definição dos requisitos. Em seguida, na seção 6.2, mostramos como esta etapa foi executada, e, por fim, os resultados são mostrados na seção 6.3 e discutidos na seção 6.4.

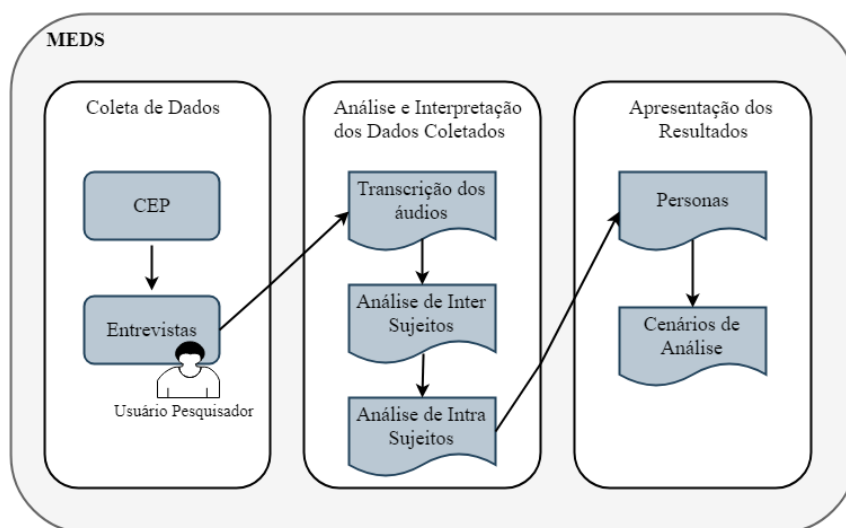
### 6.1. Considerações Metodológicas

Para responder as questões específicas QE1 e QE2, foi realizado um estudo exploratório aprofundado, através da condução de entrevistas semiestruturadas, aplicando o Método de Explicitação do Discurso Subjacente (MEDS) (NICOLACI-DA-COSTA; LEITÃO; ROMÃO-DIAS, 2004). O MEDS é um método de investigação qualitativa, que permite a realização de uma análise sistemática de material discursivo. Tal método não parte de uma hipótese pré-definida, e sim de uma pergunta aberta, sendo seu uso mais adequado do que um método quantitativo, quando se deseja entender uma questão ainda

desconhecida. Por ser qualitativo, ele também tem como característica o fato de buscar maior profundidade e detalhes do que um método quantitativo, que privilegia o conhecimento generalizado. Para atingir essa profundidade, ele envolve um processo trabalhoso, e por isso utiliza amostras pequenas (NICOLACI-DA-COSTA; LEITÃO; ROMÃO-DIAS, 2004).

Os passos estabelecidos pelo MEDS foram rigorosamente aplicadas nesta etapa da pesquisa, para a coleta e análise e interpretação dos dados, como apresenta a Figura 5. Estas atividades são descritas a seguir.

Figura 5- Passos para a Identificação das Necessidades e Definição dos Requisitos



Fonte: próprio autor

### 6.1.1. Coleta de Dados

Para a coleta de dados dos usuários, a fim de identificar as suas necessidades e requisitos de IHC, foram realizadas entrevistas semiestruturadas. Nesta técnica, o entrevistador possui um roteiro básico de orientação e a entrevista flui como um bate papo, com uma flexibilidade maior para se obter as informações necessárias, diminuindo as chances de dúvidas ou informações não coletadas (MARCONI e LAKATOS, 2003).

As entrevistas foram conduzidas de forma individual e *online*, com nove participantes<sup>21</sup>, com o perfil do usuário desejado para o *design* proposto neste estudo, ou seja, pessoas que realizam pesquisas científicas a partir de dados extraídos de mídias sociais. Os

<sup>21</sup> De acordo com Prates e Barbosa (2007), nos estudos de avaliação de interface e propostas de *design*, é recomendada a participação de cinco a nove usuários.

participantes foram selecionados a partir da rede de contatos da autora desta pesquisa com pesquisadores que trabalham com coleta e análise de dados de mídias sociais, em diferentes instituições de ensino superior no Brasil. Todos deveriam conhecer as etapas de coletar e minerar textos em redes sociais e e já ter atuado na realização dessas atividades em suas pesquisas científicas. Tal exigência foi estabelecida como uma forma de homogeneizar o perfil dos participantes e reduzir os fatores externos que poderiam influenciar suas respostas, seguindo-se as recomendações sobre perfis de alta definição, descritas pelo MEDS (NICOLACI-DA-COSTA; LEITÃO; ROMÃO-DIAS, 2004).

Antes de iniciar as entrevistas com os participantes selecionados para o estudo, foi realizado um teste piloto com um professor universitário, experiente e que atendia aos quesitos escolhidos para os perfis dos participantes, quando se pôde verificar a qualidade do material utilizado nas entrevistas. A participação de todos foi voluntária, com aceite prévio ao convite encaminhado, e todos leram e assinaram o Termo de Consentimento e Livre Esclarecimento (TCLE) antes do início da entrevista.

O roteiro da entrevista semiestruturada, disponível no Apêndice A, abordou 20 questões, abertas e fechadas, divididas em três blocos temáticos: o primeiro trouxe questões sobre informações de perfil educacional e profissional do participante; o segundo bloco levantou informações sobre a sua experiência computacional; e, por fim, o terceiro bloco buscou levantar a experiência e a satisfação dos usuários no processo de mineração de textos em RSOs, com foco na coleta de dados provenientes dessas mídias.

Também nesse bloco havia questões sobre como os participantes dão prosseguimento com os dados coletados no processo de mineração de textos e quais as suas dificuldades na realização dessas atividades. As questões foram voltadas para identificar as necessidades dos participantes e recomendações para a interface de uma nova ferramenta que permitisse realizar as suas atividades de forma mais eficiente e eficaz.

### **6.1.2. Análise e Interpretação dos Dados Coletados**

Após a realização das entrevistas, como preparação para análise dos dados, os áudios das entrevistas foram transcritos. A etapa de análise dos dados, conforme orientação do MEDS (NICOLACI-DA-COSTA; LEITÃO; ROMÃO-DIAS, 2004), foi realizada através da *análise de inter-sujeitos* e *intra-sujeitos*. A primeira consistiu em analisar as respostas dadas pelo grupo como um todo, onde todas as respostas de todos os participantes a cada uma das



perguntas foram analisadas sistematicamente e rigorosamente. A segunda consistiu em analisar cada entrevista individualmente, com as respostas de cada um dos entrevistados a todas as perguntas sendo abordadas como um único conjunto, dentro do qual foram analisados possíveis conflitos de opinião, sentimentos contraditórios, inconsistência entre as respostas e outros.

O MEDS propõe as categorias de análise para os resultados a partir das recorrências nos depoimentos dos participantes (NICOLACI-DA-COSTA, LEITÃO E ROMÃO-DIAS, 2004). Ou seja, diferentemente do caso de outras metodologias, que partem de categorias de análise previamente estabelecidas, no MEDS essas categorias emergem dos discursos dos entrevistados. Assim, após realizar de forma iterativa essas duas análises mencionadas anteriormente, a *inter-sujeitos* e a *intra-sujeitos*, foi possível realizar a identificação de recorrências comuns nos depoimentos dos participantes.

### 6.1.3. Apresentação dos Resultados

Os dados coletados e analisados, que refletem o conhecimento adquirido a partir dos usuários sobre as necessidades e requisitos para o *design*, precisam ser organizados, de forma que possam ser apresentados como resultados desta etapa.

Diversas representações são encontradas na literatura para auxiliar a registrar esse conhecimento adquirido dos usuários. Nesta pesquisa, optou-se por utilizar duas técnicas de representação: Personas e Cenários de Análise, descritas a seguir.

#### 6.1.3.1. Personas

Personas consistem em um “personagem fictício, arquétipo hipotético de um grupo de usuários reais, criado para descrever um usuário típico”. Rogers *et al.* (2013, p. 360) as definem como uma “síntese de vários usuários reais que estão envolvidos na coleta de dados” (CHOMA, 2015, p.12).

Uma persona pode ser definida pelas seguintes características (COURAGE E BAXTER, 2005, p.38):

- Identidade – informações para tornar a persona o mais realista possível, como nome e sobrenome, idade e foto;

- Status – nível de interesse do usuário ao sistema e o quanto ela pode deve influenciar as decisões do projeto, sendo classificada como primária, secundária, stakeholder ou antiusuário;
- Objetivos – listagem de todos o objetivos da persona, não somente os relacionados ao produto em específico, mas também os pessoais e corporativos;
- Habilidades – especialidade da persona, incluindo educação, treinamento e competências em relação ao produto em específico;
- Tarefas – atividades básicas e/ou críticas a persona realiza, incluindo a frequência de importância de cada uma delas;
- Relacionamentos – identificação de quais as pessoas e possíveis stakeholders a persona se relaciona;
- Requisitos – necessidades da persona; e
- Expectativas - a forma como a persona acredita que o produto deveria funcionar e como as informações no seu domínio ou trabalho deveriam ser organizadas.

Considerando tais características, as personas devem ser definidas com rigor e detalhes, pois “quanto mais específicas forem as personas, mais eficientes elas serão como ferramentas de *design*” (BARBOSA; SILVA, 2010, p. 177).

#### **6.1.3.1.1. Cenários de Análise**

Também conhecido como Cenários do Problema, estes consistem em uma história sobre pessoas realizando uma atividade, descrevendo o comportamento e as experiências dos atores. “É uma narrativa textual ou pictórica, concreta, rica em detalhes contextuais, de uma situação atual de uso da aplicação, envolvendo usuários, processos e dados reais ou potenciais” (BARBOSA e SILVA, 2010, p. 174).

Os elementos característicos de um cenário são, segundo Rosson e Carroll (2002):

- Ambiente ou contexto – são os objetivos, ações e reações dos atores motivados por alguma situação;
- Atores – pessoas que utilizam o sistema ou elementos do ambiente, normalmente são representados pelas Personas previamente determinadas;

- Objetivos – efeitos que motivam as ações realizadas pelos atores;
- Planejamento – comportamento ou conjunto de ações organizados a partir dos objetivos
- Ações – comportamento observável;
- Eventos – ações externas ou reações do computador ou de outras características do ambiente, ocultas ou não para o ator, mas que interferem no cenário; e
- Avaliação – atividade mental para interpretar uma situação.

Neste trabalho, foi elaborada uma narrativa sobre o contexto de uso, funções e tarefas dos usuários desta pesquisa, sendo os atores representados pelas personas anteriormente propostas.

## **6.2. Execução**

As entrevistas para a coleta de dados para esta etapa, realizadas em março de 2019, de forma *online*, contaram com nove participantes e tiveram duração média de 34 minutos. Para o recrutamento dos participantes, foi utilizada a técnica de amostragem bola de neve (COSTA, 2018), em que foram especificadas as características que os membros da amostra deveriam conter, no caso, trabalharem, possuírem interesse e/ou realizarem pesquisas coletando dados de redes sociais e aplicando a mineração de textos. Participantes com essas características foram convidados por meio da indicação de participantes anteriores. Dessa forma, a amostra foi autogerada, contando com a colaboração voluntária dos participantes iniciais e dos subsequentes. Os perfis dos entrevistados são descritos na Tabela 1.

Tabela 1- Perfil dos participantes do estudo

<b>ID</b>	<b>Idade</b>	<b>Ocupação</b>	<b>Formação</b>
P1	50	Professor de Ensino Superior	Computação
P2	34	Analista de Tecnologia da Informação	Computação
P3	33	Professora de Ensino Médio/Técnico	Educação
P4	28	Desenvolvedor	Computação
P5	28	Bolsista de Apoio Técnico	Comunicação Social
P6	28	Desenvolvedor	Computação
P7	34	Professor de Ensino Superior	Comunicação Social
P8	26	Bolsista de Iniciação Científica	Jornalismo
P9	40	Professor de Ensino Superior	Computação

Fonte: próprio autor

Todos os participantes eram docentes ou discentes de diferentes instituições de ensino do Brasil e atendiam a exigência previamente estabelecida de terem interesse e/ou trabalharem com pesquisa envolvendo coleta e análise de dados de redes sociais *online*. Tal exigência foi estabelecida como uma forma de homogeneizar o perfil dos participantes (LEITÃO e PRATES, 2017) e reduzir os fatores externos que poderiam influenciar suas respostas, seguindo as recomendações sobre perfis de alta definição, descritas pelo MEDS.

### 6.3. Resultados

O MEDS se baseia no pressuposto de que quanto mais recorrente for uma ideia identificada, maior a probabilidade de ter sido construída socialmente. Logo, as categorias para análise são estabelecidas a partir das recorrências comuns apresentadas nos discursos dos participantes. Assim sendo, foram identificadas cinco categorias, apresentadas a seguir.

#### **Categoria 1 – As interfaces das ferramentas existentes para coleta de dados de mídias sociais são difíceis de usar**

Buscou-se saber qual a percepção e satisfação dos usuários em relação às ferramentas de coleta existentes, com o intuito de verificar se os entrevistados as utilizavam

para a realização de coletas de dados em RSOs, e como era a sua experiência de uso. Assim, observou-se que os participantes apresentam, de forma geral, dificuldades não apenas para utilizar tais ferramentas, mas também para instalá-las e configurá-las, o que leva a um uso não efetivo das mesmas.

A maioria dos participantes relatou que teve dificuldade em utilizar as ferramentas devido às suas interfaces não serem amigáveis, como pode ser observado na fala do participante P03: *“a dificuldade existe porque todas as ferramentas carecem de uma interface gráfica mais amigável, de forma que o usuário consiga identificar onde exatamente ele pode fazer as buscas e as coletas de forma mais objetiva”*. O participante P06, mesmo sendo um profissional da área de computação o que lhe atribui a característica de possuir maior familiaridade com as funcionalidades apresentadas, concorda ao afirmar que: *“a interface das ferramentas no geral eram somente campos de busca e filtro com alguns botões soltos, não documentados e bastante confusos”*.

Logo, a utilização dos recursos disponibilizados pelas ferramentas de coleta fica comprometida, isso porque pouca informação é transmitida aos usuários através de sua interface, e, quando o são, muitas vezes é de maneira confusa.

Alguns respondentes chegaram a falar sobre isso, mostrando sua insatisfação, como pode ser observado nos dizeres do entrevistado P04: *“Nas ferramentas gratuitas que utilizei, não tem uma parte da interface que eu consiga te dizer que gostei, porque eram interfaces simples, mas não tão limpas e intuitivas assim. Existem algumas opções que ficam bem confusas na tela da coleta, não possuindo uma opção de menu”*. O participante P07 também relata essa dificuldade: *“acho que também podem causar algumas barreiras para nossa utilização é a confusão de informações. Assim, a gente tem uma série de informações de abas que podem ser acessadas, então a informação está muito solta ali na tela e a gente tem que ver exatamente o que funciona para o nosso caso, porque nem sempre todos os recursos ali dispostos serão utilizados na sua coleta, então essa a maior dificuldade.”*.

As dificuldades ocorrem também na instalação e configuração das ferramentas. O participante P06 expressa esta percepção: *“a interface não é amistosa, assim gera um distanciamento, uma tensão quando você abre ela, porque você precisa de um servidor, precisa conectar e configurar este servidor e saber como configurar isso, então fico muito perdida quando eu estou nele sozinha, as informações dela não te conduzem, na verdade você que tem que se preocupar em saber onde você vai inserir cada coisa ali pra você fazer a coleta.”* Esta dificuldade fica clara também nas falas do participante P07: *“a principal dificuldade que eu enfrentei, e inclusive precisei de auxílio do colega X nisso, é a*

*configuração. Porque como muitas vezes somos de áreas que não tem tanto conhecimento na área de programação, a gente acaba ficando refém das configurações básicas da plataforma, e quando a gente precisa fazer qualquer tipo de modificação ou adaptação para o nosso projeto, temos uma dificuldade enorme em resolver isso”.*

Sabe-se que ferramentas de coletas de dados apresentam funcionalidades complexas, e tal complexidade é reconhecida pelos participantes. Entretanto, a baixa qualidade das interfaces dessas ferramentas pode impedir os usuários de executarem a coleta de dados de forma eficiente, deixando-os frustrados e insatisfeitos, como pode ser visto na fala do participante P07: *“eu entendo o grau de complexidade das plataformas e que por isso a gente precisa realmente ter ali uma ajuda no processamento destes dados, mas acho muito complicado e frustrante quando você pega uma ferramenta que tem que ter ali uns dois tutoriais complementares para você conseguir entender o funcionamento e realmente conseguir chegar ao resultado que gostaria. A interface da ferramenta deveria me fornecer mais informações, ser mais intuitiva”.*

Esta dificuldade de uso pode contribuir para a não efetividade da coleta de dados nas mídias sociais, ao não permitir que o usuário utilize a ferramenta para atingir plenamente o seu objetivo. Isto pode ser observado na fala do participante P03: *“O usuário não sabe o que a ferramenta oferece, coisa que se tivesse a interface amigável possibilitaria o uso de todos os recursos que ela disponibiliza”.*

### **Categoria 2 - A coleta de dados em mídias sociais é uma tarefa voltada mais para profissionais da área de Computação, com conhecimento em programação de computadores**

Ao serem questionados sobre quais as ferramentas mais utilizam para realização das coletas, a maioria dos participantes relatou que preferem muitas vezes não utilizar nenhuma. Assim, o desenvolvimento de scripts foi mencionado como uma das formas mais utilizadas pelos participantes para a coleta de dados provenientes das mídias sociais.

Como a criação de scripts exige conhecimentos específicos relacionados à lógica e linguagens de programação, usuários leigos em computação, mas que precisam coletar dados de RSOs para realizar suas pesquisas, ficam prejudicados na execução desse tipo de tarefa.

Os próprios participantes manifestaram o seu entendimento de que a realização de coleta de dados de mídias sociais é atualmente mais voltada para pessoas que possuem conhecimento no mínimo intermediário em programação de computadores e outras abordagens da ciência da computação. Isto pode ser confirmado pelo discurso do participante P09: *“Acredito que o que fiz durante o doutorado não seria viável para uma pessoa que não*

*tivesse a formação e os conhecimentos em computação e programação (...) seria necessário criar uma interface e formas mais simples de um leigo em computação realizar as coletas*". A fala de P05 também reflete a necessidade de conhecimentos específicos em programação para fazer uso dos scripts na coleta de dados: *"no trabalho com scripts temos problemas típicos de scripts, como instalar biblioteca, pedaços de código faltando(...) enfim, coisas de programação mesmo que nem todos têm conhecimento"*.

A partir do discurso dos participantes, pudemos observar que a utilização de scripts para a coleta de dados é consequência justamente da baixa qualidade de uso das ferramentas existentes. P05 destaca isso em sua fala: *"Então, no final, o grosso da pesquisa eu acabei utilizando um punhado de scripts em Python que eu mesmo desenvolvi com um amigo, porque precisava dados que a ferramenta não me fornecia"*.

Alguns participantes foram além da coleta de dados, ao mencionarem que utilizam scripts também nas demais etapas do processo de mineração de textos, como pode ser visto na fala de P09: *"Os dados eram coletados através por scripts e também através deles eu realizava a limpeza dos dados coletados inválidos, incompletos, que não nos atendiam ou tinham relação com a pesquisa. Depois utilizava alguma técnica para tratar os dados, como Aprendizagem de Máquina, Mineração de Textos ou de Dados, Processamento de Linguagem Natural e outras, mas em nenhuma destas eu utilizei ferramenta, fazia tudo através de código"*.

### **Categoria 3 - O uso de diferentes ferramentas para as diferentes etapas do processo de, mineração de textos em mídias sociais o torna mais complexo**

São utilizadas ferramentas distintas para a coleta, mineração, tratamento, visualização e análise dos dados provenientes das mídias sociais, a fim de que se possa extrair conhecimento válido a partir das informações coletadas.

Neste processo de mineração de textos fazendo uso de diferentes ferramentas, os participantes relataram que encontram várias dificuldades. Além dos problemas em relação à baixa qualidade das interfaces dessas ferramentas, há ainda incompatibilidade entre as diferentes ferramentas utilizadas no processo de mineração de texto, com diferentes formatos dos dados a serem exportados e importados por elas, o que restringe a execução do processo a profissionais experientes, com conhecimentos específicos em Computação, voltando ao que foi exposto na categoria anterior.

O participante P01 apresenta claramente esta dificuldade em sua fala: *"tratávamos os dados, mas não em ferramentas; coletávamos e fazia isso através de planilhas ou ferramentas externas, separadas. Caímos naquele problema de termos que usar*

*muitas ferramentas para se chegar a um resultado.”. O participante P06, que tem formação na área de computação e atua apoiando outros pesquisadores no processo de mineração de textos em RSOs, revela esta dificuldade em sua fala: “Converter os dados tabulados em dados que pudessem ser importados e visualizados exigiam um conhecimento técnico muito grande que algumas pessoas que não são da área de computação ou exatas tinham bastante dificuldade pra poder realizar, então dependia muito de mim converter estes dados para que eles pudessem partir para o que eles tinham conhecimento”.*

Tais dificuldades apontam a importância de se criar ferramentas únicas, que permitam ao pesquisador executar todos os passos do processo de mineração de textos, ou seja, coleta, mineração, tratamento, visualização e análise dos dados provenientes das mídias sociais, de forma integrada e em um único local. Tal possibilidade é vislumbrada na fala de P01: *“Acho essa parte essencial, podermos dentro da mesma ferramenta que a gente coleta ver os resultados é muito interessante. Agrupar tudo em um único local (...) mesmo a gente que possui conhecimento, agregar todas estas ferramentas e daí tirar um resultado é muito difícil”.*

**Categoria 4 - Novas funcionalidades deveriam ser adicionadas às ferramentas para proporcionar maior precisão na coleta de dados e demais atividades do processo de mineração de textos, além de maior flexibilidade aos usuários**

Quando perguntados sobre o que consideram essencial em ferramentas que dão suporte ao processo de coleta e mineração de textos em RSOs, a maioria dos entrevistados apontou a inclusão de funcionalidades que os atenderiam de forma mais plena em relação a suas pesquisas nesses ambientes.

A funcionalidade principal e mais mencionada pela maioria dos respondentes foi relacionada à temporalidade das coletas, ou seja, a possibilidade de se realizar coletas em tempo real e coletas retroativas em uma mesma ferramenta. Isto é mostrado no discurso do participante P01: *“Para a área de coleta em redes sociais a funcionalidade que eu acho essencial (...) é a questão do tempo de coleta, em tempo real ou retroativa. Esta funcionalidade é muito importante”.* Em complemento a essa funcionalidade, a possibilidade de disponibilizar um preview dos dados para análise, ainda em tempo de coleta, também foi bastante mencionada. O entrevistado P06, por exemplo, fala sobre isso: *“(...) forneceria alguns tipos de análise já em tempo real, por exemplo exibição de gráficos ou número de palavras”.*

Outra característica bastante mencionada para uma nova ferramenta foi a possibilidade de realizar diversas configurações para a coleta, a partir de opções oferecidas ao



usuário sobre os dados que ele almeja coletar, proporcionando-lhe maior flexibilidade nesse processo. O participante P07, por exemplo, aponta utilizar vários tipos de termos diferentes para coleta: *“uso muito a extração via palavras-chave ou então hashtags, então é importante demais eu conseguir dentro de uma mesma coleta eu coletar várias opções configuráveis de hashtags e palavras-chave”*. O participante P09, por sua vez, menciona a possibilidade de se trabalhar com diferentes idiomas: *“queria que me oferecesse a opção de escolha da língua do texto que eu quero coletar, por exemplo, somente twittes em Português ou Inglês, ou em qualquer língua”*. P07 vislumbra a possibilidade de realizar pesquisas mais amplas: *“pra mim é importante demais eu conseguir dentro de uma mesma coleta eu coletar várias hashtags e várias palavras chave, eu ter a possibilidade de colocar uma interseção ou união entre os termos de coleta”*. Já o participante P05 fala sobre a possibilidade do usuário poder direcionar a coleta para um subconjunto específico dos dados: *“(...) você pode selecionar, por exemplo, que parte do banco de dados você quer, isso é bem interessante, porque te permite não utilizar o banco de dados inteiro, caso não precise”*.

Além dessa flexibilidade na coleta de dados, outra necessidade apontada pelos participantes para uma nova ferramenta foi a inclusão de pré-processamento dos termos de coleta e tratamento dos dados coletados, a fim de proporcionar uma maior precisão na coleta e análise dos dados.

O pré-processamento dos termos da coleta diz respeito a obter versões desses termos sem acentuação, cedilha, etc., para então submetê-los na coleta. O participante P04 destaca esta necessidade: *“tratar os termos para a coleta, como um pré-processamento, por exemplo, acentuação de palavras, re-twittes, cedilha, til e outros, já serem resolvidos durante a coleta”*. Durante as entrevistas, a funcionalidade mais mencionada em relação ao tratamento dos dados coletados foi a organização dos dados no momento da extração, como mostra a fala do participante P06: *“(...) extrair esses dados tabulados e também extrair outros, como nuvem de palavras e diversos tipos de grafos, ou até mesmo organizados em diretórios”*.

### **Categoria 5 - Melhorias na interface das ferramentas de mineração de textos de mídias sociais para proporcionar uma melhor interação com o usuário**

Quando perguntados sobre que melhorias fariam na interface das ferramentas que dão apoio ao processo de mineração de textos em RSOs, os participantes foram unânimes ao afirmarem que as tornariam de fácil uso, limpas e intuitivas, características estas essenciais para que possam ser mais efetivos em suas pesquisas. Isto se deu tendo em vista as dificuldades apontadas pelos participantes, e descritas na categoria 5.1, para utilizarem as ferramentas de coleta de dados existentes. Os dizeres dos participantes P03 e P04 resumem o

que foi dito pelos demais entrevistados: *“as melhorias seriam realmente uma interface que facilitasse o uso da ferramenta”*(P03) e *“melhorias relativas ao design da tela, a deixaria mais limpa e intuitivo, com menos caminhos a se seguir para conseguir chegar a uma opção de configuração ou da realização da própria coleta em si”* (P04).

O participante P09 menciona uma característica interessante, que diz respeito a ser disponibilizado ao usuário opções de ajuda e documentações melhor elaboradas: *“nas interfaces existentes praticamente não tem ícones, mas também com certa frequência quando colocam ícones fica difícil de entender o propósito dele. A algumas melhorias estéticas seriam interessantes, já que algumas são bem esquisitas de usar; também alguma documentação um pouco melhor”*.

#### **6.4. Discussão**

A partir da análise dos dados coletados com as entrevistas, conseguimos identificar as principais dificuldades e necessidades dos participantes em relação às ferramentas utilizadas no processo de mineração de textos em RSOs.

Inicialmente, foi possível observar que os participantes do estudo conhecem as ferramentas disponíveis e entendem o quanto elas são importantes para apoiá-los no processo de coleta, que envolve um grande volume dados a serem processados para se conseguir obter informação e conhecimento sobre eles. Percebeu-se que todos os participantes sabem o que desejam coletar e o que pretendem analisar em um conjunto de dados, o que os leva a ter uma visão mais crítica sobre tais ferramentas.

Dessa forma, sobre as dificuldades enfrentadas pelos pesquisadores de dados durante a mineração de textos em RSOs, tomando como ponto de partida a coleta de dados, identificamos que tais dificuldades estão relacionadas primordialmente à baixa qualidade de uso das interfaces das ferramentas, que impactam negativamente a interação dos usuários.

Tal fato leva os pesquisadores de dados a até mesmo desistirem de utilizar tais ferramentas, optando muitas vezes por utilizarem *scripts* para realizarem suas atividades. Não que este caminho seja fácil, principalmente para aqueles que não possuem formação na área de computação, mas nossos entrevistados relatam que acham que o custo de aprender sobre linguagens de programação e alguns aspectos específicos de computação, necessário para utilizarem scripts, embora seja alto, ainda é menor do que o custo relacionado ao uso das ferramentas existentes.

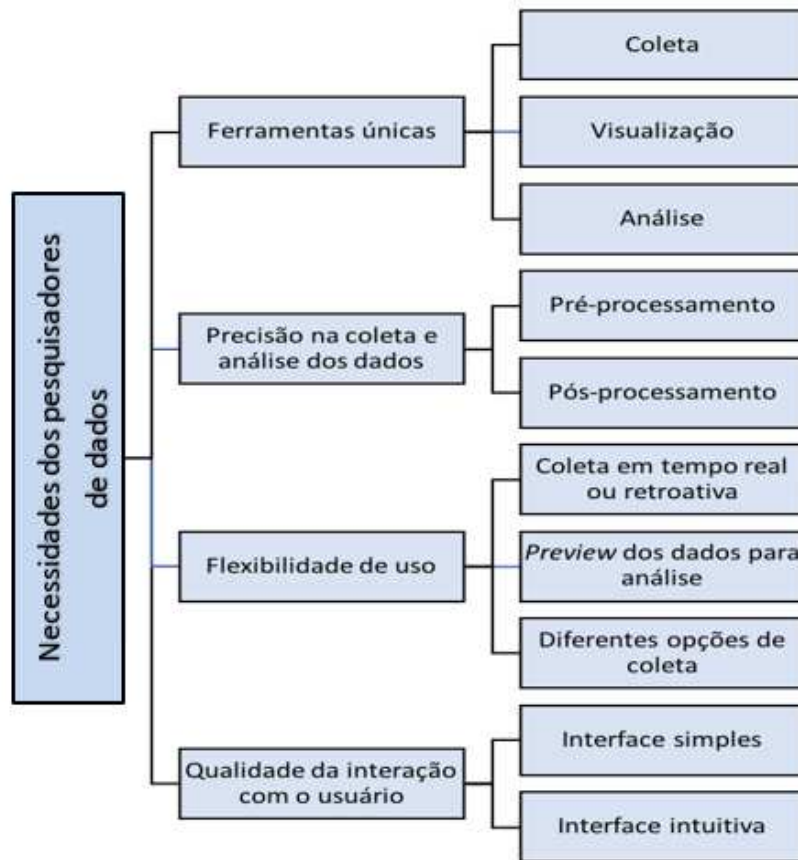
Isto também pode fazer com que as atividades do processo de mineração de textos nas mídias sociais fiquem muitas vezes restritas a profissionais com conhecimentos mais específicos de programação e Computação. Por outro lado, fica mais difícil a realização de tais atividades por parte de profissionais de outras áreas, como é o caso de jornalistas e profissionais de publicidade, dentre outros, que possuem a necessidade de analisar dados provenientes das mídias sociais.

Além disso, considerando-se todo o processo de mineração de textos em RSOs, que envolve, além, da coleta, o tratamento, a visualização e análise dos dados coletados, problemas de compatibilidade entre as diferentes ferramentas utilizadas para esses propósitos dificulta ainda mais a interação dos usuários.

Tais dificuldades frustram os pesquisadores de dados, impactando negativamente na efetividade de suas pesquisas nas mídias sociais e gerando, conseqüentemente, insatisfação em relação à forma como coletam e analisam os dados. A fim de superar essa insatisfação, no sentido de amenizar suas dificuldades no processo de coleta e análise de dados, buscou-se identificar suas necessidades durante o processo de mineração de textos em RSOs, tomando como ponto de partida a coleta de dados.

A partir do discurso dos participantes deste estudo, identificamos quatro necessidades, mostradas na Figura 6, que devem ser atendidas a fim de tornar as atividades do processo de mineração de textos em RSOs mais efetivas: (1) Utilizar ferramentas únicas para tratar, de maneira integrada, todas as etapas do processo de mineração de textos; (2) Proporcionar maior precisão na coleta e análise dos dados; (3) Proporcionar maior flexibilidade aos pesquisadores de dados no processo de coleta e análise dos dados de mídias sociais; e (4) Qualidade da interação com o usuário.

Figura 6- Necessidades dos pesquisadores de dados para ferramentas de coletas de dados de mídias sociais



Fonte: próprio autor

Podemos observar que tais necessidades refletem a importância de levar em consideração aspectos de qualidade de uso no *design* de ferramentas a serem utilizadas pelos pesquisadores no processo de mineração de textos.

De forma específica, o desenvolvimento de ferramentas únicas para atender todas as atividades desse processo, a precisão na coleta e análise dos dados e a flexibilidade proporcionada aos pesquisadores durante a coleta dos dados remetem à “eficiência de uso”, relacionada à produtividade do usuário ao fazer uso do sistema para realizar as suas atividades (NIELSEN, 1993)

Assim, ao oferecer ao usuário uma única ferramenta que o permita executar todas as tarefas da mineração de textos, ou seja, coleta, tratamento, visualização e análise de dados de forma integrada, ele poderá atingir os seus objetivos de forma mais rápida, uma vez que não terá que fazer uso de diferentes ferramentas e, conseqüentemente, não terá o custo de aprender a utilizá-las e nem de converter dados para tratar questões de incompatibilidade entre elas.

Além disso, ao oferecer aos usuários formas automatizadas de pré-processamento dos termos de coleta, e tratamento dos dados coletados, a precisão na coleta e análise dos dados será assegurada pela própria ferramenta, poupando o usuário de ter que realizar essas tarefas de forma manual ou utilizando outras ferramentas externas. Tal fato permitirá que o usuário também realize suas tarefas de forma mais ágil e eficiente.

Já ao proporcionar maior flexibilidade ao pesquisador de dados em relação à configuração da forma como será realizada a coleta, a ferramenta está dando um maior controle ao usuário no sentido de obter os dados de forma mais ampla ou específica, conforme necessário. Assim, o usuário já obtém diretamente os dados desejados, sem ter que manipulá-los ou fazer coletas adicionais para atender às suas necessidades específicas.

A qualidade da interação com o usuário dessas ferramentas, proporcionada por interfaces simples e intuitivas, remete aos fatores de usabilidade “facilidade de aprendizado” e “facilidade de lembrar como se usa” (NIELSEN, 1993). Tal fato é bastante evidente, pois ao interagir com uma interface amigável, o usuário não terá dificuldades para aprender a utilizar ferramenta e nem de se lembrar como realizar o acesso às suas funcionalidades, após ficar um tempo sem utilizá-la.

Por fim, o atendimento das necessidades dos pesquisadores em relação às ferramentas que o auxiliem em seu trabalho de mineração de textos em RSOs, mostradas na Figura 12, irá proporcionar-lhes uma maior satisfação de uso. Assim, a qualidade de uso irá levar a uma positiva experiência de usuário, relacionada às emoções e sentimentos despertados no usuário enquanto interage com o sistema (SHARP; ROGERS; PREECE, 2007).


A importância de se levar em consideração a experiência de usuário no *design* de ferramentas que apoiam as atividades do processo de mineração de textos em RSOs pode ser justificada pela identificação, a partir do discurso dos participantes deste estudo, de que as dificuldades por eles enfrentadas no uso das ferramentas existentes são o principal fator desmotivador para que possam efetivamente adotá-las em suas pesquisas.

Para a apresentação dos resultados obtidos a partir da análise realizada, foram criadas duas personas para representar os grupos de usuários das ferramentas que apoiam o processo de mineração de textos em RSOs, de acordo com suas percepções, experiências de uso e necessidades. A persona 1, apresentada na figura 7, representa pesquisadores de dados provenientes da área de Computação, com conhecimento em linguagens de programação e afins. Estes possuem conhecimento sobre as ferramentas computacionais que apoiam o processo de mineração de textos em redes sociais, mas, na maioria das vezes, preferem

desenvolver scripts para realizar suas atividades. Suas principais necessidades estão relacionadas a ter uma maior precisão na coleta e análise dos dados e uma maior flexibilidade de uso ao executarem essas atividades. Entretanto, também desejam ter uma interação de qualidade com as ferramentas que os apoiam em seu trabalho.

A persona 2, mostrada no figura 8, representa pesquisadores que possuem pouco ou nenhum conhecimento em Computação, mas possuem necessidade de obter conhecimento a partir das mídias sociais para realizarem o seu trabalho. Estes possuem dificuldade em utilizar as ferramentas existentes e precisam da ajuda de profissionais da área de Computação para realizar a instalação e configuração das ferramentas, e inclusive para auxiliá-los no processo de coleta, tratamento e visualização dos dados. Suas necessidades são basicamente todas aquelas mostradas na Figura 6.

Figura 7- Persona 1: Pesquisadores com conhecimentos na área de computação




Nome: Caetano Leite  
 Idade: 36 anos

Caetano faz doutorado no Programa de Pós-graduação em Ciência da Computação, na Universidade de Diamantina. Ele realiza pesquisas a partir de base de dados textuais, coletando e analisando dados em redes sociais. Por ser bolsista do laboratório do programa, ele ainda apoia colegas e professores da instituição, que trabalham com esta mesma finalidade. As coletas são realizadas, na maioria das vezes, através de linguagens de comandos (scripts) que ele mesmo desenvolve, pois, a seu ver, embora a utilização de ferramentas com interface gráfica pudesse lhe proporcionar uma experiência de uso mais positiva, elas atualmente não o apoiam de maneira satisfatória. Caetano prefere aprender fazendo do que assistindo cursos ou lendo extensos manuais. A cada coleta, ele precisa de deslocar ao laboratório da universidade, seja para acompanhar algum processo de coleta em andamento, seja para desenvolver scripts ou fazer alguma configuração em diferentes ferramentas para visualizar os dados coletados e tratá-los, para então conseguir obter alguma informação relevante a partir deles. Caetano gostaria que todas estas atividades pudessem ser realizadas em uma única ferramenta. Ele entende a complexidade das atividades nesta área, e não se importa em ter que realizar configurações pontuais para adequar à sua necessidade, mas gostaria de ter ferramentas com uma interface mais simples e intuitiva, com orientações sobre as funcionalidades disponibilizadas, permitindo a realização das atividades de forma eficiente, satisfatória e agradável.

Fonte: próprio autor

Figura 8- Persona 2: Pesquisadores que não possuem conhecimento na área de computação




Nome: Paula Braga  
 Idade: 32 anos

Paula possui graduação em Jornalismo e atualmente é estudante de pós-graduação em comunicação social. Ela realiza pesquisas a partir de base de dados textuais, coletando e analisando dados em redes sociais. Paula utiliza algumas ferramentas para o processo de coleta e mineração dos dados, mas muitas vezes suas necessidades não são atendidas. Às vezes, ela as utiliza apenas para auxiliá-la no processo de análise, o que considera um empecilho, pois acredita que a utilização de ferramentas com interface gráficas, em todas as etapas do processo de coleta e mineração, seria mais eficiente para as suas necessidades e que, se possuísse formação em uma área que lhe desse suporte computacional, suas dificuldades seriam minimizadas. A cada coleta, ela precisa se deslocar ao laboratório da universidade, seja para acompanhar algum processo de coleta em andamento, seja para encontrar algum colega que apoie no suporte computacional de configuração das ferramentas, servidores ou até mesmo no desenvolvimento de scripts para que ele consiga prosseguir com sua pesquisa. Ela realmente fica frustrada por ter que depender de outras pessoas da área de computação. Assim, gostaria de ter tempo para aprender a configurar e utilizar as ferramentas existentes, porque acredita que poderia concluir sua pesquisa do mestrado com maior agilidade e confiança no que vem realizando.

Fonte: próprio autor

Foram criados também cenários de análise para as tais personas, a partir do conhecimento obtido sobre como as atividades dos usuários são realizadas atualmente, como pode ser visualizado nas figuras 9 e 10 a seguir.

Figura 9- Cenários de Análise dos pesquisadores que possuem conhecimento na área de computação (Persona 1)



Caetano Leite, discente do Programa de Pós-graduação em Educação da Universidade de Federal de Diamantina – UFDtna, precisa coletar dados em redes sociais para o desenvolvimento de sua pesquisa de mestrado e, por ser bolsista, recebeu a função de também coletar para apoiar outras pesquisas científicas desenvolvidas por colegas e professores da instituição. Caetano precisa escolher quais ferramentas existentes poderão apoiá-lo nestas tarefas. Por já conhecer a área de coleta e mineração em redes sociais, ele não perde muito tempo e já começa a planejar quais etapas do processo ele poderá utilizar tais ferramentas, quais ferramentas possuem suporte aos formatos dos arquivos gerados em casa etapa do processo e em quais etapas ele precisará desenvolver scripts para executar suas atividades de maneira mais eficiente.

Assim, em um primeiro momento, ele se preocupa em conseguir instalar e configurar as ferramentas que irão realizar as coletas, pois sabe das dificuldades para este processo. Passados alguns dias, Caetano então realiza, através de seus scripts, as suas primeiras coletas em redes sociais, no caso iniciou os trabalhos utilizando o *Twitter*. Porém, os dados coletados foram exportados completamente desorganizados e sem sentido algum, impossibilitando sua interpretação. Neste momento, ele decide então desenvolver mais algumas linhas de comandos para solucionar este problema e começar a mineração de textos em sua base de coleta, seja a partir de algum princípio matemático, estatístico, computacional, lógico ou outros.

Assim que termina esta etapa, consegue visualizar os dados minerados de uma forma mais organizada, porém, devido à grande quantidade de informações, ele ainda não consegue obter conhecimento a partir deles. Logo, ele instala e configura ferramentas específicas de visualização e análise de dados, sempre se preocupando em utilizar ferramentas sejam compatíveis com aquelas utilizadas nas etapas anteriores do processo de mineração de textos, para evitar mais transtornos no processo.

Ele então, após muitos empecilhos e trabalho árduo, consegue analisar os dados coletados e continuar sua pesquisa de mestrado. Porém Caetano continua apoiando os colegas pesquisadores, repetindo todo o processo ao adequar às necessidades específicas de cada um deles.

Fonte: próprio autor

Figura 10- Cenários de Análise dos pesquisadores que possuem conhecimento na área de computação (Persona 2)

Paula Braga, formada em Jornalismo e discente do Programa de Pós-graduação em Educação da Universidade de Federal de Diamantina – UFDtna, precisa coletar dados em redes sociais para o desenvolvimento de sua pesquisa de mestrado. Como nunca havia trabalhado com ferramentas de coleta, Paula precisa pesquisar quais ferramentas irá apoiá-la nesta tarefa. Logo que inicia suas atividades, ela descobre que não existem ferramentas completas que irão apoiá-la em todo o processo de coleta, mineração e visualização dos dados.



Ela decide então ir por partes, se preocupando inicialmente em conseguir instalar e configurar as ferramentas que poderão ajudá-la em algum momento neste processo. Passados alguns dias, percebe que esta é uma tarefa árdua e como segunda opção, para não perder mais tempo, decide estudar sobre como desenvolver scripts, aprendendo uma linguagem de programação e um pouco sobre sistemas de computação, para conseguir implementar suas próprias linhas de comandos para realizar as coletas que necessita; sua sorte é ter amigos que são programadores e se colocaram à disposição para ajudá-la. Após muitos dias de tentativas e erros, ela consegue instalar as ferramentas de coleta de dados e começa a testá-las, mas



logo desiste, pois percebe que suas interfaces são confusas e nada intuitivas, atrasando ainda mais sua pesquisa.

Após alguns meses do início do trabalho na área, Paula então realiza, através de seus scripts, as suas primeiras coletas em redes sociais, no caso iniciou os trabalhos no *Twitter*. Porém os resultados obtidos não foram os esperados: os dados coletados foram exportados completamente desorganizados, impossibilitando sua interpretação. Neste momento, ela decide então desenvolver mais algumas linhas de comandos para solucionar este problema. Assim que conseguiu visualizar os dados de uma forma mais organizada, ele começa a se preocupar em como realizará a mineração de textos em sua base de coleta. Paula começa então a estudar princípios matemáticos e alguns algoritmos e decide desenvolver algumas linhas de comando para realizar tal atividade, porém possui muita dificuldade e recorre a colegas da área de computação para auxiliá-la neste processo.

Passados algum tempo, após realizada a mineração, Paula precisa então exportar os arquivos minerados para alguma ferramenta que a auxiliará no processo de visualização e análise. Ela configura e instala ferramentas com esta finalidade, contudo tão logo que começa a utilizá-las percebe que compreendê-las não será uma tarefa fácil. Paula entende a complexidade da área que está pesquisando e a dificuldade de interpretar os dados coletados, mas não imaginava ter tantas dificuldades justamente por não possuir formação que o ampare nesta tarefa. Paula então, durante mais alguns meses, desenvolve scripts para analisar os dados que coletou anteriormente e obter, a partir deles, os conhecimentos necessários. Ela então, após muitos empecilhos e trabalho árduo, consegue dar prosseguimento em sua pesquisa de mestrado.

Fonte: próprio autor

## 7. DESIGN DO FRAMEWORK ORÁCULO

Este capítulo apresenta o projeto da intervenção, ou seja, o *design do framework Oráculo*. As atividades de *design* começam de forma iterativa, após os requisitos terem sido estabelecidos (ROGERS; SHARP; PREECE, 2013). Na etapa anterior, o foco era a análise da situação atual. Já esta atividade “concentra-se no projeto de intervenção, realizado através do *design* do sistema computacional interativo, visando apoiar melhor os usuários no alcance de seus objetivos” (BARBOSA; SILVA, 2010, p.208).

A seguir, na seção 7.1, apresentamos as considerações metodológicas abordadas na criação do *design* conceitual e do *design* físico do *framework*. Em seguida, na seção 7.2, são descritos os artefatos produzidos em cada uma dessas atividades.

## 7.1. Considerações Metodológicas

Para Rogers *et al.* (2013), o *design* é uma atividade central na criação de um produto interativo, que pode ser desmembrada em duas subatividades: *design* conceitual, que envolve a produção de um modelo conceitual do produto, que descreve o que este deve fazer e como deve se comportar; e o *design* físico, que considera com o que o sistema deverá se parecer e quais os detalhes do *design*, como as estruturas de tela e menu, ícones e gráficos.

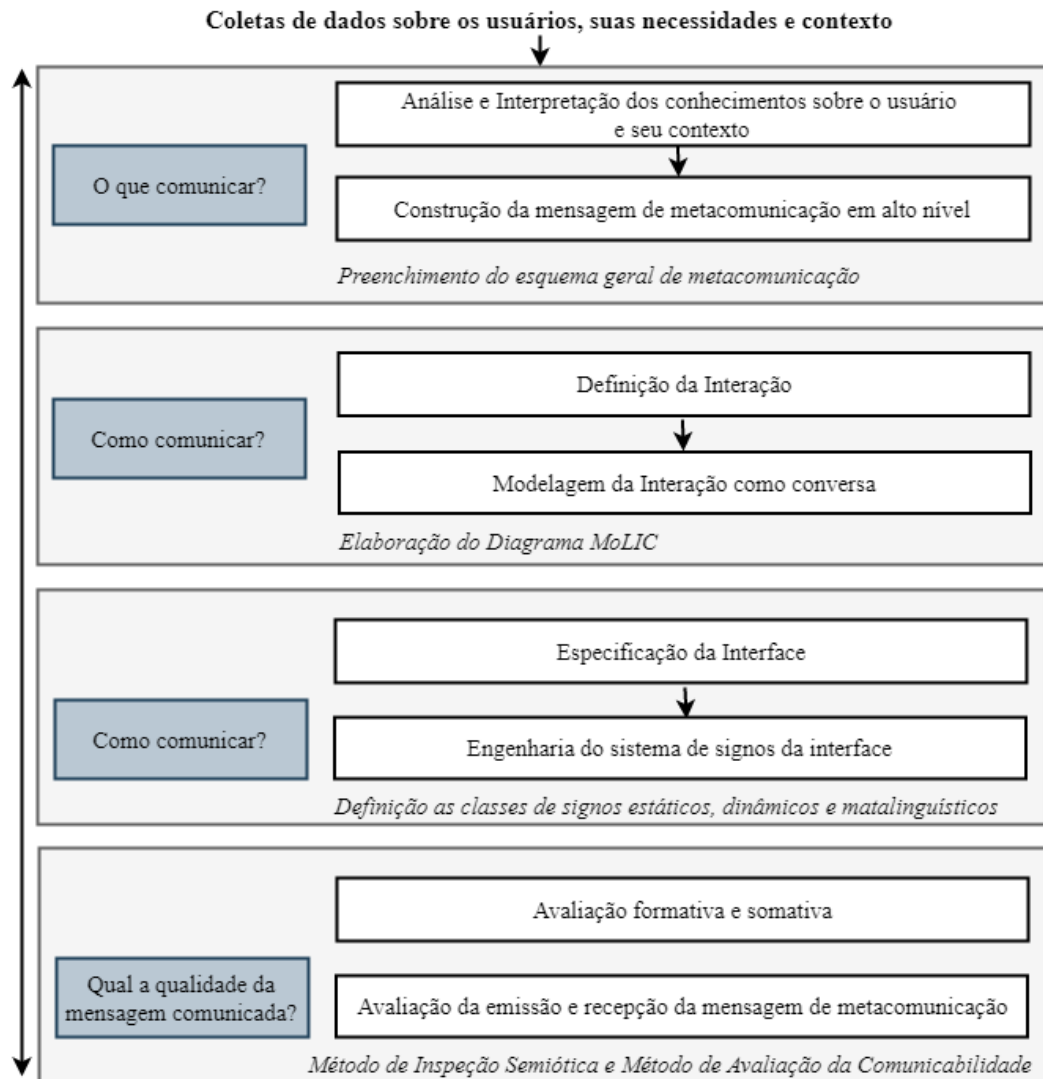
A abordagem de *design* utilizada neste trabalho, tanto para o *design* conceitual quanto para o *design* físico, é o *Design Centrado na Comunicação* (DE SOUZA, CLARISSE S., 2005), que ressalta a importância do papel ativo do *designer* na interação usuário-sistema. Tal abordagem é descrita a seguir.

### 7.1.1. *Design Centrado na Comunicação* (DCC)

O *Design Centrado na Comunicação* (DCC) tem como base teórica a Engenharia Semiótica (DE SOUZA, 2005). Assim, o objetivo do DCC é elaborar uma solução de *design* que transmita a metacomunicação do *designer* de forma eficiente e eficaz, produzindo um sistema interativo com alta comunicabilidade. Para isso, esse processo coloca o *designer* como um dos interlocutores das conversas que ocorrem durante a interação do usuário com o sistema (BARBOSA; SILVA, 2010)

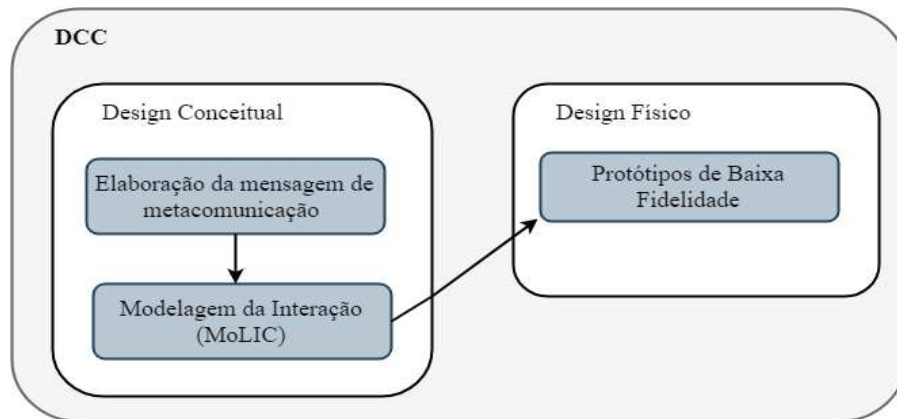
No DCC, é responsabilidade do *designer* comunicar aos usuários sua visão de *design*, dando-lhes melhores condições de entender e aprender sobre o sistema projetado. O DCC faz uso de conceitos, métodos e modelos propostos pela Engenharia Semiótica em suas etapas, conforme mostrado na Figura 11.

Figura 11- Etapas do Design Centrado na Comunicação



Fonte: próprio autor. Adaptado de Leitão *et al.*, (2013)

Assim, dentro da etapa de *design*, no presente trabalho, foram realizadas as sub-etapas de Construção da mensagem de metacomunicação de alto nível e a Modelagem da Interação como uma Conversa, no *design* conceitual, e Engenharia de Sistemas de Signos da Interface, no *design* físico, como mostra a Figura 12.

Figura 12- Atividades de *Design*

Fonte: próprio autor

Estas atividades são detalhadas na subseções a seguir.

#### 7.1.1.1. *Design* Conceitual

O *design* conceitual consiste na criação do modelo conceitual para o produto. Tal modelo é uma abstração que descreve o que as pessoas podem fazer com o produto e quais conceitos necessários para entender como interagir com ele (ROGERS; SHARP; PREECE, 2013). Tal descrição emerge dos requisitos identificados na etapa de identificação das necessidades dos usuários e definição dos requisitos, e os conceitos dependem de quem será o usuário, que tipo de interação e interface serão utilizadas.

O primeiro passo na obtenção de uma visão concreta de um modelo conceitual consiste em submergir nos dados coletados sobre os usuários e seus objetivos e tentar compreendê-los. Após o desenvolvimento de um modelo conceitual inicial, suas ideias iniciais devem ser pensadas com mais detalhes e expandidas, antes de serem prototipadas e testadas com usuários. Assim, nesse momento, deve-se tomar decisões sobre quais funções o produto irá executar (ou quais o usuário irá executar), como essas funções estão relacionadas e que informações são necessárias para lhes dar suporte. Essas decisões são tomadas inicialmente como tentativas e podem mudar após a prototipação e avaliação (ROGERS; SHARP; PREECE, 2013). Assim, no presente estudo, o *design* conceitual consistiu das seguintes etapas do DCC: Construção da mensagem de metacomunicação de alto nível e Modelagem da Interação como uma Conversa, detalhadas nas subseções a seguir.

#### 7.1.1.1.1. Construção da mensagem de metacomunicação de alto nível

O conteúdo da mensagem de metacomunicação de alto nível é estruturado considerando o conhecimento obtido a partir da coleta de dados sobre os usuários, suas necessidades e contexto, realizado na etapa de identificação das necessidades e requisitos de IHC.

Assim, deve ser preenchido o esquema geral da metacomunicação (*template* de metacomunicação do *designer* para o usuário), apresentado a seguir, privilegiando-se a abrangência do conteúdo em detrimento do detalhamento:

*“Esta é a **minha** interpretação sobre quem **você** é, o que **eu** entendi que **você** quer ou precisa fazer, de que formas prefere fazê-lo e por quê. Este é portanto o sistema que **eu** projetei para **você**, e esta é a forma que **você** pode ou deve usá-lo para atingir objetivos alinhados com a **minha** visão”* (DE SOUZA, 2005, p.6)

Para se construir a metamensagem, o *designer* deve responder as seguintes perguntas: (a) Quem são os usuários e o que querem?; (b) De que maneiras ele quer ou precisa fazer?; (c) Como o sistema pretende atender às demandas e necessidades dos usuários?; (d) Quais as crenças e valores estarão embutidos no sistema?

#### 7.1.1.1.2. Modelagem da Interação como Conversa

Uma vez construída a mensagem de metacomunicação de alto nível do sistema, é o momento de criar o seu modelo de interação. No *Design Centrado no Usuário*, “a modelagem da interação humano-computador ocorre como um conjunto de conversas que os usuários podem (ou devem) travar com o preposto do *designer* para atingir seus objetivos” (BARBOSA; SILVA, 2010, p. 229). Tal modelagem é realizada através da linguagem MoLIC (*Modeling Language for Interaction as Conversation*)(PAULA, 2003; BARBOSA e PAULA, 2003, que foi, projetada para apoiar os *designers* no planejamento da interação.

O diagrama de interação representa como os objetivos dos usuários poderão ser atingidos durante a interação e, assim como os cenários, o diagrama MoLIC serve como ponte entre a definição dos objetivos dos usuários e o projeto da interface propriamente dita (BARBOSA; SILVA, 2010).

A construção do diagramas MoLIC é realizada em duas etapas. Na primeira etapa, o *designer* define os tópicos de todas as possíveis conversas usuário-sistema e as trocas de turno entre o usuário e o preposto do *designer* que encadearão os tópicos dessas conversas. Na

segunda etapa, os tópicos das conversas são detalhados, ocorrendo a definição dos diálogos e signos envolvidos nas trocas comunicativas que correspondem a cada tópico. O diagrama MoLIC detalhado consiste em um recurso importante para o projeto da interface de usuário concreta, nas etapas posteriores do processo de *design* (BARBOSA; SILVA, 2010).

#### 7.1.1.1.2.1. Definição da Estrutura da Conversa

Na etapa de definição da estrutura da conversa, inicialmente, cada objetivo do usuário é mapeado para um tópico em uma cena. Cenas representam conversas sobre um determinado tópico e são graficamente representadas por retângulos com bordas arredondadas. O tópico representa a ação dentro da cena, e pode ser lido, do ponto de vista do preposto, como: “Neste momento, você (usuário) pode (ou deve) <tópico>”. Graficamente, o tópico é representado pelo texto, dentro da cena, que inicia com o verbo que indica a ação no infinitivo.

Mudanças de tópico são representadas por falas de transição, do usuário ou do preposto, e são representadas por uma linha direcionada, indicando pelo menos o enunciador da fala (“u:” para usuário e “d:” para o preposto do *designer*) e seu conteúdo.

Pode ocorrer que, para que uma mudança de tópico aconteça, uma determinada condição deva ser atendida. Nesse caso, uma expressão precedida pela palavra-chave “precond” deve ser adicionada à fala de transição, além do rótulo que a identifica.

Além das mudanças de tópicos, o usuário deve poder iniciar uma nova conversa em direção a um objetivo, em qualquer momento da interação. Para isso, a MoLIC possui acessos ubíquos, que são representados por uma cena anônima de fundo cinza mais a fala de transição do usuário pra a cena de destino.

Nos diagramas de interação, é preciso também iniciar e terminar a conversa usuário-preposto. Para isso, são definidos os pontos de abertura, ou seja, as cenas que permitem começar a interação quando o usuário entra no sistema, e os pontos de encerramento, que ocorrem quando o usuário sai do sistema. Pontos de abertura são representados por círculos preenchidos na cor preta, e pontos de encerramento por círculos na cor preta, circunscrito a um círculo branco.

Em alguns momentos da interação, ao invés da fala de transição do usuário levar diretamente a uma cena de destino, por ocorrer uma troca de turno, quando o preposto do *designer* deve interpretar (processar) o que o usuário disse e responder adequadamente. Dois

elementos da MoLIC indicam a troca de turno e a resposta do preposto: processo de sistema e fala de transição (fala de troca de turnos). Um processo de sistema apresenta a interpretação da fala do usuário pelo sistema, que decide então o rumo da conversa, de acordo com o que o usuário disse. A representação gráfica do processo de sistema em um diagrama MoLIC consiste em um quadrado com o fundo preto, indicando a “caixa preta” do processamento, que não são mostradas para os usuários.

Existem duas formas de comunicação do preposto para o usuário sobre um processamento: consecutiva, que é logo após o processamento; e síncrona, que é durante o processamento. A comunicação consecutiva é representada como uma fala de transição do preposto para uma cena e é rotulada como “d: resposta”. Já a comunicação síncrona é projetada para comunicar ao usuário o progresso do processamento ou seus estados intermediários e é utilizada geralmente em processos longos. Para representar esse tipo de comunicação, utiliza-se uma cena acoplada à caixa preta por um canal de comunicação, que pode apresentar falas do *designer* sincronizadas com o processamento do sistema e sobre ele.

A MoLIC também permite a representação de falas de recuperação de ruptura, que consiste em oportunidades explicitamente projetadas pelo *designer* para o usuário se recuperar de uma conversa acidental. Quando emitida pelo usuário, a fala de recuperação de ruptura indica um momento em que o usuário pode mudar de ideia e, conseqüentemente, o rumo da conversa. Quando emitida pelo preposto do *designer*, tal fala indica que o preposto não conseguiu interpretar uma ou mais falas do usuário adequadamente, sendo necessário que o usuário as retifique. Essas falas são representadas graficamente por linhas tracejadas, com o rótulo tendo o prefixo “u:” ou “d:”, respectivamente.

Também é possível, na MoLIC, representar situações que o *designer* identifica como sendo rupturas em potencial, mas o diagnóstico final cabe ao usuário. Assim, há as cenas em que o preposto comunica um alerta ou captura de erro. Uma cena de alerta é representada com uma linha tracejada. Uma cena de captura de erro é representada da mesma forma, mas não requer que o preposto de *designer* projete, a partir dela, uma fala de transição para recuperação da ruptura.

### 7.1.1.1.2.2. Detalhamento dos diálogos e falas

N a segunda etapa da construção de um diagrama MoLIC, os *designers* detalham a conversa sobre cada tópico, especificando os diálogos, as falas e os signos de cada cena. A conversa sobre um tópico pode ser composta por diversos diálogos, que são compostos por falas sobre signos. Os diálogos são representados em um segundo compartimento da cena, na forma “verbo + objeto”. Pode haver também precondições para que um diálogo ocorra, representado pela palavra-chave “precond:” + precondição, entre parênteses, após a frase que representa o diálogo.

A partir da definição dos diálogos, pode-se detalhar os signos envolvidos em cada um deles, que são representados entre chaves, à direita da frase que representa o diálogo. Signo emitido apenas pelo preposto do *designer* é precedido por “d:”, enquanto que signo emitido por preposto do *designer* e usuário, quando o *designer* dá oportunidade para o usuário falar sobre o signo, é precedido por “d+u:”.

A Figura 13, ilustra a notação diagramática básica dos elementos da MoLIC, descritos acima, em uma conversa para efetuar uma transferência bancária.

Figura 13- Elementos da MoLIC

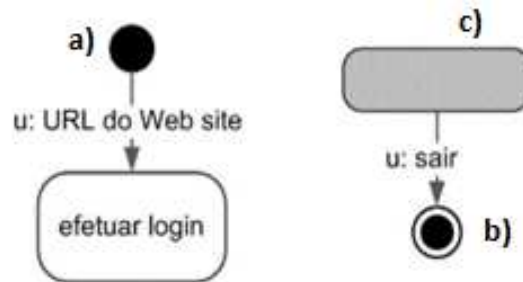


Fonte: (BARBOSA; SILVA, 2014, p. 11). Adaptado.



A Figura 14, apresenta as representações dos (a) ponto de abertura, (b) ponto de encerramento e (c) acesso ubíquo, apresentados pela MoLIC.

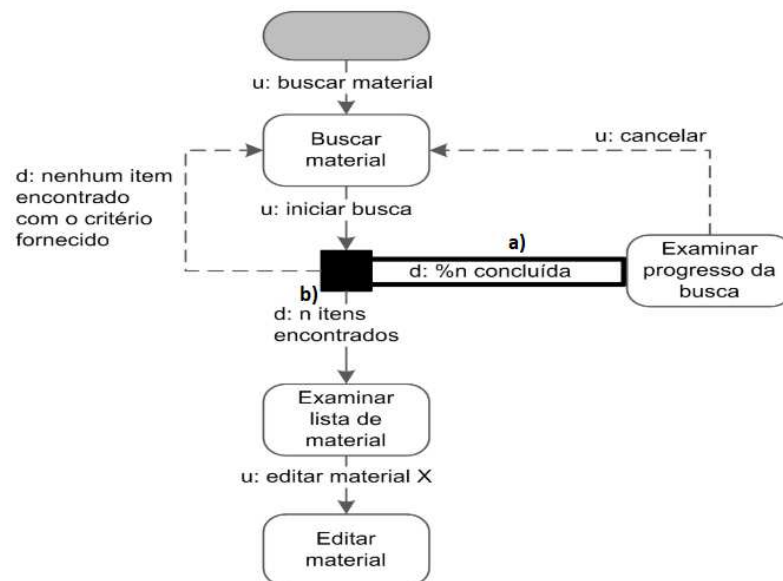
Figura 14- Modelagem dos Pontos de Abertura e Encerramento e acesso ubíquo na MoLIC



Fonte: BARBOSA; SILVA, 2010, p. 233. Adaptado.

A Figura 15 apresenta o processo do sistema e comunicação síncrona (a) e consecutiva (b).

Figura 15- Processo do sistema e comunicação síncrona e consecutiva



Fonte: BARBOSA; SILVA, 2010, p. 235. Adaptado.

### 7.1.1.2. Design Físico

Uma vez criado o *design* conceitual, é o momento do *designer* definir a interface propriamente dita, que consiste na parte física com a qual o usuário irá manter contato, considerando detalhes como, por exemplo, cores, sons e imagens a serem utilizadas.

No presente estudo, o *design* físico consistiu na Engenharia de Sistemas de Signos da Interface, do *Design* Centrado na Comunicação. Assim, a definição da interface foi realizada basicamente em duas etapas: (1) escolha dos estilos de interação para o sistema, e (2) representação a interface, em diferentes níveis de abstração (BARBOSA; SILVA, 2010).

A definição da interface do sistema apresentado no presente estudo foi feita a partir da modelagem de interação como uma conversa, representada pelo seu diagrama MoLIC, seguindo as regras propostas em Silva *et al.* (2005), esboçados na Tabela 2:

Tabela 2- Elementos da Molic e suas formas de mapeamento

<b>Elementos da MoLIC</b>	<b>Forma de mapeamento na interface</b>
<b>Acessos Ubíquos</b>	Barras de navegação ou itens de menu
<b>Cenas</b>	Unidades de apresentação, como telas, janelas ou páginas, dependendo da plataforma para a qual o sistema será desenvolvido
<b>Grupos de diálogos</b>	Quadros ou contêineres de outros elementos de interface, marcados visivelmente, ou definidos por um grid
<b>Signo</b>	Elemento da interface ( <i>widget</i> ) <i>i</i>
<b>Fala de transição do usuário</b>	Link, botão ou item de menu <ul style="list-style-type: none"> <li>• Caso haja uma condição associada à fala, pode haver uma mudança de estado (ativo ou inativo) ou na visibilidade (visível ou oculto) do elemento da interface correspondente;</li> </ul>
<b>Falas do preposto do <i>designer</i></b>	Mensagens de erro e de <i>status</i> <ul style="list-style-type: none"> <li>• Essas mensagens podem ser representadas de duas formas: embutidas na unidade de apresentação ou em unidades de apresentação independentes.</li> </ul>

Fonte: SILVA *et al.*, 2005. Adaptado.

Para a representação da interface, foi utilizado o *software* Balsamiq Mockups<sup>22</sup>. Esta ferramenta gera protótipos com aparência de desenho manual por meio de uma tela com *layout* extremamente simples. Assim a elaboração das telas ocorre de forma rápida e

<sup>22</sup>Disponível em: <https://balsamiq.com/wireframes/>

funcional. Logo, como produto desta etapa, foi obtido o protótipo de baixa fidelidade do Oráculo.

## 7.2. Artefatos de *Design*

A seguir serão descritos os artefatos de *design* desenvolvidos nesta pesquisa, tanto no que se refere ao *design* conceitual quanto ao *design* físico.

### 7.2.1. *Design* Conceitual

A partir das necessidades dos usuários e requisitos de IHC relacionados à ferramenta proposta neste trabalho, foi inicialmente produzido o seu modelo conceitual. Tal modelo descreve o que os usuários poderão fazer com o *framework* Oráculo e quais os conceitos necessários para entender como interagir com ele.

#### 7.2.1.1. Construção da mensagem de metacomunicação de alto nível

Após estudar os usuários, suas atividades e seu ambiente, foi possível conceber uma visão ampla sobre como todo o processo é realizado atualmente, e sobre como ele pode ou deve mudar com a introdução de uma nova tecnologia. Esta visão foi expressa, no âmbito da Engenharia Semiótica, através da elaboração da metagemagem de alto nível para o *framework* Oráculo, apresentada na figura 16.

Figura 16- Metagemagem elaborada para o Oráculo

Este é o meu entendimento, como *designer*, (a) **de quem você, usuário é**, um pesquisador de uma instituição de ensino superior; (b) **aprendi que você precisa** coletar dados de redes sociais *online* e aplicar técnicas de mineração de textos e analisar estes dados coletados para que você possa obter algum conhecimento que o apoie em suas pesquisas científicas; (c) dado a sua realidade atual, aprendi **que você prefere** realizar suas atividades com o apoio de uma ferramenta de fácil uso, com interface limpa e intuitiva, (d) **por que** as ferramentas que você possui ao seu alcance atualmente são difíceis de instalar, configurar e

usar, implicando na sua insatisfação; estas possuem interface não amigável, o que lhe exige muitas vezes desenvolver *scripts* para concluir o processo, além de ser ferramentas distribuídas, exigindo que você se preocupe com o aprendizado de como instalá-las e utilizá-las, e também com a formatação dos arquivos a exportar e importar para serem compatíveis com as demais ferramentas a serem utilizadas para realizar todo o processo. Este, portanto, (e) é o Oráculo, o **sistema que projetei** para você, desenvolvido com o *design* da interface voltado para as suas necessidades. **Com ele você pode**, (f) em uma única ferramenta, executar os passos do processo de mineração de textos, ou seja, coleta, pré-processamento, indexação, mineração, visualização e análise dos dados provenientes das mídias sociais, de forma integrada e em um único local, além de realizar a filtragem de bots em suas coletas. Mais especificamente, o Oráculo lhe permite realizar coletas em tempo real e retroativo, você pode configurar algumas opções para os termos de coleta, adaptar as funcionalidades de coleta e mineração, e exportar os seus dados em diferentes formatos, conforme sua necessidade, além de poder salva-los de forma organizada. Além disso, esta ferramenta foi desenvolvida com algumas opções de ajuda e explicações sobre os campos disponíveis, pensando justamente em lhe proporcionar uma experiência de uso mais agradável e intuitiva. Assim, com ele, espero que você possa alcançar uma gama de objetivos que se encaixam nesta visão.

Fonte: próprio autor

### 7.2.1.2. Modelagem da Interação Como Conversa

A partir das personas desenvolvidas e dos cenários de análise, apresentados anteriormente, no capítulo 6, foi elaborado o diagrama de interação MoLIC. Tal diagrama representa como os objetivos do usuário poderão ser atingidos durante a interação, através das possíveis trocas comunicativas entre o este e o preposto de *designer*, no caso, a interface do *framework* Oráculo.

É interessante ressaltar que inicialmente foi desenvolvido o diagrama MoLIC para o perfil dos pesquisadores que possuem conhecimento na área de computação (Persona 1), com perfil experiente. Contudo, ao elaborar a MoLIC para o perfil dos pesquisadores que não possuem conhecimento na área de computação (Persona 2), com perfil inexperiente, observamos uma tendência de as modelagens ficarem praticamente idênticas, diferenciando apenas os itens de ajuda, não contemplados nas abordagens desta modelagem. Assim, por

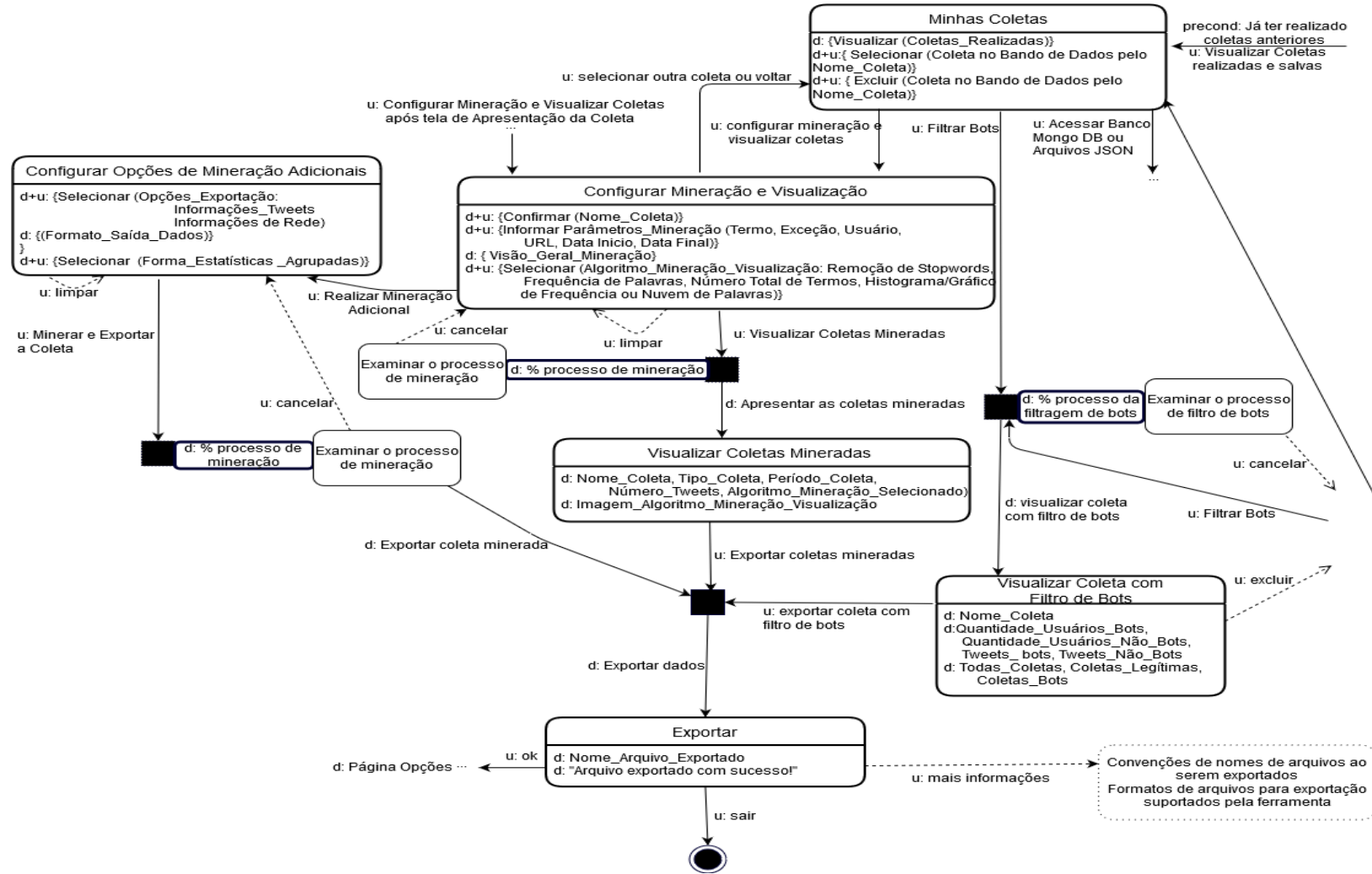
possuírem os mesmos objetivos, optou-se por elaborar um diagrama da interação para ambas as personas.

O diagrama MoLIC, disponibilizado no Google Drive<sup>23</sup>, foi elaborado em duas etapas: na primeira, foram desenvolvidos os tópicos de todas as possíveis conversas entre usuário-Oráculo e as trocas de turno que encadearão os tópicos destas possíveis conversas. Na segunda etapa, os tópicos foram detalhados e os diálogos e signos envolvidos nas trocas comunicativas foram definidos. A Figura 17 (Parte I e II) mostra o diagrama de interação do usuário do *framework* Oráculo.

---

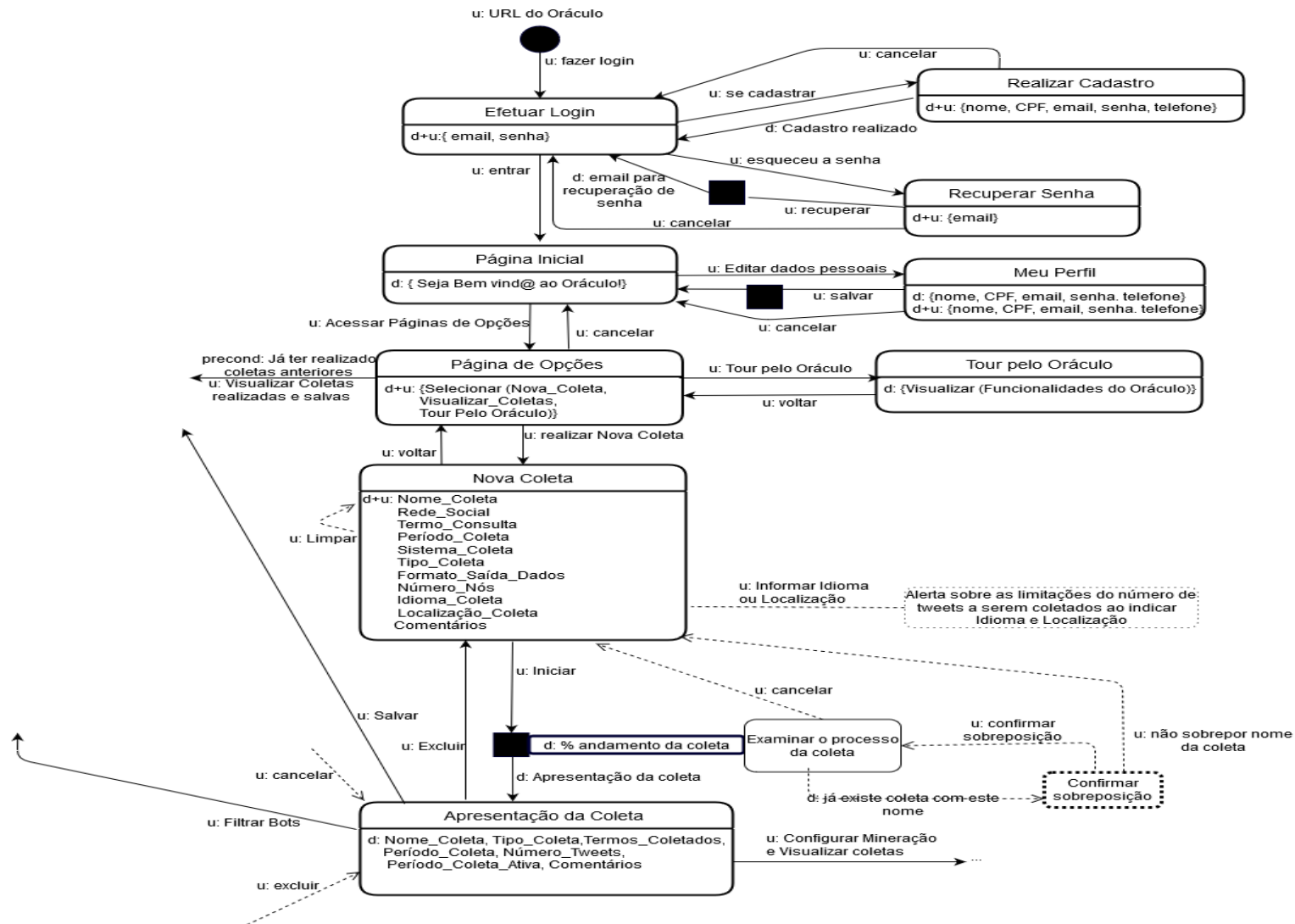
<sup>23</sup> [https://drive.google.com/drive/folders/1BlGaXCBpyKva7K\\_d-KBklfeCD4PHm3as?usp=sharing](https://drive.google.com/drive/folders/1BlGaXCBpyKva7K_d-KBklfeCD4PHm3as?usp=sharing)

Figura 17- MoLIC – Diagrama de interação do framework Oráculo – Parte I



Fonte: próprio autor

Figura 17- MoLIC – Diagrama de interação do framework Oráculo – Parte II



Fonte: próprio autor

### 7.2.2. Design Físico

A partir da modelagem da Interação, realizada utilizando a MoLIC, foi feito o projeto da interface de usuário do *framework* Oráculo, através do mapeamento dos elementos da MoLIC em elementos da interface.

A definição da interface iniciou com a escolha do estilo de interação do sistema, que foi WIMP (*Windows, Icons, Menus, and Pointers*), adotado nos ambientes baseados em janelas. Em seguida, foram construídos os protótipos de baixa fidelidade das telas do *framework* Oráculo.

É interessante ressaltar que o *design* físico para o Oráculo foi proposto, acompanhando a modelagem da interação proposta (MoLIC) sem realizar distinção evidente entre as duas personas estipuladas anteriormente, acrescentando opções de ajuda para os usuários inexperientes. Esta escolha foi a julgada como a melhor opção na etapa anterior e justificada neste momento, uma vez que ao desenvolver o *design* físico para a persona inexperiente a qualidade da ferramenta foi comprometida ao retirarmos algumas de suas funcionalidades.

Todas as telas do protótipo de baixa fidelidade do *framework* Oráculo estão disponibilizadas no Google Drive<sup>24</sup>. A seguir serão mostradas as telas que remetem às principais funcionalidades do sistema.

Pode-se observar, em algumas das telas, informações em notas em amarelo. Estas representam a parte da interface do Oráculo que será desenvolvida com foco na Persona (inexperiente), onde serão elaboradas telas de ajuda ou explicações dos campos com as funcionalidades existentes.

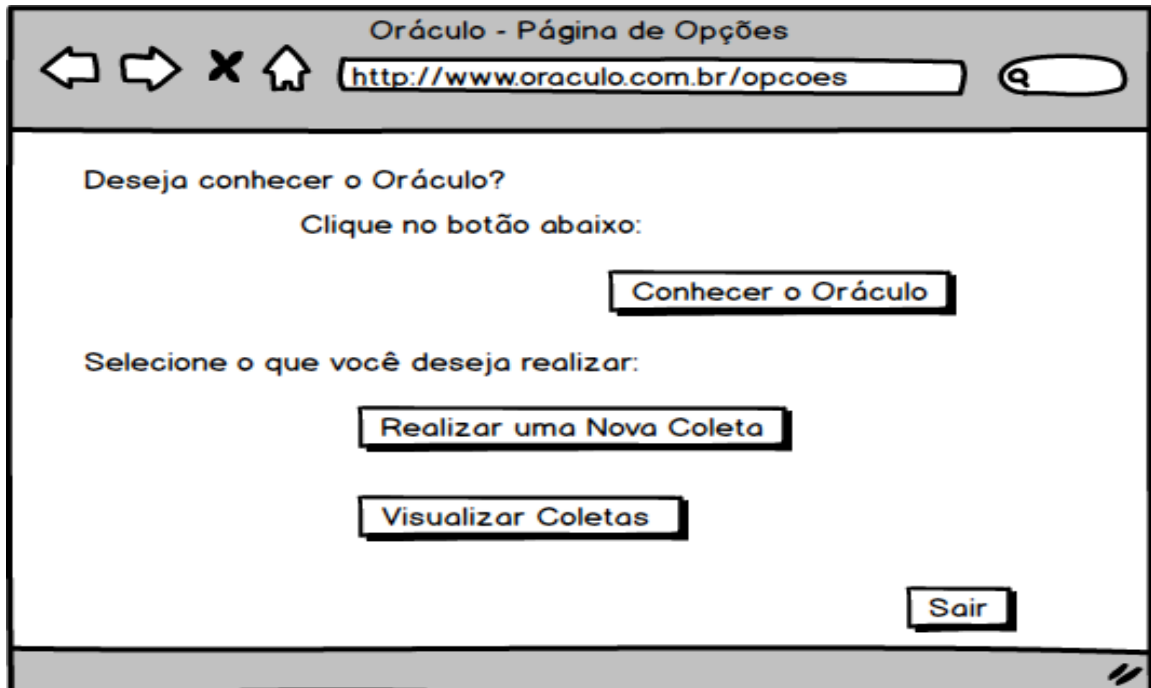
A Tela de Opções (Figura 18) possibilita ao usuário Conhecer o Oráculo, sendo direcionado a uma tela com um vídeo apresentando a ferramenta (Figura 19). O usuário pode ainda solicitar a realização de uma nova coleta (Figura 20) ou a visualizar as suas coletas realizadas anteriormente (Figura 29).

---

<sup>24</sup> <https://drive.google.com/drive/folders/1VIOmZ42jaxgrozZJqjZG53CBcLDZjCW?usp=sharing>

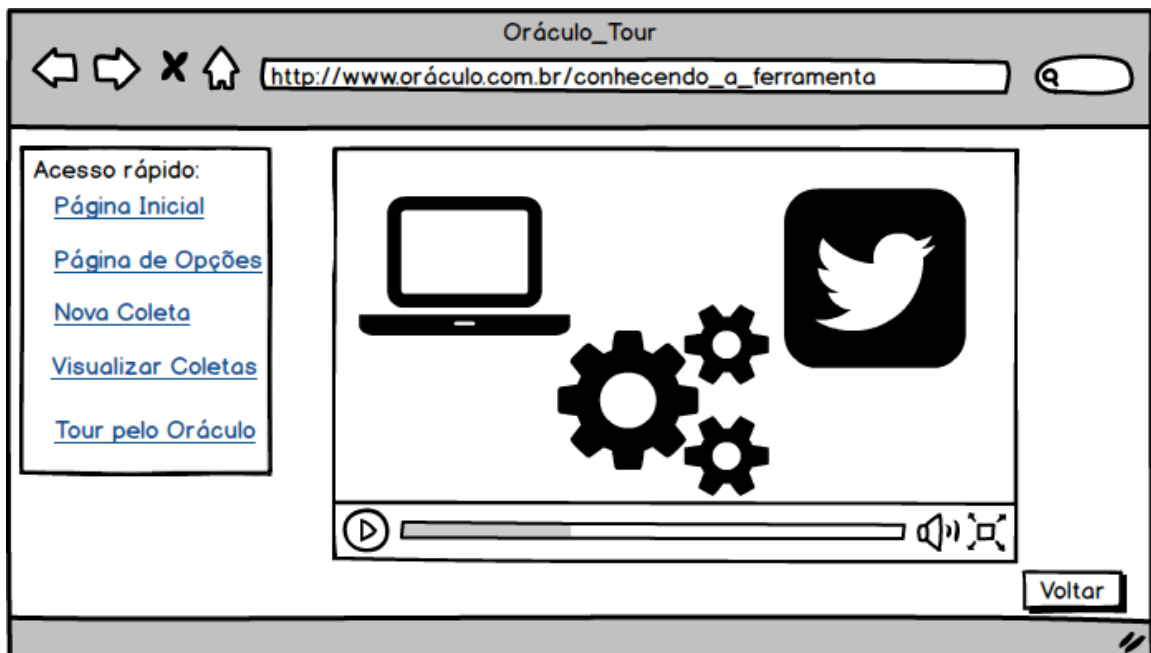


Figura 18- Tela de Opções



Fonte: próprio autor

Figura 19- Tela de Tour pelo Oráculo



Fonte: próprio autor

As Figuras 21 a 27 contemplam os passos com as várias opções de configuração para a realização de uma coleta, atendendo as necessidades dos pesquisadores. A Figura 20, é o Passo 1 desta configuração e nota-se que, atendendo a um dos requisitos levantados, existe no

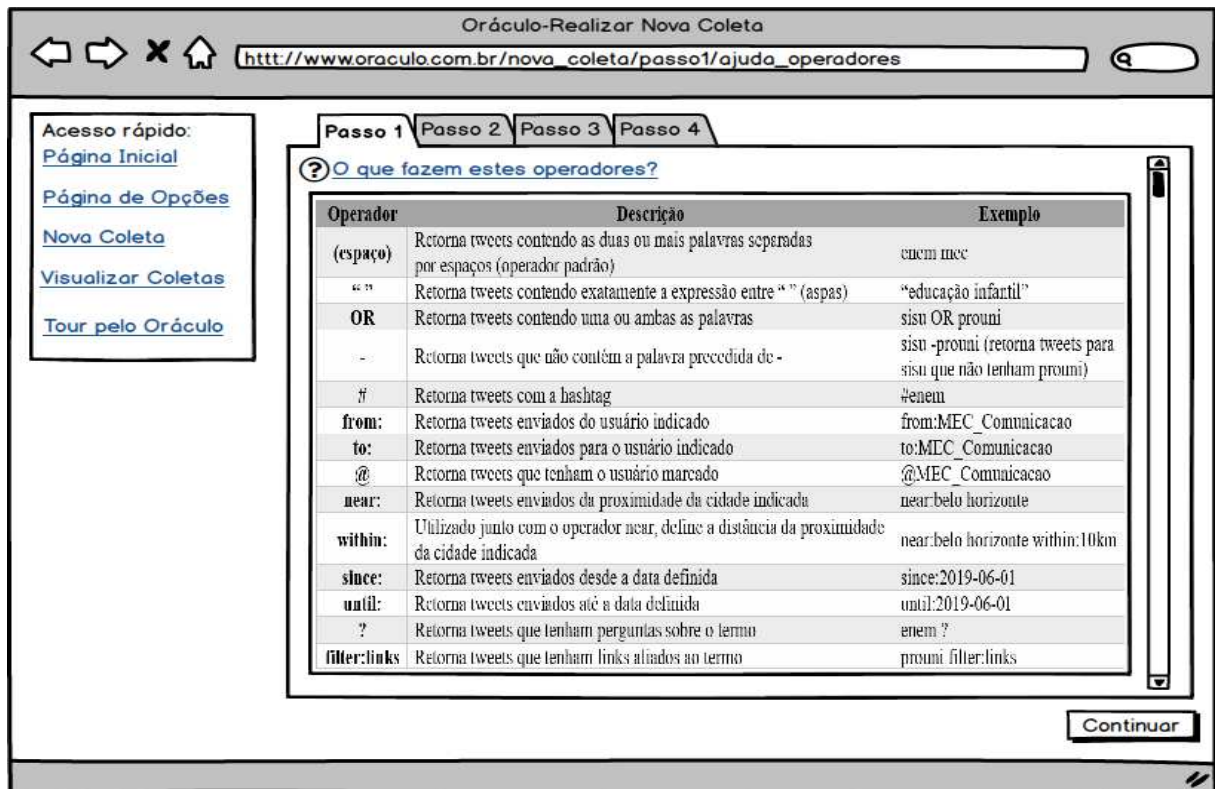
Oráculo a possibilidade de inclusão de vários termos de busca associados a diferentes operadores. A princípio, no Oráculo, a coleta será realizada na Rede Social *Twitter*.

Figura 20- Tela de Realizar Nova Coleta – Passo 1

Fonte: próprio autor

A Figura 21 é uma tela de ajuda aos usuários que não possuem conhecimento nas áreas de computação, representados pela Persona 2, estes possuem um conhecimento mais geral em relação ao uso de tecnologia. Neste caso, são descritos os operadores para associação aos termos para a realização das coletas.

Figura 21- Tela de Realizar Nova Coleta – Passo 1 – Ajuda com Operadores



Fonte: próprio autor

A Figura 22, é o Passo 2 de configuração para a realização da coleta. Nesta tela é possível o usuário escolher o tema de coleta, o formato para saída dos seus dados coletados, o período de realização da coleta e qual o sistema de coleta necessita. Por meio do *link* "mais informações", o usuário será direcionado a um menu de ajuda com mais informações sobre os campos existentes (explicados neste protótipo de baixa fidelidade através dos bilhetes em amarelo).

Figura 22- Tela de Realizar Nova Coleta – Passo 2

The screenshot shows a web browser window titled "Oráculo-Realizar Nova Coleta" with the URL "http://www.oraculo.com.br/nova\_coleta/passo2". The interface is divided into four steps: Passo 1, Passo 2 (active), Passo 3, and Passo 4.

**Form Fields:**

- Tempo de Coleta\*:** Radio buttons for "Tempo Real" and "Retroativo".
- Saída de Dados\*:** Radio buttons for "Banco Mongo DB" and "Arquivos JSON".
- Período da Coleta\*:** Fields for "Data de Início" and "Data de Fim" (with calendar icons), and "Hora de Início" and "Hora de Fim" (with time pickers).
- Sistema de Coleta\*:** Radio buttons for "API Twitter" and "Crawler".
- Buttons:** "Limpar" and "Continuar".
- Link:** "? Mais informações".

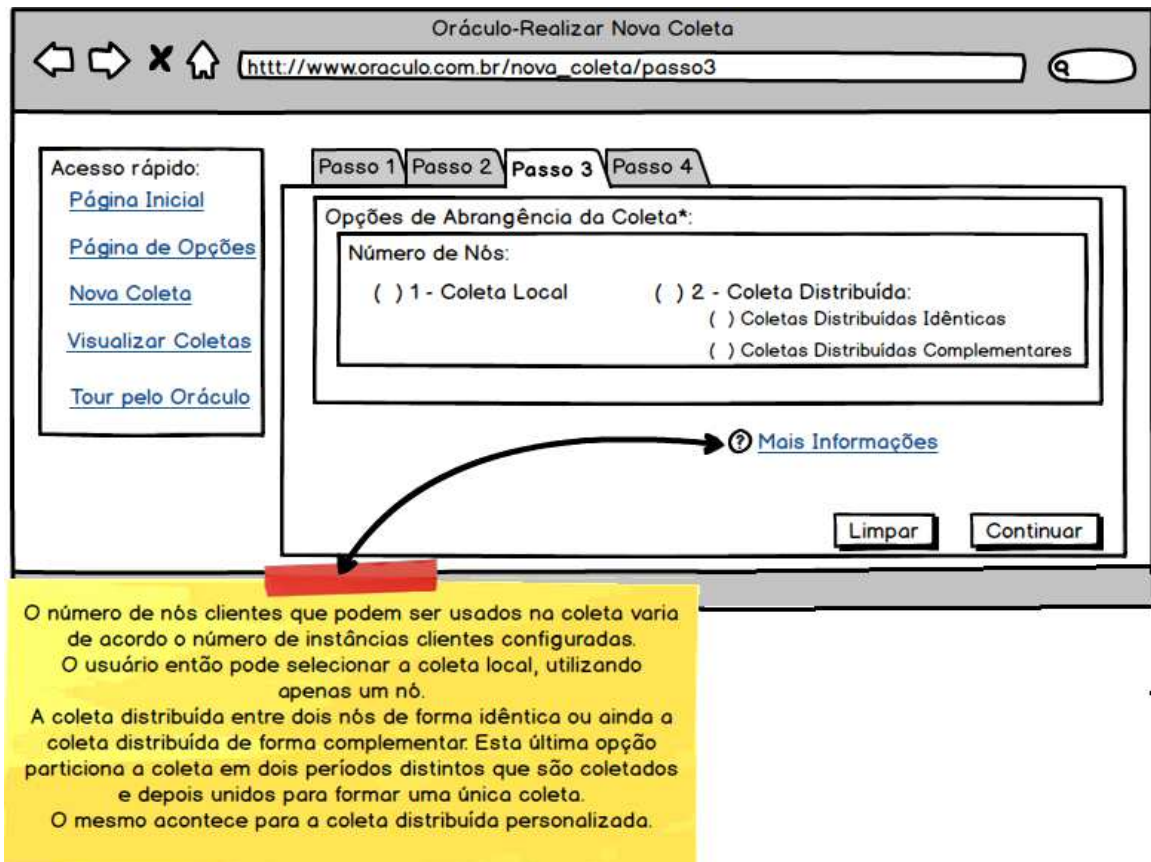
**Callout Boxes:**

- Tempo Real -** retorna os tweets em realtime, ou seja, que forem postados enquanto a coleta está sendo executada.
- Retroativo -** retorna os tweets postados anteriores a coleta, a partir do período da coleta estipulada pelo usuário, assim, é possível acessar os dados históricos dos tweets postados na rede social. Esta opção está disponível apenas para o Sistema de Coleta - API do Twitter
- Os Sistemas de Coleta disponíveis são:**
  - API Twitter -** Conjunto de procedimentos e padrões, ou Application Programming Interface - API, para que seja possível acessar os tweets, status e dados dos usuários. Assim, as coletas de textos realizadas com esta opção estão restritas aos limites impostos por esta rede social.
  - Crawler -** são sistemas que navegam e baixam automaticamente páginas da Web seguindo os hiperlinks de maneira metódica e automatizada. Está é uma opção viável para contornar as limitações da API do Twitter, porém retorna menos campos de retorno em sua coleta.
- Alerta:** O Crawler realiza coletas apenas em tempo real e a API do Twitter coleta em Tempo Real e Retroativo

Fonte: próprio autor

A Figura 23 representa o Passo 3 no processo de realização de coleta, onde o usuário pode selecionar a quantidade de computadores (nós) que realizarão as coletas. A explicação destes campos estarão disponíveis aos usuários em uma tela de ajuda ao clicar no link “mais informações”.

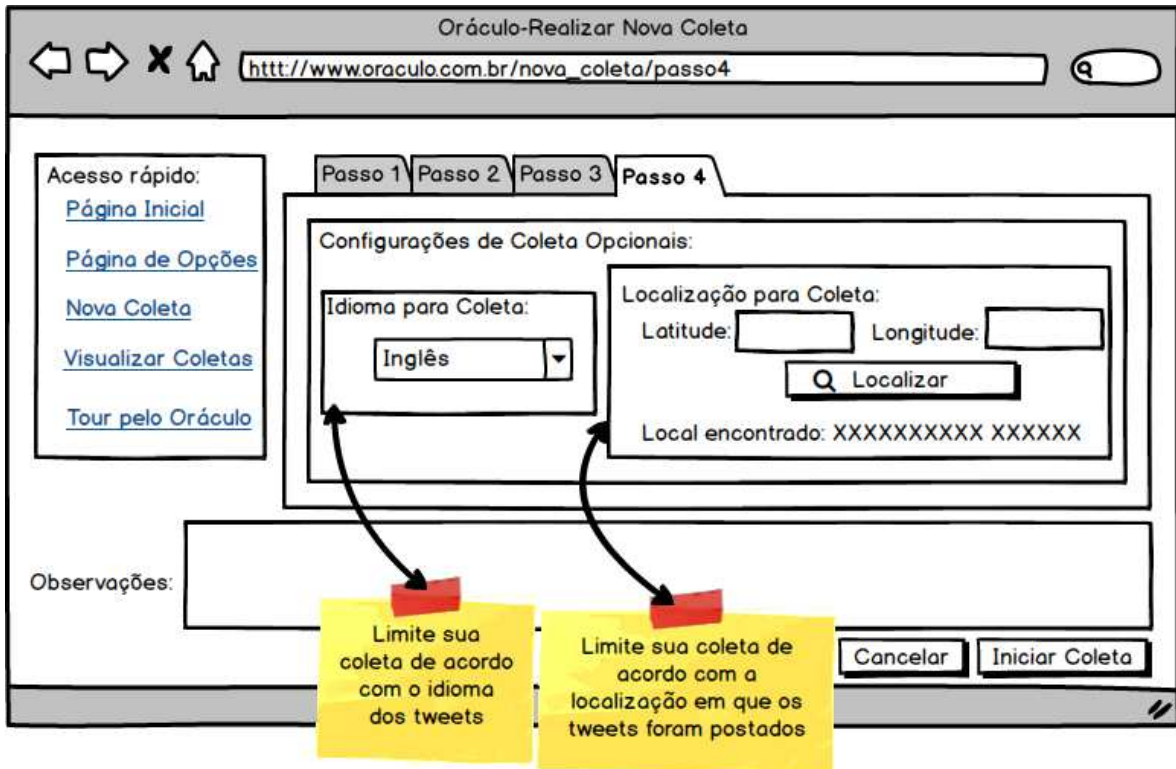
Figura 23- Tela de Realizar Nova Coleta – Passo 3



Fonte: próprio autor

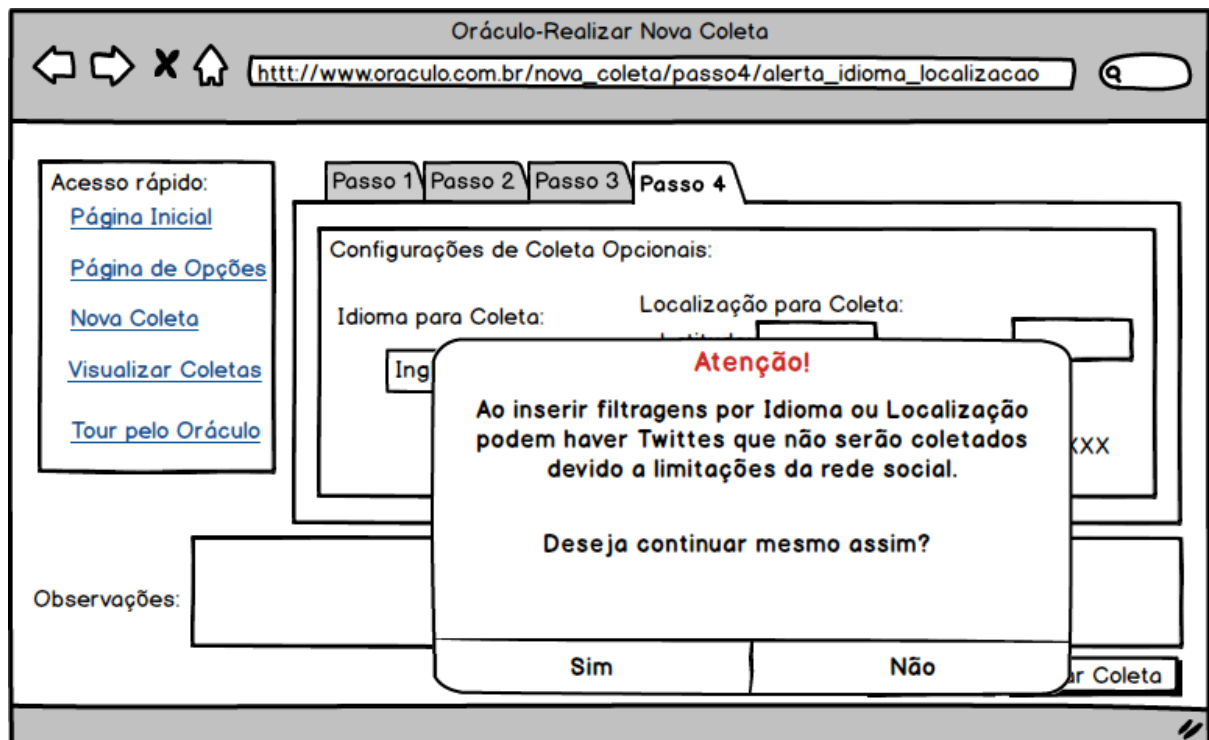
A Figura 24 representa o Passo 4 da configuração para realização da coleta. Nesta o usuário irá selecionar, caso seja de seu interesse as limitações de coleta entre idioma e localização. A Figura 25 apresenta uma alerta ao usuário caso ele preencha esses campos, pois dados podem não ser coletados devido a limitações da rede social. A explicação aparecerá quando o usuário passar o *mouse* sobre os campos (modo Hint).

Figura 24- Tela Realizar Nova Coleta – Passo 4



Fonte: próprio autor

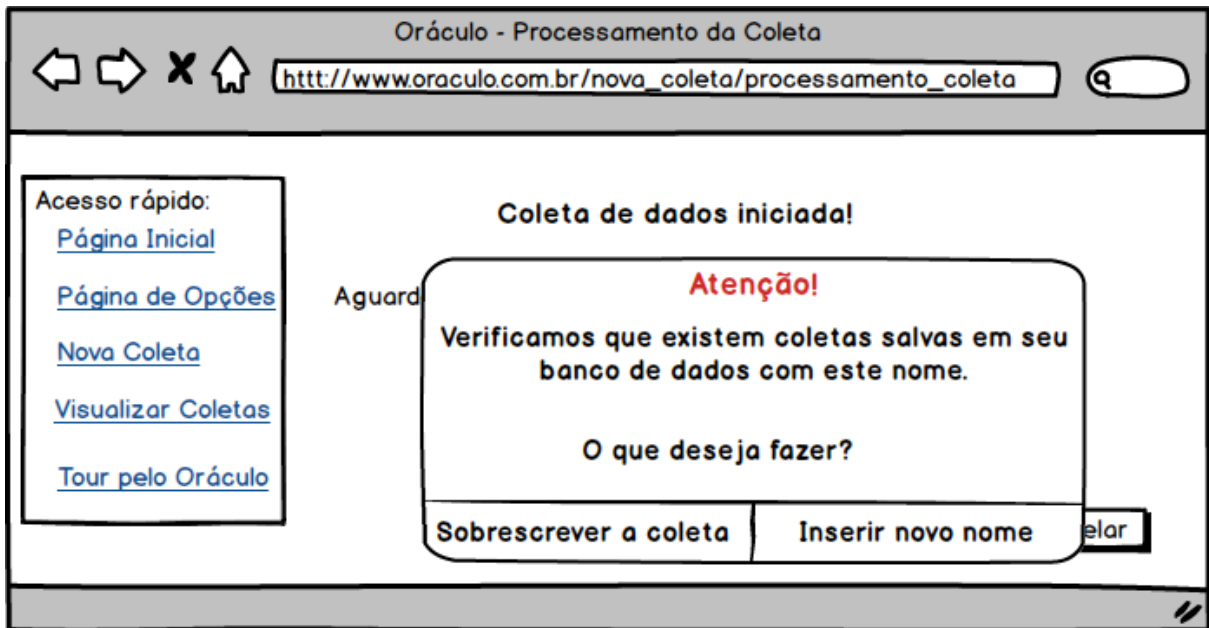
Figura 25- Tela de Realizar Nova Coleta – Passo 4 – Confirmação Idioma e Localização



Fonte: próprio autor

A Figura 26 é uma alerta, como prevenção de ruptura comunicativa, em que a interface irá solicitar ao usuário que confirme o andamento do processo caso haja coincidência de nome da coleta atual com uma realizada anteriormente.

Figura 26- Tela de Processamento da Coleta – Alerta Sobreposição de Nome

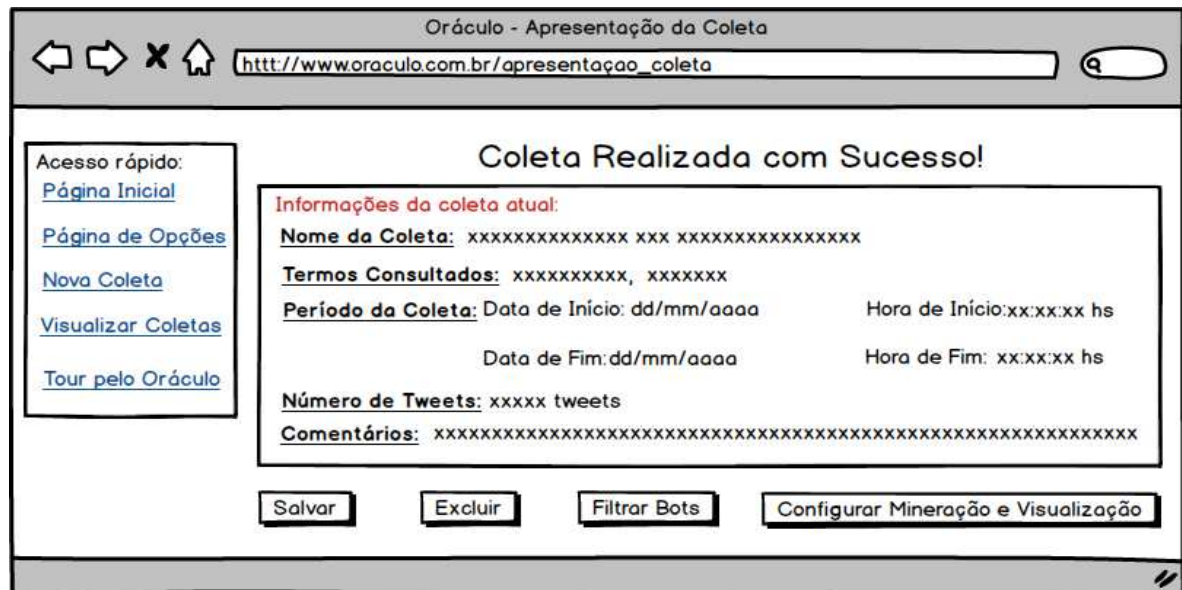


Fonte: próprio autor

Assim que o processamento da coleta é concluído, o usuário irá clicar no botão continuar e a tela representada pela Figura 27 será apresentada constando um resumo da coleta que o usuário acabou de realizar.

Nesta tela, o usuário poderá: “salvar” a coleta realizada, sendo direcionado a Tela de “Minhas Coletas”, representada pela Figura 29; “excluir” a coleta realizada; realizar o “Filtro de Bots”, sendo direcionado para a Tela de Configuração de Filtro de Bots, representada pela Figura 28; e poderá configurar mineração de textos sobre esta coleta ao clicar no botão “Configurar Mineração e Visualização”, sendo direcionado para a tela representada pela Figura 24.

Figura 27- Tela de Apresentação da Coleta

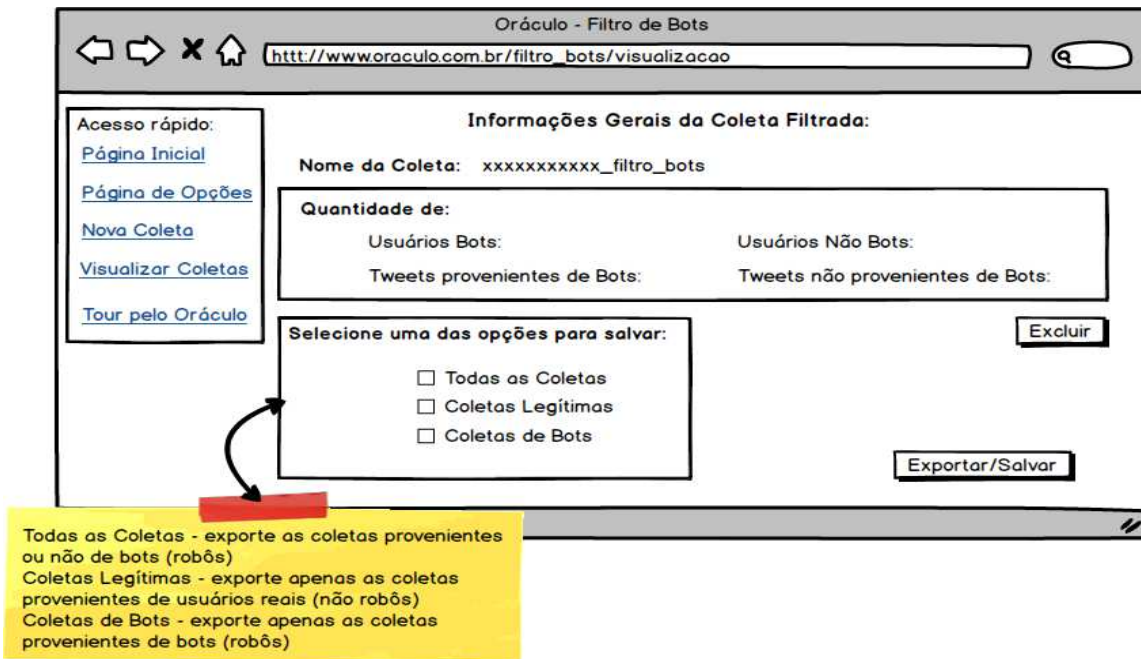


Fonte: próprio autor

Através da Figura 28 o usuário poderá realizar as configurações para realização do filtro de bots sobre sua coleta. Esta funcionalidade estará disponível ao usuário apenas se o “Sistema de Coleta”, selecionado na Figura 22, tiver sido “API do *Twitter*”. Isso se deve ao fato de que o Sistema de Coleta “Crawler” inviabiliza a identificação de bots por retornar poucas informações sobre os usuários da rede social. A explicação dos campos nesta tela, aparecerão ao usuário quando ele passar o *mouse* sobre os campos (modo Hint).



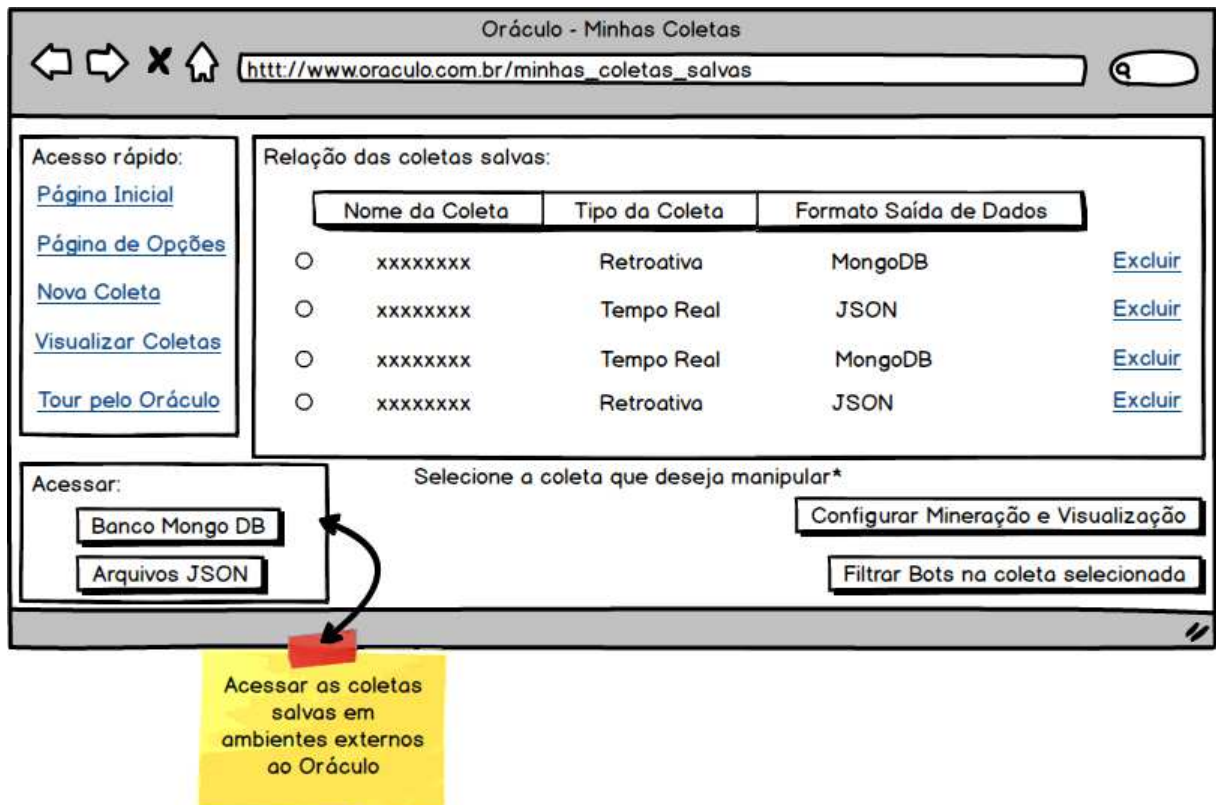
Figura 28- Tela de Configurar Filtro de Bots



Fonte: próprio autor

A Figura 29 representa a tela onde o usuário poderá visualizar as coletas realizadas anteriormente e que foram salvas, constando algumas informações sobre as mesmas. A partir do momento que selecionar uma das coletas, ele poderá “Configurar Mineração e Visualização” ou “Filtrar Bots”. A explicação dos campos sobre acessar os arquivos em ferramentas externas, aparecerão ao usuário quando ele passar o *mouse* sobre os campos (modo Hint).

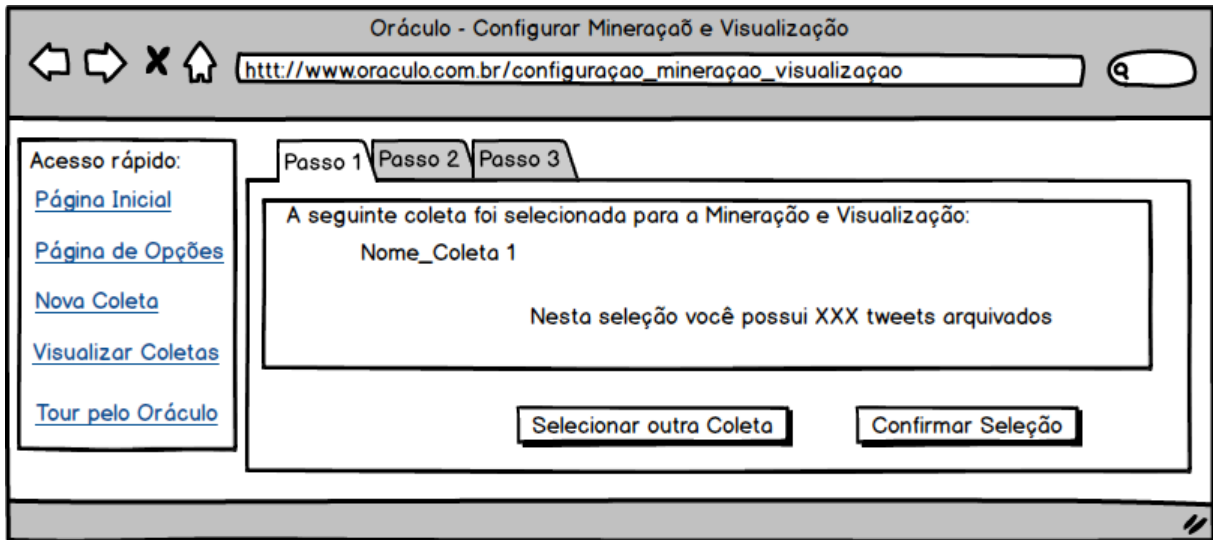
Figura 29- Tela de Minhas Coletas



Fonte: próprio autor

As telas apresentadas nas Figuras 30 a 37 representam os passos com as opções de configuração da mineração de textos a serem aplicadas às coletas. A Figura 30 apenas solicita ao usuário confirmar que aquela é a coleta selecionada para realização da mineração. Caso não seja, ele poderá clicar no botão “Selecionar outra Coleta” e será direcionado a tela de coletas realizadas e salvas, representada pela Figura 29 -“Minhas coletas”.

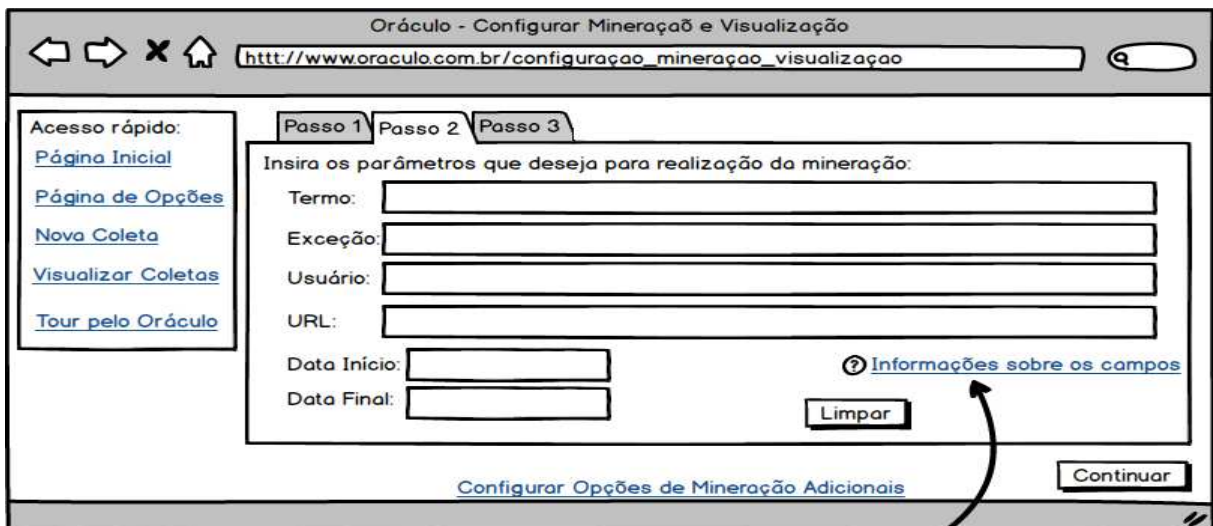
Figura 30- Tela de Configurar Mineração e Visualização – Passo 1



Fonte: próprio autor

A tela a seguir, Figura 31, representa o Passo 2 do processo de mineração. Seus campos serão detalhados em uma tela de ajuda, solicitada pelo usuário ao clicar no *link* “Informações sobre os campos”.

Figura 31- Tela de Configurar Mineração e Visualização – Passo 2

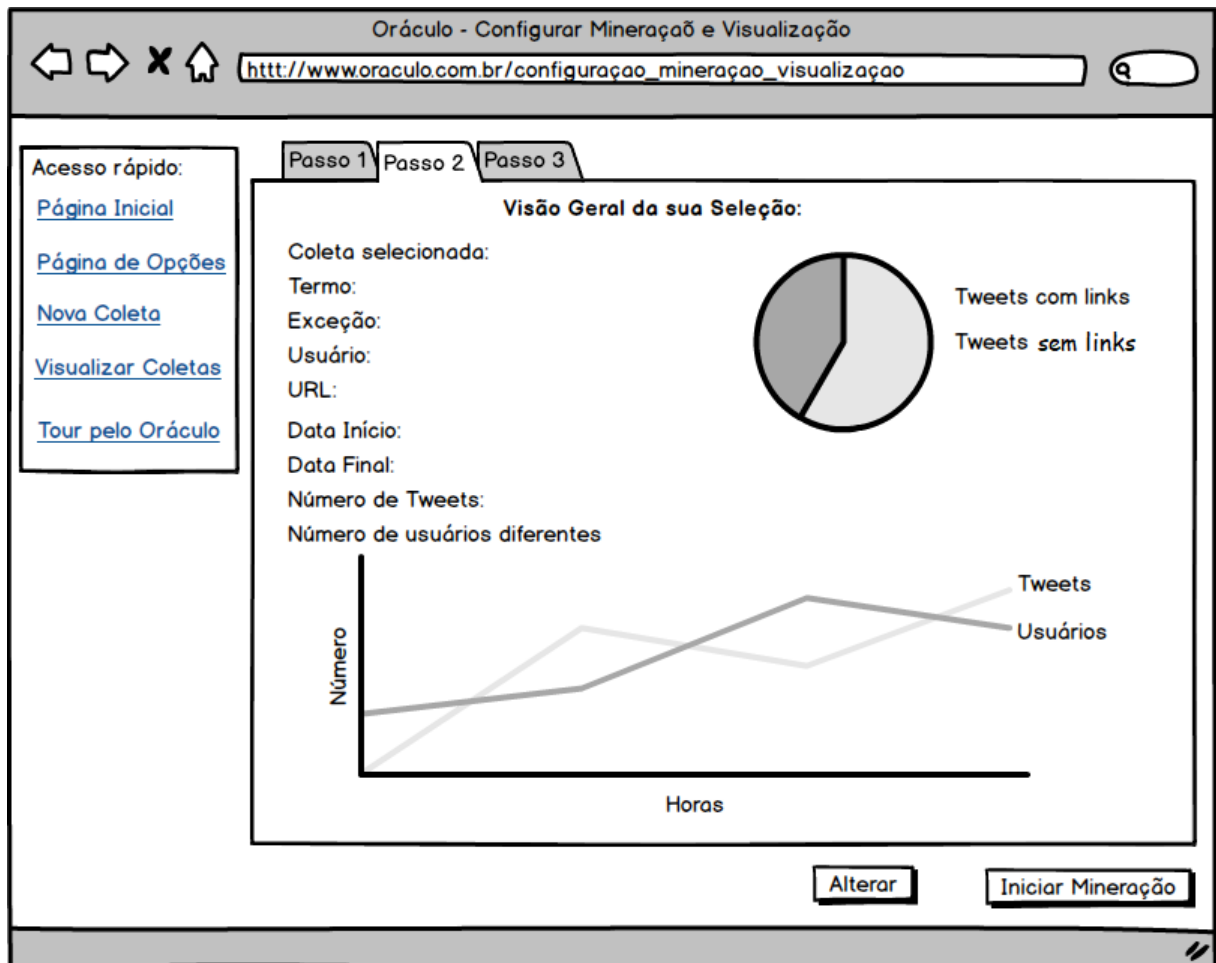


Termo - se vazia, a mineração ocorrerá contendo qualquer texto  
 Excluir - se vazia, a mineração ocorrerá sem excluir nada da coleta selecionada  
 Do usuário - se vazia, a mineração ocorrerá para qualquer usuário  
 URL (ou parte do URL) - se vazia, a mineração ocorrerá contendo qualquer URL  
 Data Início - dd/mm/aaaa para início da mineração  
 Data final - dd/mm/aaaa para finalização da mineração

Fonte: próprio autor

Assim que o usuário seleciona o botão “Continuar”, ele será destinado a uma tela, Figura 32, em que conterà o resumo de suas escolhas na tela anterior, com uma prévia em dois gráficos: um contendo a porcentagem dos Tweets que contém ou não links, e o outro, relaciona o número de Tweets e usuários em determinado espaço de tempo. Informações estas consideradas úteis aos pesquisadores antes da efetivação da mineração.

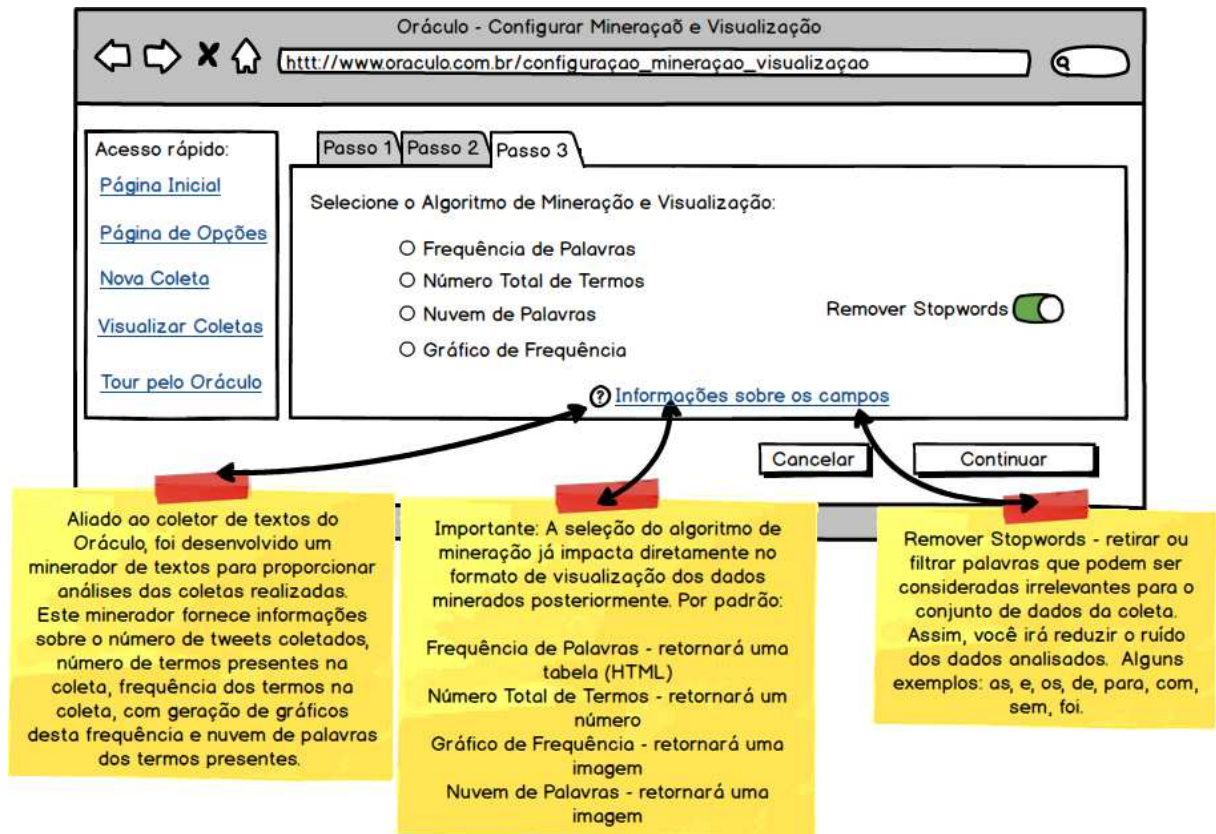
Figura 32- Tela de Configurar Mineração e Visualização – Passo 2 – Visão Geral



Fonte: próprio autor

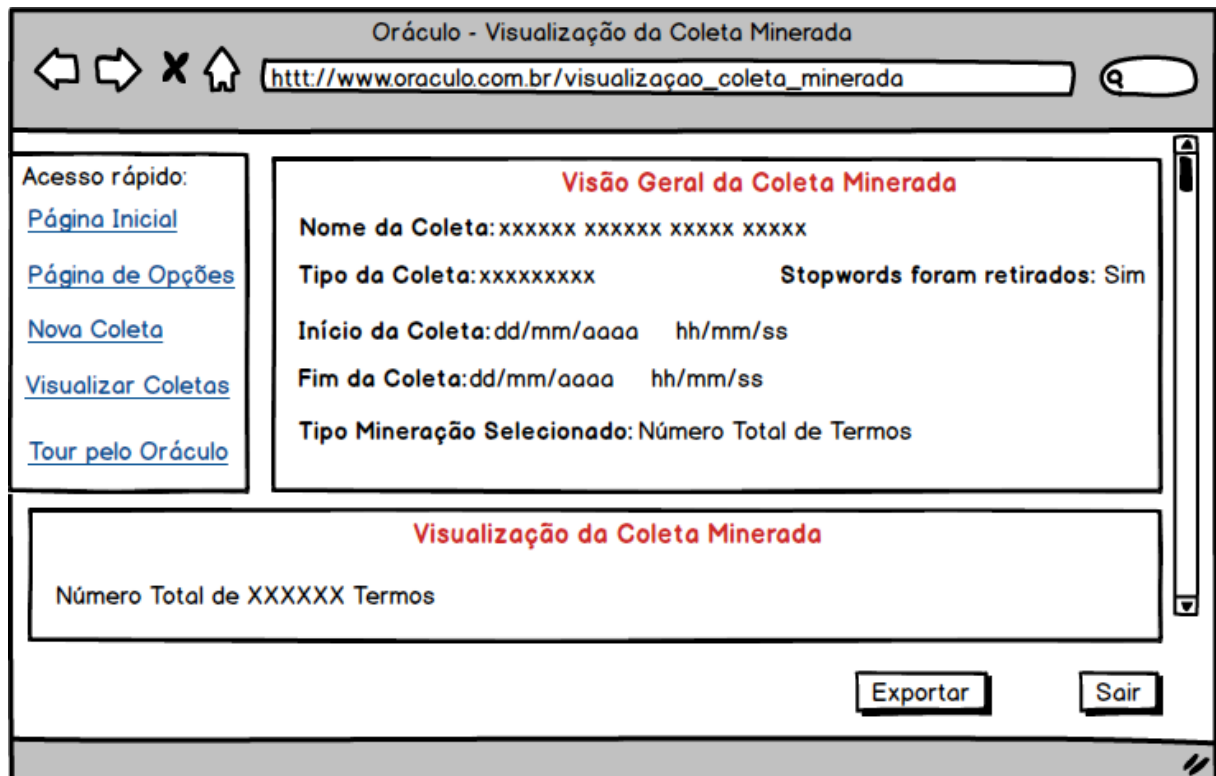
Na tela representada pela Figura 33, o usuário selecionará a técnica de visualização que aplicará a sua coleta. É importante ressaltar que a visualização dos dados, disponível até o momento no Oráculo, depende diretamente desta escolha, onde: “Frequência de Palavras” retornará uma tabela, sua tela está representada na Figura 34; “Número Total de Termos” retornará um número, sua tela está representada na Figura 35; “Nuvem de Palavras” retornará uma imagem, sua tela está representada na Figura 36; e o “Gráfico de Frequência” retornará um gráfico, sua tela está representada na Figura 37.

Figura 33- Tela de Configurar Mineração e Visualização – Passo 3



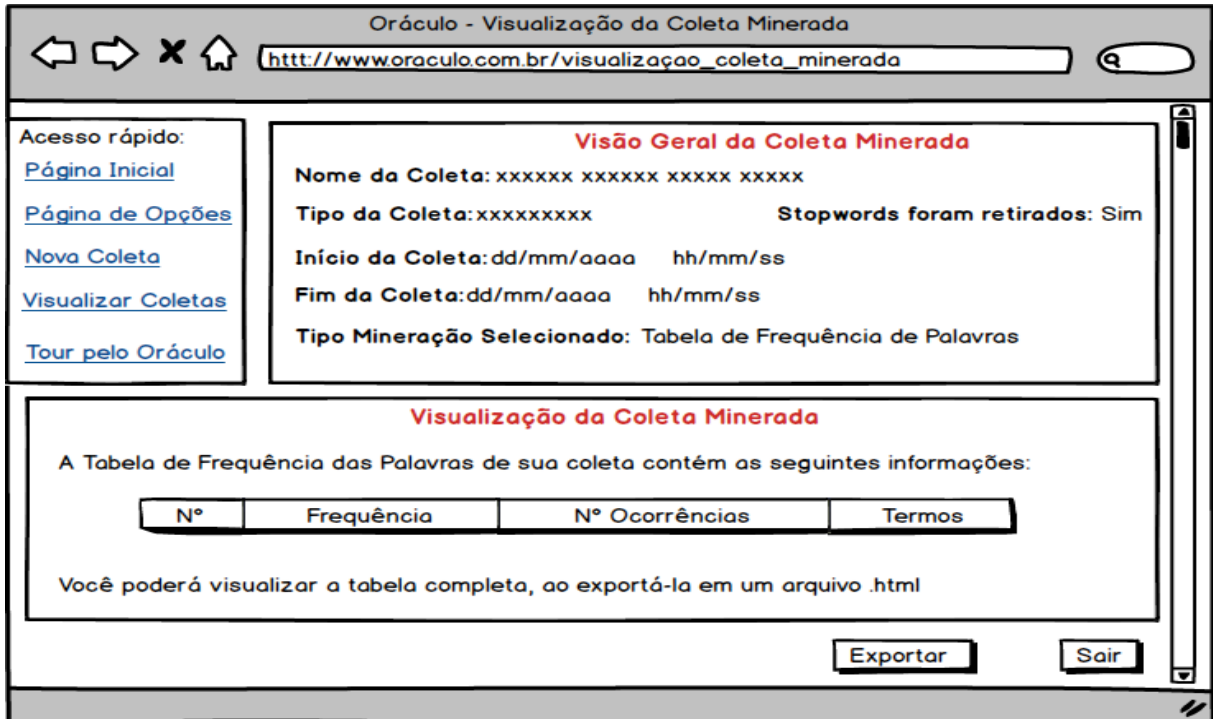
Fonte: próprio autor

Figura 34- Tela de Visualização da Coleta Minerada – Número Total de Termos



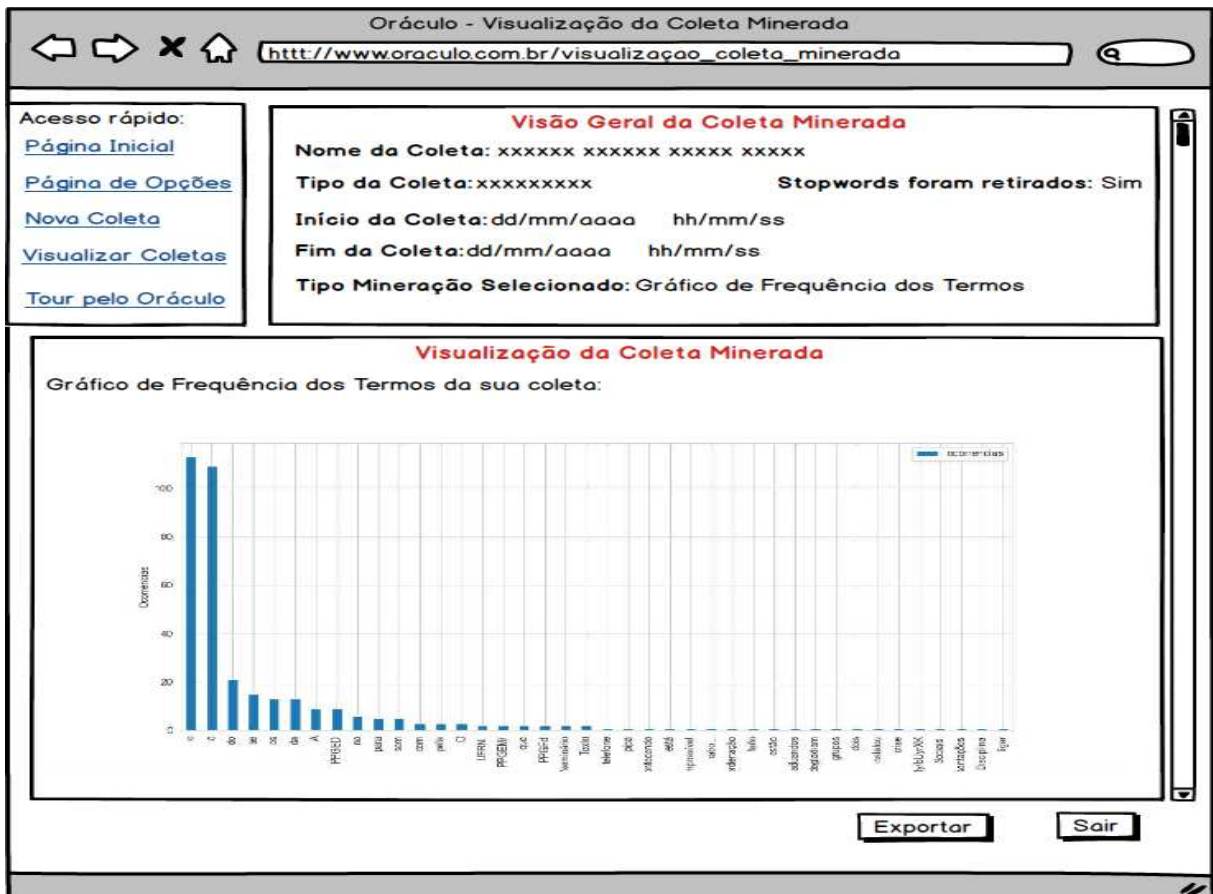
Fonte: próprio autor

Figura 35- Tela de Visualização da Coleta Minerada – Tabela de Frequência de Palavras



Fonte: próprio autor

Figura 36- Tela de Visualização da Coleta Minerada – Gráfico de Frequência de Termos



Fonte: próprio autor

Figura 37- Tela de Visualização da Coleta Minerada – Nuvem de Palavras

Oráculo - Visualização da Coleta Minerada

http://www.oraculo.com.br/visualizacao\_coleta\_minerada

**Visão Geral da Coleta Minerada**

Nome da Coleta: xxxxxx xxxxxx xxxxx xxxxx

Tipo da Coleta: xxxxxxxx Stopwords foram retirados: Sim

Início da Coleta: dd/mm/aaaa hh/mm/ss

Fim da Coleta: dd/mm/aaaa hh/mm/ss

Tipo Mineração Selecionado: Nuvem de Palavras

**Visualização da Coleta Minerada**

Nuvem de Palavras da sua coleta:

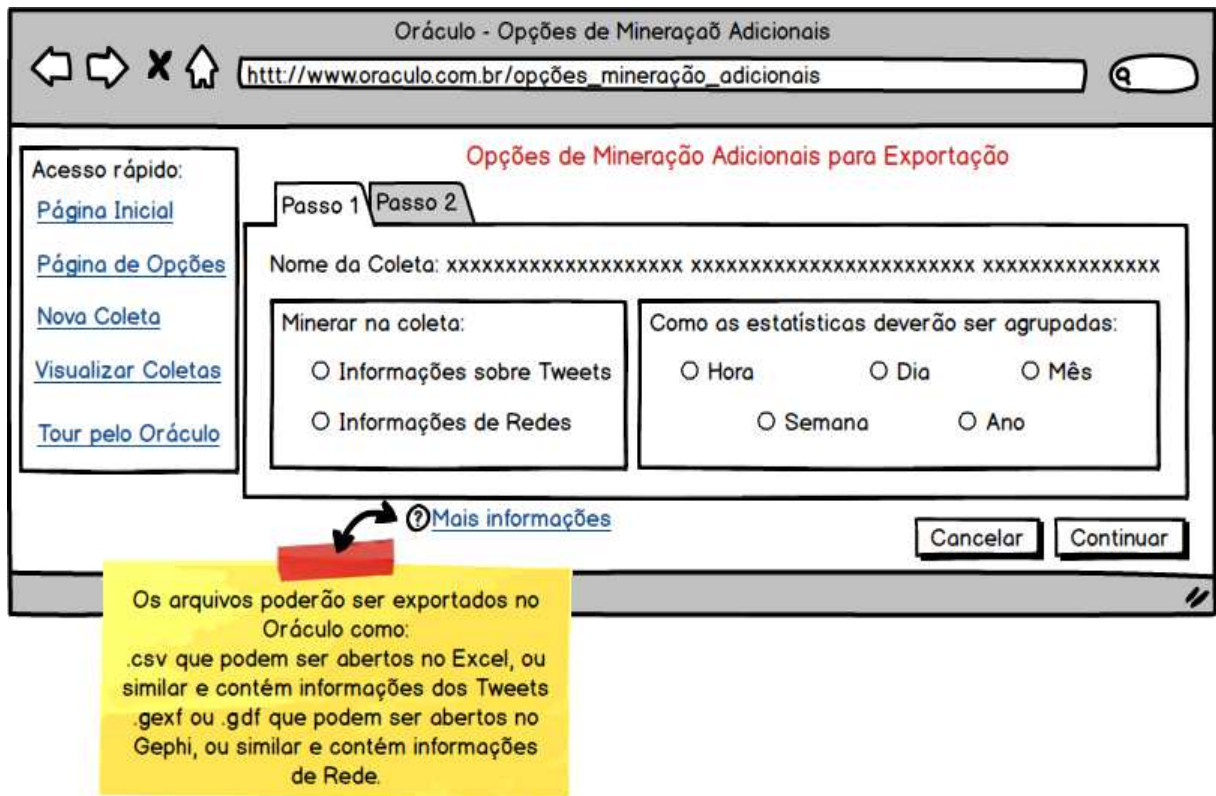
Exportar Sair

Fonte: próprio autor

O Oráculo disponibiliza ao usuário quatro tipos de algoritmos de mineração e visualização, como apresentados anteriormente. Porém sabe-se que existem infinitas necessidades dos pesquisadores. Assim, a ferramenta disponibiliza também outras formas de mineração de textos, porém sem as opções de visualização. A partir desta escolha, configuráveis nas telas representadas pelas Figuras 38 a 40, os usuários podem realizar a mineração em suas coletas e exportar seus resultados em formatos compatíveis com outras ferramentas existentes para visualização e análise de dados.

Na tela de Opções de Mineração Adicionais, passo 1, representada pela Figura 38, o usuário poderá escolher o grupo de informações que deseja minerar, sendo Tweets ou sobre Redes, e a temporalidade que deseja agrupar tais resultados. Nesta tela também ele poderá, caso ache necessário, a tela de ajuda com informações sobre os campos.

Figura 38- Tela de Opções de Mineração Adicionais – Passo 1

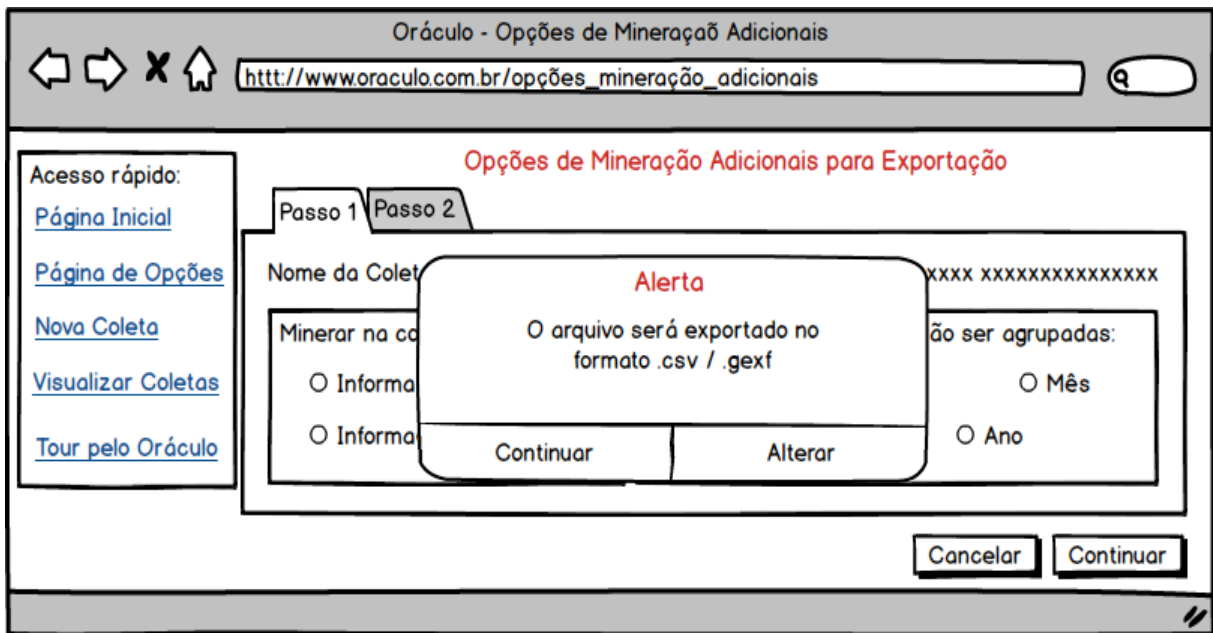


Fonte: próprio autor

A tela a seguir, Figura 39, informa ao usuário o formato do arquivo que será exportado conforme sua escolha sobre quais informações ele deseja minerar, no caso: “Informações sobre Tweets” retornam arquivos a serem exportados em .csv, e “Informações sobre Rede” retornam arquivos a serem exportados em .gexf.



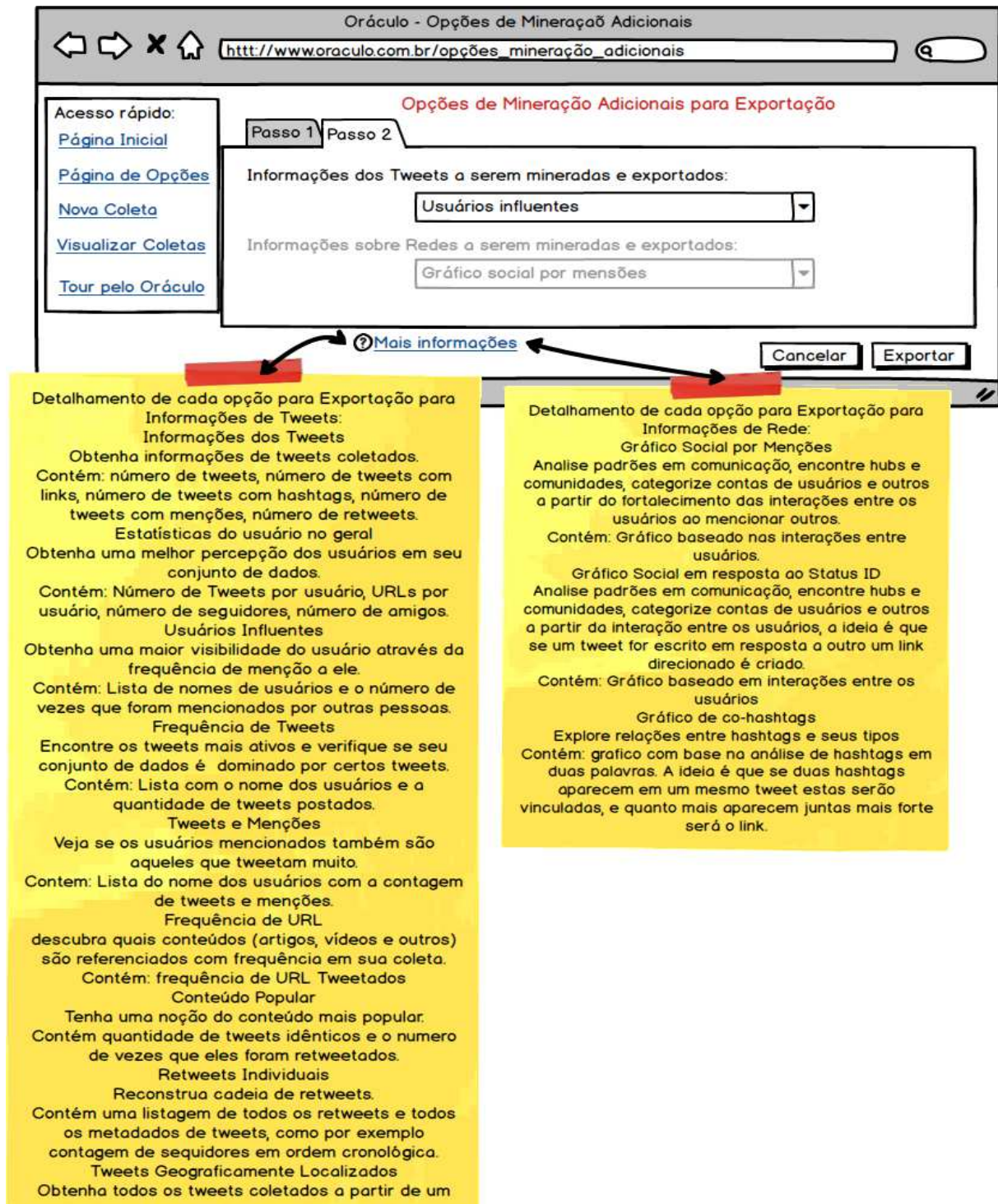
Figura 39- Tela de Opções de Mineração Adicionais – Passo 1 – Alerta Formato Arquivo



Fonte: próprio autor

O passo 2 do processo de “Mineração Adicionais” é representado pela Figura 40, e conforme a escolha do usuário sobre quais informações ele deseja minerar, estará disponível um conjunto de opções em uma caixa de combinações. A explicação sobre cada uma destas aparecerá ao usuário assim que ele passar o *mouse* sobre os campos (modo hint).

Figura 40- Tela de Opções de Mineração Adicionais – Passo 2



Fonte: próprio autor

O protótipo de baixa fidelidade do *framework* Oráculo foi mostrado a dois dos participantes desta pesquisa, para que pudesse ser validado, antes de se passar para a próxima etapa do processo de *design*. Assim, o *design* da ferramenta foi aprovado, e então pôde-se dar início à construção do protótipo funcional de alta fidelidade.

## 8. CONSTRUÇÃO DE UMA VERSÃO INTERATIVA

Este capítulo apresenta a proposta de uma versão interativa do *framework* Oráculo, através da construção de um protótipo funcional de alta fidelidade, para que a qualidade de uso do *design* possa ser avaliada junto ao seus potenciais usuários.

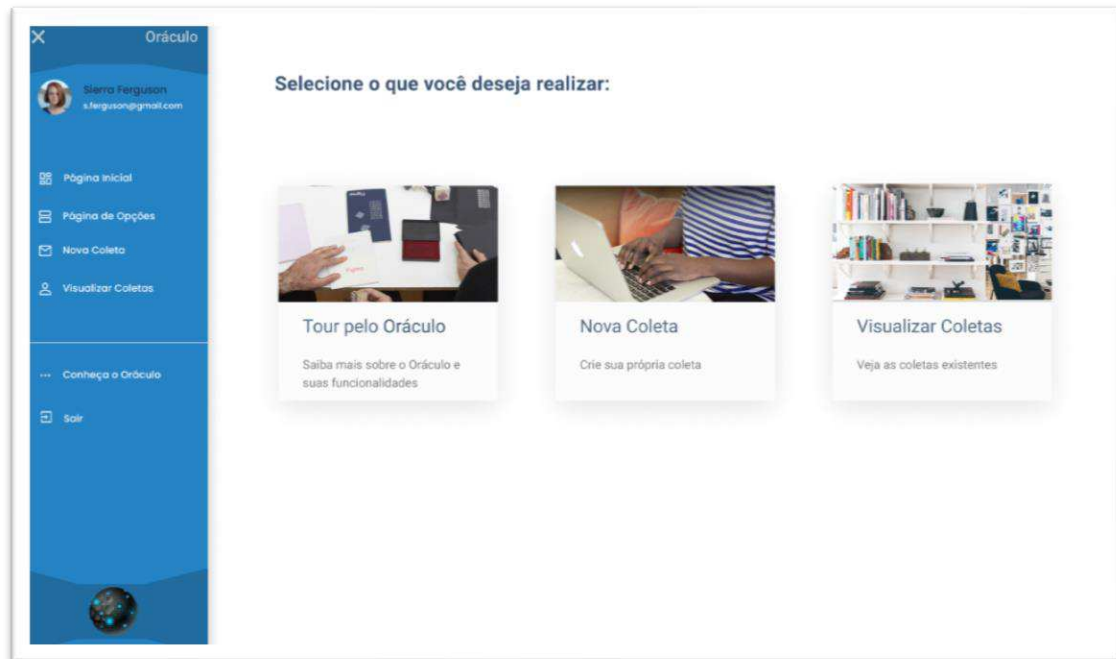
Para a construção de tal protótipo, foram levados em consideração os seguintes princípios e diretrizes para o *design* de IHC (BARBOSA E SILVA, 2010): (1) Correspondência com a expectativa do usuário; (2) Simplicidade nas Estruturas das Tarefas; (3) Equilíbrio entre Controle e Liberdade do Usuário; (4) Visibilidade e reconhecimento; (5) Consistência e Padronização; (6) Promovendo a Eficiência do Usuário; (7) Antecipação; (8) Visibilidade; (9) Conteúdo relevante e expressão adequada; e (10) Projeto para erros.

Foi utilizada a ferramenta Figma para a construção do protótipo, que é, dentre as ferramentas gratuitas e/ou livres existentes no mercado hoje para esta finalidade, uma das mais utilizadas e recomendadas pelos especialistas da área. O protótipo funcional de alta fidelidade encontra-se disponível em:

[https://www.figma.com/file/y7PvgH7NuASG0S70SXa4Md/Or%C3%A1culo\\_MTPLNAM\\_Original?node-id=158%3A0](https://www.figma.com/file/y7PvgH7NuASG0S70SXa4Md/Or%C3%A1culo_MTPLNAM_Original?node-id=158%3A0)

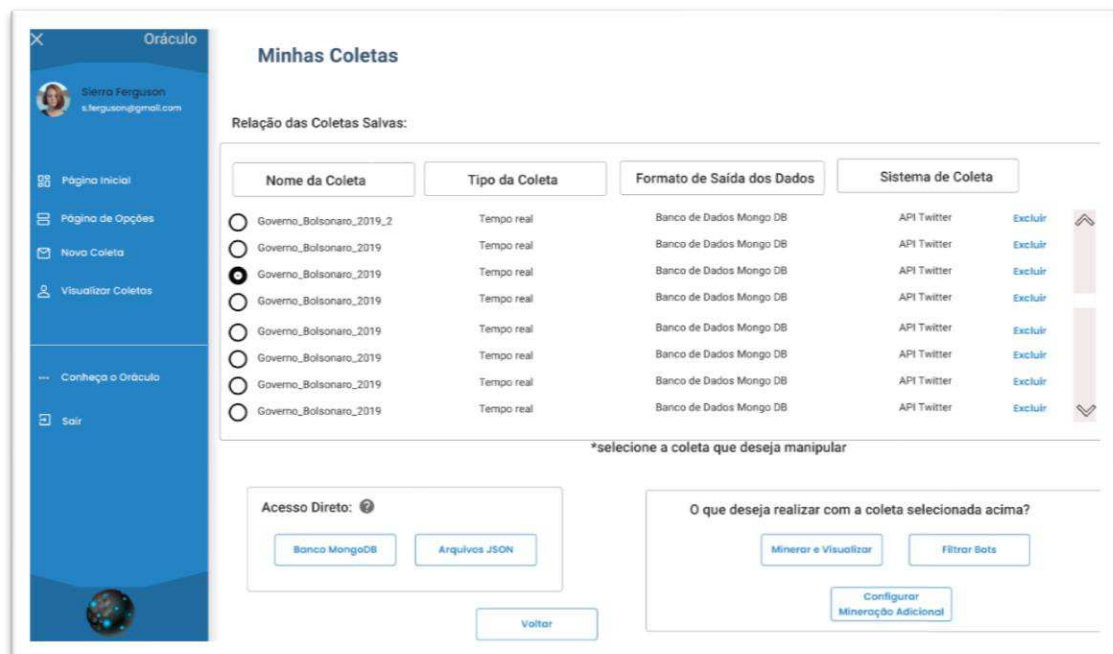
As Figuras 41 e 42, a seguir mostram a tela principal do sistema, que mostra as opções de interação do usuário com o Oráculo, e a tela “Minhas Coletas”.

Figura 41- Tela de Opções do Oráculo



Fonte: próprio autor

Figura 42- Minhas Coletas



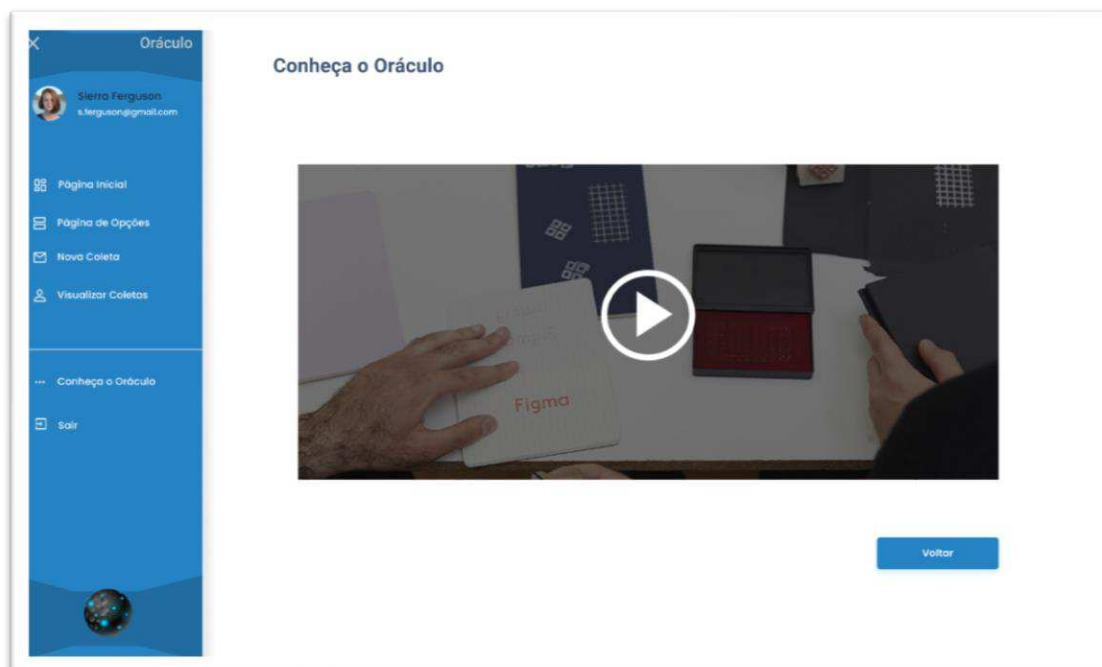
Fonte: próprio autor

As Figuras 43 e 44 representam, respectivamente, as configurações iniciais para realização de uma Nova Coleta e a tela que apresenta ao usuário a possibilidade de conhecer o Oráculo.

Figura 43- Nova Coleta (Passo 1)

Fonte: próprio autor

Figura 44- Conheça o Oráculo



Fonte: próprio autor

## 9. AVALIAÇÃO

Este capítulo apresenta a avaliação preliminar do *design* do *framework* Oráculo, que foi realizada a partir do protótipo funcional de alta fidelidade, bem como os resultados obtidos a partir dela. Tal avaliação teve o propósito de investigar a qualidade da interação proporcionada pelo *design* proposto aos seus usuários, focando especificamente na experiência de uso, utilidade e facilidade de uso percebida.

Para isso, foram realizadas avaliações em ambientes controlados, envolvendo potenciais usuários do Oráculo, ou seja, pesquisadores que trabalham com coleta e análise de dados provenientes de redes sociais *online*. Em seguida, foram realizadas investigações sobre a experiência de uso, bem como a utilidade e facilidade de uso percebidas pelo usuários, fazendo uso dos questionários SUMI (Software Usability Measurement Inventory)<sup>25</sup> (KIRALOWSKI; CORBETT, 1993) e TAM (Technology Acceptance Model)<sup>26</sup> (DAVIS, 1989), respectivamente. Os dados coletados com a avaliação foram analisados e permitiram apreciar a qualidade do *design* proposto.

Devido aos prazos para finalização desta pesquisa, optou-se por se utilizar métodos de avaliação que nos exigisse menor tempo, não sendo possível a utilização de métodos da Engenharia Semiótica que nos permitisse avaliar a comunicabilidade do *design* proposto, sendo esta uma sugestão acrescentada como de trabalhos futuros.

A seguir, na seção 9.1, apresentamos as considerações metodológicas da avaliação. Em seguida, na seção 9.2, mostramos como esta etapa foi executada, e, por fim, os resultados obtidos são mostrados e discutidos na seção 9.3.

### 9.1. Considerações Metodológicas

Para a avaliação da interação do *framework* Oráculo, foi projetado um teste com usuários. Foi solicitado que eles realizassem tarefas relacionadas à coleta de dados, mineração e visualização de dados e filtro de bots.

Após a observação dos usuários interagindo com o sistema, em ambiente controlado, foi aplicada uma versão reduzida do questionário SUMI (KIRALOWSKI;

---

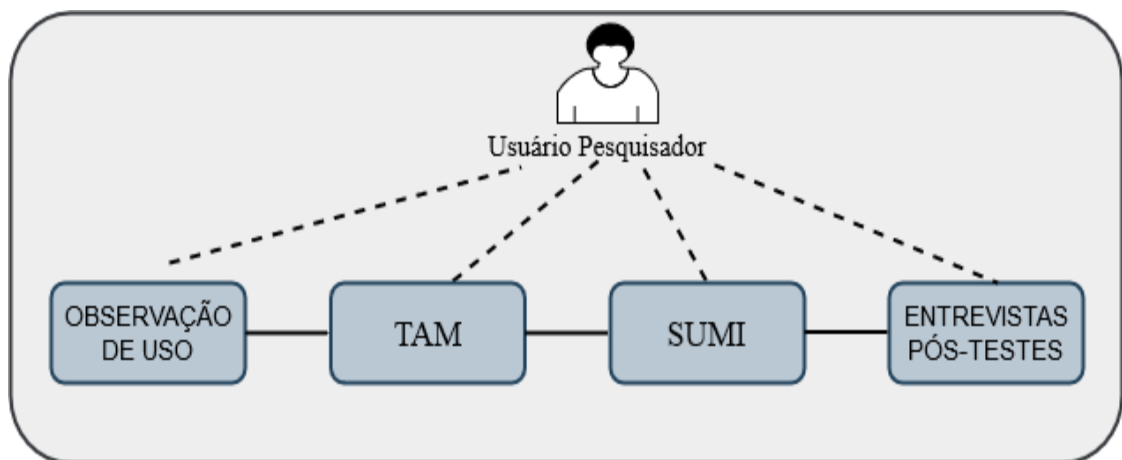
<sup>25</sup> Disponível em: <http://sumi.ucc.ie>

<sup>26</sup> Disponível em: <https://pubsonline.informs.org/doi/abs/10.1287/mnsc.35.8.982>

CORBETT, 1993), para medir qualidade do software a partir do ponto de vista do usuário, considerando a sua experiência de uso. Além desse, o TAM (DAVIS, 1989) foi aplicado com o objetivo medir a aceitação dos usuários em relação ao Oráculo. Ao final, foi realizada uma entrevista pós-teste, com o objetivo de coletar as opiniões dos participantes sobre a ferramenta, bem como sugestões de melhoria.

Seguindo os modelos dos passos metodológicos aplicados até o momento, as três formas de avaliação que foram utilizadas nesta pesquisa são esboçadas na Figura 45.

Figura 45- Passos para realização da Avaliação



Fonte: próprio autor

A seguir será apresentada uma descrição dos questionários SUMI e TAM, aplicados na avaliação.

### 9.1.1. Software Usability Measurement Inventory (SUMI)

O SUMI é um questionário de avaliação, desenvolvido pelo *Human Factors Group* (HFC) da *University College*, para medir a satisfação do usuário durante a interação e a qualidade do *software* sob o ponto de vista do usuário, podendo detectar falhas de usabilidade (KIRALOWSKI e CORBETT, 1993).

Ele é composto por 50 sentenças sobre as quais o usuário avaliador deve analisar se concorda, não concorda ou se ela não se aplica. As questões englobam aspectos de satisfação, eficiência, aprendizagem, ajuda e controle. O preenchimento do formulário SUMI leva cerca de cinco minutos e uma amostra de cinco participantes já provê resultados úteis para análise (MOTTI, 2009).

Para sua aplicação nesta pesquisa, o questionário SUMI original foi traduzido e adaptado para reduzir o número de sentenças. Assim, o questionário de satisfação abordou 19 questões afirmativas, que possibilitaram uma análise sobre a utilização do Oráculo, através de questões agrupadas em quatro fatores de usabilidade: satisfação (S), aprendizagem (A), ajuda (D), e controle que a interface oferece ao longo da interação (C). Através das afirmações, os participantes deveriam selecionar apenas uma resposta entre as três opções disponíveis: 1. Concordo; 2. Indeciso; 3. Discordo.

### 9.1.2. **Technology Acceptance Model (TAM)**

O *Technology Acceptance Model* (TAM) (DAVIS, 1989) consiste em um instrumento de avaliação que tem como objetivo medir a aceitação dos usuários em relação a uma tecnologia.

Davis (1989) define os cinco princípios determinantes do modelo TAM: (1) Utilidade percebida - grau em que uma pessoa acredita que o uso de um sistema particular pode melhorar o seu desempenho; e (2) Facilidade de uso percebida - grau em que uma pessoa acredita que o uso de um sistema de informação será livre de esforço; (3) Atitude, definido como um sentimento individual em relação a determinado comportamento; (4) Intenção Comportamental de Uso, definido como o grau em que uma pessoa tem a intenção de desempenhar determinado comportamento; e (5) Comportamento ou Uso real do sistema. Para refletir o grau de concordância e discordância do usuário a afirmações relacionadas a estes princípios, é utilizada uma escala Likert de sete pontos.

Assim, o TAM normalmente é utilizado para entender o porquê do usuário aceitar ou rejeitar a tecnologia de informação e como melhorar a aceitação, oferecendo um suporte para prever e explicar a aceitação.

Para sua aplicação nesta pesquisa, o TAM original foi adaptado, com foco na Facilidade de Uso Percebida e na Intenção Comportamental de Uso.

O questionário de aceitação foi então elaborado contendo 11 questões que possibilitam uma análise sobre a utilização do Oráculo, através das questões agrupadas em dois princípios básicos: a Facilidade de Uso Percebida (FUP) e a Intenção Comportamental de Uso (ICU). Através das afirmações, os usuários finais, identificados como P01 a P05 marcaram apenas uma resposta entre as sete opções disponíveis: 1. Discordo inteiramente; 2. Discordo em grande



parte; 3. Discordo Parcialmente; 4. Neutro; 5. Concordo Parcialmente; 6. Concordo em grande parte; e 7. Concordo inteiramente.

## 9.2. Execução

Para a avaliação com potenciais usuários do framework Oráculo, cinco pesquisadores foram convidados para participar, sendo estes selecionados a partir da rede de contatos da autora desta pesquisa com pesquisadores que trabalham com coleta e análise de dados de mídias sociais em suas pesquisas científicas. Como queríamos investigar a qualidade da interação do *design*, era importante que todos eles desenvolvessem pesquisas na área de mineração de textos em redes sociais *online*, ou tivessem interesse no tema.

A avaliação com potenciais usuários foi realizada em uma única sessão, no dia 26 de novembro de 2019, no Laboratório do Grupo de Pesquisa MTPLNAM, com duração aproximada de uma hora, e consistiu de quatro etapas: apresentação do framework Oráculo, execução das tarefas, resposta aos questionários SUMI e TAM adaptados e entrevistas pós-teste. Cabe salientar que, a estrutura aqui mostrada para a avaliação foi obtida após a execução de um teste piloto, realizado com a finalidade de verificar a exequibilidade das tarefas solicitadas durante o teste com usuários, além da compreensibilidade das questões dos questionários e entrevista pós-teste, a fim de detectar possíveis problemas, e então refiná-los.

Assim, na primeira etapa, foi informado aos participantes o objetivo da avaliação, e coletado o consentimento dos mesmos, através da assinatura do termo de consentimento (veja Apêndice B- Termo de Consentimento Livre e Esclarecido). Ainda nesta etapa, os participantes responderam a um questionário (veja Apêndice B – Questionário Pré-Teste), com questões sobre o levantamento do perfil, formação profissional e a sua experiência de uso de ferramentas de coleta de dados e mineração de textos. A primeira etapa foi encerrada com uma apresentação sobre o *framework* Oráculo que teve duração aproximada de 5 minutos.

Em seguida, na segunda etapa da avaliação com potenciais usuários, cada participante recebeu material com as instruções de uso para o teste proposto, as limitações que encontrariam ao interagirem com um protótipo de alta fidelidade e não uma versão interativa implementada e solicitando que executassem as seguintes tarefas:

- Tarefa 1: Conhecer o Oráculo ao explorar a ferramenta livremente por cinco minutos;

- Tarefa 2: Realizar uma Nova Coleta de dados utilizando um termo de busca;
- Tarefa 3: Realizar uma Nova Coleta de dados utilizando dois termo de busca;
- Tarefa 4: Aplicar o Filtro de Bots em uma coleta previamente realizada;
- Tarefa 5: Realizar a Mineração em uma coleta previamente realizada e Visualizar os dados minerados a partir da seleção de Nuvem de Palavras;
- Tarefa 6: Realizar a Mineração em uma coleta previamente realizada e Visualizar os dados minerados a partir da seleção de Frequência de Palavras;
- Tarefa 7: Realizar a Mineração em uma coleta previamente realizada e Visualizar os dados minerados a partir da seleção de Número Total de Termos;
- Tarefa 8: Realizar a Mineração em uma coleta previamente realizada e Visualizar os dados minerados a partir da seleção de Gráfico de Frequência.

Na terceira etapa da avaliação com potenciais usuários, os participantes responderam a questões tanto do questionário SUMI quanto do TAM, em um formulário *online* no Google Forms (veja Apêndice C – Questionário pós-teste), expressando sua satisfação com o *design*, bem como as percepções sobre as intenções e facilidade de uso da ferramenta.

Por fim, na quarta etapa, os usuários participaram de uma entrevista sobre o teste (veja Apêndice C – Roteiro da Entrevista pós-teste), com três participando da entrevista em grupo, e dois entrevistados de forma individual. Nesse momento, os participantes tiveram oportunidade de falar sobre suas impressões relacionadas às principais dificuldades encontradas para a execução das tarefas propostas e se estas foram concluídas com facilidade. Também foram exploradas, durante as entrevistas, as opiniões dos participantes sobre a aparência proposta no *design* e quais as sugestões de melhorias eles conseguiram pontuar.

Os seguintes materiais foram coletados durante o teste com potenciais usuários: (1) termo de consentimento assinado pelos participantes; (2) questionários pré-teste preenchido pelos participantes; (3) gravação da interação durante realização das tarefas; (4) respostas ao questionário SUMI; (5) respostas ao questionário TAM e, finalmente, (6) o áudio das entrevistas, que foram transcritos para que pudessem ser devidamente analisados, juntamente com todo o material gerado.

### 9.3. Resultados e Discussão

Nesta seção, apresentamos e discutimos sobre os resultados gerados a partir do teste realizado com os usuários, no sentido de investigar a qualidade da interação proporcionada pelo *design* proposto durante a execução das tarefas apresentadas na seção 8.2. Sendo assim, na subseção 9.3.1, mostramos os resultados relacionados a aplicação do questionário SUMI, focamos na investigação da satisfação dos usuários durante a interação com o *design* proposto, e levantamos questões relacionadas à qualidade da ferramenta sob o ponto de vista dos participantes, bem como detectamos possíveis falhas de usabilidade. Na subseção 9.3.2, mostramos os resultados relacionados à aceitação do usuário a partir a aplicação do questionário TAM, focando na facilidade e intenção comportamental de uso e a pelo uso da ferramenta. Em seguida, na seção 9.3.3, reforçamos os resultados obtidos com as respostas a esses questionários, a partir da interpretação das respostas coletadas na entrevista oral, apresentando os principais argumentos relacionados à percepção dos participantes em relação ao *design* proposto, e sugestões de possíveis melhorias para a interface.

Conforme já mencionado, a avaliação envolveu cinco participantes, todos considerados potenciais usuários do framework Oráculo. Todos os participantes relataram realizar pesquisas envolvendo a coleta e análise de dados de redes sociais *online*, sendo que a maioria realiza suas atividades de coleta e mineração através de scripts. Durante a utilização do Oráculo, foi possível verificar o interesse e a motivação dos participantes em utilizar todos os recursos oferecidos. Os participantes também anotaram erros encontrados, o que contribuiu para fornecerem sugestões de melhorias para o *design* após o uso, na entrevista pós-teste.

#### 9.3.1. A satisfação do usuário durante a interação e a qualidade do *design* proposto

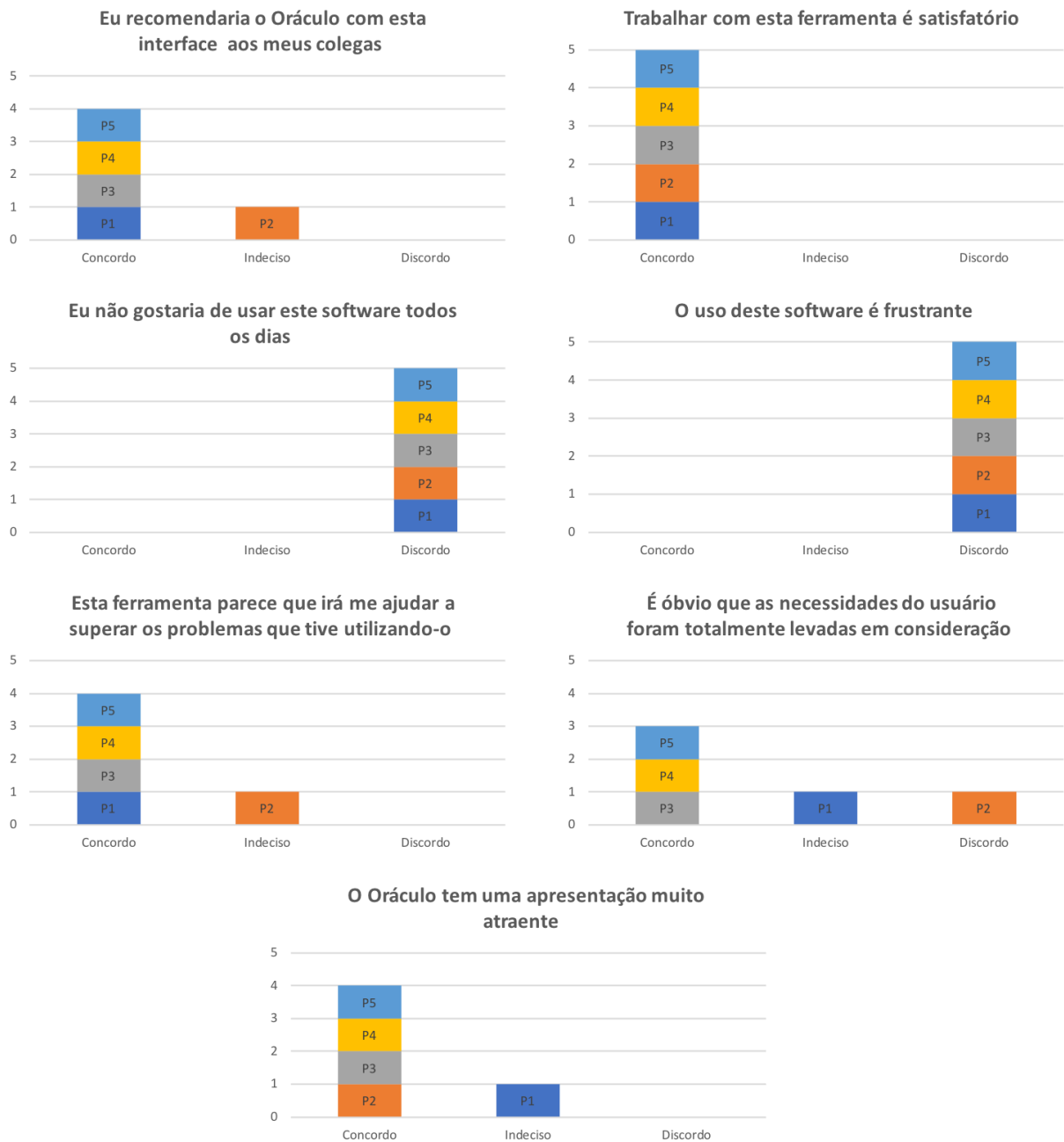
Os indicadores referentes aos resultados mostrados nesta subseção foram obtidos a partir da aplicação do questionário SUMI, bem como nas respostas dos participantes às questões relacionadas à satisfação do usuário e qualidade do *design*, na entrevista pós-teste. A análise dos dados quantitativos coletados foi realizada através primeiramente da tabulação das respostas e pela exportação dos gráficos gerados automaticamente pelo Google Forms, fornecendo uma imagem clara, que nos auxiliou no processo de interpretação dos resultados.

A seguir são apresentados os resultados referentes à avaliação do questionário de satisfação por categoria, e discussões correspondentes ocorridas na durante a entrevista oral em grupo e dos dois participantes entrevistados individualmente.

#### **9.3.1.1.Satisfação (S)**

O Gráfico 46 indica a visão dos participantes do estudo em relação à categoria “satisfação” do questionário SUMI.

Gráfico 1- Satisfação



Fonte: próprio autor

Pôde-se perceber que todos os usuários se manifestaram satisfeitos com o *design* proposto, sensação esta reforçada pela unanimidade ao discordar da afirmação de que o seu uso é frustrante ou que não gostaria de utiliza-lo todos os dias. Ao serem questionados, na entrevista pós testes, se tivessem que utilizar o Oráculo em suas respectivas pesquisas se estariam satisfeitos, todos os entrevistados responderam que “sim”, e concordaram que a

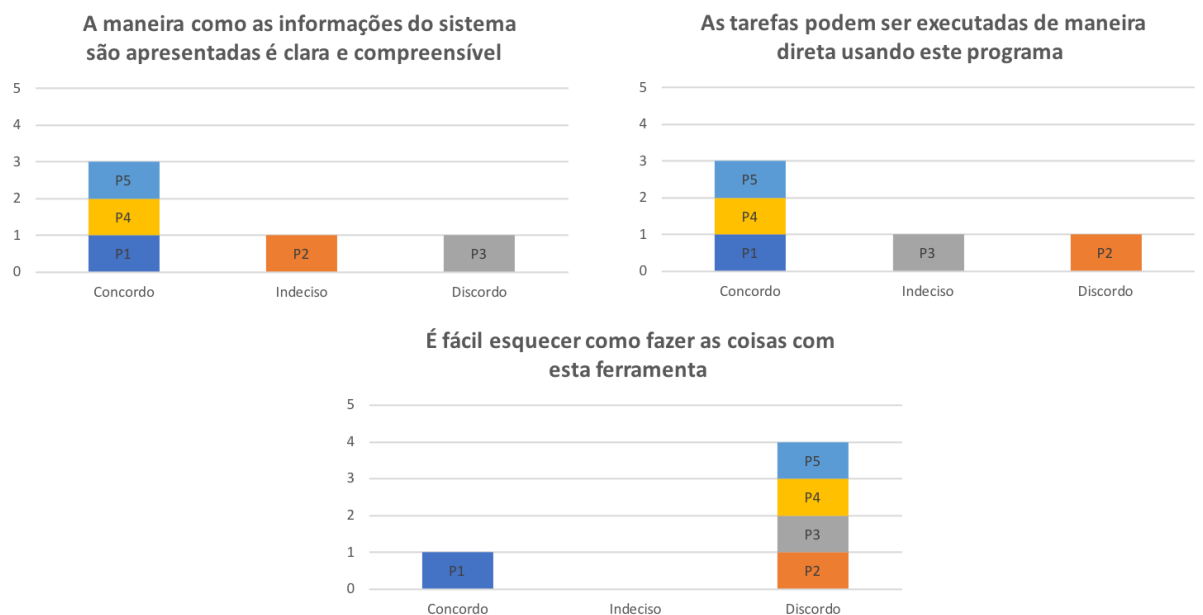
aparência proposta para o Oráculo é agradável, como pode ser percebido na fala de P02: “*sim porque gostei muito e achei muito interessante a proposta do interface para a ferramenta*”.

Nas demais afirmações referentes a esta categoria, temos algumas divergências de opiniões que precisam ser levadas em consideração. Por exemplo, um dos usuários mostrou-se indeciso sobre a afirmativa de que a ferramenta lhe ajudaria a solucionar os problemas que possuía. O participante P02 menciona, que “*não estou completamente satisfeito atualmente, tem muitas questões que precisam ser melhoradas, mas eu o utilizaria sim atualmente e vejo que posso ficar satisfeito futuramente*”. Isto mostra o reconhecimento da importância do Oráculo e da interface proposta para ele, mas ressalta a necessidade de ajustes para aprimoramento da ferramenta. Esta mesma percepção pode ser observada em respostas de indecisão e de discordância obtidas na afirmação sobre todas as necessidades dos usuários terem sido levadas em consideração.

### 9.3.1.2. Aprendizagem (A)

O Gráfico 47 indica a visão dos participantes do estudo em relação à categoria “aprendizagem” do questionário SUMI.

Gráfico 2- Aprendizagem



Fonte: próprio autor

As duas primeiras afirmativas desta categoria são praticamente complementares e acredita-se que por este motivo os resultados obtidos foram similares de indecisão entre os participantes. Um dos participantes não concorda, outro se mostrou indeciso e os demais, que são a maioria, concordaram que a maneira como as informações estão apresentadas é clara e por este motivo as tarefas podem ser executadas de maneira direta.

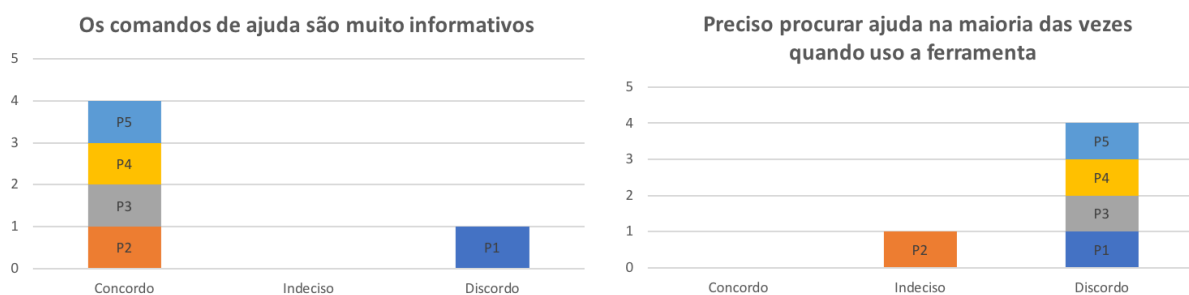
Acreditamos que o fato da indecisão e discordância dos participantes seja pelo uso de algumas nomenclaturas ineficientes na interface, o que é reforçado na fala do participante P02: *“teve algumas coisas das tarefas que eu não encontrei e teve coisas que não gostei do nome e não entendi...mas assim, são coisas que eu esperava fazer uma coisa e acabou fazendo outra, como aconteceu em uma das telas que estava escrito ‘sair,’ sendo que pensei que deveria ser ‘voltar’, ai quando cliquei saiu do sistema...”* Apesar da divergência de opiniões, a categoria de aprendizagem é bem avaliada pelos usuários, mesmo que necessite de melhorias, pois a maioria dos participantes discordaram com a afirmação de que é fácil esquecer como fazer as coisas com esta ferramenta.

Uma outra questão levantada pelos participantes foi a dificuldade de aprender ou descobrir que o usuário realiza, além da coleta, a mineração dos dados previamente coletados. O participante P05 reforça esta questão ao afirmar que *“achei meio escondido a função de minerar. Quando bati o olho na página inicial você não tem lá esta opção, então você não sabe que o Oráculo faz até ele fazer”*.

### 9.3.1.3. Ajuda (D)

O Gráfico 48 indica a visão dos participantes do estudo em relação à categoria “ajuda” do questionário SUMI.

Gráfico 3- Ajuda



Fonte: próprio autor

Observa-se que as soluções de ajuda estipuladas para o Oráculo são boas opções de auxílio aos usuários, inclusive o P04 teve esta percepção e enfatizou que *“Inclusive toda a ajuda eu acho que vai ajudar muito mesmo pra que não sabe nada de nada da área”*, demonstrando a importância do desenvolvimento do *design* com foco também nos usuários inexperientes, que não são da área de computação.

Apenas um dos participantes discordou que os comandos de ajuda são muito informativos e este mesmo participante, se diferenciando dos demais, afirma que concorda que é preciso procurar ajuda na maioria das vezes quando usou a ferramenta, porém este não demonstrou esta opinião durante a entrevista oral. Já o participante P02 dá um exemplo dos itens de ajuda que julgou super importante: *“Inclusive tem uma tela lá com os operadores, que na hora que eu vi eu falei ‘nossa isso aqui tá em inglês cara’, porque pra quem é da área faz sentido, mas pra quem, sei lá, é da área de linguística, por exemplo, que quer exportar tweets, mas no Oráculo vi depois que está muito bem explicado o que é cada operador, por mais que esteja em inglês tem lá a descrição e até um exemplo de como usa. Mas resolveu por que na hora que eu bati olho, pensei que seria um problema, mas aí tem a ajuda que resolveu tudo”*.

#### **9.3.1.4. Controle sobre a interface (C)**

O Gráfico 49 indica a visão dos participantes do estudo em relação à categoria “controle sobre a interface” do questionário SUMI.



Gráfico 4- Controle sobre a interface



Fonte: próprio autor

Nesta categoria é unânime a opinião dos participantes sobre a utilidade dos botões de controle da ferramenta, o que justifica a opinião da maioria discordar com a afirmação de que a ferramenta não faz o que gostaria, ou concordar com a afirmação de que é fácil fazer com que a ferramenta faça exatamente o que deseja.

Assim, os resultados apontam um certo domínio dos usuários em relação a interface, mostrando que conseguem comandar as tarefas que necessitam conforme suas necessidades. Já as dificuldades que tiveram quanto a estes aspectos estão, na maioria das vezes, relacionadas às limitações da interface por estarem interagindo com o protótipo funcional, como pode ser verificado na fala do participante P03: *“as dificuldades que tive foram mais*

*por conta de ser um protótipo...você está acostumado a você comandar a ferramenta, e pelo protótipo tem coisa que seria você que insere e tem outras que já estão ali disponíveis, então não sabia o que de fato estava ou não desabilitado”. Acreditamos que esta percepção do usuário P03 também justifica as respostas de discordo da afirmativa sobre não saber as vezes o que fazer em seguida com a ferramenta.*

Uma outra questão mencionada na entrevista oral que diz respeito ao controle a interface foi enfatizado pelo participante P04 ao afirmar que *“uma outra coisa que fiquei pensando, achei muito interessante do jeito que está porque são realmente etapas diferentes você coletar, filtrar e minerar. Mas fica muito passo a passo, quebradas as atividades, mas é interessante porque se eu quiser fazer somente uma eu também faço. E está tudo aqui, nessa tela de minhas coletas”.*

### **9.3.2. A aceitação do usuário durante a interação**

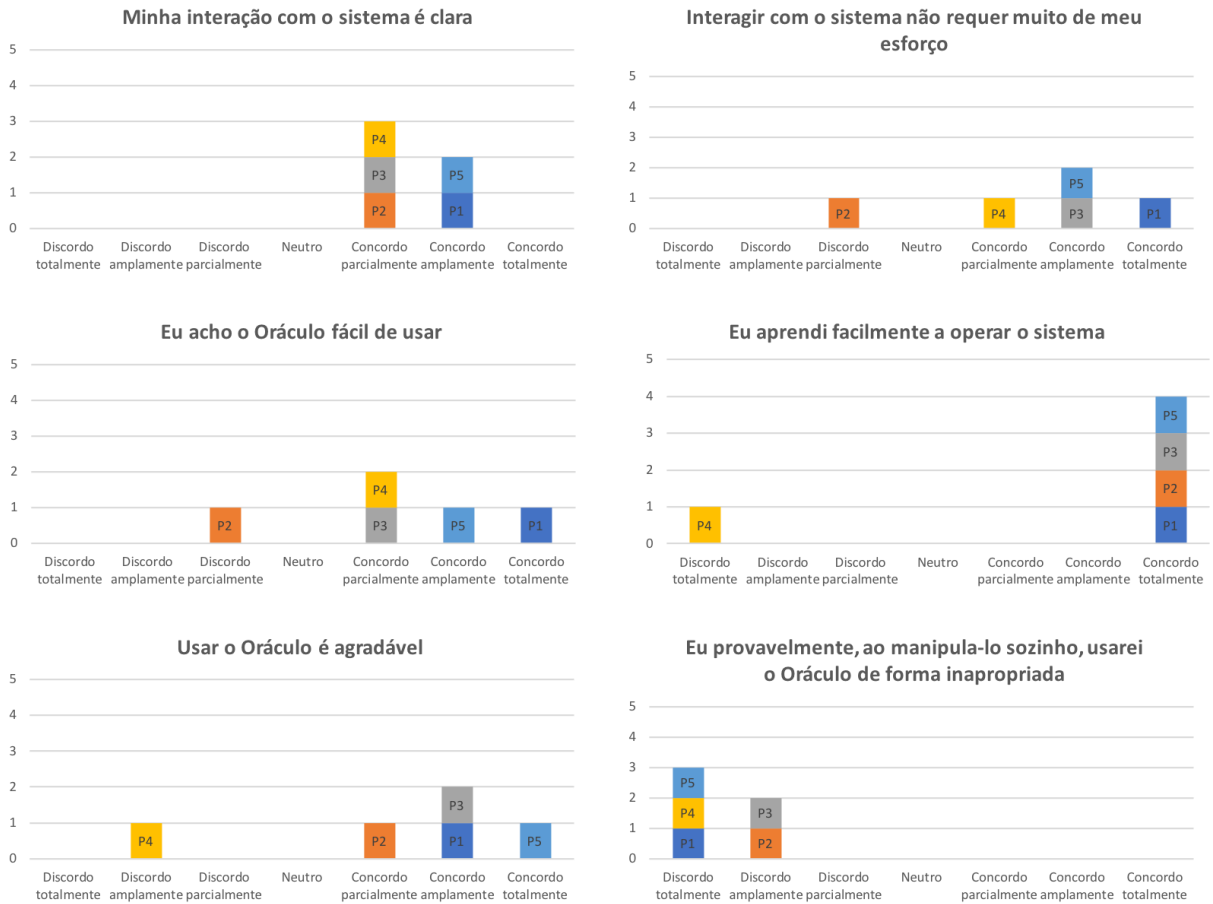
Os indicadores referentes aos resultados mostrados nesta subseção foram obtidos a partir da aplicação do questionário TAM. A análise dos dados quantitativos coletados foi realizada através primeiramente da tabulação das respostas e pela exportação dos gráficos gerados automaticamente pelo Google Forms, fornecendo uma imagem clara, que nos auxiliou no processo de interpretação dos resultados.

A seguir são apresentados os resultados referentes à avaliação do questionário de aceitação por categoria, e discussões correspondentes ocorridas durante a entrevista oral em grupo e dos dois participantes entrevistados individualmente.

#### **9.3.2.1. Facilidade de uso percebida (FUP)**

O fator de Facilidade de uso percebida aborda questões voltadas para o aprendizado e para a operação do *design*. A seguir, o Gráfico 50 indica a visão dos participantes do estudo em relação à categoria “Facilidade de uso percebida” do questionário TAM.

Gráfico 5- Facilidade de Uso Percebida



Fonte: próprio autor

Os resultados demonstram que os participantes percebem a facilidade de uso do Oráculo, visto que os resultados mostram que estes concordam com todas as questões que tendem a aceitação da ferramenta. A questão que possui uma afirmativa com sentido contrário às anteriores, obteve os conceitos “discordo inteiramente” e “discordo em grande parte”, o que demonstra que os participantes provavelmente usariam o Oráculo de forma apropriada, mesmo que sem as orientações das tarefas propostas.

Assim, pode-se afirmar que os respondentes perceberam uma facilidade e uso relevante do *design* do Oráculo, usando e compreendendo o que podem realizar com esta ferramenta. O participante P03 inclusive enfatiza a maior facilidade de realização das coletas e mineração de textos através de uma interface gráfica em comparação com o uso de scripts: “hoje em dia pra mim que faço com scripts é muito mais fácil né? Você vai ali, escolhe as coisinhas que você quer, exporta...já está fácil de visualizar ali, com o arquivo de exportação então é mão na roda”.

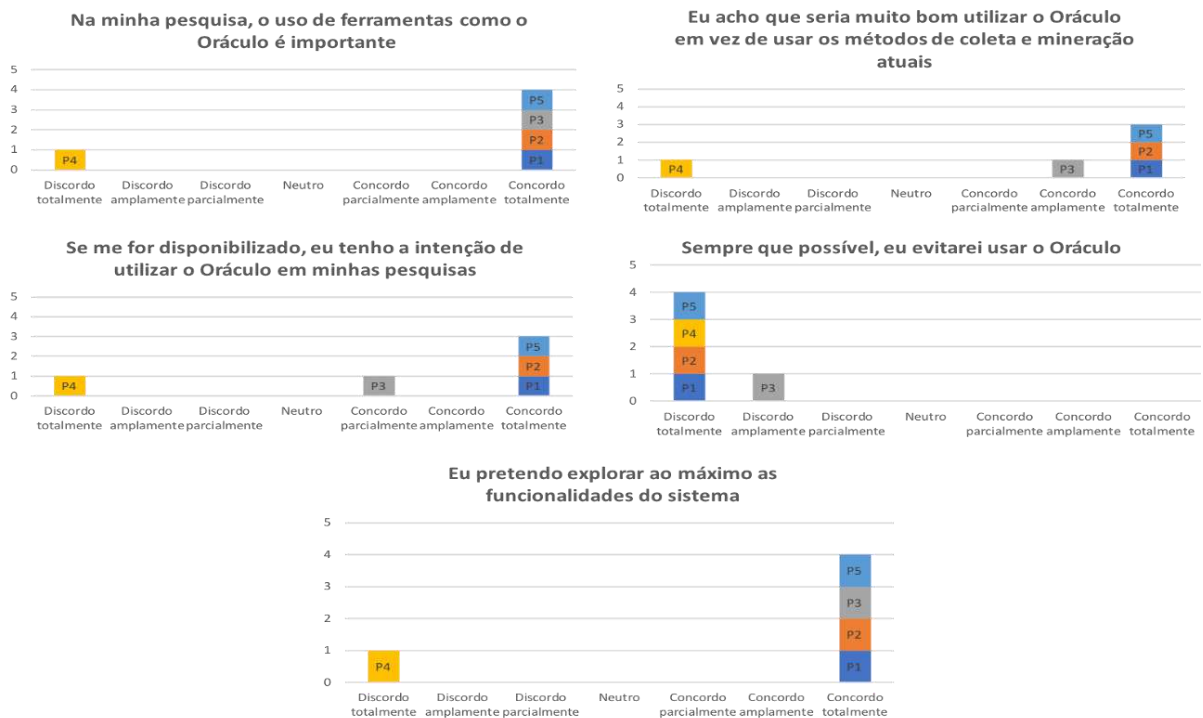
Contudo, vale ainda discutir sobre os possíveis motivos que levaram os participantes a não “concordarem totalmente” com as afirmativas. Durante a entrevista oral em grupo os participantes pontuaram algumas questões que remetem a FUP, como, por exemplo, P02, ao afirmar que “*uma das únicas coisas que não fez muito sentido pra mim foi chamar de ‘Página de Opções’, talvez sei lá, ‘Menu de Atividades’ seria mais claro perceber o intuito era aquela tela ali com as funcionalidades principais*”. Já P05 completa afirmando que “*para mim, o problema nem é somente o nome, acho que a tela que você chamou de ‘Tela Inicial’ é desnecessária, a sua tela de ‘página de opções’ deveria ser a primeira logo depois que você loga*”.

**9.3.2.2. Intenção Comportamental de Uso (ICU)**

Davis *et al.* (1989) afirma que se a facilidade de uso de uma ferramenta for percebida, o indivíduo desenvolverá intenção de uso e, conseqüentemente, utilizará a tecnologia desenvolvida, e em nossos resultados podemos perceber exatamente esta afirmação.

O fator de Intensão Comportamental de Uso (ICU) aborda questões voltadas para a pretensão de utilização da ferramenta. A seguir, o Gráfico 51 indica a visão dos participantes do estudo em relação à categoria “Intensão Comportamental de Uso” do questionário TAM.

Gráfico 6- Intenção Comportamental de Uso



Fonte: próprio autor.

As análises demonstraram que existe uma aceitabilidade satisfatória pelos usuários em relação ao *design* proposto, refletindo em uma real vontade dos usuários de utilizarem o Oráculo em suas pesquisas, visto que, os resultados para todas as questões que tendem a aceitação da ferramenta ao “concordo parcialmente” e “concordo em grande parte” em todas as afirmativas favoráveis ao *design*. A questão não favorável representa um conceito entre “discordo inteiramente” e “discordo em grande parte”, o que demonstra que os participantes provavelmente não evitarão a utilização do Oráculo quando obtiverem outros recursos às suas disposições.

Vale ainda discutir sobre os possíveis motivos que levaram a um dos participantes avaliar o *design* com valores tão discrepantes em relação aos demais, quando se diz respeito a aceitação da ferramenta. Durante a entrevista oral em grupo, o entrevistado P04 levantou algumas questões que remetem à sua ICU, como as opções limitadas de configuração para a realização da mineração e visualização dos dados: *“essa coisa de selecionar a mineração e automaticamente selecionar a mineração, isso é muito confuso, porque para mim uma coisa é a mineração e outra coisa é a visualização. A mineração envolve várias técnicas e eu não vi nenhum nome de técnica de mineração ali, vi opções de visualização”*. Esta colocação do participante é interessante ser levada em consideração para um planejamento de melhoria da interface em relação as opções de mineração futuras que o Oráculo irá realizar.

Até o momento, o Oráculo foi desenvolvido utilizando um único algoritmo de mineração que faz, em sequência, o carregamento de tweets coletados, contagem de tweets na base de dados, remoção de pontuações, remoção de stopwords, normalização de textos, contagem dos termos e realiza o cálculo da frequência dos termos, opções estas não configuráveis, com exceção da remoção de stopwords. A partir deste algoritmo, temos os produtos derivados, elaborados pelo *design* proposto, que, chamamos na interface de opções de mineração e visualização da coleta minerada. A partir da reflexão da colocação do participante P01, é realmente necessário, como ele mesmo diz, a realização de *“um estudo mais aprofundado sobre quais as técnicas de mineração existentes, quais são, o que elas retornam, o que o usuário pode obter com isso”*. E o mesmo se aplica à visualização de dados, para se conseguir disponibilizar, na interface, um leque de possibilidades ao usuário para que possa customizar a mineração e visualização, conforme suas necessidades.

Ainda sobre as possíveis considerações levantadas sobre o ICU diz respeito a inclusão da possibilidade de configurações do Oráculo para realização do Filtro de Bots. Os

participantes P01, P02 e P04 afirmam que sabem que a interface ainda será aprimorada a medida que novas funcionalidades forem sendo desenvolvidas, porém eles já sugerem que “*seria interessante mostrar quais critérios de filtro, por exemplo inserir quais são os perfis que foram filtrados, sei lá, citar quantos, quais e o porque foram considerados ou não bots*” (P03). É interessante ressaltar que o processo de identificação de bots é bastante complexo e que, como afirma o participante P04, “*muitas contas genuínas serão filtradas como bots*”, sendo assim, o entrevistado P03 sugeriu que “*o usuário deveria escolher o nível de exigência de sua filtragem para identificação dos bots*”, uma sugestão que seria interessante, pois passa para o usuário a opção de controlar a ferramenta no sentido de escolher a melhor opção conforme as suas necessidades.

### 9.3.3. As impressões dos usuários e propostas de melhorias das telas

As análises das respostas dos participantes aos questionários SUMI e TAM demonstraram, respectivamente, que os participantes, em sua maioria, estão satisfeitos com desing proposto e pretendem utilizar o Oráculo como apoio de suas pesquisas. Contudo, é importante ressaltar que existem algumas dificuldades e necessidades que precisam ser melhor analisadas, e estas foram também identificadas pelos participantes da pesquisa, como mostra a Tabela 3 a seguir.

Tabela 3- Sugetões de melhorias

DEFICIÊNCIA IDENTIFICADA	SUGESTÃO DE MELHORIA
Nome não intuitivo para a página com os menus	Colocar um nome mais intuitivo para a tela, até então denominada Página de Opções
Desnecessário uma tela para desejar boas vindas ao usuário	Retirar a “Página Inicial” e colocar o link para o “Meu Perfil” e as Boas Vindas na “Página de Opções”
Confusão de signos em Nova Coleta ao adicionar novo termo	Rever uma melhor estrutura para adicionar mais de um termo a coleta e acrescentar o operador de relação entre os termos
Limitado a exportação dos dados	Acrescentar a opção de exportar todos os formados de dados visualizados após a realização da mineração.
Filtro de Bots não retorna dados após exportar	Elaborar uma tela contendo as informações sobre o filtro realizado, como um feedback ao usuário
Configurações de Filtro de Bots limitado	Estudar e propor as funcionalidades e signos que

		comportem a configuração do filtro de bots no Oráculo
Configuração de Mineração e Visualização juntos		Após o desenvolvimento de novos algoritmos de mineração, propor na tela algumas opções de seleção de qual algoritmo utilizar.
Opção de Mineração está escondida ao existir somente na tela de Minhas Coletas		Acrescentar opção de Mineração e Visualização de Dados na página de opções.

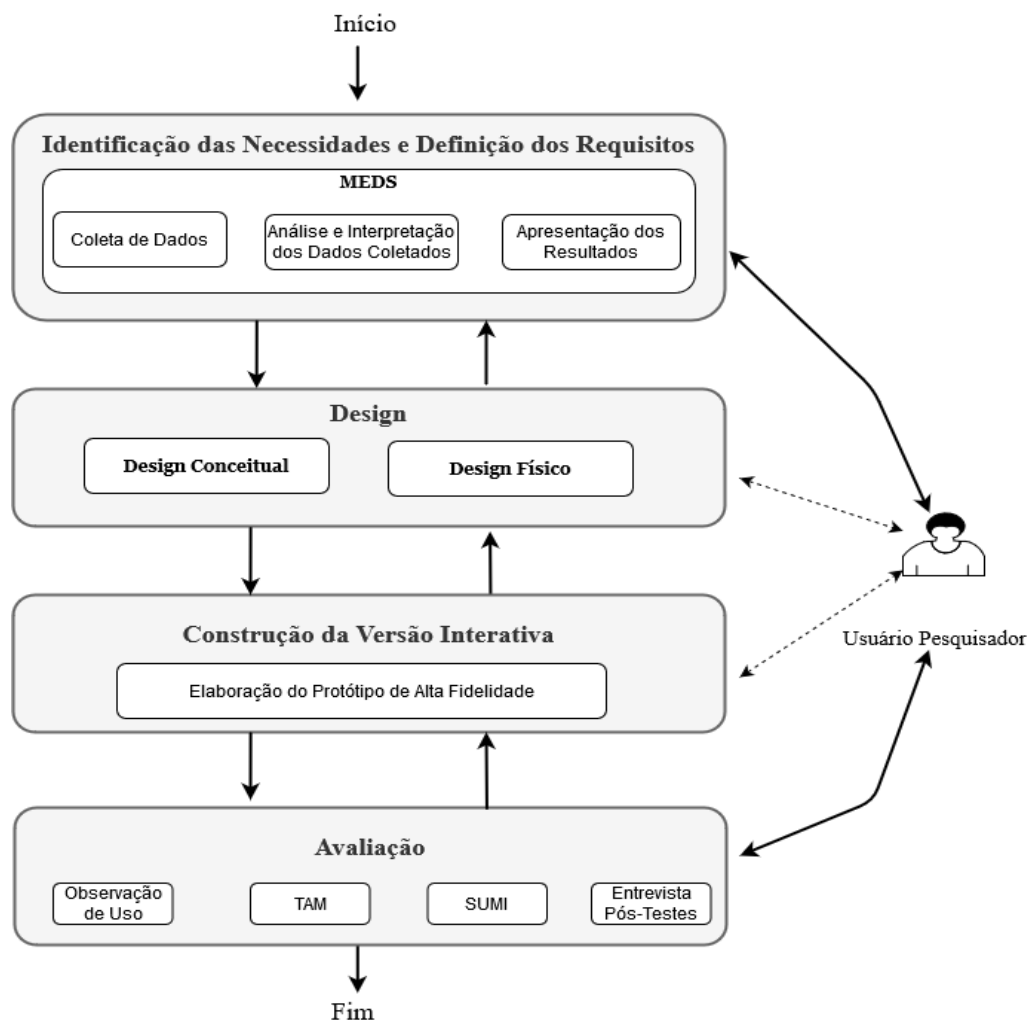
---

Fonte: próprio autor

## 10. SÍNTESE DO MODELO DE *DESIGN* DA INTERAÇÃO DO *FRAMEWORK* ORÁCULO

A partir da pesquisa desenvolvida, foi possível então propor um modelo, representado pela figura 46, que esboça as atividades para a elaboração do *design* da interação para o *Framework* Oráculo.

Figura 46- Síntese do modelo da pesquisa



Fonte: próprio autor

A atividade de Identificação das Necessidades e Definição dos Requisitos é orientada pelo MEDS, sendo desenvolvida em três etapas principais:

- Coletas de Dados – se preocupa com as questões éticas e com a elaboração e aplicação de entrevistas com os participantes da pesquisa.



- Análise e Interpretação dos Dados Coletados – os dados coletados nas entrevistas são transcritos e são aplicadas análises de intra e inter sujeitos, a partir das recorrências nos depoimentos dos participantes.
- Apresentação dos Resultados – o modelo propõe a utilização de Cenários de Análise e a utilização de Personas.

A segunda etapa, *Design*, implica no projeto de intervenção, sendo criados:

- *Design* Conceitual – elaboração da modelagem conceitual para o produto. Assim, é construída a mensagem de metacomunicação de alto nível, e a partir desta, é criado o modelo da interação, representado pelo Diagrama MoLIC.
- *Design* Físico – atividade em que o *designer* deve se preocupar com a definição da interface. Este modelo propõe a constituição da Engenharia de Signos da Interface do DCC, em que os elementos da MoLIC são mapeados na interface.

A terceira etapa, Construção da Versão Interativa, consiste no desenvolvimento da versão interativa do *framework* Oráculo, através da construção do protótipo de alta fidelidade.

Na quarta e ultima etapa, de Avaliação, o modelo propõe a aplicação de quatro técnicas para a investigação da qualidade da interação proporcionada pelo *design*. São elas:

- Observação de uso – em um ambiente controlado, propor tarefas para que os participantes da pesquisa executem enquanto são observados conforme tópicos de interesse do observador;
- TAM – aplicação do questionário TAM para verificação da aceitação do usuário em relação a nova tecnologia.
- SUMI – aplicação do questionário SUMI para verificação da satisfação do usuário durante a interação e a qualidade do *design* proposto.
- Entrevista Pós-Testes – aplicação de uma entrevista ao final do processo de avaliação, para que os participantes possam esboçar oralmente suas facilidades e dificuldades.

## 11. CONSIDERAÇÕES FINAIS

O crescente uso das redes sociais *online* tem despertado interesse dos pesquisadores, de diferentes áreas de conhecimento. Estes estão extraindo e analisando dados de mídias sociais como fonte de informações para suas pesquisas científicas, aplicando técnicas e utilizando ferramentas de mineração de textos. Porém existe, até o momento, pouca atenção por parte dos desenvolvedores destas ferramentas no que diz respeito a suas interfaces, o que as tornam muitas vezes complexas e de difícil aprendizado.

Neste contexto, o principal objetivo da pesquisa foi desenvolver o *design* da interação do *Framework* Oráculo, uma ferramenta de apoio ao processo de mineração de textos em redes sociais *online*, com foco em sua qualidade de uso. Assim, o trabalho foi desenvolvido com foco na seguinte questão norteadora: Como apoiar os pesquisadores de instituições de ensino superior no processo de mineração de textos em redes sociais *online*, por meio do *design* da interação de uma ferramenta para coleta e análise de dados provenientes dessas mídias?

Para isso, inicialmente foi feita a identificação das necessidades e definição dos requisitos de pesquisadores em relação à coleta e análise de dados em redes sociais *online* e, em seguida, foi feito o *design* do *framework*, seguindo a abordagem do *Design* Centrado na Comunicação. Foi criado o *design* conceitual e físico, a partir da construção da metamensagem do *designer*, do modelo de interação e do protótipo de baixa fidelidade do *framework* Oráculo.

A partir da primeira etapa da pesquisa, “Identificação das Necessidades e Definição dos Requisitos”, obteve-se uma maior compreensão sobre este universo de pesquisa e as necessidades dos pesquisadores que trabalham com mineração de textos em RSOs, em relação às ferramentas que os apoiem nesse processo. Os resultados indicaram que há uma insatisfação por parte destes profissionais em relação às ferramentas existentes. Tal fato leva-os muitas vezes a utilizarem estratégias para a coleta e análise de dados que exigem um conhecimento mais específico na área de Computação, que nem todos eles possuem. Além disso, o estudo destaca também as necessidades dos pesquisadores de dados no sentido de se criar uma ferramenta, cujo *design* lhes possa proporcionar uma maior efetividade durante a coleta e mineração dados de redes sociais *online*.

Na segunda etapa da pesquisa, focamos no projeto de intervenção, propondo um *design* para o *Framework* Oráculo que acreditamos atender as necessidades dos usuários e os requisitos identificados na etapa anterior. Nesta etapa propomos o modelo conceitual, sendo

descritos o que os usuários poderão fazer com tal ferramenta e quais conceitos são necessários para entender como realizar tal interação. Posteriormente, a partir do modelo conceitual, realizado utilizando a MoLIC, propomos o *design* físico, desenvolvendo o projeto da interface de um usuário do Oráculo ao esboçar protótipos de baixa fidelidade.

A partir do *design* físico elaborado, foi desenvolvido um protótipo funcional de alta fidelidade, para posterior avaliação do *design* proposto. Tal avaliação ocorreu com potenciais usuários para o Oráculo e com o propósito de investigar a qualidade da interação proporcionada pelo *design* proposto aos seus usuários, focando especificamente na experiência de uso, e facilidade de uso percebida e aceitabilidade.

Os resultados apontaram para uma facilidade de uso e aceitabilidade relevantes do *design* proposto. Os usuários se manifestaram satisfeitos com a interface com a qual interagiram, e que, apesar de exigir melhorias pontuadas, reconhecem a importância do Oráculo, refletindo em uma real vontade de utilizá-lo em suas pesquisas.

A partir dos resultados obtidos no trabalho, nas seguintes seções serão apresentadas contribuições e possíveis trabalhos futuros relacionados a esta pesquisa.

### **11.1. Contribuições**

Com o *design* da interação do *Framework* Oráculo, acredita-se que foi possível contribuir com os pesquisadores de instituições de ensino ao lhes proporcionar qualidade de uso ao *framework*, proporcionando uma experiência de uso positiva no processo de mineração de textos.

Quanto as áreas de IHC e mineração de textos, a pesquisa contribui ao trazer considerações sobre aspectos relevantes para o *design* de ferramentas de coleta e mineração de dados provenientes de RSOs.

Na área de pesquisa de IHC, o presente trabalho contribui ao propor um processo de *design* de interação, com atividades específicas para o *design* de uma ferramenta de mineração de textos em redes sociais *online*.

Por fim, foi possível também contribuir com os pesquisadores que utilizam dados provenientes de redes sociais *online* em suas pesquisas, ao lhes oferecer o *design* de uma ferramenta interativa de coleta de dados que lhes proporcione uma experiência de uso positiva no alcance de seus objetivos.

## 11.2. Trabalhos Futuros

Os resultados deste trabalho sugerem alguns caminhos de trabalho futuro:

- o desenvolvimento da interface proposta para ser agregado às demais camadas do *Framework* Oráculo;
- a elaboração e desenvolvimento do *design* para o Oráculo atendendo especificidades de outras redes sociais, diferentes do Twitter apresentado nesta pesquisa;
- o desenvolvimento do sistema de ajuda do Oráculo bem elaborado e fundamentado a partir de métodos de IHC;
- aprimoramento das funções de mineração e visualização de dados, incluindo maiores possibilidades de configurações ao usuário;
- correção dos erros nas telas propostas na seção 9.3.3;
- desenvolver a interface para as telas de filtro de bots após a seu desenvolvimento com o grupo MTPLNAM;
- anexar ao Oráculo a camada de análise dos dados, incluindo o desenvolvimento do módulo e a interface se adaptar às estas possibilidades.
- utilizar métodos da Engenharia Semiótica, que avaliam a comunicabilidade do modelo proposto, uma vez que este está fundamentado no *design* centrado na comunicação.

## 12. REFERÊNCIAS

- AHMED, Ashir; SCHEEPERS, Helana; STOCKDALE, Rosemary. Social Media Research: A Review of Academic Research and Future Research Directions. *Pacific Asia Journal of the Association for Information Systems*, v. 6, n. 1, p. 21–37, 2014.
- ARANHA, Christian; PASSOS, Emmanuel. A Tecnologia de Mineração de Textos. *RESI-Revista Eletrônica de Sistemas de Informação*, v. 2, p. 1–8, 2006. Disponível em: <<https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CC4QFjAA&url=http://revistas.facecla.com.br/index.php/reinfo/article/view/171/0&ei=2aTpUbCeF5Gx4AOM3YH4BQ&usg=AFQjCNGw9NhonODwp7diAmhkYwU374FABg&sig2=1hL7ctx6DQlnxZ1fPFp7Rg&bv>>.
- ARAÚJO, Ellen Thallita Hill *et al.* Use of social networks for data collection in scientific productions in the health area: Integrative literature review. *Aquichan*, v. 19, n. 2, p. 1–12, 2019.
- BARBOSA, Simone D. J.; SILVA, Bruno S. Da. *Interação Humano-computador*. Rio de Janeiro: Campus, 2010.
- BARBOSA, Simone Diniz Junqueira; SILVA, Bruno Santana. Design da Interação Humano-Computador com MoLIC. *IHC '14 Companion Proceedings of the 13th Brazilian Symposium on Human Factors in Computing Systems*, p. 109–138, 2014. Disponível em: <<http://dl.acm.org/citation.cfm?id=2738165.2738199>>.
- BARBOSA, Simone Diniz Junqueira; SILVA, Bruno Santana Da. *Interação Humano-Computador*. Rio de Janeiro: [s.n.], 2010.
- BARION, Eliana Cristina Nogueira; LAGO, Decio. Mineração de textos. *Revista de Ciências Exatas e Tecnologia*, v. 3, n. 3, p. 123–140, 2015. Disponível em: <<http://revista.pgsskroton.com.br/index.php/rcext/article/view/2372>>.
- BENEVENUTO, F. Redes Sociais *Online*: Técnicas de Coleta, Abordagens de Medição e Desafios Futuros. *Tópicos em Sistemas Colaborativos, Interativos, Multimídia, Web e Banco de Dados*, p. 41–70, 2010. Disponível em: <<http://www.dcc.ufmg.br/~fabricio/download/mini-curso-swib10.pdf>>.
- BISTA, Krishna. Is Twitter a Pedagogical Tool in Higher Education? Perspectives of Education Graduate Students. *Journal of the Scholarship of Teaching and Learning*, v. 15, n. 2, p. 83–102, 2015.
- BORRA, E *et al.* *Programmed method: Developing a toolset for capturing and analyzing tweets*. [S.l.: s.n.], 2014.
- BROOKER, Phillip; BARNETT, Julie; CRIBBIN, Timothy. Doing social media analytics. *Big Data & Society*, v. 3, n. 2, p. 205395171665806, 2016.
- BROOKS, Michael. *Human Centered Tools for Analyzing Online Social Data*. 2015. 218 f. University of Washington, 2015.
- BRUNS, Axel; LIANG, Yurian Eugene. Tools and methods for capturing Twitter data during natural disasters. BRUNS, A.; LIANG, Y. E. Tools and methods for capturing twitter data during natural disasters. *First Monday*, University of Chicago, v. 17, n. 4, p. 1–8, 2012., v. 17, n. 4, p. 1–8, 2012.
- CARVALHO, Jof De. O papel da interação humano-computador na inclusão digital. *Transinformação*, v. 15, p. 75–89, 2012. Disponível em: <<http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/1461>>.

CARVALHO, Marcelo Batista De; TSUNODA, Denise Fukumi. Análise de dados em artigos recuperados da *Web of Science* ( WoS ). *Revista eletrônica de biblioteconomia e ciência da informação*, (VIJAYARANI; MUTHULAKSHMI, 2013, v. 23, p. 112–125, 2018).

CASTRO, Lenadro Nunes De; FERRARI, Daniel Gomes. *Introdução à Mineração de Dados. Conceitos Básicos, Algoritmos e Aplicações*. 1. ed. São Paulo: [s.n.], 2016.

CASTRO, Fernanda Meneses de Miranda; SPINOLA, Carolina de Andrade. Metodologia de pesquisas na internet: breves considerações sobre uma pesquisa qualitativa em turismo nas redes sociais. *Revista Iberoamericana de Turismo – RITUR*, v. 5, n. 51, p. 170–188, 2015.

CHEN, Nan-Chen *et al.* Lariat: A Visual Analytics Tool for Social Media Researchers to Explore Twitter Datasets. 2017, [S.l.: s.n.], 2017.

CHEN, Pai-Lin; CHENG, Yu-Chung; CHEN, Kung. Analysis of Social Media Data: An Introduction to the Characteristics and Chronological Process BT - Big Data in Computational Social Science and Humanities. In: CHEN, SHU-HENG (Org.). *Big Data in Computational Social Science and Humanities*. Cham: Springer International Publishing, 2018. p. 297–321.

CHOMA, Joelma. *ScrumUX: uma abordagem para integrar design de interação do usuário ao processo scrum*. 2015. 144 f. Universidade Federal de São Carlos, 2015.

COSTA, Barbara Regina Lopes. Bola de neve virtual: o uso das redes sociais virtuais no processo de coleta de dados de uma pesquisa científica. *RIGS - Revista Interdisciplinar de Gestão Social*, v. 7, n. 1, p. 15–37, 2018. Disponível em: <<https://portalseer.ufba.br/index.php/rigs/article/viewFile/24649/16131>>.

DAVIS, Fred D. Preparation of Rutile TiO<sub>2</sub> Films by RF Magnetron Sputtering Related content Role of He Gas Mixture on the Growth of Anatase and Rutile TiO<sub>2</sub> Films in RF Magnetron Sputtering Kunio Okimura and Akira Shibata - Deposition of High-Quality TiO<sub>2</sub> Films by RF M. *Japanese Journal of Applied Physics Kunio Okimura et al Jpn. J. Appl. Phys*, v. 34, n. September, 1989.

DE SOUZA, Clarisse S. *The semiotic engineering of human computer interaction*. MIT Press, Cambridge MA. Cambridge, MA: MIT Press, 2005.

DE SOUZA, Clarisse Sieckenius. *The Semiotic Engineering of Human-Computer Interaction*. Kindle Edi ed. [S.l.]: The MIT Press (February 11, 2005), 2005.

FERNEDA, Edberto. Recuperação de Informação: Análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. *Text*, v. 2, n. 1, p. 137, 2003. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/pt-br.php>>.

FERREIRA, Daniel Vitor Costa. *Lean Communication-Centered Design : Um processo leve de design centrado na comunicação*. 2015. 159 f. Pontífica Universidade Católica do Rio de Janeiro, 2015.

GIL, Antonio Carlos. *Como Elaborar Projetos de Pesquisa*. 4<sup>a</sup> ed. São Paulo: [s.n.], 2002.

GOMES, Gracielle Mendonça Rodrigues; CENDÓN, Beatriz Valadares. Análise da integração da recuperação da informação, information search behaviour e interação humano computador para avaliação de sistemas de recuperação da informação. *TransInformação*, Usar na Introdução como Justificativa e como motivação, v. 27, n. 3, p. 277–284, 2015. Disponível em: <<http://dx.doi.org/10.1590/0103-37862015000300009>>.

JORDAN, Katy. From finding a niche to circumventing institutional constraints: Examining the links between academics' *online* networking, institutional roles, and identity-trajectory. *International*

*Review of Research in Open and Distance Learning*, v. 20, n. 2, p. 96–112, 2019.

KAKUSHI, Luciana Emi; ÉVORA, Yolanda Dora Martinez. As redes sociais na educação em enfermagem: Revisão integrativa da literatura. *Revista Latino-Americana de Enfermagem*, v. 24, 2016.

KAYNNE, Mayara *et al.* Mineração de dados do twitter através do r. p. 1–8, 2015.

KIRALOWSKI, Jurek; CORBETT, Maru. SUMI: the Software Usability Measurement Inventory. *British Journal of Educational Technology*, Software Usability Measurement Inventory. British Journal of Educational Technology, 24: 210–212., v. 24, n. 3, p. 210–212, 1993.

LEITÃO, Carla Faria; PRATES, Raquel Oliveira. A Aplicação de Métodos Qualitativos em Computação. *Jornadas de Atualização em Informática*, p. 1–48, 2017. Disponível em: <<http://csbc2017.mackenzie.br/public/files/all/livro-jai.pdf>>.

LEITÃO, Carla Faria; SILVEIRA, Milene Selbach; SOUZA, Clarisse Sieckenius De. Uma Introdução à Engenharia Semiótica: Conceitos e Métodos. *Proceedings of the 12th Brazilian Symposium on Human Factors in Computing Systems. Brazilian Computer Society*, LEITÃO, Carla Faria; SILVEIRA, Milene Seibach; DE SOUZA, Clarisse Sieckenius. Uma introdução à engenharia semiótica: conceitos e métodos. In: *Proceedings of the 12th Brazilian Symposium on Human Factors in Computing Systems. Brazilian Computer Society*, 2013. p. 356-358. Disponível em: <http://www3.serg.inf.puc-rio.br/docs/Leitao-Silveira-deSouza-Tutorial-IHC2013.pdf>, p. 356–358, 2013. Disponível em: <<http://dl.acm.org/citation.cfm?id=2577101.2577200>>.

MARCONI, Maria; LAKATOS, Eva. *Fundamentos de metodologia científica*. [S.l: s.n.], 2003.

MARQUES, Lidiany Kelly da Silva; VIDIGAL, Frederico. Prosumers e redes sociais como fontes de informação mercadológica : uma análise sob a perspectiva da intelligence in Brazilian companies. *TransInformação*, v. 30, n. 1, p. 1–14, 2018.

MARTINS, Mariane Denise. Redes Sociais Virtuais: uma eflexão sobre o lugar de aprendizagem. v. 4, 2015.

MARTINS, RF; PEREIRA, Adriano; BENEVENUTO, Fabrício. An Approach to Sentiment Analysis of Web Applications in Portuguese. *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web*, p. 105–112, 2015. Disponível em: <<http://doi.acm.org/10.1145/2820426.2820446>><<http://dl.acm.org/citation.cfm?id=2820446>>.

MATHEUS MOREIRA DA SILVA. Mineração De Dados No Twitter: Uma Ferramenta Prática Para Extração E Análise Dos Resultados. p. 62, 2017. Disponível em: <[http://www.monografias.ufop.br/bitstream/35400000/325/1/MONOGRAFIA\\_MineraçãoDadosTwitter.pdf](http://www.monografias.ufop.br/bitstream/35400000/325/1/MONOGRAFIA_MineraçãoDadosTwitter.pdf)>.

MOTTI, Vivian Genaro. *Design Centrado no Usuário de um Ambiente de Reunião Instrumentado*. 2009. 106 f. 2009.

MUITOS - BARBOSA E SILVA. *Cap 5 - Livro do Simposio IHC 2014*. [S.l: s.n.], 2014.

NICOLACI-DA-COSTA, Ana Maria; LEITÃO, Carla Faria; ROMÃO-DIAS, Daniela. Como conhecer usuários através do Método de Explicitação do Discurso Subjacente (MEDS). 2004, [S.l: s.n.], 2004. p. 47–56.

NICOLCI-DA-COSTA, Ana Maria; LEITÃO, Carla Faria; ROMÃO-DIAS, Daniela. Como conhecer usuários através do Método de Explicitação do Discurso Subjacente (MEDS). *VI Simpósio sobre Fatores Humanos em Sistemas Computacionais - Mediando e Transformando o Cotidiano.*, p. 47–56,

2004. Disponível em: <[www2.serg.inf.puc-rio.br/docs/MEDS-IHC2004.pdf](http://www2.serg.inf.puc-rio.br/docs/MEDS-IHC2004.pdf)>.

NIELSEN, Jakob. *Usability Engineering*. 1. ed. [S.l.: s.n.], 1993.

OLIVEIRA, Hercules Batista De. *FRAMEWORK ORÁCULO: Camada de Coleta e Mineração de Textos para o Twitter*. 2019. 73 f. Universidade Federal dos Vales do Jequitinhonha e Mucuri, 2019.

PAULA, Maíra G. *Projeto da Interação Humano-Computador Baseado em Modelos Fundamentados na Engenharia Semiótica : Construção de um Modelo de Interação*. 2003. 76 f. 2003.

PEREIRA JUNIOR, M. *UM MODELO PARA APOIAR PROJETISTAS DE SISTEMAS COLABORATIVOS NA ANTECIPAÇÃO DE CENÁRIOS*. 2016. 168 f. Universidade Federal de Minas Gerais, 2016.

PERRIN, Andrew. 65% of Adults Now Use Social Networking Sites – a Nearly Tenfold Jump in the Past Decade. *Pew Research Center*, n. October, p. 2005–2015, 2015.

POLONI, Katia Maria; TOMAÉL, Inês Maria. Coleta De Dados Em Plataformas De Redes Sociais: Estudo De Aplicativos. *Anais do II Workshop de Pesquisa em Ciência da Informação*, 2014.

PRATES, Raquel Oliveira; BARBOSA, Simone Diniz Junqueira. Introdução à Teoria e Prática da Interação Humano Computador fundamentada na Engenharia Semiótica. *XXVII Congresso da Sociedade Brasileira de Computação. Jornadas de Atualização em Informática JAI/SBC*, p. 263–326, 2007.

PREECE, J.; ROGERS, Y.; SHARP, H. *DESIGN DE INTERAÇÃO - Além da IHC*. 3. ed. Porto Alegre: [s.n.], 2013.

PREECE, Jenny; SHARP, Helen; ROGERS, Yvonne. *Design da Interação: Além da Interação Humano-Computador*. 1. ed. Porto Alegre: [s.n.], 2007.

PREECE, Jenny; SHARP, Helen; ROGERS, Yvonne. *Interaction design: beyond human-computer interaction*. *Book*, v. 11, p. 773, 2005. Disponível em: <<http://www.id-book.com/secondedition/>>.

RECUERO, Raquel. Contribuições da Análise de Redes Sociais para o Estudo das Redes Sociais na Internet: O caso da hashtag #Tamojuntodilma e #CalaabocaDilma. *Revista Fronteiras (Online)*, v. 16, n. 1, p. 22, 2014. Disponível em: <<http://www.raquelrecuero.com/fronteirasrecuero2014.pdf>>.

ROGERS, Yvonne; SHARP, Helen; PREECE, Jennifer. *Design de Interação: Além da interação humano-computador*. 3. ed. Porto Alegre: Bookman, 2013.

SANTOS, Caroline Q.; SILVEIRA, Milene S. What Do Social Media Data Analysts Want? p. 1–4, 2018a.

SANTOS, Caroline Q.; SILVEIRA, Milene S. What do social media data analysts want? An analysis from the perspective of data visualization. *IHC 2018*, 2018b, New York, NY, USA: ACM, 2018. p. 54:1--54:4.

SANTOS, Caroline Queiroz; SILVEIRA, Milene Selbach; MANSSOUR, Isabel Harb. Visualization and Social Media Data Analysis: preliminary studies about data analysts' perception. p. 10–13, 2016. Disponível em: <[http://delivery.acm.org/10.1145/3040000/3033737/a35-Santos.pdf?ip=103.5.183.42&id=3033737&acc=ACTIVE SERVICE&key=69AF3716A20387ED.F0EE3E682E94A1DD.4D4702B0C3E38B35.4D4702B0C3E38B35&CFID=798089274&CFTOKEN=47699165&\\_\\_acm\\_\\_=1502883209\\_77ef5d4da9359994ada464b](http://delivery.acm.org/10.1145/3040000/3033737/a35-Santos.pdf?ip=103.5.183.42&id=3033737&acc=ACTIVE SERVICE&key=69AF3716A20387ED.F0EE3E682E94A1DD.4D4702B0C3E38B35.4D4702B0C3E38B35&CFID=798089274&CFTOKEN=47699165&__acm__=1502883209_77ef5d4da9359994ada464b)>.



SHARP, Helen; ROGERS, Yvonne; PREECE, Jenny. *Interaction Design: Beyond Human-Computer Interaction*. 2. ed. New York, NY, USA: [s.n.], 2007.

SILVA, T; STABILE, M. *Monitoramento e Pesquisa em Mídias Sociais*. [S.l: s.n.], 2016. Disponível em: <<http://tarciziosilva.com.br/blog/livro-monitoramento-e-pesquisa-em-midias-sociais-ja-esta-disponivel/>>.

SILVA, Leandro Augusto *et al.* Mineração de Dados em publicações de Fóruns de Discussões do Moodle como geração de Indicadores para aprimoramento da Gestão Educacional. *Anais dos Workshops do IV Congresso Brasileiro de Informática na Educação (CBIE 2015) Mineração*, n. Cbie, p. 1084, 2015. Disponível em: <<http://br-ie.org/pub/index.php/wcbie/article/view/6220>>.

SILVA, Tarcízio. Pesquisa baseada em Dados Sociais Digitais: mapeamento de ferramentas e táticas de coleta de dados no Intercom. *Razón y palabra*, n. 90, p. 51, 2015.

SILVA, Tarcízio. Pesquisa baseada em Dados Sociais Digitais: mapeamento de ferramentas e táticas de coleta de dados no Intercom. *Razon y Palabra*, 2015. , v. 111, n. 479, p. 1009–1010.

SIMONE DINIZ JUNQUEIRA BARBOSA; PAULA, Maíra Greco. Using an interaction model to support communication among hci *design* team members from multidisciplinary backgrounds. 2003, Hotel Bourbon, Curitiba: <https://www.researchgate.net/publication/228857967/download>, 2003. p. 155–162.

TAN, Ah-Hwee. Text mining: The state of the art and the challenges. 1999, [S.l.]: sn, 1999. p. 65–70.

VERAS, RUTH PRAXEDES DE PAIVA. *Estudo de Caso Sobre o Processo de Redesign e Gamificação do Sistema de Levantamento de Necessidades de Capacitação da Pró-Reitoria de Gestão de Pessoas (UFRN)*. 2017. 163 f. Universidade Federal do Rio Grande do Norte, 2017.

## APÊNDICE A - Roteiro da Entrevista e Questionário

### ROTEIRO DE ENTREVISTA PARA IDENTIFICAÇÃO DOS REQUISITOS, PERFIL E EXPERIÊNCIA DOS PARTICIPANTES

Participante N°: \_\_\_\_\_

Data: \_\_\_\_/\_\_\_\_/\_\_\_\_ Hora Início: \_\_\_\_:\_\_\_\_hs Hora Término: \_\_\_\_:\_\_\_\_hs

Local: \_\_\_\_\_

#### Informações de Perfil, Educacionais e Profissionais

1. Qual a sua idade? \_\_\_\_\_ anos
2. Sexo: \_\_\_\_\_
3. Qual o seu grau de instrução?
  - a. 2º grau completo
  - b. 3º grau completo
  - c. Pós-graduação – Mestrado
  - d. Pós-graduação – Doutorado
  - e. Outros.  
Especifique qual a formação e em que área \_\_\_\_\_
4. Em qual local você trabalha ou já trabalhou com coleta de dados?  
\_\_\_\_\_
5. Qual a sua profissão? \_\_\_\_\_
6. Há quanto tempo se encontra nesta profissão? \_\_\_\_\_ anos

#### Experiência Computacional

1. Há quanto tempo você utiliza o computador?
  - a. Menos de 1 ano
  - b. Entre 1 e 3 anos
  - c. Entre 3 e 5 anos
  - d. Mais de 5 anos
2. Em que local você utiliza o computador? (Pode-se marcar mais de uma alternativa)
  - a. No trabalho
  - b. Em casa
  - c. Na universidade
  - d. Nos três lugares e mais outros

3. Aproximadamente, quantas horas por semana você utiliza o computador?
  - a. Até 5 horas
  - b. Entre 5 e 10 horas
  - c. Mais de 10 horas

### **Experiência e Satisfação com softwares de coleta de dados em redes sociais *online***

1. Há quanto tempo você utiliza ou utilizou ferramentas de coleta de dados em suas pesquisas?
  - a. Menos de 1 ano
  - b. Entre 1 e 2 anos
  - c. Mais de 2 anos
2. Com que frequência você utiliza ou utilizava este tipo de ferramenta?
  - a. Uma vez por semana
  - b. Entre 2 e 4 vezes por semana
  - c. Mais de 4 vezes por semana
3. Em que local você acessa ou acessava as ferramentas de coleta de dados? (Pode-se marcar mais de uma alternativa)
  - a. No trabalho
  - b. Em casa
  - c. Na Universidade
  - e. Nos três lugares e mais outros. Quais? \_\_\_\_\_
4. Você utiliza ou utilizava ferramentas de coleta de dados para qual finalidade?
  - a. Uso pessoal/curiosidade
  - b. Pesquisa Científica
  - c. Trabalho
  - d. Outra. Qual? \_\_\_\_\_
5. Qual (ais) ferramentas/softwarees você utiliza ou utilizava em suas pesquisas para coleta de informações/dados?
6. Quais dificuldades você enfrenta ou já enfrentou ao utilizar estas ferramentas de coleta de coleta?
7. Você está satisfeito (a) com a (s) interface (s) da (s) ferramenta (s) de coleta de dados que utiliza? Justifique a sua resposta.
8. Quais as principais funcionalidades as ferramentas que utiliza ou utilizava possui? (Quais funções delas você acha essencial?)
9. Sobre as ferramentas de coleta que utiliza ou utilizava, que partes da interface você mais gosta e/ou acha essencial?
10. Como você trata os dados que coleta ou coletou com estas ferramentas? Você utiliza ou utilizava alguma ferramenta para tratar os dados? Se sim, qual?
11. Diante da sua experiência, quais melhorias, em relação a interface e sua interação com ela, você realizaria na (s) ferramenta (s) de coleta que você conhece?

## TERMO DE CONSENTIMENTO E LIVRE ESCLARECIMENTO I



MINISTÉRIO DA EDUCAÇÃO  
Universidade Federal dos Vales do Jequitinhonha e Mucuri  
Comitê de Ética em Pesquisa



### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Você está sendo convidado (a) a participar de uma pesquisa intitulada: **“APRIMORANDO O PROCESSO DE PESQUISA EM EDUCAÇÃO: Uma proposta de *Design* de Interação para uma ferramenta de coleta de dados em redes sociais *online*”**, em virtude de você ser um potencial usuário do ambiente referido acima, uma vez que atua como pesquisador na em uma instituição de ensino superior, já trabalha com este tipo de software e/ou realiza pesquisas na área de mineração de textos. Tal projeto é coordenado pela aluna do Programa de Pós-graduação em Educação da UFVJM, **Jésyka Milleny Azevedo Gonçalves** e contará ainda com acompanhamento e avaliação dos seguintes pesquisadores orientadores: Prof<sup>ª</sup>. Dra. Maria Lucia Bento Villela e Prof. Dr. Marcus Vinicius Carvalho Guelpeli e, além do apoio de execução pelo grupo de pesquisa Mineração de Textos, Processamento em Linguagem Natural e Aprendizado de Máquina (MTPLNAM) pertencente ao Departamento de Computação (DECOM) da UFVJM.

A sua participação não é obrigatória sendo que, a qualquer momento da pesquisa, você poderá desistir e retirar seu consentimento. Sua recusa não trará nenhum prejuízo para sua relação com o pesquisador e/ou com a UFVJM.

O objetivo desta pesquisa é o desenvolvimento do *Design* da Interação para uma ferramenta de coleta de dados em redes sociais *online*, denominado Oráculo. Este mescla as principais funcionalidades existentes nas ferramentas existentes hoje, que possuem a mesma finalidade, e as disponibiliza em uma única, com uma interface com alta qualidade de uso.

Caso você decida aceitar o convite, será submetido (a) a duas etapas distintas, conforme o desenvolvimento desta pesquisa: Na primeira etapa você será submetido a uma entrevista, com o intuito de obter informações para o desenvolvimento do *design* mencionado; na segunda etapa, você utilizará a ferramenta desenvolvida e será submetido a uma nova entrevista, que terá o objetivo de avaliar a interface proposta. Em ambas a etapas não serão coletados dados sobre você, usuário, que possa identifica-lo, somente sobre os aspectos que deverão ser considerados no ambiente do *Framework* Oráculo, para que ele possa cumprir efetivamente o seu propósito.

Neste primeiro encontro serão requeridas informações referentes as suas sensações e opiniões sobre a forma que realiza suas pesquisas e sobre as ferramentas que você utiliza para realização destas. A entrevista foi agendada previamente, conforme sua disponibilidade, e terá duração aproximada de quarenta minutos. Durante a entrevista somente o áudio será gravado, o qual será posteriormente transcrito para análise nesta pesquisa.

O risco relacionado com sua participação é mínimo, e consiste em algum constrangimento ou desconforto que você poderá sentir, caso não saiba ou não queira responder alguma pergunta. Tal risco será minimizado pelo fato de tanto a entrevista quanto o teste de aplicação serem conduzidos individualmente, sem a presença de outras pessoas que não seja o pesquisador e um auxiliar. Além disso, caso mesmo assim tal constrangimento ou desconforto aconteça, a coleta de dados durante a entrevista será interrompida, para que você se recupere e possa decidir se deseja continuar ou não.

Os benefícios com a sua participação estão relacionados ao uso que você irá fazer do ambiente do *Framework* Oráculo quando ele for concluído, tirando proveito da sua proposta de realizar coletas em redes sociais *online* com uma interface com alta qualidade de uso.

Os resultados desta pesquisa poderão ser apresentados em seminários, congressos e similares, entretanto, os dados/informações obtidos por meio da sua participação serão confidenciais e sigilosos, não possibilitando sua identificação.

Você não terá nenhum gasto e/ou prejuízo com a sua participação no estudo, uma vez que os encontros ocorrerão em horários e locais conforme sua disponibilidade com agendamento prévio. Também não há remuneração com sua participação, bem como a de todas as partes envolvidas. Não está previsto indenização por sua participação, mas em qualquer momento se você sofrer algum dano, comprovadamente decorrente desta pesquisa, terá direito à indenização.

Você receberá uma cópia deste termo onde constam o telefone e o endereço do pesquisador principal, podendo tirar suas dúvidas sobre o projeto e sobre sua participação agora ou em qualquer momento.

Você pode fazer todas perguntas que julgar necessárias.

---

**Coordenadora do Projeto: Jésyka Milleny Azevedo Gonçalves**  
**End.: Monsenhor Celso de Carvalho, 78 – Vila Arraiolos - Diamantina – MG**  
**E-mail: [jesykamilleny@gmail.com](mailto:jesykamilleny@gmail.com) - Celular: (38) 988137472**

Diante do exposto acima, declaro que entendi os objetivos, a forma de minha participação, riscos e benefícios da mesma e aceito o convite para participar.

Autorizo a publicação dos resultados da pesquisa, a qual garante o anonimato e o sigilo referente à minha participação.

---

**Nome do participante da pesquisa**

---

**Assinatura do participante da pesquisa**

Informações – Comitê de Ética em Pesquisa da UFVJM  
Rodovia MGT 367 - Km 583 - nº 5000 - Alto da Jacuba - Diamantina/MG CEP39100-000 - Tel.: (38)3532-1240  
Coordenadora: Prof.ª Simone Gomes Dias de Oliveira  
Secretária: Cristina de Figueiredo Vieira - Email: [cep.secretaria@ufvjm.edu.br](mailto:cep.secretaria@ufvjm.edu.br)

## APÊNDICE B - Questionários Pré Teste e TCLE

### TERMO DE CONSENTIMENTO DE LIVRE ESCLARECIMENTO II



MINISTÉRIO DA EDUCAÇÃO  
Universidade Federal dos Vales do Jequitinhonha e Mucuri  
Comitê de Ética em Pesquisa



#### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Eu, Jéyika Milleny Azevedo Gonçalves, orientada pela Profa. Dra. Maria Lúcia Bento Villela, desenvolvi como parte da minha pesquisa de mestrado o *Design* da Interação para o *Framework* Oráculo, uma ferramenta de coleta e mineração de textos em redes sociais *online*, desenvolvida pelo Grupo de Pesquisa Mineração de Textos, Processamento de Linguagem Natural e Aprendizagem de Máquina (MTPLNAM).

Diante disso, você está sendo convidado (a) por mim, e pelo grupo supracitado, para participar do presente estudo, que tem como objetivo avaliar o *design* proposto.

Neste contexto, gostaria de solicitar que você manifeste o seu consentimento para participar deste estudo, realizando as seguintes atividades:

- Responder um questionário sobre o seu perfil pessoal e profissional, além da sua experiência com ferramentas de coleta e mineração de textos em redes sociais *online*;
- Ouvir uma explicação sobre o Oráculo e o *design* proposto;
- Realizar tarefas previamente elaboradas utilizando a interface proposta;
- Preencher o questionário de verificação de aceitação do usuário em relação a ferramenta proposta (TAM);
- Preencher o formulário para verificação de sua satisfação ao utilizar a interface da ferramenta proposta (SUMI);
- Participar de uma breve entrevista pós-teste, oral, que será realizada em grupo com os demais participantes presentes.

O estudo completo terá duração aproximada de duas horas, incluindo todas as atividades descritas acima. É importante você saber que:

1. Os dados coletados durante o estudo serão utilizados **estritamente** no contexto acadêmico e de pesquisa.

2. A equipe envolvida neste estudo tem o compromisso de publicar os resultados de suas Pesquisas, Entretanto, a publicação é baseada em nosso respeito à **privacidade e anonimato** dos participantes. Assim, a sua identidade e a sua participação nesta pesquisa serão mantidas em sigilo e os dados divulgados pela pesquisa não conterão nomes ou quaisquer outras informações que permitam identificá-lo(a).
3. O consentimento para participar deste estudo é uma **escolha livre** de sua parte, realizada a partir do esclarecimento de todas as suas dúvidas e questões sobre a pesquisa.
1. Você **pode interromper a sua participação** neste estudo a qualquer momento, sem sofrer nenhuma penalidade. Neste caso, todos os seus dados e resultados parciais serão descartados.

Eu, Jesyka, responsável pela condução do presente estudo, estou **disponível** para contato pelo telefone (38) 988137472 e pelo e-mail [jesykamilleny@gmail.com](mailto:jesykamilleny@gmail.com).

De posse dessas informações, gostaríamos que você se pronunciasse acerca de sua participação no referido trabalho:

- ( ) Declaro que entendi os objetivos, riscos e benefícios de minha participação no trabalho e desejo participar voluntariamente, podendo desistir de colaborar a qualquer momento.
- ( ) Não desejo participar do trabalho.

---

**Coordenadora do Projeto: Jésyka Milleny Azevedo Gonçalves**  
**End.: Monsenhor Celso de Carvalho, 78 – Vila Arraiolos - Diamantina – MG**  
**E-mail: [jesykamilleny@gmail.com](mailto:jesykamilleny@gmail.com) - Celular: (38) 988137472**

---

**Nome do participante da pesquisa**

---

**Assinatura do participante da pesquisa**

Informações – Comitê de Ética em Pesquisa da UFVJM  
Rodovia MGT 367 - Km 583 - nº 5000 - Alto da Jacuba - Diamantina/MG CEP39100-000 - Tel.: (38)3532-1240  
Coordenadora: Prof.ª Simone Gomes Dias de Oliveira  
Secretária: Cristina de Figueiredo Vieira - Email: [cep.secretaria@ufvjm.edu.br](mailto:cep.secretaria@ufvjm.edu.br)

## QUESTIONÁRIO PRÉ TESTE

### Identificação do Perfil dos Participantes da Avaliação

1. Qual o seu ano de nascimento?
  - Antes de 1960
  - 1960 a 1975
  - 1975 a 1990
  - Depois de 1990
  
2. Sexo:
  - Feminino
  - Masculino
  - Não declarar
  
3. Até o momento, qual o seu grau de formação concluído?
  - 2º Grau completo
  - 3º Grau completo
  - Pós-graduação - Mestrado
  - Pós-graduação – Doutorado
  - Outros
  - 4. Especifique a sua área de formação profissional  
\_\_\_\_\_
  


---

5. Você realiza, possui interesse ou planeja realizar pesquisas utilizando as informações providas de redes sociais?
  - Sim
  - Não
  
6. Se você respondeu sim e já atua na área, nos informe, por favor, há quanto tempo você coleta dados em redes sociais para obter informações para suas pesquisas?
  - Menos de 1 ano
  - Entre 1 e 2 anos
  - Entre 2 e 3 anos
  - Mais de 3 anos
  - 7. Se você respondeu sim e já atua na área, nos informe resumidamente qual o foco da sua pesquisa.  
\_\_\_\_\_
  


---

8. Se você respondeu sim e já atua na área, nos informe como você obtém dados das redes sociais atualmente.
  - Através de ferramentas pagas disponíveis no mercado
  - Através do desenvolvimento de scripts
  - Através de ferramentas disponíveis para fins de pesquisa
  - Outro: \_\_\_\_\_
  - \_\_\_\_\_



## APÊNDICE C - Tarefas, Questionários e Entrevista Pós Teste

### QUESTIONÁRIO PÓS TESTE SUMI

Prezad@ Participante, Boa tarde!

As informações fornecidas serão mantidas em confidencialidade e nenhuma informação que possa identificá-lo será armazenada. Este questionário possui 19 (dezenove) questões. Por favor, responda cada uma delas. Para cada afirmação existem três opções de resposta (CONCORDA, INDECISO, DISCORDA). Ao marcar as opções CONCORDA ou DISCORDA, você não está necessariamente indicando forte concordância ou desacordo, mas apenas o seu sentimento geral na maioria das vezes. Quanto a opção INDECISO, a considere como um sentimento de meio-termo ou imparcialidade.

Desde já agradeço pela sua colaboração!

	Concordo	Indeciso	Discordo
Eu recomendaria o Oráculo com esta interface aos meus colegas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
As instruções e comandos são úteis.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Às vezes não sei o que fazer em seguida com esta ferramenta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Trabalhar com esta ferramenta é satisfatório.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A maneira como as informações do sistema são apresentadas é clara e compreensível.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Os comandos de ajuda são muito informativos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Eu não gostaria de usar este software todos os dias.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
As tarefas podem ser executadas de maneira direta usando este programa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O uso deste software é frustrante.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Esta ferramenta parece que irá me ajudar a superar os problemas que tive utilizando-o.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
É óbvio que as necessidades do usuário foram totalmente levadas em consideração.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Houve momentos em que usei o Oráculo e me senti bastante tenso.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A organização dos menus ou listas de informações parece bastante lógica.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
As mensagens de prevenção de erros não são adequadas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
É fácil fazer com que a ferramenta faça exatamente o que eu desejo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O Oráculo nem sempre fez o que eu esperava.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O Oráculo tem uma apresentação muito atraente.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
É fácil esquecer como fazer as coisas com esta ferramenta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Preciso procurar ajuda na maioria das vezes quando uso a ferramenta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## INSTRUÇÕES E TAREFAS DE UTILIZAÇÃO DO *DESIGN*

### Instruções de Uso

Bem-vindo ao estudo da interface da Ferramenta Oráculo!

O objetivo deste trabalho é realizar testes com o *design* proposto para o *Framework* Oráculo, identificando principalmente a satisfação dos usuários, a comunicabilidade do sistema e a aceitação de você, participante da pesquisa e futuro usuário do Oráculo, em relação a esta nova ferramenta de coleta e mineração em redes sociais.

Durante o teste, você deverá **utilizar o Oráculo para realizar uma série de tarefas específicas**, que serão detalhadas a seguir, a fim de nos ajudar a obter um melhor entendimento acerca da ferramenta. Você terá disponível lápis, caneta e papel para que possa responder aos questionários e fazer anotações. Uma impressora a jato de tinta está conectada ao computador.

Peço que **reserve suas observações**, se possível as anote, para que na etapa de conversa em grupo, ao fim desta avaliação, você as exponha. Gostaria de ressaltar que **suas opiniões e sentimentos, tanto em relação ao sistema quanto em relação ao teste em si, são mais do que bem-vindas**. O que nós queremos avaliar aqui é a interface do sistema e não o seu desempenho como usuário. Portanto, por favor, não tenha receio de nos dizer o que você realmente pensa.

Registrarei em vídeo sua interação com o sistema através de uma ferramenta de gravação de tela. Estarei presente no laboratório durante os testes para realizar anotações, mas para não influenciar a sua interação com a ferramenta, não poderei responder perguntas relacionadas à execução das tarefas após o início do teste.

Por fim, gostaria de observar que não estamos colocando limite para a duração máxima do teste, mas se você quiser parar antes, sinta-se à vontade. Não há problema algum.

### Orientações Gerais

Lembre-se que você interagirá com um protótipo interativo de alta fidelidade, portanto este possui algumas limitações. Sendo assim:

- Os campos não são editáveis com textos e alguns tópicos não poderão ser selecionados, assim alguns campos com informações que deveriam ser inseridas pelo usuário já estão disponíveis na tela e neste documento estão destacados entre *\*xxx\**;

- Os botões de “Limpar dados”, em sua maioria das vezes, estão desabilitados para esta avaliação, justamente porque alguns campos não poderão ser inseridos pelo usuário;
- A qualquer momento você pode reiniciar as atividades no Oráculo, voltando para a sua “Tela de Inicial” ao clicar “R” em seu teclado.
- Nas telas que representam um processamento do sistema em andamento (no caso 48%) dê um clique para que o processamento seja concluído.

Mais uma vez, obrigado pela participação! Para nós, a sua colaboração é essencial!

Até o momento, você tem alguma dúvida, estou a disposição antes de iniciarmos as tarefas.

Então vamos começar a utilizá-lo?

Você deverá realizar as tarefas descritas nas páginas seguintes.

### TAREFA I

Conheça o Oráculo!

1. A sua primeira tarefa é passar cinco minutos explorando livremente a ferramenta.

### TAREFA II

A primeira tarefa é Realizar Nova Coleta utilizando 1(um) termo de busca

1. Insira o nome *\*Governo\_Bolsonaro\_2019\** para a sua coleta;
2. Insira o termo *\*Bolsonaro\**
3. Exclua o termo inserido e o insira novamente para a sua busca
4. Selecione o Tempo de Coleta como *Retroativo*
5. Selecione o Sistema de Coleta como *Crawler*
6. Selecione a Saída de Dados como *Arquivos JSON*
7. Insira informações sobre o período da coleta: *16/11/2019 das 16h às 16:30hs*
8. Limpe dos dados dos Passos 4 a 7
9. Agora, volte a tela anterior e repita o processo de 1 a 7.
10. Selecione a abrangência de sua coleta como *\*Coleta Local\**
11. Adicione as observações sobre sua coleta: *\*Coleta com tweets referentes ao governo Bolsonaro no ano de 2019\**
12. Adicione informações de Idioma que deseja coletar para *\*Inglês\**
13. Adicione informações de Localização que deseja coletar – confirmando se a coleta será realizada na região de *\*Belo Horizonte\**
14. Salve a coleta realizada inserindo o novo nome, sem sobrescrever uma coleta anteriormente realizada. Nome = *\*Governo\_Bolsonaro\_2019\_2\**
15. Volte para a Página de Opções do Oráculo

### TAREFA III

A segunda tarefa é Realizar Nova Coleta utilizando 2 (dois) termo de busca

1. Insira o nome *\*Governo\_Bolsonaro\_2019\** para a sua coleta
2. Insira os termos *\*Bolsonaro\** e *\*Governo\**, utilizando o operador (espaço)
3. Exclua o termo 2 inserido e o insira novamente para a sua busca

4. Selecione o Tempo de Coleta como *Retroativo*
5. Selecione o Sistema de Coleta como *Crawler*
6. Selecione a Saída de Dados como *Arquivos JSON*
7. Insira informações sobre o período da coleta: *16/11/2019 das 16h às 16:30hs*
8. Limpe dos dados dos Passos 4 a 7
9. Agora, volte a tela anterior e repita o processo de 1 a 7.
10. Selecione a abrangência de sua coleta como *\*Coleta Local\**
11. Adicione as observações sobre sua coleta: *\*Coleta com tweets referentes ao governo Bolsonaro no ano de 2019\**
12. Adicione informações de Idioma que deseja coletar para *\*Inglês\**
13. Adicione informações de Localização que deseja coletar – confirmando se a coleta será realizada na região de *\*Belo Horizonte\**
14. Salve a coleta sobrescrevendo uma coleta anteriormente realizada.
15. Volte para a Página de Opções do Oráculo

#### TAREFA IV

A quarta tarefa é Realizar o Filtro de Bots em uma de suas coletas

1. Selecione uma de suas coletas já realizadas
2. Realize a filtragem dos bots na coleta selecionada
3. Observe as informações da coleta filtrada
4. Selecione a opção *\*Coletas de Bots\** para salvar
5. Salve sua coleta e a exporte para manuseá-la posteriormente
6. Vá para a Página de Opções do Oráculo.

#### TAREFA V

Na quinta tarefa você deverá realizar a Mineração em sua coleta e Visualizar o dados minerados

1. Selecione uma de suas coletas já realizadas
2. Visualize as informações sobre a mineração a ser realizada
3. Confirme a coleta selecionada
4. Insira as seguintes informações de Parâmetro:

- a. Termo = *\*Bolsonaro\**
  - b. Mineração a ser realizada no *dia 15/11/2019, de 14hs às 18hs*
5. Selecione a opção de mineração e visualização que deseja realizar = *“Nuvem de Palavras”*
  6. Visualize a coleta minerada e a exporte.

#### TAREFA VI

Na sexta tarefa você deverá realizar a Mineração em sua coleta e Visualizar o dados minerados

1. Realize o mesmo processo da Tarefa 4, porém agora para selecionar a opção de mineração e visualização = *“Frequência de Palavras”*

#### TAREFA VII

Na quinta tarefa você deverá realizar a Mineração em sua coleta e Visualizar o dados minerados

1. Realize o mesmo processo da Tarefa 4, porém agora para selecionar a opção de mineração e visualização = *“Número Total de Termos”*

#### TAREFA VIII

Na quinta tarefa você deverá realizar a Mineração em sua coleta e Visualizar o dados minerados

1. Realize o mesmo processo da Tarefa 4, porém agora para selecionar a opção de mineração e visualização = *“Gráfico de Frequência”*

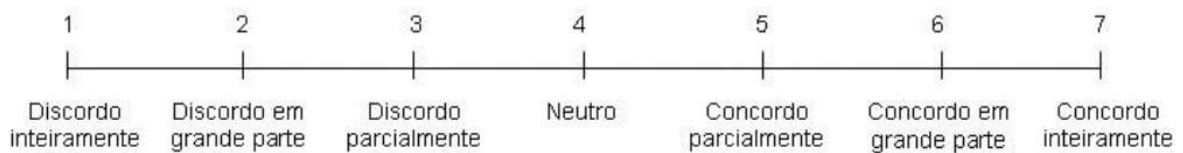
## QUESTIONÁRIO PÓS TESTE TAM

Este questionário objetiva utilizar o modelo TAM para indicar a sua aceitação em relação ao Oráculo, uma nova ferramenta disponibilizada para as suas atividades de coleta e mineração de textos em redes sociais *online*. Por favor, caso necessite realizar algum comentário sobre o questionário ou tenha alguma dúvida, fique a vontade para retirá-las comigo durante a avaliação.

Ao responder essas perguntas, use os números nas escalas para indicar até que ponto você concorda ou discorda de cada afirmativa, utilizando escalas similares à que está reproduzida abaixo "Escala TAM".

Estou a disposição! Desde já agradeço sua participação!

### Escala TAM



1 - Minha interação com o sistema é clara.

	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente

2 - Interagir com o sistema não requer muito de meu esforço.

	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente

3 - Eu acho o Oráculo fácil de usar.

	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente

4 - Eu aprendi facilmente a operar o sistema.

	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente

5 - Na minha pesquisa, o uso de ferramentas como o Oráculo é importante.

	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente

6 - Eu acho que seria muito bom utilizar o Oráculo em vez de usar os métodos de coleta e mineração atuais.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente
7 - Usar o Oráculo é agradável.								
	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente
8 - Se me for disponibilizado, eu tenho a intenção de utilizar o Oráculo em minhas pesquisas.								
	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente
9 - Sempre que possível, eu evitarei usar o Oráculo.								
	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente
10 - Eu pretendo explorar ao máximo as funcionalidades do sistema.								
	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente
11 - Eu provavelmente, ao manipula-lo sozinho, usarei o Oráculo de forma inapropriada.								
	1	2	3	4	5	6	7	
Discordo Inteiramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo Inteiramente



### **ROTEIRO PARA ENTREVISTA PÓS-TESTE**

A intenção desta etapa do teste é que vocês exponham as observações realizadas durante a execução das tarefas propostas e fiquem a vontade para dar sugestões de melhorias para a interface do Oráculo.

1. Sobre as tarefas propostas:
  - a. vocês conseguiram realizá-las com facilidade?
  - b. Alguma tarefa em específico foi mais difícil ou confusa de ser realizada?
2. Se vocês tivessem que utilizar o Oráculo para as pesquisas de vocês, vocês estariam satisfeitos?
3. Vocês acharam a aparência do Oráculo agradável?
4. Quais os pontos positivos e negativos vocês identificaram no Oráculo?
5. Quais as dificuldades identificadas durante a interação?
6. Ao ficar os 5 minutos conhecendo livremente o Oráculo e posteriormente terem executado as tarefas propostas, vocês conseguiram identificar funcionalidades não abordadas nas tarefas? Se sim, quais?
7. Quais são suas sugestões de melhoria para o Oráculo?