



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CAMPUS DE RUSSAS**  
**CURSO DE GRADUAÇÃO EM ENGENHARIA DE SOFTWARE**

**LAVÍNIA MATOSO FREITAS**

**UMA METODOLOGIA DE GERAÇÃO AUTOMÁTICA DE PERSONAS A PARTIR  
DE DADOS DOS USUÁRIOS EM SISTEMAS SOCIAIS**

**RUSSAS**

**2018**

LAVÍNIA MATOSO FREITAS

UMA METODOLOGIA DE GERAÇÃO AUTOMÁTICA DE PERSONAS A PARTIR DE  
DADOS DOS USUÁRIOS EM SISTEMAS SOCIAIS

Trabalho de Conclusão de Curso apresentado ao  
Curso de Graduação em Engenharia de Software  
do Campus de Russas da Universidade Federal  
do Ceará, como requisito parcial à obtenção do  
grau de bacharel em Engenharia de Software.

Orientadora: Profa. Dra. Marília Soares  
Mendes

RUSSAS

2018

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

F936m Freitas, Lavínia Matoso.

Uma metodologia de geração automática de personas a partir de dados dos usuários em sistemas sociais / Lavínia Matoso Freitas. – 2018.

123 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Russas, Curso de Engenharia de Software, Russas, 2018.

Orientação: Profa. Dra. Marília Soares Mendes.

1. Personas. 2. Geração Automática. 3. Sistemas Sociais. 4. Dados públicos. I. Título.

CDD 005.1

---

LAVÍNIA MATOSO FREITAS

UMA METODOLOGIA DE GERAÇÃO AUTOMÁTICA DE PERSONAS A PARTIR DE  
DADOS DOS USUÁRIOS EM SISTEMAS SOCIAIS

Trabalho de Conclusão de Curso apresentado ao  
Curso de Graduação em Engenharia de Software  
do Campus de Russas da Universidade Federal  
do Ceará, como requisito parcial à obtenção do  
grau de bacharel em Engenharia de Software.

Aprovada em:

BANCA EXAMINADORA

---

Profa. Dra. Marília Soares Mendes (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Profa. Dra. Anna Beatriz dos Santos Marques  
Universidade Federal do Ceará (UFC)

---

Profa. Dra. Maria Elizabeth Sucupira Furtado  
Universidade de Fortaleza (UNIFOR)

À minha avó, Terezinha Matoso.

À minha mãe, Kátia Matoso.

## AGRADECIMENTOS

A Deus, por ter me guiado durante todas as escolhas e dificuldades encontradas.

À minha mãe, Kátia Matoso, razão da minha existência e minha inspiração. Obrigada por ser a melhor pessoa que eu conheço e por sempre fazer o impossível por mim. Te amo!

À minha família, meus maiores incentivadores. Obrigada por tudo que vocês já me proporcionaram. Espero conseguir retribuir, pelo menos metade, de tudo que vocês já fizeram/fazem por mim. Amo vocês!

À minha orientadora, Marília Mendes, minha mãe de orientação. Obrigada por acreditar em mim, e por todo o tempo e esforço depositado desde o segundo semestre. Obrigada por, diretamente ou indiretamente, ter me incentivado a entrar e permanecer no curso. Você é a melhor orientadora que eu poderia ter!

Ao meu namorado, Thiago Oliveira, o melhor presente que a UFC me deu. Obrigada por dividir todos os momentos comigo, toda atenção e carinho, durante todo esse processo. Além de parceiro na vida, contribuí com meu desenvolvimento na pesquisa desde 2015.

Aos melhores amigos desses últimos quatro anos, Liana Mara, Gabriel Gonçalves e, em especial, Alex Felipe. Meus melhores momentos na UFC foram vividos com vocês. Obrigada por todos os conselhos, risadas, memes, sofrimento coletivo e desespero na realização de trabalhos no último dia. Juntos, formamos o melhor grupo que eu já tive!

Ao meu melhor amigo, quase irmão, Alex Felipe (FUFinho). Sem você, eu acho que nem teria chegado até o final. Obrigada por todos os dias vividos, por rir de todas as piadas bestas comigo e por compartilhar todos os momentos de desespero juntos. Você é o melhor!

À minha amiga de infância, Beatriz Ramalho. Obrigada por, mesmo com toda a distância, se fazer presente, me apoiar e incentivar a seguir meus sonhos.

Aos meus amigos, Beatriz Marques, Rodrigo, Bárbara Feijão e Isabelly Lima, por todos os momentos de descontração e felicidade. Ao meu array, Yan Vancelis e Marcos Alencar, por juntos organizarmos a III SESCOOP, me apresentando um universo além da pesquisa. Espero sempre ter todos vocês por perto.

Ao Laboratório INterdisciplinar de Computação e Engenharia de Software – LINCE, especialmente ao grupo do Projeto MALTU, o qual faço parte desde de 2015. Onde me foi dada a oportunidade de ingressar no mundo da pesquisa. À todos os meus colegas de pesquisa do projeto: Afonso, Douglas, Isaac, Isaias, Paloma e Thiago. Por terem contribuído de forma direta ou indireta na realização deste trabalho. E a professora Elizabeth Furtado, que contribuiu com

ideias e auxiliou na escrita de um artigo relacionado a esta pesquisa.

À FUNCAP, pelo apoio financeiro com a manutenção da bolsa de iniciação científica durante dois anos e no auxílio para publicação dos meus trabalhos científicos, juntamente à UFC.

Enfim, agradeço a todos que contribuíram na realização deste trabalho e na minha graduação, seja de forma direta ou indireta. Gratidão a todos!

*“Deixe suas janelas quebradas, abertas, e a luz adentra lentamente. E você terá uma cabeça, uma cabeça cheia de sonhos.”*

(Chris Martin, Guy Berryman,  
Johnny Buckland, Will Champion)



## RESUMO

Conhecer o usuário final de um sistema é essencial tanto para seu desenvolvimento como para sua avaliação. Existe uma técnica de modelagem de usuários chamada Personas, cujo objetivo é aproximar os desenvolvedores e os usuários do sistema, construindo modelos de pessoas, seus comportamentos e motivações de uso dos usuários finais. Alguns estudos vêm propondo técnicas para gerar personas, coletando dados a partir de entrevistas e questionários. No entanto, este tipo de geração pode exigir uma grande análise manual para aumentar a precisão, tornando o processo exaustivo e demorado. Para solucionar tais problemas, recursos computacionais podem ser utilizados para automatizar o processo de extração e geração das personas. Este estudo tem como objetivo apresentar uma proposta de geração automática de Personas para auxiliar no processo de avaliação de sistemas, utilizando uma metodologia de avaliação textual chamada MALTU. Este trabalho foi executado em três etapas: i) mapeamento sistemático sobre geração automática de personas, a fim de entender as técnicas; ii) definição de uma proposta; e iii) experimento utilizando dados do sistema social Twitter. Com a realização do mapeamento sistemático, não foi identificada nenhuma proposta de geração automática de personas que tenha sido aplicada no contexto de avaliação de sistemas. Com isso, este estudo apresenta uma proposta, que utiliza dados públicos de usuários, juntamente com os resultados obtidos com a execução da avaliação textual. Para complementar os resultados do mapeamento, foi realizado um estudo da API do twitter a fim de entender quais dados podem ser extraídos e como podem ser analisados. Por fim, foi realizada uma validação da proposta, a partir de um experimento com postagens do Twitter. Este experimento mostrou a validade da proposta.

**Palavras-chave:** Personas. Geração Automática. Sistemas Sociais. Dados públicos.

## ABSTRACT

Knowing the end user of a system is essential both for its development and for its evaluation. There is a user modeling technique called Personas, whose goal is to bring developers and users together, building people models, behaviors, and motivations for end-users. Some studies have proposed techniques to generate personas, collecting data from interviews and questionnaires. However, this type of generation may require a large manual analysis to increase accuracy, making the process exhaustive and time consuming. To solve such problems, computational resources can be used to automate the process of extraction and generation of personas. This study aims to present a proposal of automatic generation of Personas to assist in the process of system evaluation, using a textual evaluation methodology called MALTU. This work was executed in three steps: i) systematic mapping on automatic generation of personas, in order to understand the techniques; ii) definition of a proposal; and iii) experiment using data from the social system Twitter. With the accomplishment of the systematic mapping, no proposal of automatic generation of personas that has been applied in the context of systems evaluation was identified. With this, this study presents a proposal, which uses public data of users, together with the results obtained with the execution of textual evaluation. To complement the results of the mapping, a study of the twitter API was performed in order to understand what data can be extracted and how they can be analyzed. Finally, a proposal validation was performed, based on an experiment with Twitter posts. This experiment showed the validity of the proposal.

**Keywords:** Personas. Automatic Generation. Social Systems. Public data.

## LISTA DE FIGURAS

Figura 1 – Exemplo de uma persona . . . . .	21
Figura 2 – Fluxo de atividades para a geração de personas de Jung <i>et al.</i> (2018b). . . . .	25
Figura 3 – Tela de cadastro de contexto de uso e definição dos usuários. . . . .	29
Figura 4 – Artigos por base de dados. . . . .	45
Figura 5 – Ano de publicação dos artigos aceitos. . . . .	46
Figura 6 – WordCloud dos autores dos 25 artigos. . . . .	46
Figura 7 – Técnicas X Quantidade de estudos relacionados . . . . .	47
Figura 8 – Etapas automáticas X Quantidade de estudos relacionados . . . . .	48
Figura 9 – Local de Extração dos dados. . . . .	49
Figura 10 – Dados extraídos. . . . .	49
Figura 11 – Quantidade de estudos que apresentam ferramenta. . . . .	52
Figura 12 – Ferramentas utilizadas para geração das personas. . . . .	52
Figura 13 – Contexto de aplicação das Personas X Quantidade de estudos relacionados. . . . .	54
Figura 14 – Número de artigos que realizaram validação das personas. . . . .	55
Figura 15 – Técnicas de validação utilizadas X Quantidade de estudos relacionados. . . . .	55
Figura 16 – Aspectos avaliados na validação das personas. . . . .	56
Figura 17 – Exemplo de persona gerada com a MASC. . . . .	58
Figura 18 – Exemplo de Persona Indivíduo gerada com o FMK@PE. . . . .	58
Figura 19 – Exemplo de Persona Empresa gerada com o FMK@PE. . . . .	59
Figura 20 – Exemplo de persona gerada por Guimarães. . . . .	60
Figura 21 – Exemplo de persona gerada com a técnica APG. . . . .	62
Figura 22 – Fluxo de etapas da metodologia. . . . .	65
Figura 23 – Demonstração da Ferramenta Face++. . . . .	71
Figura 24 – Persona do Grupo 1. . . . .	73
Figura 25 – Persona do Grupo 2. . . . .	74

## LISTA DE TABELAS

Tabela 1 – Comparativo dos resultados por base. . . . .	34
Tabela 2 – Quantidade de artigos por etapa. . . . .	37
Tabela 3 – Detalhamento dos artigos encontrados no Fluxo Primário. . . . .	38
Tabela 4 – Número de citações por artigo. . . . .	41
Tabela 5 – Quantidade de artigos por etapa (Fluxo Secundário). . . . .	41
Tabela 6 – Detalhamento dos artigos encontrados no Fluxo Secundário. . . . .	42
Tabela 7 – Avaliação da qualidade dos artigos. . . . .	45
Tabela 8 – Matriz de relação dos dados utilizados para geração das personas. . . . .	50

## LISTA DE QUADROS

Quadro 1 – Questões de Pesquisa do Mapeamento Sistemático. . . . .	32
Quadro 2 – Bases de dados utilizadas no Mapeamento Sistemático. . . . .	33
Quadro 3 – Checklist de Avaliação da Qualidade. . . . .	36
Quadro 4 – Formulário de Extração . . . . .	36

## LISTA DE ABREVIATURAS E SIGLAS

API	<i>Application Programming Interface</i>
IHC	Interação Humano-Computador
MALTU	Metodologia para Avaliação da interação em sistemas sociais a partir da Linguagem Textual do Usuário
PRUs	Postagens Relacionadas ao Uso
SS	Sistema Social

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	17
<b>1.1</b>	<b>Motivação</b>	18
<b>1.2</b>	<b>Objetivos</b>	19
<i>1.2.1</i>	<i>Objetivo Geral</i>	19
<i>1.2.2</i>	<i>Objetivos Específicos</i>	19
<b>1.3</b>	<b>Procedimentos Metodológicos</b>	19
<b>1.4</b>	<b>Organização do Trabalho</b>	20
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	21
<b>2.1</b>	<b>Personas</b>	21
<i>2.1.1</i>	<i>Criação de Personas</i>	22
<i>2.1.2</i>	<i>Geração Automática de Personas</i>	23
<b>2.2</b>	<b>Avaliação Textual</b>	25
<i>2.2.1</i>	<i>Usuários da Avaliação Textual</i>	27
<i>2.2.2</i>	<i>Ferramenta UUX-Posts</i>	28
<b>3</b>	<b>MAPEAMENTO SISTEMÁTICO</b>	30
<b>3.1</b>	<b>Identificação da necessidade de uma revisão</b>	30
<b>3.2</b>	<b>Planejamento do Mapeamento Sistemático: Definição do Protocolo</b>	30
<i>3.2.1</i>	<i>Formulação das Questões de Pesquisa</i>	31
<i>3.2.1.1</i>	<i>Definição do escopo aplicando a técnica PICOS</i>	31
<i>3.2.1.2</i>	<i>Questões de Pesquisa</i>	32
<i>3.2.2</i>	<i>Estratégias para busca dos estudos</i>	32
<i>3.2.2.1</i>	<i>Expressão geral de busca</i>	33
<i>3.2.2.1.1</i>	<i>Elaboração da String de Busca - Versão 1</i>	33
<i>3.2.2.1.2</i>	<i>Elaboração da String de Busca - Versão 2</i>	33
<i>3.2.2.1.3</i>	<i>Elaboração da String de Busca - Versão Final</i>	34
<i>3.2.2.1.4</i>	<i>Comparação dos resultados por base das versões da string</i>	34
<i>3.2.3</i>	<i>Critérios e Procedimentos para Seleção dos Estudos</i>	34
<i>3.2.3.1</i>	<i>Critérios de Exclusão</i>	34
<i>3.2.3.2</i>	<i>Critérios de Inclusão</i>	35
<i>3.2.4</i>	<i>Definição do procedimento de seleção</i>	35

3.2.5	<i>Seleção das publicações a partir de critérios de qualidade</i>	35
3.2.6	<i>Seleção dos Dados Relevantes</i>	36
3.2.7	<i>Síntese dos dados</i>	36
3.3	<b>Execução do Mapeamento Sistemático</b>	37
3.3.1	<i>Fluxo principal - Mapeamento Sistemático</i>	37
3.3.2	<i>Fluxo secundário - Forward Snowballing</i>	40
3.3.3	<i>Avaliação da qualidade dos estudos</i>	44
3.4	<b>Documentação</b>	44
3.4.1	<i>Apresentação dos estudos</i>	44
3.4.2	<i>Resultado das questões</i>	47
3.4.2.1	<i>Quais são as técnicas e métodos utilizados para a geração automática de personas?</i>	47
3.4.2.2	<i>Quais os dados extraídos para a geração das personas?</i>	48
3.4.2.3	<i>Existem ferramentas que geram personas de forma automática?</i>	52
3.4.2.4	<i>Quais os contextos de aplicação das personas geradas de forma automática?</i>	54
3.4.2.5	<i>Existem técnicas para validar as personas geradas?</i>	54
3.5	<b>Conclusão</b>	56
4	<b>TRABALHOS RELACIONADOS</b>	57
5	<b>UMA PROPOSTA DE GERAÇÃO AUTOMÁTICA DE PERSONAS A PARTIR DOS DADOS PÚBLICOS DOS USUÁRIOS EM SISTEMAS SOCIAIS</b>	63
5.1	<b>Estudo da API do Twitter</b>	63
5.2	<b>Metodologia proposta</b>	64
5.3	<b>Discussão</b>	68
5.4	<b>Conclusão</b>	69
6	<b>VALIDAÇÃO DA METODOLOGIA A PARTIR DE UM EXPERIMENTO NO TWITTER</b>	70
6.1	<b>Extração dos dados</b>	70
6.2	<b>Análise dos dados</b>	70
6.2.1	<i>Análise dos dados demográficos</i>	72
6.3	<b>Geração das personas</b>	72
6.4	<b>Apresentação das personas</b>	73



<b>6.5</b>	<b>Discussão</b> . . . . .	<b>74</b>
<b>6.6</b>	<b>Conclusão</b> . . . . .	<b>75</b>
<b>7</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS</b> . . . . .	<b>76</b>
<b>7.1</b>	<b>Contribuições do trabalho</b> . . . . .	<b>76</b>
<b>7.2</b>	<b>Trabalhos Futuros</b> . . . . .	<b>76</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>78</b>
	<b>APÊNDICES</b> . . . . .	<b>83</b>
	<b>APÊNDICE A</b> – Formulário de extração - Informações sobre a técnica . .	<b>83</b>
	<b>APÊNDICE B</b> – Formulário de extração - Informações sobre os dados . .	<b>89</b>
	<b>APÊNDICE C</b> – Formulário de extração - Informações sobre a ferramenta	<b>95</b>
	<b>APÊNDICE D</b> – Formulário de extração - Informações sobre o contexto .	<b>101</b>
	<b>APÊNDICE E</b> – Formulário de extração - Informações sobre a validação .	<b>105</b>
	<b>APÊNDICE F</b> – Classificação dos estudos de acordo com os Critérios de Qualidade . . . . .	<b>111</b>
	<b>APÊNDICE G</b> – Dados utilizados para a geração de personas . . . . .	<b>114</b>
	<b>ANEXOS</b> . . . . .	<b>117</b>
	<b>ANEXO A</b> – Relatório de seleção dos estudos . . . . .	<b>117</b>
	<b>ANEXO B</b> – Resultado de extração de uma postagem utilizando a API do Twitter . . . . .	<b>122</b>

## 1 INTRODUÇÃO

Utilizar conceitos de usabilidade em um sistema tem mostrado resultados no crescimento da produtividade do usuário e um decréscimo em seus erros (MARCUS, 2005). A usabilidade, em sua definição, assegura que os produtos sejam fáceis de usar, eficientes e agradáveis a seus usuários (ROGERS; SHARP; PREECE, 2013). Para isso é importante conhecer o usuário final de um sistema, tanto para seu desenvolvimento como para sua avaliação.

No contexto de avaliação de sistemas é necessário definir os participantes do teste, que são os usuários do sistema, ou seja, deve ser definido o perfil desses usuários. Para isso, os participantes devem corresponder ao perfil do público-alvo do sistema a ser avaliado, podendo existir diversos grupos com perfis diferentes (PRATES; BARBOSA, 2003). Uma forma de identificar essa variedade é representando por tipos específicos de indivíduos com necessidades específicas (COOPER; REIMANN; CRONIN, 2007). Para isso, pode ser utilizada a técnica de criação de personas.

Personas é uma técnica de design de interação que visa auxiliar no processo de conhecimento dos usuários, aproximando o desenvolvedor do sistema ao seu usuário final (NIELSEN, 2012). Uma persona é uma representação do usuário final baseada em dados, que apresenta um nome, uma breve descrição, suas características, comportamento, objetivos, preferências e motivações ao usar um aplicativo (NIELSEN; HANSEN, 2014).

Sua criação vem sendo aplicada em diversas áreas da criação do software, como para o design e desenvolvimento, e para atingir diversos objetivos, como auxiliar criadores de conteúdo digital a entender seus usuários e distribuir melhor seus produtos em plataformas online (AN *et al.*, 2016). Além disso, a utilização dessas personas vem trazendo diversos benefícios para empresas, ajudando a priorizar os usuários e manter o foco neles, auxiliando a defender suas necessidades diante a equipe de desenvolvimento (NIELSEN; HANSEN, 2014).

Um exemplo de avaliação de sistemas utilizando a definição do usuário com base em personas é apresentada no estudo de Ang (2016). É realizada uma avaliação de um sistema que visa auxiliar no aconselhamento de estudantes de graduação e pós-graduação do ensino superior. Para conhecer melhor os usuários, foram elaboradas 3 (três) personas com base nos tipos de usuários do sistema, e com base nas personas geradas foram selecionados os 48 (quarenta e oito) participantes para o teste.

Uma forma de avaliação que vem mostrando bons resultados é a avaliação textual que tem como base a opinião textual dos usuários, utilizando suas postagens em Sistemas Sociais

(Sistema Social (SS)). A Metodologia para Avaliação da interação em sistemas sociais a partir da Linguagem Textual do Usuário (Metodologia para Avaliação da interação em sistemas sociais a partir da Linguagem Textual do Usuário (MALTU)) (MENDES, 2015) utiliza esse tipo de postagens denominadas de Postagens Relacionadas ao Uso (Postagens Relacionadas ao Uso (PRUs)) e vem apresentando bons resultados, como a identificação de problemas de suporte (FREITAS; SILVA; MENDES, 2016) que em outros tipos de avaliação seriam mais difíceis de ser identificados. Em sua primeira etapa, a MALTU permite que haja uma definição do contexto de avaliação, sendo necessário definir quem são os usuários e o contexto de uso do sistema. Essa definição tem o propósito de relacionar os resultados de uma avaliação com as preferências, restrições e intenção de uso dos usuários e seu contexto. Atualmente, essa etapa é realizada utilizando outros estudos ou contendo uma definição de forma superficial (FREITAS; SILVA; MENDES, 2016; SILVA; FREITAS; MENDES, 2017).

Deste modo, por que não trabalhar precisamente com os usuários do sistema, ou seja, aqueles que escreveram tais postagens? Mendes (2015), em sua metodologia, propõe o uso de personas para a definição dos usuários buscando saber o tempo de uso, faixa etária, entre outras informações. A partir dessas personas, podem ser metodologias melhorias ao sistema, com base no modelo mental das personas, além de possibilitar o entendimento dos usuários para justificar seus elogios e/ou críticas quanto ao uso do sistema avaliado. Por exemplo, em um estudo anterior (FREITAS; SILVA; MENDES, 2016) foram encontradas críticas quanto a funcionalidade de alteração da foto do perfil do sistema Spotify, porém se 90% dos usuários fossem personas caracterizadas como usuários do sistema há 3 (três) anos, pode-se chegar a uma conclusão que se trata de um problema sério de aprendizado.

Este estudo tem como objetivo auxiliar o processo de conhecimento do usuário para realizar uma avaliação textual. Para isso, será apresentada uma metodologia de geração de personas, podendo ser realizada de forma automática, com base nos dados reais dos usuários extraídos dos sistemas que serão avaliados. Como contribuições, espera-se obter resultados mais precisos na avaliação, levando em consideração os usuários do sistema.

## **1.1 Motivação**

Este trabalho faz parte de um projeto de pesquisa chamado: avaliação da interação em sistemas sociais a partir da linguagem textual do usuário, coordenado pela Profa. Dra. Marília Soares Mendes, iniciado em 2015 e financiado pela Fundação Cearense de Apoio ao

Desenvolvimento Científico e Tecnológico (FUNCAP) no período de 2016 a 2018. Este projeto tem por objetivo estudar e implementar novas técnicas de avaliação da interação em sistemas a partir da linguagem textual do usuário.

A autora deste TCC participa do projeto desde seu início, tendo experiência com toda a metodologia de avaliação textual, desde a fase de definição do contexto à extração e classificação de postagens. Com tais experiências, em conjunto com a orientadora e outro participante do projeto, teve dois artigos aceitos e publicados no Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (Interação Humano-Computador (IHC)).

A partir da realização de diversas avaliações utilizando a metodologia MALTU, foi percebido que todos os alunos investigavam apenas as etapas de extração e classificação das postagens. Com isso, foi vista a necessidade de entender melhor a etapa inicial da metodologia e principalmente os usuários de tais sistemas.

## **1.2 Objetivos**

### ***1.2.1 Objetivo Geral***

Apresentar uma metodologia de criação de personas a partir dos dados públicos dos usuários em Sistemas Sociais.

### ***1.2.2 Objetivos Específicos***

- Apresentar um estado da arte sobre a geração automática de personas.
- Apresentar uma metodologia de criação automática de personas de um Sistema Social.
- Validar a metodologia executando um experimento.

## **1.3 Procedimentos Metodológicos**

Nesta seção são apresentados e descritos os procedimentos metodológicos utilizados para alcançar os objetivos desta pesquisa. Foram realizados: (1) mapeamento sistemático da literatura; (2) elaboração de uma metodologia de criação de personas; e (3) validação da metodologia a partir de um experimento em um sistema social.

O mapeamento sistemático foi realizado com base no processo definido por Kitch-

nam e Charters (2007), cujo principal objetivo é encontrar uma avaliação justa de um tema de pesquisa, utilizando uma metodologia rigorosa e confiável. Para encontrar estudos que servissem como base para a execução desta pesquisa, esse mapeamento seguiu as três fases metodologias por Kitchenham e Charters (2007), sendo elas: planejamento, condução e documentação. Como procedimento adicional, a condução foi realizada em dois fluxos. O fluxo primário consistiu na execução do mapeamento sistemático em si e o fluxo secundário na aplicação da técnica *Forward Snowballing*.

A elaboração da metodologia teve como base os resultados do mapeamento sistemático e de um estudo da *Application Programming Interface* (API) do Twitter, a fim de investigar quais dados públicos dos usuários poderiam ser extraídos. Já a validação se deu por um experimento que foi realizado da seguinte forma: 1) configuração da API do Twitter; 2) implementação de um algoritmo de extração para coleta destes dados; 3) simulação de uma persona a partir dos dados coletados.

#### **1.4 Organização do Trabalho**

Este trabalho está organizado e dividido em 7 capítulos:

- (a) no capítulo dois, é apresentada a Fundamentação Teórica deste estudo, explicando os principais conceitos, como Personas e Avaliação Textual;
- (b) no capítulo três, é apresentado o mapeamento sistemático realizado neste estudo;
- (c) no capítulo quatro, os artigos encontrados na revisão sistemática são apresentados, juntamente com os demais trabalhos relacionados a este;
- (d) no capítulo cinco, é apresentada uma metodologia de criação de personas metodologia por este estudo;
- (e) no capítulo seis, é apresentado um experimento realizado para validar a metodologia de geração de personas apresentada;
- (f) no capítulo sete, são apresentadas as considerações finais e trabalhos futuros.


## 2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção são apresentados os principais conceitos para a definição deste estudo. Por se tratar de uma proposta de geração automática de personas, são abordados os seguintes temas: 1) personas: criação e geração automática de personas; e 2) avaliação textual: usuários de uma avaliação textual.

### 2.1 Personas

Personas é uma técnica de design de interação cujo objetivo é aproximar os desenvolvedores e os usuários do sistema (PRUITT; GRUDIN, 2003). Essa técnica fornece um modelo mental para ajudar a entender o usuário, de forma a colocar a equipe em seu lugar (JANSEN; MECHELEN; SLEGGERS, 2017), a partir de representações de usuários fictícios, mostrando uma descrição de seus objetivos, motivação e relacionamento com o sistema (NIELSEN, 2012). Um exemplo de criação de personas pode ser visto na Figura 1, criada no estudo de Holanda (2010).

Figura 1 – Exemplo de uma persona

Júnior	
	<b>Conhecendo o Júnior</b> Idade: 10 anos Estado Civil: Solteiro Nível de Escolaridade: Ensino Fundamental (incompleto) Classe Social Familiar: D (Brasileiro); D (Italiano) Ocupação Profissional: Estudante Tipo de Usuário: Usuário Primário Categoria: Intermediária
	<b>Uso da TV</b> Tempo Dedicado: Mais de 5h por dia Preferência de Programas: Júnior gosta de assistir filmes e programas infantis Sentimentos quando está assistindo TV: Se sente relaxado, porque a TV é pura diversão
	<b>Uso do Sistema Interativo</b> Preferências: Rápido acesso a informação e rápida resposta do sistema Necessidades: Entretenimento Dificuldades: Navegação quando existe o predomínio de texto, ao invés de imagens Finalidades de Uso: Jogos
	<b>Internet, Telefone celular e Uso do Computador</b> Aquisição: Júnior não tem um computador Local de Acesso: Ele acessa a Internet em uma Lan House e na escola Familiaridade com a Tecnologia: Júnior não tem nenhuma dificuldade em jogar jogos de computador ou de celular

Fonte: Holanda (2010).

Desde sua criação, a técnica de Personas, proposta por Cooper (1999), vem sendo utilizada para atingir os mais diversos objetivos, aptidões e interesses do usuário. Por ser uma poderosa e polivalente ferramenta de design, sua utilização possui diversos benefícios, como: auxílio na determinação do produto; comunicação com stakeholders, desenvolvedores e outros

designers; construção de um consenso e comprometimento do design; medição da eficácia de um design; até a contribuição para outros produtos que estejam relacionados (COOPER; REIMANN; CRONIN, 2007).

A aplicação da técnica tem se mostrado eficiente para o entendimento de diferentes tipos de usuários e suas necessidades, sendo a melhor maneira de acomodar a diversidade que há em sistemas (COOPER; REIMANN; CRONIN, 2007). Além da aplicação para análises de marketing, as Personas podem auxiliar no fornecimento de informações para testes com usuários e para prototipação focada nos participantes de um projeto (GRUDIN; PRUITT, 2002).

A utilização de personas aumenta o foco e conhecimento sobre os usuários, diminuindo o número de suposições realizadas sobre a sua opinião (PRUITT; GRUDIN, 2003). Portanto, a criação de personas pode auxiliar no processo de avaliação do sistema, devido ao maior conhecimento das necessidades e premissas dos usuários.

### ***2.1.1 Criação de Personas***

Segundo Pruitt e Adlin (2010) há uma necessidade de criar personas a partir de dados reais dos usuários, ou seja, a partir de dados qualitativos e quantitativos obtidos por meio de observações em campo, questionários, entrevistas com usuários, entre outros. Porém, essas personas devem ser vistas de forma diferente ao perfil do usuário. Os perfis dos usuários fornecem apenas um nome e uma imagem anexada a uma breve descrição demográfica, juntamente com um breve parágrafo fictício que pode não ser útil como ferramenta de design (COOPER; REIMANN; CRONIN, 2007). As personas possuem mais informações que precisam ser identificadas, tendo que ser focadas na relação do usuário com o sistema (NIELSEN, 2012) e tais informações fictícias devem ser apresentadas apenas para dar uma maior vida e proximidade do usuário para/com a equipe de desenvolvimento do produto (COOPER; REIMANN; CRONIN, 2007).

Este trabalho adota a posição de Cooper, Reimann e Cronin (2007), pois acredita-se que quanto mais dados coletar dos usuários, mais próximo ao usuário final o designer vai estar. Portanto, a fidelidade aos usuários será essencial para sua associação com os resultados da avaliação textual.

Alguns trabalhos (HOLANDA, 2010; NÓBREGA; FURTADO, 2011; GUIMARÃES, 2012) têm coletados dados específicos dos usuários, como:

- tipo de usuário: diz respeito ao nível de conhecimento do usuário sobre tecnologia;

- perfil sociodemográfico: está relacionado a vida pessoal e social do usuário, sendo composto por dados relativos aos perfis sociais, demográficos e geográficos;
- contexto de uso: dados relativos ao ambiente em que ocorre a utilização do sistema;
- necessidades, dificuldades e atitudes: informações associadas aos objetivos e necessidades de uso do produto, além de suas dificuldades e atitudes;
- características webográficas: informações referentes ao conhecimento da utilização da internet pelo usuário;
- características tecnológicas: conhecimento do usuário sobre os componentes tecnológicos, como computador e celular;
- perfil psicográfico: condizem ao comportamento do usuário, ou seja, informações sobre sua personalidade e autoimagem;
- características do uso: correspondem as informações do uso do produto, como a frequência de uso, intensidade e facilidade de acesso à informação; e
- feedback: informações manifestadas pelos usuários sobre sua percepção de uso.

Em trabalhos de sistemas de redes sociais, quando os dados foram coletados por meio de entrevistas ou questionários, os dados podem ser mais detalhados, utilizando os dados descritos acima, quando coletados das informações disponibilizadas pelos usuários nos sistemas, geralmente coletam-se a idade, gênero e localização dos usuários (SALMINEN *et al.*, 2017).

Na seção de trabalhos relacionados são apresentados alguns trabalhos que sugeriram métodos de criação de personas utilizando os dados descritos acima.

### **2.1.2 Geração Automática de Personas**

Geração Automática de Personas consiste em construir personas de forma automática, a partir de dados fornecidos ou coletados sobre os usuários. Essa geração automática vem sendo aplicada em estudos como o de Masiero, Ferreira e Jr (2012) e Watanabe *et al.* (2017). Esses trabalhos utilizam dados dos usuários coletados a partir de respostas de questionários online. Uma outra forma de obter dados dos usuários é utilizar seus dados provenientes do perfil de redes sociais, podendo resultar em personas mais precisas devido à grande escala de dados.

A técnica Geração Automática de Personas (em inglês *Automatic Persona Generation*, APG) é apresentada nos estudos de An, Kwak e Jansen (2016), An *et al.* (2016), An, Kwak e Jansen (2017), An *et al.* (2018), Jansen, Mechelen e Slegers (2017), Jansen *et al.* (2017b),



Jansen *et al.* (2017a), Salminen *et al.* (2018b), Salminen *et al.* (2017) e outros, e demonstra a geração dessas personas a partir de dados de usuários extraídos de sistemas sociais como o Youtube, Facebook e Twitter. Seu principal objetivo é auxiliar no processo de marketing para novos empreendedores que utilizam as mídias sociais para divulgação de seus produtos.

Jung *et al.* (2018b) apresentaram quatro etapas para a geração de personas, sendo elas: 1) configuração; 2) coleção; 3) geração; e 4) interação. Para a etapa de configuração são realizadas as principais configurações da Interface de Programação de Aplicativos (em inglês *Application Programming Interface*, API) de extração do sistema em questão. Uma API de extração é uma biblioteca que estabelece um conjunto de rotinas e padrões para utilizar os serviços de um determinado aplicativo, por exemplo definições de como extrair os dados do SS. Na segunda etapa, são coletados os dados dos usuários da rede social definida, utilizando as configurações definidas na primeira etapa. Na terceira etapa, geração, é seguido o fluxo apresentado na Figura 2. Tal fluxo possui as seguintes atividades:

- identificar os diferentes padrões da interação do usuário. Para isso, é utilizada uma matriz que representa a interação dos usuários com um determinado tipo de conteúdo;
- fazer um link entre os padrões da interação com os grupos encontrados a partir dos dados demográficos dos usuários. A partir da matriz identificada na etapa anterior, buscam-se padrões que identifique um grupo demográfico de usuários com maior coeficiente;
- identificar grupos de usuários que tenham maior impacto a partir dos dados demográficos extraídos;
- criar personas shell, ou seja, criar personas apenas com dados demográficos, como nome, idade, gênero e foto de perfil. Para isso os grupos encontrados na etapa anterior são gerados com informações como nome, foto, idade e país;
- enriquecer as personas criadas utilizando as informações obtidas na primeira etapa, adicionando uma descrição sobre os principais tópicos de interesse para o usuário.

Por fim, na etapa de interação, as personas geradas podem ser visualizadas em uma ferramenta, também chamada de APG.

Neste trabalho, diferente dos autores da APG, as personas geradas não terão objetivo de marketing para divulgação de produto, elas serão utilizadas para auxiliar no processo de

Figura 2 – Fluxo de atividades para a geração de personas de Jung *et al.* (2018b).



Fonte: a autora.

conhecimento do usuário para uma avaliação textual, como será descrito na próxima seção.

## 2.2 Avaliação Textual

Devido ao crescimento de aplicações web, que permitem uma maior colaboração, comunicação e interação entre os seus usuários, há a necessidade de considerar a utilização de novos tipos de avaliações da Usabilidade e Experiência do Usuário (UX) em sistemas, que sejam diferentes dos tradicionais (PEREIRA; BARANAUSKAS; SILVA, 2010). Uma forma de utilizar os dados provenientes de aplicações web é realizando avaliações textuais. Muitos trabalhos convidam os usuários a escreverem sobre seu uso, porém a espontaneidade do relato é importante para a integridade da avaliação (MENDES, 2015), pois quando um avaliador faz perguntas aos usuários, ele pode influenciar na experiência de uso e/ou tendenciar as respostas dos usuários (ROGERS; SHARP; PREECE, 2013).

Visando preencher as lacunas citadas, Mendes (2015) propôs a metodologia MALTU,

cujo principal objetivo é guiar profissionais de Interação Humano-Computador (IHC) em uma avaliação da UUX de um sistema a partir de um conjunto de Postagens Relacionadas ao Uso (PRUs). PRUs são postagens públicas realizadas de forma espontânea pelos usuários de SS que se referem ao uso do próprio sistema ao utilizá-lo, como exemplo, tem-se a seguinte postagem extraída da plataforma Google Play sobre o aplicativo MyFitnessPal: “*Bastante completo em termos de alimentos disponíveis, mas a interface é bem menos compreensível ao usuário que outros aplicativos semelhantes.*”.

A metodologia MALTU possui cinco etapas de avaliação, sendo elas: 1) definição do contexto de avaliação; 2) extração de PRUs; 3) classificação das PRUs; 4) interpretação dos resultados; e 5) relato dos resultados. A seguir, as etapas são descritas de forma mais detalhada, apresentando o seu objetivo e atividades a serem realizadas.

A etapa 1, definição do contexto de avaliação, tem como objetivo a definição do contexto de uso do sistema, seu domínio e os objetivos da avaliação. Para a definição do contexto de uso, devem ser indicados quais os usuários, a plataforma e o seu ambiente social e cultural. Para a definição do domínio do sistema, os termos mais prováveis a serem utilizados para falar sobre o sistema, assim como suas funcionalidades devem ser estabelecidos.

A extração de PRUs, etapa 2, consiste em obter PRUs do sistema a ser avaliado, podendo ser realizada de forma manual ou automática. Para realizar a extração, devem ser utilizados padrões de extração, proposto pela metodologia. Na forma manual, as postagens são coletadas com a utilização dos campos de buscas, utilizando tais padrões. Já na forma automática, é utilizada uma ferramenta de extração UUX-Posts (MENDES, 2015; MENDES; FURTADO, 2017).

Na etapa 3, as PRUs extraídas são classificadas em categorias, sendo elas: a) tipo; b) intenção; c) análise de sentimentos; d) funcionalidade; e) critérios de qualidade de uso; e f) artefato. A classificação deve ser realizada por, pelo menos, dois classificadores e um validador. Em seguida, é realizada a etapa 4, interpretação dos resultados de forma a organizar os dados de uma forma que possa evidenciar a relação entre eles. Por fim, é realizado o relato dos resultados, etapa 5, de forma a gerar um relatório com as informações sobre a avaliação realizada.

Em avaliações realizadas no *Waze* e *Google Maps* nos sistemas operacionais *Android* e *Windows Phone* (SILVA; FREITAS; MENDES, 2017), por exemplo, foram obtidos os seguintes resultados de avaliação:

O *Waze* para *Android* apresentou os melhores resultados da avaliação. Na classificação por análise de sentimentos, 70,8% de suas sentenças classificadas

possuíam sentimento positivo. Na classificação por facetas de UUX, o aplicativo obteve uma predominância de 71,1% em “Satisfação”. Mesmo assim, o aplicativo apresentou problemas relacionados a “Eficácia” (8,8%), “Eficiência” (4,5%) e “Aprendizado” (0,6%) (SILVA, FREITAS e MENDES, 2016, p. 7).

Também foram identificadas as funcionalidades que os usuários mais criticaram, como por exemplo a funcionalidade de “Cálculo de rota” e “Contabilização dos pontos”. Para o sistema operacional Windows Phone foram obtidos os resultados:

O Waze para Windows Phone apresentou o pior resultado da avaliação. Na classificação por Tipo, o aplicativo obteve 24,9% de sentenças do tipo “Elogio”, enquanto 71,8% das sentenças foram do tipo “Crítica”. Na classificação por análise de sentimentos, 44,5% de suas sentenças foram de polaridade “Neutra”, 34,2% “Negativa” e 21,3% “Positiva”. Na classificação por facetas de UUX, o aplicativo obteve apenas 22,1% de sentenças relacionadas à “Satisfação”. O aplicativo apresentou problemas de “Eficácia” (51,7%), “Eficiência” (3%) e “Aprendizado” (3,6%) (SILVA, FREITAS e MENDES, 2016, p. 7).

Além da identificação das funcionalidades mais criticadas, como por exemplo, “Navegação por voz” e “Cálculo de rota”, foi possível identificar as funcionalidades que causaram dúvidas no usuário, como, por exemplo, a “Instalação do aplicativo”.

### **2.2.1 *Usuários da Avaliação Textual***

Mendes (2015), na metodologia MALTU, aponta que, durante a primeira etapa: definição do contexto de avaliação, os usuários do sistema a ser avaliado devem ser definidos. Furtado (2012, apud Mendes, 2015) afirma que o usuário é o indivíduo que realiza tarefas no sistema interativo com o intuito de atingir os mais diversos objetivos. Para representar esses usuários é recomendada a técnica de criação de personas, levando em consideração características reais dos usuários de Sistemas Sociais, como idade, sexo, formação acadêmica e tempo de uso no sistema.

Em estudos anteriores (FREITAS; SILVA; MENDES, 2016; SILVA; FREITAS; MENDES, 2017), aplicando a metodologia MALTU (MENDES, 2015), os usuários foram definidos por meio de pesquisas prévias, não realizadas pelos autores, que continham informações dos principais usuários dos sistemas a serem avaliados.

Essas pesquisas podem ser por meio de artigos publicados, dados geográficos ou informações disponibilizadas no próprio sistema. Na avaliação do sistema Spotify (FREITAS; SILVA; MENDES, 2016), por exemplo, os usuários do sistema foram definidos da seguinte forma: “Os usuários do Spotify são compostos por 70% de jovens entre 15 a 24 anos”. Essa definição foi obtida a partir de uma referência a um outro trabalho realizado pelo Spotify (2016 apud FREITAS, SILVA e MENDES, 2016). Na avaliação realizada com os aplicativos Waze e

Google Maps, Silva, Freitas e Mendes (2017) definiram os usuários da seguinte forma: “o Waze na Google Play possui entre 100 e 500 milhões de instalações, no Windows Phone os dados não foram encontrados, já o Google Maps possui entre 1 e 5 bilhões”. Foi definida, também, a motivação dos usuários para o usar o aplicativo, sendo ela:

Além disso, pessoas que moram em grandes cidades precisam se locomover diariamente para diversos lugares, porém, mesmo que elas conheçam o caminho, muitas vezes, há a necessidade de encontrar o mais rápido, considerando fatores como a intensidade do trânsito e acidentes pelo trajeto. (SILVA, FREITAS, MENDES, 2017, p. 3).

Tais dados foram coletados a partir da página do aplicativo nas Apps Stories, tendo sido definida a motivação dos usuários a partir da página do aplicativo do Google Maps na plataforma Google Play, e o número de usuários de cada aplicativo levou em consideração o número de instalações nas plataformas.

Nestes casos apresentados, foi possível ter uma percepção dos usuários de tais sistemas, porém, a possibilidade de construção de personas estando diretamente relacionadas às postagens e seus autores poderiam fornecer um resultado mais preciso de avaliação do sistema. Por exemplo, a causa de muitas críticas ao sistema Waze (SILVA; FREITAS; MENDES, 2017) na plataforma Android e Windows Phone foi a dependência de internet para o cálculo de rotas e reprodução das coordenadas utilizando a navegação por voz, tais problemas podem estar relacionados à localização do usuário quanto ao sinal local da sua operadora.

### **2.2.2 Ferramenta UUX-Posts**

A ferramenta UUX-Posts<sup>1</sup>, atualmente em sua segunda versão, foi proposta por Mendes (2015) com o intuito de apoiar a metodologia MALTU, fornecendo suporte nas etapas de extração e classificação das postagens.

Na versão atual, a ferramenta oferece dois tipos de serviços. Para usuários cadastrados na ferramenta, apresenta um suporte para toda a avaliação, se realizada de forma manual. A sequência de atividades da ferramenta segue as etapas da metodologia, sendo possível definir o contexto de avaliação, inserir as postagens que serão analisadas, classificar tais postagens de forma manual, obter relatórios sobre a interpretação dos dados obtidos com a classificação e relatar os resultados. Para usuários não cadastrados, possibilita a extração de PRUs a partir de um site escolhido pelo avaliador, por exemplo o site do SS Twitter, ou pelo envio de uma planilha de postagens, utilizando o modelo booleano. Quanto a etapa de classificação, a ferramenta apresenta

<sup>1</sup> Acessado em: 10 de novembro de 2018. Disponível em: <http://uuxposts.russas.ufc.br/v3/>

um suporte na definição de padrões para identificar o que será avaliado no sistema, por exemplo, em facetas de UUX, realizando a classificação durante a extração dessas postagens.

Este estudo irá utilizar o processo de extração de postagens, a partir do SS Twitter. Além disso, o principal objetivo deste estudo é contribuir nas atividades condizentes com a definição do contexto de uso, pois, nessa etapa são definidos os usuários do sistema. A Figura 3 apresenta a tela atual onde são cadastradas essas informações. O cadastro dos usuários está inserido nesta definição de contexto de uso. O avaliador, por meio de uma pesquisa prévia tem a possibilidade de cadastrar essas informações. Além disso, essas informações não são mostradas no decorrer da avaliação, apenas na geração do relatório final, após a realização do relato dos resultados, não relacionando seus dados com os resultados obtidos.

Figura 3 – Tela de cadastro de contexto de uso e definição dos usuários.

The screenshot displays the 'UUX-Posts' application interface. On the left is a dark sidebar menu with options: 'Início', 'Editar perfil', and a section titled 'AVALIAÇÕES' containing 'Criar uma nova avaliação', 'Avaliações em andamento', and 'Avaliações concluídas'. The main content area is titled 'UUX-Posts Versão 2.0' and shows the evaluation context for 'Avaliação do sistema My Fitness Pal para Android'. Below this, it prompts the user to 'Informe o contexto da avaliação (opcional)'. The form includes several fields: 'Ambiente físico' (Example: 'Os usuários usam o aplicativo em casa'), 'Ambiente social' (Example: 'em família, no trabalho'), 'Ambiente cultural' (Example: 'universitários, crianças, motoristas'), 'Faixa etária' (dropdown: 'Selecione'), 'Sexo' (dropdown: 'Selecione'), 'Formação acadêmica' (dropdown: 'Selecione'), 'Tempo de uso do sistema' (Example: '2 anos'), and 'Experiência tecnológica' (Example: 'Usa com frequência'). A blue 'Voltar' button is located at the bottom of the form. At the very bottom of the screen, a footer reads: 'UUX-Posts - Uma ferramenta para extração e classificação de PRUs de Sistemas Sociais'.

Fonte: UUX-Posts. Acessado em: 12 de novembro de 2018. Disponível em <<http://uuxposts.russas.ufc.br/>>.

### 3 MAPEAMENTO SISTEMÁTICO

Um mapeamento sistemático tem como objetivo caracterizar uma área de pesquisa, encontrando uma avaliação justa de um tema de pesquisa, utilizando uma metodologia rigorosa e confiável (KITCHENHAM; CHARTERS, 2007). Este mapeamento se difere de uma revisão sistemática quanto aos estudos selecionados, pois, por ser uma área nova, ela não possui muitos estudos primários. Para isso, este mapeamento foi realizado seguindo os preceitos de uma revisão sistemática, quanto ao caráter de organização e planejamento.

Com base no protocolo de Kitchenham e Charters (2007), três fases foram realizadas, sendo elas: planejamento, condução e documentação. Dentre essas fases, foram realizadas as oito etapas propostas: 1) identificação da necessidade de uma revisão; 2) desenvolvimento de um protocolo de revisão; 3) identificação da pesquisa; 4) seleção dos estudos; 5) avaliação da qualidade dos estudos; 6) extração dos dados; 7) apresentação dos estudos; e 8) resultado das questões.

Para obter uma cobertura mais completa dos resultados, foi realizado, em adicional na fase de execução, um fluxo secundário, denominado *Forward Snowballing*. Esse fluxo é um procedimento manual, cujo objetivo é garantir que trabalhos que podem ser interessantes para a pesquisa, mas não foram retornados pelo fluxo principal, possam fazer parte dos resultados deste mapeamento (WOHLIN, 2014).

Nas próximas seções, será apresentado o detalhamento do mapeamento sistemático realizado neste estudo.

#### 3.1 Identificação da necessidade de uma revisão

Apesar de alguns trabalhos apresentarem um processo para criação de personas, não foram encontrados trabalhos que apresentassem um estado de arte sobre a geração automática de personas, identificando informações sobre o que é e como funciona.

#### 3.2 Planejamento do Mapeamento Sistemático: Definição do Protocolo

O planejamento do mapeamento sistemático compreende à definição do protocolo, onde são detalhadas as etapas metodológicas a serem realizadas, com o intuito de assegurar a validade de seus resultados. Além disso, o protocolo permite a redução do risco de comprometimento dos dados, promovendo a transparência das decisões tomadas no decorrer da

execução.

Nas seções seguintes, serão apresentadas as etapas propostas por Kitchenham e Charters (2007), sendo elas: definição das questões de pesquisa, definição da estratégia de busca, definição dos critérios de Inclusão/Exclusão, definição dos procedimentos de seleção, definição dos procedimentos de extração e síntese dos dados. Para auxiliar no processo de seleção, foi realizada também uma definição dos critérios de avaliação da qualidade dos estudos.

### **3.2.1 *Formulação das Questões de Pesquisa***

A formulação de questões de pesquisa corretas e relevantes para uma pesquisa pode se tornar um processo complexo e demorado (LIBERATI *et al.*, 2009). Uma forma de solucionar este problema é utilizando a técnica PICOS (em inglês *Population, Intervention, Comparison, Outcome e Study*) (KITCHENHAM; CHARTERS, 2007). Essa técnica propõe o auxílio na definição correta das questões de pesquisa, definindo a população foco da pesquisa, qual sua intervenção, finalidade, resultados esperados e o tipo de estudo a serem considerados. Para isso, deve ser executada com base na necessidade identificada na seção anterior. A definição do escopo é apresentada a seguir.

#### **3.2.1.1 *Definição do escopo aplicando a técnica PICOS***

- População: usuários de sistemas sociais.
- Intervenção: uma proposta para geração automática de personas a partir de dados públicos em SS.
- Controle: revisões sistemáticas e mapeamentos sistemáticos em engenharia de software são de caráter exploratório, destinando-se a caracterizar uma linha de pesquisa específica. Portanto, não se faz necessário ter um grupo de controle.
- Resultados: técnicas para geração automática de personas, compreendendo processos ou metodologias; identificação dos dados extraídos de SS; requisitos para uma ferramenta; e técnicas para a validação das personas geradas automaticamente.
- Estudos: artigos que apresentem uma proposta de geração automática de personas, ferramentas e/ou a validação dessas personas.
- Finalidade: elaborar e desenvolver uma proposta de geração automática de personas a partir dos dados públicos dos usuários em SS.



### 3.2.1.2 Questões de Pesquisa

Com base nas características identificadas com a técnica PICOS, as questões de pesquisa deste mapeamento sistemático são apresentadas a seguir (Quadro 1).

Quadro 1 – Questões de Pesquisa do Mapeamento Sistemático.

ID	Pergunta	Justificativa
QP1	Quais são as técnicas e métodos utilizados para a geração automática de personas?	Descobrir se existem e quais as técnicas e métodos adotados para a geração automática de personas, assim como seu funcionamento
QP2	Quais os dados extraídos para a geração das personas?	Descobrir quais os dados, normalmente, são coletados para gerar as personas e como são aplicados na geração da persona
QP3	Existem ferramentas que geram personas de forma automática?	Identificar ferramentas que gerem personas automaticamente a fim de levantar requisitos para adicionar na ferramenta UUX-Posts
QP4	Quais os contextos de aplicação das personas geradas de forma automática?	Identificar o contexto de aplicação das personas quando geradas automaticamente
QP5	Existem técnicas para validar as personas geradas?	Descobrir como as personas são validadas, qual a técnica utilizada e se validam o uso das personas ou os dados extraídos

Fonte: elaborado pela autora.

### 3.2.2 Estratégias para busca dos estudos

Para definir a estratégia de busca e seus recursos, foram selecionados três itens fundamentais, sendo eles:

- *Palavras-chave*: seguem o princípio do objetivo deste mapeamento, o qual é encontrar as técnicas de geração automática de personas. Para aumentar a abrangência, foram utilizados sinônimos da palavra geração, como construção, criação, especificação e elaboração.
- *Idioma*: para obter um maior número de resultados, a estratégia de busca foi aplicada no idioma Inglês.
- *Fontes de busca*: foram utilizados seis, dos sete principais mecanismos de busca acadêmico, segundo Buchinger, Cavalcanti e Hounsell (2014). A *Springer Link* não foi considerada, pois grande parte de seus artigos são de difícil acesso, sendo necessário pagar, mesmo com o acesso pelo proxy da Universidade Federal do Ceará (UFC). Tais mecanismos são apresentados no Quadro 2, com seus respectivos links de acesso.
- *Data de publicação*: não foi levada em consideração na busca, pois não se sabe

Quadro 2 – Bases de dados utilizadas no Mapeamento Sistemático.

Mecanismo de busca	Link
Web of Knowledge	<a href="https://www.webofknowledge.com">https://www.webofknowledge.com</a>
Engineering Village	<a href="https://www.engineeringvillage.com">https://www.engineeringvillage.com</a>
Scopus	<a href="https://www.scopus.com">https://www.scopus.com</a>
IEEE Xplorer	<a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a>
ACM Digital Library	<a href="https://dl.acm.org/dl.cfm">https://dl.acm.org/dl.cfm</a>
Science Direct	<a href="https://sciencedirect.com">https://sciencedirect.com</a>

Fonte: elaborado pela autora.

ao certo quando foi criada a primeira técnica de geração automática de personas. Portanto, para que fosse obtido o maior número de técnicas, não foi estabelecido um período de publicação dos artigos.

### 3.2.2.1 Expressão geral de busca

Esta seção descreve as tentativas de refinamento da string de busca, onde todas foram submetidas as bases de dados eletrônicas selecionadas, apresentadas anteriormente.

#### 3.2.2.1.1 Elaboração da String de Busca - Versão 1

Inicialmente, utilizando a estratégia PICOS para a construção da string, foi realizado um teste com a string de busca "*Generation of Persona OR Construction of Persona*". Com a execução do mapeamento utilizando essa string, foram encontradas diversas técnicas de geração de personas, porém, nem todas eram realizadas de forma automática, obtendo técnicas que não se enquadram no escopo desta pesquisa. Além disso, foi percebido que deveriam ser incluídos mais sinônimos para a palavra *Generation*. Para sanar tais problemas, foi criada uma nova string que será apresentada no próximo tópico.

#### 3.2.2.1.2 Elaboração da String de Busca - Versão 2

Com as informações percebidas com a aplicação da primeira tentativa, uma nova string foi criada, sendo ela: "*Automatic Generation of Persona OR Automatic Persona Generation OR Automatic Construction of Persona OR Automatic Persona Construction OR Automatic Elaboration of Persona OR Automatic Persona Elaboration OR Automatic Specification of Persona OR Automatic Persona Specification OR Automatic Creation of Persona OR Automatic Persona Creation*". Porém, a aplicação dessa string não foi possível na base de dados *Science Direct* devido o número de termos. Além disso, com uma breve análise dos

resultados, foi percebido que alguns artigos não foram retornados devido a ordem das palavras.

### 3.2.2.1.3 Elaboração da String de Busca - Versão Final

Por fim, chegou-se a esta versão: "*Automatic AND Persona AND (Generation OR Construction OR Elaboration OR Specification OR Creation)*", sanando o problema de tamanho da string e ordem das palavras, mantendo os mesmos termos.

### 3.2.2.1.4 Comparação dos resultados por base das versões da string

Com as alterações realizadas na string, os resultados obtidos passaram por um processo de ajuste, até conseguir obter um bom resultado dentro do escopo da pesquisa. Essa variação pode ser vista na Tabela 1, onde os resultados da string 1 e da versão final mudam, devido a mudança no escopo da revisão e os resultados da string 2 e da versão final aumentam.

Tabela 1 – Comparativo dos resultados por base.

<b>Base de Dados</b>	<b>String 1</b>	<b>String 2</b>	<b>String final</b>
Web of Knowledge	Não aplicada	3	6
Engineering Village	141	6	25
Scopus	223	6	22
IEEE Xplorer	10	3	8
ACM Digital Library	23	2	8
Science Direct	Não aplicada	Número de termos excedido	2
<b>TOTAL</b>	<b>397</b>	<b>20</b>	<b>71</b>

Fonte: elaborado pela autora.

## 3.2.3 Critérios e Procedimentos para Seleção dos Estudos

O processo de seleção dos estudos podem estar sujeito a vieses, para isso, os critérios de inclusão e exclusão foram definidos de forma clara e explícita, com base no escopo do mapeamento sistemático (DRESCH; LACERDA; JÚNIOR, 2015). Nas seções seguintes, são apresentados os critérios de inclusão e exclusão desta revisão.

### 3.2.3.1 Critérios de Exclusão

Para realizar a exclusão dos artigos que não se enquadram neste mapeamento sistemático, foram definidos os seguintes critérios de exclusão:

**CE1** - O artigo não foi escrito no idioma Inglês ou Português;

**CE2** - Artigo pago, mesmo utilizando o proxy da UFC;

**CE3** - Artigo duplicado;

**CE4** - O artigo não possui autor;

**CE5** - O artigo não possui *abstract*;

**CE6** - O artigo está fora do escopo da pesquisa;

**CE7** - Estudos que não sejam artigos de conferência ou periódicos.

### 3.2.3.2 *Crítérios de Inclusão*

Foram definidos os seguintes critérios de inclusão, caso o artigo não atenda a nenhum critério de exclusão:

**CI1** - O artigo propõe uma técnica para geração automática de personas;

**CI2** - O artigo apresenta uma ferramenta de geração automática de personas;

**CI3** - O artigo realiza validação das personas geradas automaticamente.

### 3.2.4 *Definição do procedimento de seleção*

A execução deste mapeamento foi realizada em três fases, sendo elas: Identificação, Seleção e Extração. A etapa de Identificação consiste na aplicação da *string* de busca nas bases de dados selecionadas e agrupamento dos resultados obtidos em cada uma.

Na etapa de Seleção, é obtido um conjunto inicial de estudos, com base na leitura dos títulos e resumos de todos os artigos e seleção realizada com base nos critérios de inclusão e exclusão. Para prevenir o viés de exclusão, os artigos em que o revisor teve dúvida passaram para fase seguinte. Essa medida foi adotada pois essa seleção foi realizada por apenas um revisor.

Por fim, a etapa de extração. Os artigos incluídos na etapa de seleção foram lidos e avaliados de acordo com os critérios de inclusão e exclusão, com base em sua leitura completa. No decorrer da leitura, para os artigos incluídos, foram respondidos os formulários apresentados nas seções seguintes, referentes a extração dos dados e avaliação da qualidade.

Vale ressaltar que os artigos obtidos com a execução de cada fase, assim como os dados extraídos na etapa de extração, foram validados pela orientadora desta pesquisa.

### 3.2.5 *Seleção das publicações a partir de critérios de qualidade*

Após a leitura completa do artigo ter sido realizada, tais artigos foram caracterizados e avaliados com base no checklist apresentado a seguir, no Quadro 3. Por se tratar de um

mapeamento sistemático, onde nem todos os artigos apresentam um estudo primário completo, a utilização deste formulário caracterizou os estudos mais relevantes para o contexto desta pesquisa.

Quadro 3 – Checklist de Avaliação da Qualidade.

ID	Pergunta	Pontuação
CQ1	Apresenta uma ferramenta para a geração automática das personas?	Sim (+1) Não (0)
CQ2	Extrai dados automaticamente de redes sociais?	Sim (+1) Não (0)
CQ3	Descreve o processo de mineração para geração automática da persona?	Sim (+1) Não (0)
CQ4	Existe uma validação da persona?	Sim (+1) Não (0)

Fonte: elaborado pela autora.

### 3.2.6 Seleção dos Dados Relevantes

O processo de seleção dos dados relevantes consiste na análise dos artigos, com o intuito de extrair resultados qualitativos e quantitativos importantes para o contexto da pesquisa em questão. Para nortear a extração das informações mais relevantes, com base nas questões de pesquisa, foi utilizado o formulário descrito a seguir, apresentado no Quadro 4.

Quadro 4 – Formulário de Extração

ID	Descrição dos Dados	Relacionada a questão
D1	Nome da ferramenta? Link de acesso? Plataforma? Autores? País? Universidade? Qual o sistema utilizado na extração? Principais requisitos da ferramenta?	QP3
D2	Qual a técnica utilizada? Autores? Quais as etapas realizadas na técnica? Qual(is) a(s) etapa(s) automática(s)?	QP1
D3	Quais as características das personas geradas? Como os dados foram extraídos? Quais os dados extraídos?	QP2
D4	Como as personas foram validadas? Foi utilizada qual técnica? Como foi realizada?	QP5
D5	Em qual contexto a persona gerada é aplicada?	QP4

Fonte: elaborado pela autora.

### 3.2.7 Síntese dos dados

Para responder as questões de pesquisa deste mapeamento, foi realizada uma síntese dos dados obtidos. Seguindo a proposta de Kitchenham e Charters (2007), a síntese de dados deste mapeamento consistiu em uma análise quantitativa, a partir das publicações obtidas, e descritiva, quanto os dados extraídos a partir da leitura completa dos trabalhos. Os dados obtidos serão descritos na seção 3.4.1.

### 3.3 Execução do Mapeamento Sistemático

A execução deste mapeamento sistemático aconteceu no período de agosto a outubro de 2018, onde a etapa de identificação ocorreu no fim do primeiro mês e as demais etapas no decorrer do período. Sua realização ocorreu com base no protocolo desenvolvido e apresentado na seção 3.2 e sua execução ocorreu em dois fluxos, sendo eles o fluxo principal (Procedimento de busca automática) e fluxo secundário (Procedimento de busca manual), denominado *Forward Snowballing*.

Os trabalhos encontrados como resultados da busca realizada tiveram suas referências bibliográficas (*bibtex*) extraídas nas próprias bases de dados. O fluxo processual do mapeamento sistemático foi gerenciado pela ferramenta StArt <sup>1</sup>. Em seguida, para melhor apresentação e documentação dos estudos, foi utilizada uma Planilha Google.

#### 3.3.1 Fluxo principal - Mapeamento Sistemático

Realizando a etapa de Identificação dos estudos, aplicação da *string* nas bases de dados, foi obtido um total de 71 artigos. A quantidade de artigos por base foi apresentada na Seção 3.2.2.1.4 no Quadro 1.

O *bibtex* de todos os artigos foi inserido na ferramenta StArt, e a partir dela, os artigos foram aceitos, rejeitados ou definidos como duplicados. Os artigos duplicados foram identificados pela própria ferramenta, com exceção de alguns que precisaram ser definidos pelo revisor. Uma visão geral dos artigos obtidos neste fluxo é apresentada no Quadro 2.

Tabela 2 – Quantidade de artigos por etapa.

Etapa	Aceitos	Rejeitados	Duplicados
Identificação dos estudos	71	0	0
Seleção	21	18	32
Extração	15	6	0

Fonte: elaborado pela autora.

No final deste fluxo, foram obtidos 15 (quinze) artigos que se adequaram as questões de pesquisa deste estudo, apresentando uma técnica de geração automática de personas. Quanto ao resultado por bases, oito dos quinze artigos são da *ACM Digital Library*, cinco são da *Engineering Village* e três da *IEEE*. As demais bases não tiveram artigos selecionados.

Os quinze artigos aceitos são apresentados em seguida (Tabela 3). Nesse quadro são

<sup>1</sup> Acesso em: 29 de outubro de 2018. Disponível em: < <http://lapes.dc.ufscar.br/tools/startool> >

apresentados: o título, autores, ano de publicação, local de publicação e base de dados de cada artigo resultante da etapa final.

Tabela 3 – Detalhamento dos artigos encontrados no Fluxo Primário.

<b>Título</b>	<b>Autor(es)</b>	<b>Ano</b>	<b>Local de publicação</b>	<b>Base de Dados</b>
Findings of a User Study of Automatically Generated Personas	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	2018	CHI 2018	ACM
Algoritmos De Clusterização e Python Científico Apoiando Modelagem De Usuário	Masiero, Andrey Araujo and Ferreira, Leonardo A. and Aquino,Jr., Plinio Thomaz	2012	11th Brazilian Symposium on Human Factors in Computing Systems	ACM
Automatic Persona Generation (APG): A Rationale and Demonstration	Jung, Soon-gyo and Salminen, Joni and Kwak, Haewoon and An, Jisun and Jansen, Bernard J.	2018	Conference on Human Information Interaction and Retrieval 2018	ACM
Persona Generation from Aggregated Social Media Data	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeed and Nielsen, Lene and Jansen, Bernard J.	2017	CHI 2017	ACM
Ingimp: Introducing Instrumentation to an End-user Open Source Application	Terry, Michael and Kay, Matthew and Van Vugt, Brad and Slack, Brandon and Park, Terry	2008	SIGCHI Conference on Human Factors in Computing Systems	ACM
Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	2016	CHI 2016	ACM

What Would They Think?: A Computational Model of Attitudes	Liu, Hugo and Maes, Pattie	2004	9th International Conference on Intelligent User Interfaces	ACM
Validating social media data for automatic persona generation	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	2017	IEEE/ACS International Conference on Computer Systems and Applications	Engineering Village
Archetype-based modeling of persona for comprehensive personality computing from personal big data	Guo, Ao and Ma, Jianhua	2018	Journal Sensors (Switzerland)	Engineering Village
Scenario-based modeling of ontic personae for automatic personality perception	Guo, Ao and Ma, Jianhua	2017	SmartWorld/ SCALCOM/ UIC/ ATC/ CBD-Com/ IOP/ SCI 2017	Engineering Village
Automatically conceptualizing social media analytics data via personas	Jung, Soon-Gyo and Salmiinen, Joni and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	2018	ICWSM 2018	Engineering Village
An Integrative and Precise Approach in Personality Computing Based on Ontic Personae Modeling	Guo, Ao and Ma, Jianhua	2018	IEEE Cyber Science and Technology Congress, DASC-PICom-DataCom-CyberSciTec	Engineering Village



Towards Automatic Persona Generation Using Social Media	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohammed Z. and Jansen, Bernard J.	2016	4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	IEEE
Generating Cultural Personas from Social Data: A Perspective of Middle Eastern Users	Salminen, Joni and Sengün, Sercan and Kwak, Haewoon and Jansen, Bernard J. and An, Jisun and Jung, Soon-Gyo and Vieweg, Sarah and Harrell, D. Fox	2017	5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	IEEE
Leveraging Social Analytics Data for Identifying Customer Segments for Online News Media	Jansen, Bernard J. and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	2017	IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)	IEEE

Fonte: elaborado pela autora.

### 3.3.2 Fluxo secundário - *Forward Snowballing*

A principal motivação para a realização de um segundo fluxo, foi o baixo número de artigos relevantes encontrados no mapeamento sistemático (15 artigos). Portanto, para ampliar tais resultados, foi realizado o *Forward Snowballing*, uma técnica complementar ao mapeamento sistemático para identificar outros artigos, com base nas citações realizadas dos artigos encontrados no fluxo principal (WOHLIN, 2014).

Este fluxo utiliza o mesmo protocolo do fluxo principal, sua única diferença está na execução da fase de Identificação dos Estudos. Diferente do primeiro fluxo, essa fase foi realizada de forma manual, onde todos os artigos selecionados no fluxo principal foram estudados

utilizando o GoogleScholar<sup>2</sup>.

Essa fase ocorreu em três etapas. Na primeira etapa, busca do artigo, foi realizada uma pesquisa no *GoogleScholar* a partir do título do artigo resultante do fluxo principal. Em seguida, foi extraído o bibtex de todos os artigos que o citam e gerado um arquivo com todos eles, de forma manual. Por fim, esse arquivo foi inserido na ferramenta *StArt*, juntamente aos resultados do mapeamento. O número de citações para cada artigo pode ser visto no Quadro 4.

Tabela 4 – Número de citações por artigo.

<b>Artigo</b>	<b>Número de citações</b>
Liu e Maes (2004)	45
Terry <i>et al.</i> (2008)	39
Masiero, Ferreira e Jr (2012)	0
Zhang, Brown e Shankar (2016)	19
An <i>et al.</i> (2016)	12
An, Kwak e Jansen (2016)	7
Jung <i>et al.</i> (2017b)	16
Jansen <i>et al.</i> (2017a)	0
Guo e Ma (2017a)	0
Guo e Ma (2017b)	0
Salminen <i>et al.</i> (2017)	6
Guo e Ma (2018)	0
Jung <i>et al.</i> (2018a)	0
Jung <i>et al.</i> (2018b)	0
Salminen <i>et al.</i> (2018b)	0

Fonte: elaborado pela autora.

Após a inserção dessa nova busca, resultando em 144 artigos, as demais etapas foram realizadas seguindo o protocolo. A quantidade de artigos por etapa é apresentada no Quadro 5.

Tabela 5 – Quantidade de artigos por etapa (Fluxo Secundário).

<b>Etapa</b>	<b>Aceitos</b>	<b>Rejeitados</b>	<b>Duplicados</b>
Identificação dos estudos	144	0	0
Seleção	15	91	34
Extração	10	5	0

Fonte: elaborado pela autora.

Os dez artigos aceitos são apresentados em seguida (Tabela 6). Nesse quadro é apresentado o título, autores, ano de publicação, local de publicação e artigo origem dos resultantes da etapa final.

<sup>2</sup> Acesso em: 29 de outubro de 2018. Disponível em: <https://scholar.google.com.br/>

Tabela 6 – Detalhamento dos artigos encontrados no Fluxo Secundário.

<b>Título</b>	<b>Autor(es)</b>	<b>Ano</b>	<b>Local de publicação</b>	<b>Artigo origem</b>
Personas for content creators via decomposed aggregate audience statistics	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	2017	International Conference on Advances in Social Networks Analysis and Mining 2017	Jung <i>et al.</i> (2017b)
Viewed by too many or viewed too little: Using information dissemination for audience segmentation	Jansen, Bernard J and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	2017	Association for Information Science and Technology	Jung <i>et al.</i> (2017b)
Inferring Social Media Users' Demographics from Profile Pictures: A Face++ Analysis on Twitter Users	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Salminen, Joni and Jansen, Bernard J	2017	ICEB 2017 Proceedings	Jung <i>et al.</i> (2017b)
Fixation and Confusion: Investigating Eye-tracking Participants' Exposure to Information in Personas	Salminen, Joni and Jansen, Bernard J and An, Jisun and Jung, Soon-Gyo and Nielsen, Lene and Kwak, Haewoon	2018	Conference on Human Information Interaction and Retrieval 2018	Jung <i>et al.</i> (2017b)
"Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles	Salminen, Joni and Nielsen, Lene and Jung, Soon-gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J	2018	CHI 2018	Jung <i>et al.</i> (2017b)

Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data	An, Jisun and Kwak, Haewoon and Jung, Soon-gyo and Salminen, Joni and Jansen, Bernard J	2018	Social Network Analysis and Mining	Jung <i>et al.</i> (2017b)
ID3P: Iterative Data-Driven Development of Persona based on quantitative evaluation and revision	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Noyori, Yuki and Fukazawa, Yoshiaki and Morizuki, Aoi and Shibata, Hiroyuki and Ogawa, Kentaro and Ishigaki, Mikako and Shiizaki, Satiyo and others	2017	10th International Workshop on Cooperative and Human Aspects of Software Engineering	Zhang, Brown e Shankar (2016)
Retrospective based on data-driven persona significance in B-to-B software development	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Fukazawa, Yoshiaki and Taga, Masahiro and Matsuzaki, Akira and Suzuki, Takayoshi	2018	40th International Conference on Software Engineering: New Ideas and Emerging Results	Zhang, Brown e Shankar (2016)
Persona Perception Scale: Developing and Validating an Instrument for Human-Like Representations of Data	Salminen, Joni and Kwak, Haewoon and Santos, João M and Jung, Soon-Gyo and An, Jisun and Jansen, Bernard J	2018	CHI 2018	Zhang, Brown e Shankar (2016)

The Potential and Issues in Data-Driven Development of Web Personas	Mijac, Tea and Jadric, Mario and Cukusic	2018	MIPRO 2018	Zhang, Brown e Shankar (2016)
---	--	------	------------	-------------------------------

Fonte: elaborado pela autora.

### 3.3.3 Avaliação da qualidade dos estudos

A avaliação da qualidade dos estudos foi realizada com base no formulário representado no Quadro 3 da seção 3.2.5.

Com medidas variando entre 0 e 4, obteve-se uma média geral de 2,92, onde nenhum artigo identificado foi classificado com nota 0. No geral, 8 artigos foram classificados com nota máxima (quatro), 8 com nota três, 8 com nota dois e apenas um com nota 1. Essa classificação pode ser vista no Apêndice F.

Os artigos que ficaram abaixo da média já eram esperados, devido as perguntas voltadas para a natureza deste mapeamento sistemático. Por exemplo, artigos que apresentaram uma ferramenta de geração automática de personas são importantes para o contexto desta pesquisa, porém, se não apresentasse não ganhava um ponto na avaliação. A avaliação realizada pode ser vista na Tabela 7.

## 3.4 Documentação

Nesta seção será apresentada a fase de documentação dos resultados. Nela serão explanados os 25 artigos obtidos na fase de execução e a análise e interpretação dos resultados. Com isso, serão respondidas as cinco questões de pesquisa propostas na seção 3.2.1.2.

### 3.4.1 Apresentação dos estudos

Quanto aos resultados obtidos no fluxo principal, os artigos aceitos foram advindos de apenas três bases, como representados no Gráfico 4, sendo elas ACM (46,7%), Engineering Village (33,3%) e IEEE (20%).

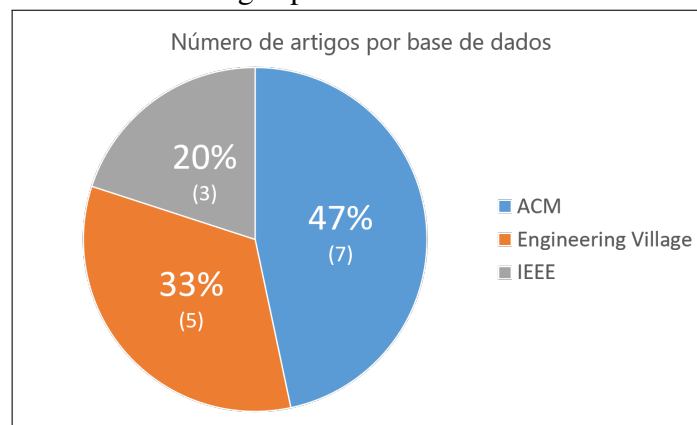
Ao unir os resultados dos dois fluxos do mapeamento sistemático, foram obtidos 25 artigos relevantes para a pesquisa. As seleções realizadas neste estudo resultaram em uma taxa

Tabela 7 – Avaliação da qualidade dos artigos.

Artigo	CQ1	CQ2	CQ3	CQ4	Total
Salminen <i>et al.</i> (2018b)	1	1	1	1	4
Liu e Maes (2004)	1	1	1	1	4
Guo e Ma (2018)	1	1	1	1	4
Guo e Ma (2017a)	1	1	1	1	4
Salminen <i>et al.</i> (2017)	1	1	1	1	4
Jansen <i>et al.</i> (2017b)	1	1	1	1	4
Salminen <i>et al.</i> (2018d)	1	1	1	1	4
An <i>et al.</i> (2018)	1	1	1	1	4
Jung <i>et al.</i> (2018b)	1	1	1	0	3
Jung <i>et al.</i> (2017b)	1	1	1	0	3
An <i>et al.</i> (2016)	1	1	1	0	3
Jansen <i>et al.</i> (2017a)	1	1	1	0	3
An, Kwak e Jansen (2017)	0	1	1	1	3
Salminen <i>et al.</i> (2018a)	1	1	0	1	3
Watanabe <i>et al.</i> (2017)	0	1	1	1	3
Watanabe <i>et al.</i> (2018)	0	1	1	1	3
Terry <i>et al.</i> (2008)	1	1	0	0	2
Zhang, Brown e Shankar (2016)	0	1	0	1	2
An, Kwak e Jansen (2016)	0	1	1	0	2
Guo e Ma (2017b)	0	0	1	1	2
Jung <i>et al.</i> (2018a)	1	1	0	0	2
Jung <i>et al.</i> (2017a)	0	1	0	1	2
Salminen <i>et al.</i> (2018c)	1	0	0	1	2
Mijač, Jadrić e Ćukušić (2018)	0	1	1	0	2
Masiero, Ferreira e Jr (2012)	0	0	1	0	1

Fonte: elaborado pela autora.

Gráfico 4 – Artigos por base de dados.



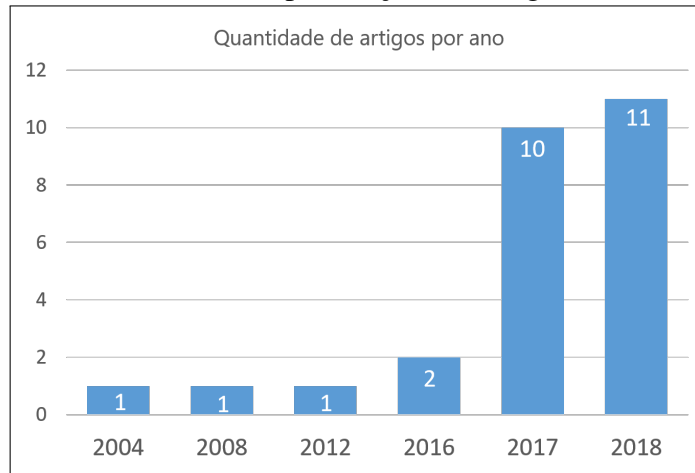
Fonte: a autora.

de aceitação dos artigos de 11,6%, 21,1% no fluxo principal e 0,69% no secundário. O processo de seleção dos estudos em cada etapa, assim como a classificação de acordo com os critérios de inclusão e exclusão podem ser vistos no Anexo A.

O primeiro artigo que retrata uma técnica de geração automática de personas é referente ao ano de 2008, afirmando que este tipo de estudo é bem novo e recente. Tal afirmação pôde ser observada, pois a data de publicação dos artigos não foi restrita durante a busca. Grande parte dos artigos foram recentemente publicados, sendo 23 artigos referentes aos anos de 2016,

2017 e 2018, como apresentado no Gráfico 5. Portanto, pode-se perceber uma tendência de crescimento, ou seja, um aumento de interesse na geração automática de personas.

Gráfico 5 – Ano de publicação dos artigos aceitos.



Fonte: a autora.

Desses 25 artigos, apenas dois foram publicações em revistas, os demais foram publicados em conferências. Os locais de publicação estão descritos nos Quadros 3 e 6. Vale ressaltar que a "*Conference on Human Factors in Computing Systems (CHI)*" foi a conferência com o maior número de artigos relacionados ao tema (cinco artigos).

Dentre esses artigos, 117 autores foram envolvidos nos 25 estudos. A Figura 6 apresenta uma *WordCloud* com o nome desses autores, onde o tamanho da fonte é o reflexo da quantidade de trabalhos em que o autor está envolvido. A figura mostra que o maior número de artigos foi escrito por Kwak, Jansen e Jung, tais autores aparecem juntos em todos os artigos que escreveram.

Figura 6 – WordCloud dos autores dos 25 artigos.



Fonte: a autora.

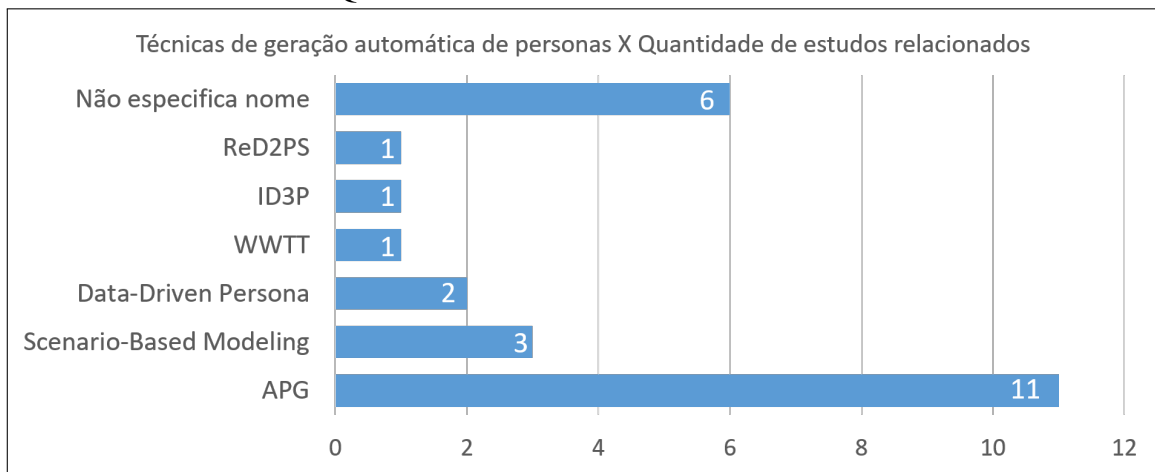
### 3.4.2 Resultado das questões

A seguir, são descritos os resultados das questões de pesquisa, com base nos dados extraídos na etapa de extração. Tais dados podem ser vistos nos Apêndices A, B, C, D e E.

#### 3.4.2.1 Quais são as técnicas e métodos utilizados para a geração automática de personas?

A partir dos 25 artigos selecionados foi possível identificar 12 técnicas para geração automática de personas, sendo 6 nomeadas e 6 não nomeadas. Dentre essas técnicas destaca-se a APG (Automatic Persona Generation), identificada em 11 artigos. A quantidade de estudos relacionados a cada técnica pode ser vista no Gráfico 7.

Gráfico 7 – Técnicas X Quantidade de estudos relacionados



Fonte: a autora.

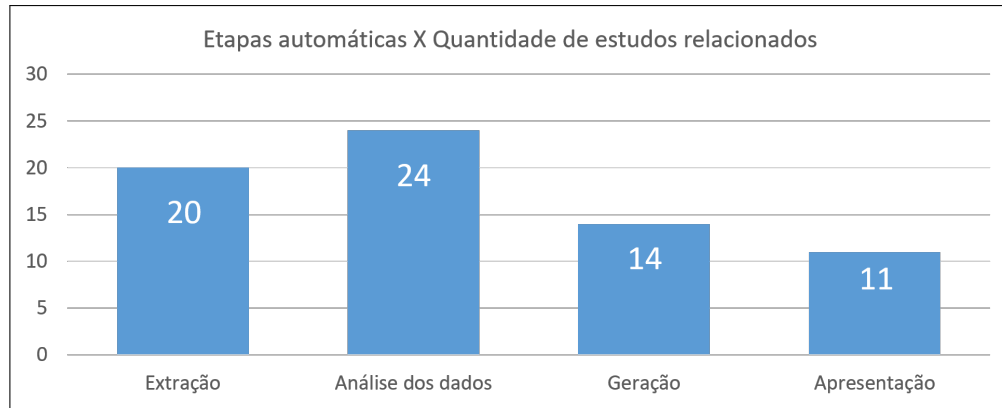
Para entender melhor como funcionava cada técnica, foram analisadas quais as etapas a serem executadas e quais eram realizadas de forma automática. No geral, as técnicas apresentaram as seguintes fases: Extração dos dados, Análise dos Resultados, Geração das Personas e Apresentação. A etapa de extração consistia em coletar os dados sobre os usuários de uma determinada aplicação. A análise, em geral, era realizada com a aplicação de algoritmos de clusterização com os dados obtidos na fase de extração. A geração consistia na especificação da persona, gerando quais dados deveriam ser apresentados. E, por fim, a apresentação das personas, normalmente realizada para técnicas que dispunham de uma ferramenta. Os resultados referentes a quantidade de estudos relacionados a cada etapa podem ser vistos no Gráfico 8.

Um resultado importante foi que 96% dos artigos focaram, principalmente, em explicar a etapa de Análise dos Resultados. Com isso, pôde-se entender melhor como diversos



dados foram utilizados para gerar a persona.

Gráfico 8 – Etapas automáticas X Quantidade de estudos relacionados



Fonte: a autora.

Conforme observado na gráfico, três técnicas se destacam por aparecer em um maior número de artigos, sendo elas: APG, Scenario-Based Modeling e Data-Driven Persona.

A técnica APG foi apresentada em 11 artigos, sendo eles: Salminen *et al.* (2018b), Jung *et al.* (2018b), Jung *et al.* (2017b), Jung *et al.* (2018a), An *et al.* (2016), Salminen *et al.* (2017), Jansen *et al.* (2017a), Jansen *et al.* (2017b), Salminen *et al.* (2018a), Salminen *et al.* (2018d) e An *et al.* (2018). A técnica é composta pelas quatro etapas identificadas no Quadro 8. Todas as etapas realizadas nesse processo foram executadas de forma automática, desde a extração até a apresentação das personas.

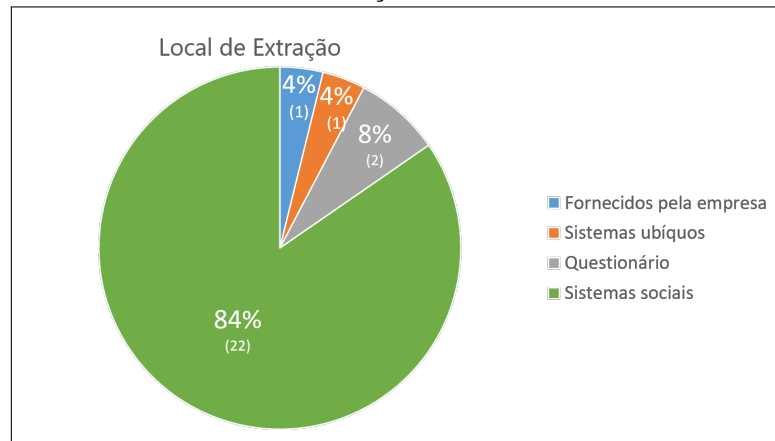
A técnica Scenario-Based Modeling foi apresentada em 3 artigos, sendo eles: Guo e Ma (2018), Guo e Ma (2017b) e Guo e Ma (2017a). A técnica é composta por duas fases automáticas, sendo elas a Análise dos Dados e a Geração.

A técnica Data-Driven Persona foi apresentada em 2 artigos, sendo eles: Zhang, Brown e Shankar (2016) e Mijač, Jadrić e Čukušić (2018). A técnica é composta por duas fases automáticas, sendo elas a Extração dos Dados e Análise.

#### 3.4.2.2 Quais os dados extraídos para a geração das personas?

O objetivo dessa questão de pesquisa foi entender quais dados são extraídos, e a partir dele, entender como as personas podem ser geradas. Para isso, é necessário entender de onde eles são coletados. O Gráfico 9 mostra os quatro locais de extração identificados nos estudos. A partir disso, é possível constatar um grande interesse em dados públicos dos usuários em SS, representando 84,6% do total.

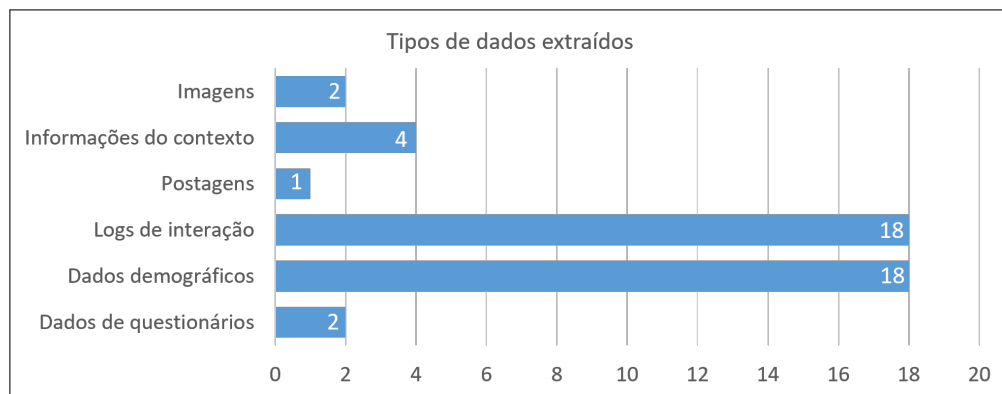
Gráfico 9 – Local de Extração dos dados.



Fonte: a autora.

Também foi possível identificar quais dados foram utilizados para a geração de personas, respondendo esta questão. Tais dados provêm de: dados demográficos dos usuários em SS, logs de interação, postagens em um fórum online, imagens de perfil, informações do contexto de uso e dados obtidos com questionários. O Gráfico 10 apresenta a quantidade de vezes em que esses dados foram utilizados nos estudos, sendo que alguns deles utilizaram mais de uma fonte. Portanto, o total encontrado no gráfico resulta em um número maior que o de artigos encontrados na revisão (25).

Gráfico 10 – Dados extraídos.



Fonte: a autora.

Grande parte dos estudos (80%) utilizou mais de um tipo de dado para gerar as personas. Em sua maioria (60%), foi realizada uma relação entre os logs de interação e dados demográficos. A Tabela 8 apresenta a relação desses dados, referente a cada estudo.

Dentre as técnicas que realizam esse tipo de relação, pode-se destacar a APG, pois em todos os seus estudos, realiza essa relação (dados demográficos e logs de interação) para gerar suas personas de forma automática.

Tabela 8 – Matriz de relação dos dados utilizados para geração das personas.

<b>Estudo</b>	<b>Dados demográficos</b>	<b>Logs de interação</b>	<b>Dados de questionários</b>	<b>Postagens</b>	<b>Informações do contexto</b>	<b>Imagens</b>
Salminen <i>et al.</i> (2018b)	X	X	-	-	-	-
Masiero, Ferreira e Jr (2012)	-	-	X	-	-	-
Jung <i>et al.</i> (2018b)	X	X	-	-	-	-
Jung <i>et al.</i> (2017b)	X	X	-	-	-	-
Terry <i>et al.</i> (2008)	X	X	-	-	X	-
Zhang, Brown e Shankar (2016)	-	X	-	-	-	-
Liu e Maes (2004)	-	-	-	X	-	X
An, Kwak e Jansen (2016)	X	X	-	-	-	-
Guo e Ma (2018)	X	-	-	-	X	-
Guo e Ma (2017b)	X	-	-	-	X	-
Jung <i>et al.</i> (2018a)	X	X	-	-	-	-
Guo e Ma (2017a)	X	-	-	-	X	-

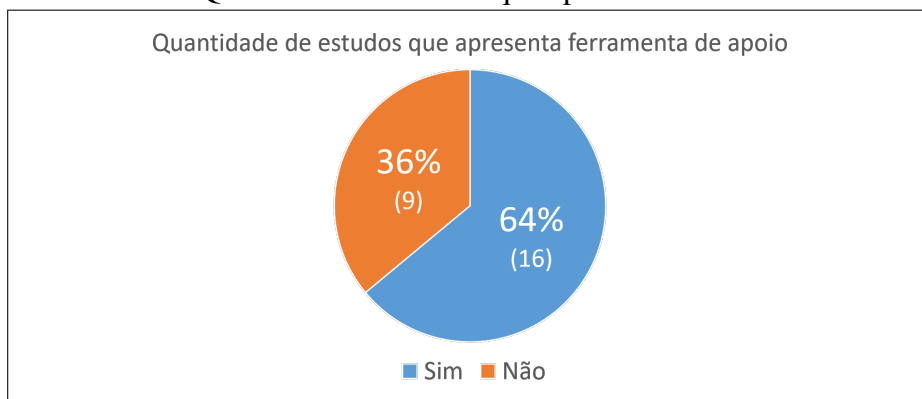
An <i>et al.</i> (2016)	X	X	-	-	-	-
Salminen <i>et al.</i> (2017)	X	X	-	-	-	-
Jansen <i>et al.</i> (2017a)	X	X	-	-	-	-
An, Kwak e Jansen (2017)	X	X	-	-	-	-
Jansen <i>et al.</i> (2017b)	X	X	-	-	-	-
Jung <i>et al.</i> (2017a)	-	-	-	-	-	X
Salminen <i>et al.</i> (2018a)	X	X	-	-	-	-
Salminen <i>et al.</i> (2018d)	X	X	-	-	-	-
An <i>et al.</i> (2018)	X	X	-	-	-	-
Watanabe <i>et al.</i> (2017)	-	X	X	-	-	-
Watanabe <i>et al.</i> (2018)	-	X	-	-	-	-
Salminen <i>et al.</i> (2018c)	-	-	-	-	-	-
Mijač, Jadrić e Ćukušić (2018)	X	X	-	-	-	-

Fonte: elaborado pela autora.

### 3.4.2.3 *Existem ferramentas que geram personas de forma automática?*

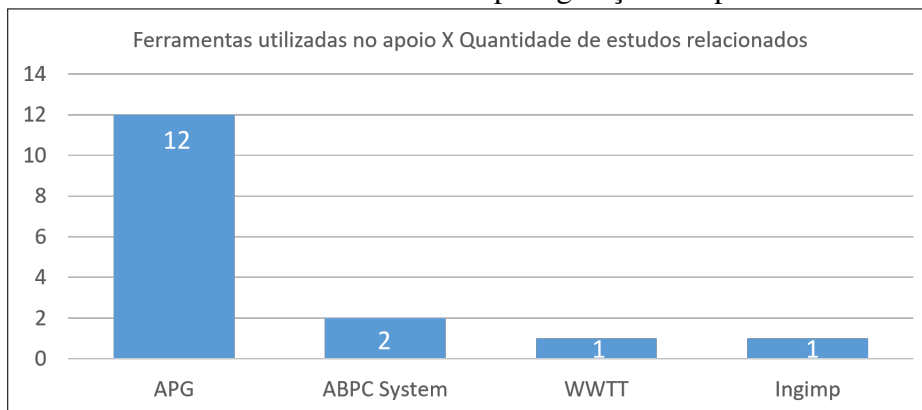
Durante a leitura completa dos artigos, foi identificado que 60,4% dos estudos apresentaram uma ferramenta para apoio da geração automática das personas, como apresentado no Gráfico 11. No total, foram encontradas quatro ferramentas que realizaram a extração, análise, geração e/ou apresentação das personas, sendo elas: APG, Ingimp, WWTT e ABPC System. O Gráfico 12 mostra a quantidade de vezes que a ferramenta foi utilizada nos estudos.

Gráfico 11 – Quantidade de estudos que apresentam ferramenta.



Fonte: a autora.

Gráfico 12 – Ferramentas utilizadas para geração das personas.



Fonte: a autora.

A fim de obter requisitos, foi realizada uma Arqueologia de Sistemas, uma técnica que extrai requisitos a partir da documentação ou implementação de um sistema legado ou concorrente (POHL; RUPP, 2012). As ferramentas e os requisitos encontrados, a partir delas, são detalhados a seguir.

A ferramenta APG <sup>3</sup>, assim como a técnica, teve o maior número de citações em

<sup>3</sup> Acessado em: 12 de novembro de 2018. Disponibilizado em: <<https://persona.qcri.org/persona/>>

artigos. Com sua análise, foi possível identificar diversos requisitos para a geração automática de personas. Esses requisitos foram:

- RF1 - A ferramenta deve dar suporte a todas as etapas da técnica;
- RF2 - A ferramenta deve apresentar um página contendo todas as informações geradas.
- RF3 - A ferramenta deve apresentar as informações das personas de forma a separar por conjunto de dados. Por exemplo, o perfil demográfico da persona deve estar em destaque e as demais informações, como tópicos de interesse, devem estar em um bloco separado.
- RF4 - Uma breve descrição sobre as personas deve ser apresentada, resumindo as informações contidas na tela.
- RF5 - Deve ser apresentada uma quantidade média de pessoas relacionadas aquela persona.
- RF6 - Antes de gerar as personas, deve ser questionado ao usuário da ferramenta, o número de personas a serem geradas.

As demais ferramentas não foram disponibilizadas. Portanto, não foi possível coletar informações sobre elas. A seguir, são apresentadas algumas informações coletadas a partir dos artigos.

A *Ingimp* é uma versão desktop da ferramenta *GNU Image Manipulation Program*, de código aberto. Essa ferramenta foi identificada como suporte para a criação das personas, pois ofereceu apoio apenas na etapa de extração dos dados, de forma automática, no estudo de Terry *et al.* (2008). Com ela, era realizada uma obtenção de logs de interação dos usuários da ferramenta.

A *WWTT*, é uma ferramenta desktop, apresentada no estudo de Liu e Maes (2004), que consiste em um painel de “consultores” na área de trabalho dos usuários. O sistema observa o usuário enquanto ele navega em páginas Web, escreve um texto ou responde um e-mail. Para a geração de personas, a ferramenta contribui com a extração de dados demográficos, logs de interação e, até mesmo, informações sobre o ambiente computacional do usuário.

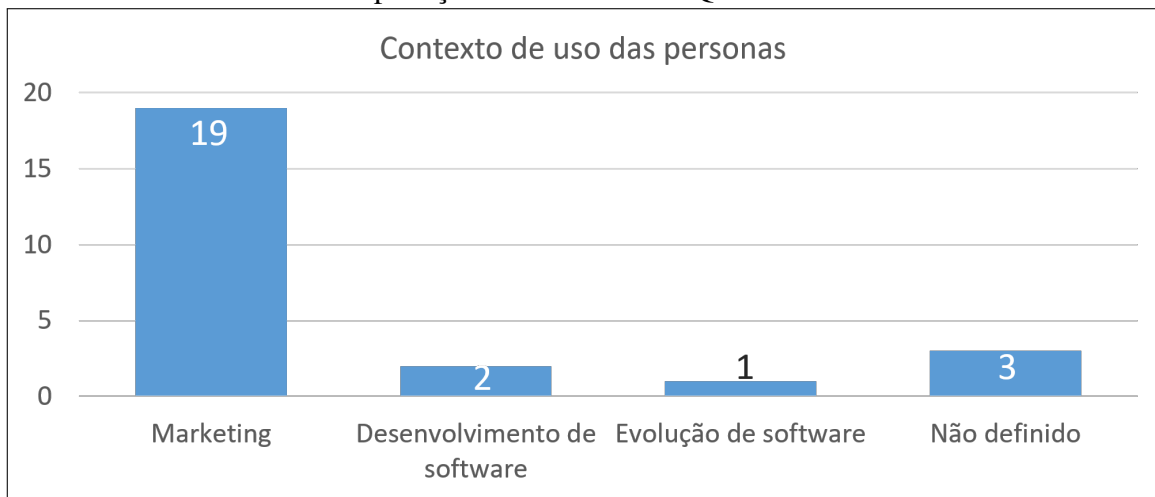
A *ABPC System*, é uma ferramenta que consiste na extração de dados dos usuários, apresentada nos estudos de Guo e Ma (2018) e Guo e Ma (2017a). Para isso, são coletados os dados demográficos e dados sobre o ambiente de uso, utilizando informações obtidas por sistemas IoT, como relógios digitais. A partir do processamento desses dados, a ferramenta

consegue aferir a personalidade das personas, a fim de enriquecer a descrição desses usuários.

#### 3.4.2.4 *Quais os contextos de aplicação das personas geradas de forma automática?*

Foram encontradas pesquisas que aplicam a geração de personas com diferentes objetivos, como auxiliar os empreendedores sociais aspirantes a conhecerem seu mercado e melhorar o entendimento dos usuários em empresas de desenvolvimento. Diante esses objetivos, foram identificados os contextos de aplicação dessas personas, para que pudesse ser entendido seu objetivo real. O Gráfico 13 mostra os contextos identificados e quantidade de estudos relacionados.

Gráfico 13 – Contexto de aplicação das Personas X Quantidade de estudos relacionados.



Fonte: a autora.

A partir do gráfico, é possível perceber que grande parte das personas são utilizadas para Marketing Digital. No que se difere do presente estudo, que visa utilizar essas personas para um contexto de avaliação de sistemas.

#### 3.4.2.5 *Existem técnicas para validar as personas geradas?*

Grande parte dos artigos realizou uma validação das personas, representando 66,7%, como apresentado no Gráfico 14. Os que validaram, utilizaram diversos tipos de técnicas. O Gráfico 15 mostra a quantidade de vezes que a técnica de validação foi utilizada nos estudos.

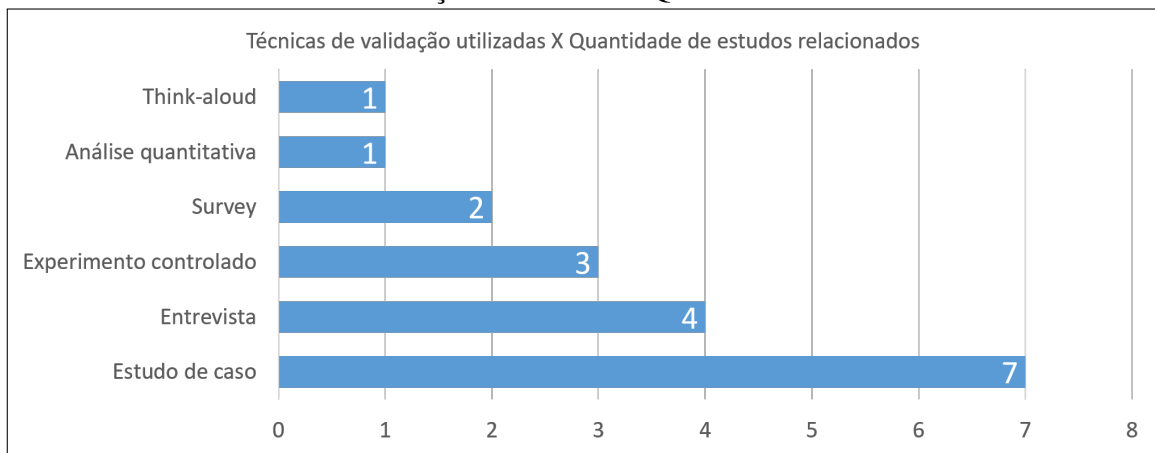
O Gráfico 16 mostra a quantidade de artigos que validaram as personas, sob duas perspectivas: quanto aos dados e quanto ao uso. Uma grande maioria (62,50%) dos estudos se preocuparam em verificar a validade da utilização dessas personas no contexto de aplicação.

Gráfico 14 – Número de artigos que realizaram validação das personas.



Fonte: a autora.

Gráfico 15 – Técnicas de validação utilizadas X Quantidade de estudos relacionados.

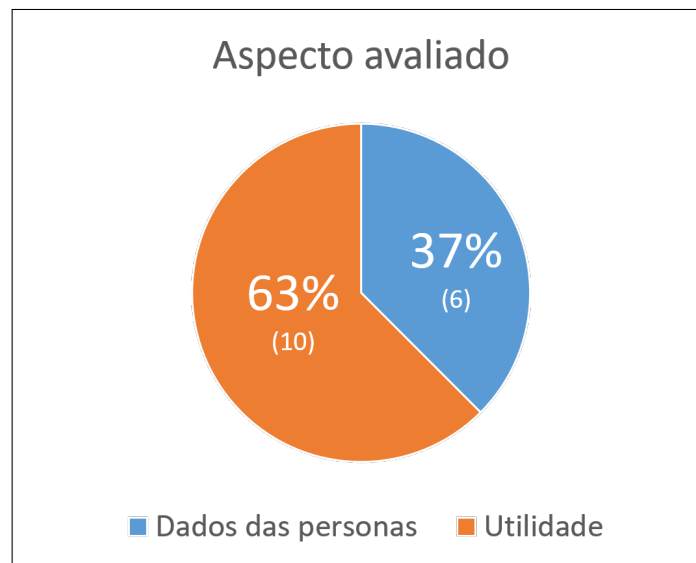


Fonte: a autora.

Para isso, foram utilizadas técnicas como estudo de caso, aplicação de questionários, entrevistas com especialistas e experimentos controlados. Os demais (37,5%) buscaram verificar a validade dos dados das personas, por exemplo, se as informações exibidas na tela da ferramenta eram importantes para a apresentação das personas. Para isso, foram realizadas análises desses dados e entrevistas com especialistas. Vale ressaltar que os dados apresentados nesse gráfico, é referente aos 66,7% de estudos que realizaram algum tipo de validação.



Gráfico 16 – Aspectos avaliados na validação das personas.



Fonte: a autora.

### 3.5 Conclusão

Este capítulo descreveu um mapeamento sistemático sobre técnicas de geração automática de personas que apresenta os estudos mais relevantes encontrados de forma descritiva e quantitativa. Foi observado que dados públicos de usuários em SS vem sendo utilizados e que há um crescimento nesse tipo de técnicas. Porém, não foi encontrada aplicação dessas personas em avaliação de SS. Portanto, os estudos encontrados servirão de base para os próximos passos a serem realizados nesta pesquisa.

#### 4 TRABALHOS RELACIONADOS

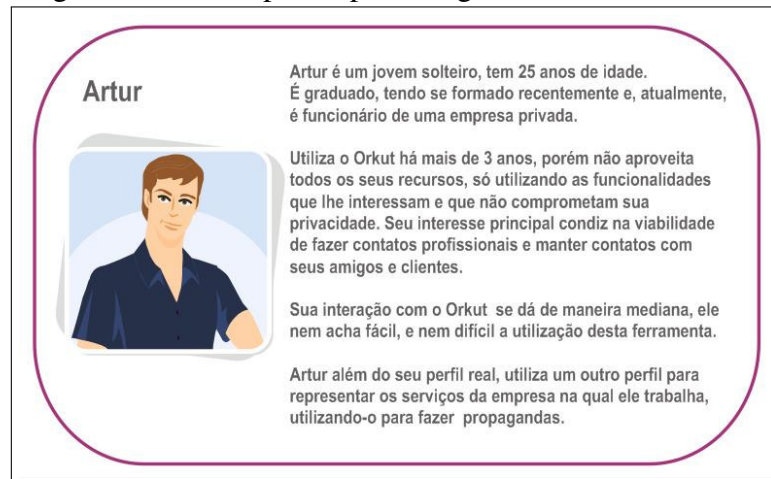
Conforme mencionado anteriormente, o objetivo deste trabalho é definir uma metodologia de criação de personas a partir dos dados públicos dos usuários em SS. Neste sentido, apresentam-se aqui trabalhos que apresentaram alguma metodologia ou modelo para criação de personas. Tais trabalhos, foram obtidos com a execução do mapeamento sistemático e com pesquisas manuais.

Holanda (2010) propôs o Framework Personas, aplicando-o em um cenário de Elicitação de Requisitos com o intuito de transmitir a representatividade do público alvo para a equipe de desenvolvimento e auxiliar na avaliação de interações em SS. Para o processo de construção das personas, foram propostas duas abordagens, uma para aplicação no desenvolvimento de qualquer produto, Metodologia para Elicitação de Requisitos (SER), e outra que considera critérios sociais, tipo de contexto e interatividade apresentada pelos usuários e o sistema, para a criação da persona, Metodologia para Avaliação de Sistemas (MASC). A metodologia SER tem o fluxo de atividades dividido em 6 (seis) etapas, sendo elas: 1) conhecer os usuários; 2) definir personas; 3) construir cenários; 4) modelar storyboards; 5) validar personas, cenários e storyboard; e 6) comunicar e acompanhar a equipe. Para realizar tais etapas, os dados são extraídos com a aplicação de técnicas de elicitação de requisitos, como: questionários, entrevistas, cenários, grupo focal e prototipagem. Um exemplo de persona criada com a PATHY pode ser vista na Figura 1, apresentada na seção Personas.

A MASC (HOLANDA, 2010) é uma atualização da metodologia SER que elenca critérios investigativos de usabilidade de um sistema, para isso foi adicionada uma nova característica: a Pesquisa de Mercado. Seu fluxo de atividade consiste na realização de 5 (cinco) etapas, sendo elas: 1) investigação contextual; 2) análise quantitativa; 3) elaboração das personas; 4) análise qualitativa do processo; e 5) integração e quantificação dos resultados obtidos. Os dados foram coletados da mesma forma que os da metodologia SER com o adicional de uma investigação de aspectos da usabilidade de um sistema, no qual os especialistas em IHC fizeram suposições sobre o comportamento dos usuários diante dos problemas encontrados. A Figura 17 apresenta um exemplo de persona gerada com a MASC.

Outra forma de gerar personas é utilizando o Framework Persona Empresa (FMK@PE), proposto por NÓBREGA e Furtado (2011). Trata-se de uma evolução do Framework Persona (HOLANDA, 2010) com aplicação direta no contexto de empresas que precisam de uma representação de seus usuários, cujas informações possam agregar valor de negócio se aplicada em uma

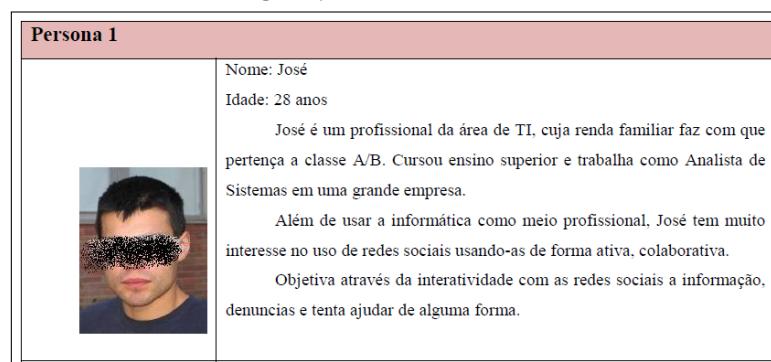
Figura 17 – Exemplo de persona gerada com a MASC.



Fonte: Holanda (2010)

rede social. Para a criação das personas são utilizadas as características definidas na metodologia MASC, apresentada anteriormente, retirando a característica “perfil psicográfico” e substituindo o “tipo de usuário” por “tipo de usuário (empresa)” quando a persona for direcionada a empresa. As personas são criadas com base na interpretação dos dados realizada por profissionais de IHC ao utilizar um template como guia. Esses dados podem ser obtidos a partir de logs de interação ou observação da interação, questionários e pesquisa de mercado. No estudo, os dados foram extraídos do perfil da empresa na rede social Twitter, tendo como exemplo o número de seguidores e mensagens enviadas para a empresa. Para criar a persona, a metodologia deve ser realizada em um fluxo de 5 (cinco) atividades: 1) categorizar usuários; 2) subcategorizar usuários; 3) elaborar esqueleto persona; 4) criar a história e, por fim, 5) criar a persona. A Figura 18 apresenta um exemplo de persona indivíduo e a Figura 19, uma persona empresa.


Figura 18 – Exemplo de Persona Indivíduo gerada com o FMK@PE.



Fonte: NÓBREGA e Furtado (2011)

Guimarães (2012) buscou entender o cenário brasileiro de uso de métodos e técnicas

Figura 19 – Exemplo de Persona Empresa gerada com o FMK@PE.

Persona 2 – EMPRESA	
	<p>Nome: ECE - Empresa de Comunicações EXE</p> <p>Tempo no mercado: 12 anos</p> <p>ECE é uma empresa que atua no meio jornalístico do estado do Ceará há 12 anos, possui certificações de qualidade ISO9001 e sua renda bruta anual chega a 15 milhões, classificando-a como uma grande empresa.</p> <p>Possui equipamentos de ponta e profissionais qualificados com vários meios de acesso a internet, e tem grande integração com redes sociais.</p> <p>Objetiva através da interatividade com as redes sociais a informação jornalística, denúncias, além de suporte e contato mais próximo com seus clientes.</p>

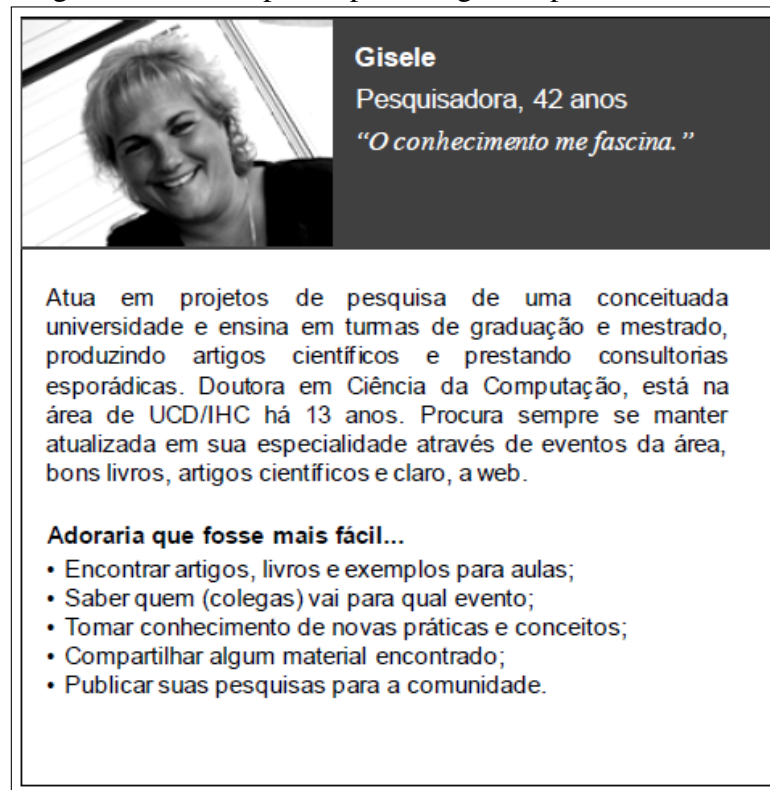
Fonte: NÓBREGA e Furtado (2011)

de IHC, para isso seu estudo modelou os interessados pela área e os conceitos conhecidos por eles utilizando a técnica de Personas. Para obter os dados para a criação das personas, foi aplicado um questionário focado em três subáreas, sendo elas: perfil do profissional (faixa etária, grau de instrução e formação acadêmica, etc.), organização em que o profissional trabalha e seu comportamento quanto seu conhecimento sobre a área de IHC. Como resultados foram obtidas 113 respostas de pessoas localizadas em todas (cinco) as regiões do Brasil. Para a construção das personas foi utilizado o Framework Personas já descrito anteriormente. Sua elaboração consistiu na execução de 5 (cinco) etapas: 1) coleta dados demográficos e psicográficos; 2) segmentar e classificar os principais perfis; 3) definir atributos relevantes; 4) definir objetivos e necessidades; e 5) montar o cartão para cada Persona. Essas personas auxiliaram no reconhecimento da comunidade brasileira de IHC, pela representação de seis personas. A Figura 20 apresenta um exemplo de persona gerada no estudo.

Também foram encontradas pesquisas que aplicam a geração de personas com diferentes objetivos, como a necessidade de promover serviços de saúde para crianças vulneráveis (WÄRNESTÅL; SVEDBERG; NYGREN, 2014), identificação de requisitos de segurança (SANTOS; RABELO; BARRETO; CONTE, 2014), auxiliar os empreendedores sociais aspirantes a conhecerem seu mercado (ORTBAL; FRAZZETTE; MEHTA, 2016), estudo sobre o design elétrico de casas (DODGE *et al.*, 2017) e para auxiliar o serviço de *coach* para pessoas obesas (JANSEN; MECHELEN; SLEGGERS, 2017).

No entanto, os trabalhos apresentados propõem a geração de personas a partir da coleta de dados dos usuários, onde especialista deve utilizar uma técnica de coleta de dados como: questionário, entrevista ou pesquisa em dados demográficos. Como o foco deste trabalho é a construção de personas de forma automática, apresentam-se a seguir, trabalhos que geram

Figura 20 – Exemplo de persona gerada por Guimarães.



Fonte: Guimarães (2012)

automaticamente personas a partir de dados extraídos dos usuários, sendo eles o estudo de Lima (2016) e os que foram obtidos com a realização do mapeamento sistemático.

Lima *et al.* (2016) apresentou, em seu estudo, uma estrutura sobre a mineração de dados de mídias sociais, explorando sua aplicabilidade para a identificação dos principais fatores psicológicos dos usuários. Para a geração da persona são utilizados dados sociais, que são: “Todos aqueles gerados por usuários de serviços web que, de alguma maneira, são compartilhados na rede, por exemplo, textos, imagens, vídeos, etc., independentemente de estrutura ou plataforma do sistema”.

Esses dados foram minerados utilizando as seguintes etapas: 1) etapa de aquisição de dados sociais; 2) etapa de pré-processamento; 3) etapa de análise; 4) etapa de validação; e 5) pré-processamento de texto. Com base nesses dados, três frameworks computacionais foram executados para construir os componentes da tríade da Persona Virtual, com base na predição da personalidade e temperamento e análise de sentimentos. A união desses componentes gera a Persona Virtual, imagem dos usuários formadas a partir do seu comportamento em mídias sociais. Cada predição pode ser definida, segundo Lima *et al.* (2016), como:

- Personalidade: consiste em determinar traços de personalidade de um indivíduo utilizando seus dados sociais, como textos, fotos e vídeos, para definir seu

comportamento;

- Sentimento: consiste em investigar as emoções e sentimentos sentidos pelo usuário analisando a forma ele como se expressa textualmente;
- Temperamento: consiste em inferir disposições do comportamento dos usuários associadas a determinantes biológicos e hereditários, sendo necessário utilizar técnicas computacionais para identificar os diferentes tipos de temperamento de *Myers-Briggs*, por exemplo.

Masiero, Ferreira e Jr (2012) apresentam um algoritmo para geração de personas, utilizando dados obtidos por meio da aplicação manual de questionários. A técnica, consiste no recebimento dos dados coletados e em sua normalização, utilizando a equação de normalização. Em seguida, um algoritmo de clusterização em Python é aplicado. Ao ser executado, o algoritmo retorna grupos de usuários. Por fim, uma combinação dos valores das variáveis é realizada, a fim de encontrar um valor em comum entre elas. O valor comum caracteriza uma persona, ou seja, um grupo de usuários.

Tal estudo se diferencia do presente trabalho por não utilizar dados demográficos dos usuários em SS, ou seja, as informações já cadastradas em um SS. Porém, artigos que apresentam as demais técnicas encontradas no mapeamento, como An, Kwak e Jansen (2016), Watanabe *et al.* (2018), Zhang, Brown e Shankar (2016) utilizam esses dados.

Vale ressaltar os artigos que apresentaram a aplicação da técnica APG, já apresentada na seção de Geração Automática de Personas deste estudo, sendo eles: An, Kwak e Jansen (2016), An *et al.* (2016), An, Kwak e Jansen (2017), An *et al.* (2018), Jansen, Mechelen e Slegers (2017), Jansen *et al.* (2017b), Jansen *et al.* (2017a), Salminen *et al.* (2018b), Salminen *et al.* (2017). Essa técnica utiliza os dados demográficos para a geração da persona e um exemplo pode ser visto na Figura 21.

Esses trabalhos apresentam uma geração automática de personas e se diferenciam do presente estudo quanto ao objetivo da criação da persona. Enquanto as personas geradas visam auxiliar na área de marketing, este estudo pretende auxiliar no processo de conhecimento do usuário de um sistema para auxiliar no entendimento dos resultados de avaliações textuais de sistemas.

Figura 21 – Exemplo de persona gerada com a técnica APG.

The screenshot displays a user profile interface for a persona named Ashley. The profile includes a header with navigation links (APG, About APG, Submit Story, View Persona) and a user ID (AJE\_test\_1). Below the header is a row of profile pictures. The main profile section features a large profile picture of a young woman with glasses and a bio that states she is an 18-year-old female living in the United States, working in management, and interested in racism, human-interest stories, and the environment. The 'Topics of Interest' section lists 'More Interested Topics' (Racism, Human-interest story, Environment) and 'Less Interested Topics' (Religion, International Affairs, Entertainment). A 'Most Viewed Videos' section on the right lists several Al Jazeera videos, including 'My Disability Won't Stop Me from Being a Mom', 'Britain's Modern Slave Trade', 'Too Young to Wed. Child Marriage in Bangladesh', 'China's Rich Girls', 'Malaysia: Babies For Sale', 'India's Slave Brides', 'Names of Manchester Arena attack victims begin to emerge', and 'China's Fake Boyfriends'.

Fonte: Jung *et al.* (2018b)

## 5 UMA PROPOSTA DE GERAÇÃO AUTOMÁTICA DE PERSONAS A PARTIR DOS DADOS PÚBLICOS DOS USUÁRIOS EM SISTEMAS SOCIAIS

Os resultados do mapeamento sistemático (Seção 3.4.2) apresentam um crescimento do interesse na geração automática de personas, a partir de dados públicos. Outras informações também foram identificadas, norteadas a construção desta proposta. Uma delas se encontra na identificação dos dados extraídos e analisados. Foi percebido que os dados utilizados para as personas, eram, principalmente, os dados demográficos (idade, gênero e localização), mas que tinham outros dados que estavam relacionados com o contexto de aplicação da persona. Por exemplo, para os estudos voltados para a área de marketing digital, em sua grande maioria, foram utilizados logs de interação dos usuários, com o conteúdo estudado.

Para auxiliar no processo de entendimento da proposta, quanto a extração de dados, foi realizado um estudo da API do *Twitter*<sup>1</sup>. Com isso, foi possível identificar quais dados podem ser extraídos.

Com base nessas informações, neste capítulo é apresentada uma metodologia de criação automática de personas, utilizando dados públicos dos usuários em SS, no contexto de avaliação textual.

### 5.1 Estudo da API do Twitter

O *Twitter*<sup>2</sup> é um SS, utilizado para compartilhamento de informações, como notícias, entretenimento e situações do cotidiano. Nele, é possível criar *tweets* e, até mesmo, *retweetar* (compartilhar um *tweet*). Para utilizar seus recursos, é necessário fazer um cadastro, disponibilizando algumas informações demográficas. Tais informações, podem ser extraídas, se estiverem no modo público.

Para extrair dados do *Twitter*, a empresa dispõe de uma API de extração. Essa API foi escolhida para ser estudada, pois sua configuração já havia sido realizada na ferramenta UUX-Posts. O principal objetivo deste estudo, foi entender quais dados podem ser utilizados para a geração de personas.

Durante o estudo, foi identificado que os dados extraídos podem ser divididos entre: dados das postagens e dados dos autores das postagens. Para os dados da postagem, podem ser coletados:

<sup>1</sup> API do Twitter. Disponível em: <<https://developer.twitter.com/en/docs.html>>. Acessado em: 10 de novembro de 2018.

<sup>2</sup> Twitter. Disponível em: <<http://twitter.com/>>. Acessado em: 10 de outubro de 2018.



- postagem;
- id da postagem;
- data de criação da postagem;
- dispositivo utilizado para realização da postagem;
- número de retweets;
- número de curtidas da postagem;
- número de hashtags mencionadas na postagem;
- número de usuários mencionados na postagem.

Para os dados relacionados aos usuários, podem ser coletados:

- id;
- nome;
- username;
- localização;
- descrição;
- número de pessoas que seguem o usuário (seguidores);
- número de pessoas que o usuário segue (seguindo);
- data de criação da conta;
- número de tweets já publicados;
- número de postagens favoritadas;
- idioma;
- imagem do Perfil.

Com a realização deste estudo, foi possível entender e definir uma proposta de utilização desses dados para a construção da persona, apresentada nas próximas seções.

## 5.2 Metodologia proposta

Como resultado do mapeamento sistemático, quanto às técnicas de geração automática de personas, apresentado na Seção 3.4.2.1, foi percebido que são realizadas quatro etapas, sendo elas: (1) Extração dos Dados; (2) Análise dos Dados; (3) Geração das Personas; e (4) Apresentação das Personas, detalhadas a seguir. Esse fluxo pode ser visto na Figura 22.

**1) Extração dos dados:** consiste em duas atividades, sendo elas: (i) configuração da API; e (ii) obtenção dos dados. Inicialmente, deve ser realizada uma configuração da API do SS a ser utilizado.

Figura 22 – Fluxo de etapas da metodologia.



Fonte: a autora.

Como a pesquisa está relacionada ao contexto de avaliação textual de sistemas por meio de PRUs, a atividade de obtenção dos dados consiste na captura dos dados dos usuários, relacionados a essas postagens, ou seja, extração dos dados fornecidos pelos autores das PRUs. Exemplos de dados que podem ser extraídos foram apresentados na seção anterior (Estudo da API do Twitter). Dentro do contexto da MALTU, essa etapa será realizada juntamente com a Extração das PRUs (Etapa 2).

**2) Análise dos dados:** nessa etapa, é realizada uma análise, para que os dados sejam úteis para as personas. Levando em consideração o estudo realizado na API do Twitter e os possíveis dados a serem obtidos, algumas características para as personas foram identificadas como informações úteis, sendo elas: tempo de uso, dados demográficos (idade, gênero e localização ou nacionalidade) e popularidade da postagem. A seguir, essas informações são detalhadas quanto sua validade e possível geração.

Informações demográficas, como idade, gênero e localização, foram encontradas em todas as personas geradas nos estudos identificados no mapeamento sistemático. Apenas a localização pode ser extraída diretamente da API do Twitter. Portanto, para adquirir idade e gênero, propõe-se a utilização da API Face++<sup>3</sup>, de forma similar ao estudo de Jung *et al.* (2017a). Essa API permite que tais dados, sejam obtidos por uma análise da imagem de perfil do usuário, sendo por meio de seu *upload* ou de uma URL onde o arquivo se encontra. Essa opção se tornou viável para esta proposta, pois é possível extrair a informação da URL da foto de perfil dos usuários.

Algumas informações podem ser adicionadas a persona, para deixá-la mais real, aproximando, ainda mais, os avaliadores de seus usuários. Tais informações podem ser relacionadas ao contexto ou ao próprio usuário. Para o contexto, pode ser realizada a identificação do

<sup>3</sup> Acessado em: 15 de novembro de 2018. Disponível em <<https://www.faceplusplus.com/>>.

dispositivo utilizado durante o uso, podendo ser extraído diretamente da API. Quanto ao usuário, a personalidade e o tempo de uso podem ser inseridos no contexto da persona.

O tempo de uso é uma importante característica para a persona. O nível de experiência do usuário com a aplicação, pode influenciar em seu uso. Com isso, os problemas relatados podem ser relacionados com o tempo de uso. Além disso, problemas de aprendizado podem ser identificados. Por exemplo, usuários que estão há 8 anos utilizando o sistema, mas relatam que não conseguem alterar sua foto de perfil. Para encontrar o tempo de uso ( $T$ ) dos usuários, dois dados serão utilizados, sendo eles: data de criação das postagens ( $Dp$ ) e data de criação da conta do usuário ( $Dc$ ). Para gerar essa informação, pode ser realizada uma subtração da  $Dp$ , com a  $Dc$  que retornará a quantidade de dias de uso, como retrata a equação:

$$T = Dp - Dc, \quad (5.1)$$

Outra informação a ser analisada é a popularidade da postagem que auxiliará em sua priorização, sendo utilizada na fase de apresentação das personas. Quanto a popularidade do usuário, não poderá ser identificada, pois temos apenas uma postagem sobre o usuário. Portanto, não conseguimos inferir sua popularidade.

Para encontrar a popularidade da postagem ( $P$ ), é proposta uma adaptação da fórmula apresentada por Leitao, Morgado e Cunha (2012). Essa fórmula utiliza três dados sobre as postagens, sendo eles: número de curtidas, número de compartilhamentos (retweets) e número de comentários (respostas). Porém, a API do Twitter não permite acesso a quantidade de comentários, e por isso, este último dado será desconsiderado, neste primeiro momento.

A fórmula proposta consistirá em um soma do número de curtidas ( $Nc$ ) e do número de compartilhamentos ( $Nr$ ). Para calcular a popularidade de um conteúdo, são definidos pesos para cada variável, baseados em sua importância ( $Wr$  e  $Wc$ ). A soma desses pesos deve ser igual a 1. Como não há como provar, neste momento, qual variável é mais importante, serão utilizados pesos 0,5 para ambas as variáveis. Portanto, a fórmula a ser utilizada será a seguinte:

$$P = (Nc * Wc) + (Nr * Wr), \quad (5.2)$$

**3) Geração das personas:** nessa etapa, serão utilizados os dados dos usuários analisados, como proposto na seção anterior, e os resultados de uma avaliação textual, como proposta na Extração dos Dados. Para isso, devem ser realizadas três atividades, sendo elas: (i) gerar grupos de usuários; (ii) definir persona shell; (iii) e relacionar grupos de usuários com

resultados da avaliação textual. A definição de tais atividades foi baseada na técnica APG (JUNG *et al.*, 2018b), descrita na Seção 2.1.2, e são detalhadas a seguir.

(i) Gerar grupos de usuários: consiste em uma análise das informações obtidas na etapa anterior, a fim de identificar padrões de usuários. Essa análise pode ser realizada de forma manual ou automática, onde devem ser identificados os usuários mais parecidos com os outros, agrupando-os com base no número de personas definido pelo avaliador do sistema. No final desta atividade, a persona já terá informações como: idade, gênero, localização, tempo de uso e dispositivo utilizado.

(ii) Definir persona shell: com os grupos identificados, algumas informações devem ser associadas as personas, com o intuito de deixá-las mais reais. Para isso, deve ser encontrado um nome e uma imagem para a persona, criando a persona shell.

(iii) Relacionar grupos de usuários com resultados da avaliação textual: essa terceira etapa da geração é responsável por deixar a persona útil, no contexto de avaliação textual, pois associa informações dos usuários com os resultados de uma avaliação textual usando a metodologia MALTU. Como por exemplo, os tipos de classificação: por funcionalidade e por tipos (elogio, crítica, sugestão, dúvida, comparação e ajuda). O agrupamento dessas informações pode ocorrer da seguinte forma: para cada usuário (U), contido em grupo (Pe), tem-se uma postagem (Po) que foi classificada (Cl) durante uma avaliação. Essa classificação dispõe de duas informações, sendo a Funcionalidade (Fun) e o tipo (Ti). Para cada tipo, será realizada uma análise de quais as funcionalidades mais citadas, de acordo com um grupo de usuários. Assim, são obtidos:

- Funcionalidades mais criticadas: aquelas que estão relacionadas a um usuário que pertence a um grupo, e que a PRU foi classificada como Crítica;
- Funcionalidades preferidas: aquelas que estão relacionadas a um usuário que pertence a um grupo, e que a PRU foi classificada como Elogio;
- Funcionalidades que causaram dúvidas: aquelas que estão relacionadas a um usuário que pertence a um grupo, e que a PRU foi classificada como Dúvida;
- Sugestões de funcionalidades: aquelas que estão relacionadas a um usuário que pertence a um grupo, e que a PRU foi classificada como Sugestão.

**4) Apresentação das personas:** com a geração das personas realizadas, devem ser representados seus resultados. Como a proposta das personas apresentada neste trabalho tem o foco na avaliação do sistema, elas devem ser apresentadas juntamente ao relatório final, emitido

pela ferramenta UUX-Posts.

A apresentação das personas será dividida em cinco blocos, sendo eles: foto para persona, número de usuários relacionados, postagens mais populares, dados da persona e relacionamento com o sistema avaliado. As informações sobre as personas, o relacionamento com o sistema avaliado, sua foto e o número de usuários relacionados, ou seja, o número de usuários que foram agrupados para a geração da persona, são resultados da etapa de geração (etapa 3). Já as postagens mais populares, consiste na exibição de postagens com os maiores valores obtidos, quanto sua popularidade. O cálculo realizado para definir as postagens mais populares, foi definido e explicado na Análise dos Dados.

### 5.3 Discussão

Nesta seção, apresenta-se uma discussão sobre a metodologia definida, suas limitações e formas de melhoria.

Um ponto positivo da metodologia é a realização da relação entre os dados dos usuários, com os resultados obtidos na avaliação textual. Com isso, há uma nova forma de apresentação dos resultados da avaliação, associando diretamente aos usuários. Acredita-se que tal forma de visualização dos resultados facilite o seu entendimento, auxiliando o avaliador a categorizar os problemas obtidos em prioridades, de acordo com o usuário representado na persona.

Novas informações podem ser adicionadas a persona, a fim de deixá-las mais reais. Por exemplo, podem ser adicionadas características da personalidade do usuário. Para isso, alguns estudos apresentam a geração da personalidade por meio da técnica *Big-Five Personality* (GOLDBERG, 1990). Nela, é afirmado que é possível inferir a personalidade dos usuários, com base em seus textos. Portanto, podem ser utilizadas as PRUs.

Mesmo com a metodologia de geração realizada, ainda não foi estabelecida uma forma automática para executá-la. Para isso, alguns algoritmos de clusterização podem ser utilizados. Por exemplo, em Terry *et al.* (2008) são utilizados dois algoritmos, o *K-means* e o *Hand-Coded*, com eles, são encontrados padrões de grupos e as personas são geradas. Da mesma forma, pode ser aplicado nesta proposta.

## 5.4 Conclusão

Este capítulo apresentou uma metodologia para geração automática de personas, no contexto de avaliação textual de SS. A metodologia apresentada, tem como base as informações encontradas no mapeamento sistemático, juntamente a um estudo da API do Twitter.

Tal metodologia foi descrita, explanando o passo-a-passo das etapas para a criação automática das personas. Para isso, foi realizada uma adaptação da técnica APG (JUNG *et al.*, 2018b). Essa adaptação foi realizada pois não foram encontradas técnicas que gerem personas no contexto de avaliação. As principais alterações encontram-se na etapa de análise dos dados e apresentação das personas, pois foram levadas em consideração características mais específicas do usuário e sua relação com a avaliação textual, e não de sua interação com conteúdos online.

## **6 VALIDAÇÃO DA METODOLOGIA A PARTIR DE UM EXPERIMENTO NO TWITTER**

Neste capítulo é apresentado um experimento realizado no SS Twitter. O objetivo deste experimento foi validar a metodologia para geração de personas apresentada no Capítulo 5. Este experimento foi executado entre os dias 10 e 25 de novembro, tendo duração de 15 (quinze) dias.

Todas as etapas da metodologia proposta foram executadas neste experimento, sendo de forma automática, apenas a primeira etapa, Extração dos Dados. Para isso foi utilizada a API do Twitter. As demais etapas foram realizadas de forma manual, a fim de entender melhor todo o processo, onde os dados obtidos foram analisados quanto sua validade e forma de modelagem. A execução das etapas da metodologia são apresentadas nas seções a seguir, juntamente com um exemplo de persona gerada e uma discussão.

### **6.1 Extração dos dados**

Para a etapa de Extração dos Dados, foi utilizada a configuração da API do Twitter, já realizada na ferramenta UUX-Posts e apresentada na seção 2.2.2. Essa ferramenta fazia a captura, apenas, do dado "postagem" do usuário. Para salvar as demais informações, foi adicionado um código para capturar os dados propostos na metodologia (Capítulo 5). Esse código foi criado em parceria com um estagiário do Projeto MALTU.

A API permitiu extrair 1500 postagens e, juntamente a elas, informações extras sobre a postagem e dados sobre os usuários (autores das postagens). Um exemplo de retorno da API pode ser visto no Anexo B.

### **6.2 Análise dos dados**

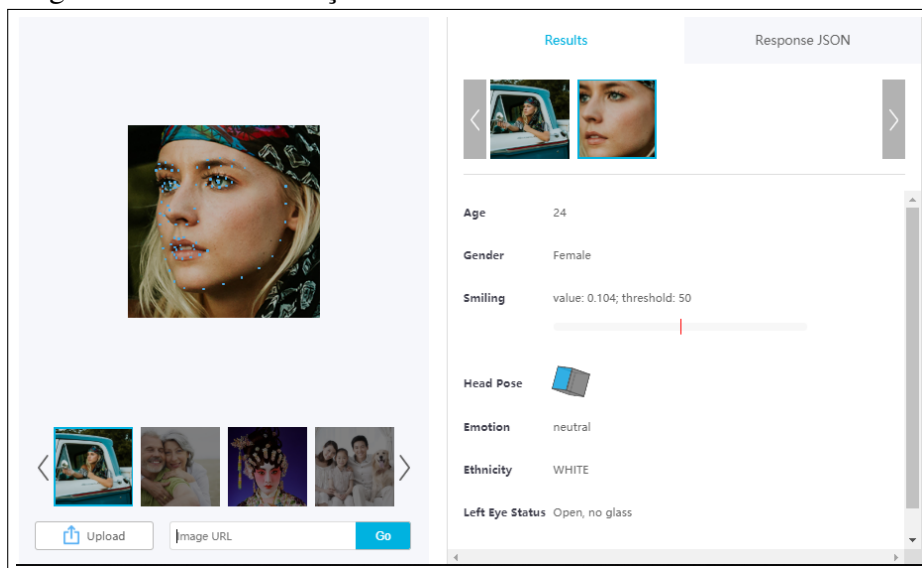
Para a etapa de Análise dos Dados, foi utilizada uma amostra de 200 postagens, referente a 13,33% da população total (1500 postagens). Com isso, foi possível realizar a execução da análise proposta, gerando as seguintes informações: (i) dispositivo; (ii) localização; (iii) popularidade; (iv) idade; (v) gênero; e (vi) tempo de uso.

Os dados utilizados para gerar as informações: dispositivo, popularidade e tempo de uso, são inseridos no perfil do usuário pelo próprio SS (*Twitter*). Portanto, há uma garantia que esses dados são reais. Porém, a localização é inserida pelo próprio usuário, onde nem sempre

são colocadas informações reais.

A idade e gênero foram dados obtidos por meio de uma análise da foto de perfil. Para isso, foi utilizada uma demo da API Face++<sup>1</sup>. As URLs das fotos foram inseridas no campo "Image URL" e quando executado, a ferramenta exibia suas informações geradas. Um exemplo pode ser visto na Figura 23. Porém, assim como a localização, tais informações são inseridas pelos usuários que nem sempre utilizam fotos reais. Conseqüentemente, tais informações não foram obtidas para todos os usuários.

Figura 23 – Demonstração da Ferramenta Face++.



Fonte: Face++. Acessado em: 15 de novembro de 2018. Disponível em <<https://www.faceplusplus.com/attributes/#demo>>.

Para a extração dos dados relacionados a avaliação textual, as PRUs foram classificadas como proposto pela metodologia MALTU. Inicialmente, as postagens foram classificadas como PRUs ou não PRUs, pois Mendes e Furtado (2017) afirmam que a extração realizada, utilizando o modelo booleano, não é tão precisa. Em seguida, as postagens foram categorizadas quanto ao seu tipo e a funcionalidade citada.

A seguir, é realizada uma análise dos dados demográficos obtidos, relacionados à sua validade.

<sup>1</sup> Acessado em: 15 de novembro de 2018. Disponível em <<https://www.faceplusplus.com/attributes/#demo>>.



### 6.2.1 *Análise dos dados demográficos*

Em relação aos dados de localização, das 200 postagens analisadas, apenas 81 são referentes às cidades existentes. Todas as localizações foram buscadas no *Google Maps*<sup>2</sup>, caso não houvesse retorno, a cidade era considerada como inexistente. Para um maior detalhamento desses casos, eles foram categorizados em: descrição do usuário, por exemplo, "*passando vergonha e divertindo*"; cidade fictícia (cidades de filmes e livros), por exemplo, "*Nárnia*" e "*Asgard*"; ou não declarado, quando o usuário não inseriu essa informação em sua biografia. Como resultados, foram obtidas 21 localizações fictícias, 33 classificadas como descrição e 64 não declaradas.

Com a análise, foi possível perceber que a localização extraída não se refere à localização onde a postagem foi publicada, mas sim, da nacionalidade do usuário. Além disso, foi possível perceber que os usuários tendem a inserir sua localização seguindo os seguintes formatos: "Cidade"(9); "Cidade, Estado"(18); "Cidade, Estado, País"(1); "Cidade, País"(32); "Estado, País"(6); e "País"(15).

Quanto aos dados obtidos com o Face++, 117 usuários conseguiram ser identificados, ou seja, 58,5% dos usuários tiveram seu gênero e idade detectados. Dentre os 83 usuários não identificados em suas fotos de perfil, foram encontrados três erros, sendo eles: a URL obtida com a API do Twitter não era válida (31,3%), a face não foi detectada (50,6%) e as fotos retratavam pessoas famosas (18,1%). Foram identificados alguns problemas que acarretaram no não reconhecimento da face, sendo eles: as fotos se tratavam de memes, personagens desenhos animados, fotos com mais de uma pessoa ou fotos em que o celular cobria o rosto do usuário.

### 6.3 **Geração das personas**

Para gerar as personas (3ª etapa da metodologia) foram utilizados os dados que seguiram os seguintes critérios: (i) todos os dados dos usuários deveriam ser válidos; e (ii) apenas dados relacionados às postagens e autores de postagens, consideradas PRUs.

Com a aplicação desses critérios, foram utilizados dados referentes a 26 postagens e 26 usuários, que podem ser vistos no Apêndice G. Com isso, foram gerados dois grupos.

Para gerar os grupos de usuários, a variável analisada foi o tempo de uso, cujo a média resultou em 4,8 anos. Portanto, os grupos foram divididos em: (1) grupo 1: em usuários com menos de 4 anos de experiência (14); e grupo 2: usuários com mais de 4 anos de experiência (12).

<sup>2</sup> Google Maps. Acessado em: 17 de novembro de 2018. Disponível em <<https://www.google.com.br/maps>>.

Para cada grupo, foi calculada a média das idades, o gênero com maior número de ocorrências, o dispositivo mais utilizado, média do tempo de uso e quanto a localização, foi escolhido o estado e país com maior ocorrência.

Para a definição da persona shell, foi utilizada a ferramenta UI Faces<sup>3</sup>. Nela, podem ser obtidas fotos de usuários com base em sua idade e gênero. Os nomes foram escolhidos pela própria autora deste estudo, pois não foi encontrado um dicionário ou ferramenta que pudesse gerar de forma automática.

Para agrupar os dados dos usuários, com os obtidos com a avaliação, foram analisadas as principais funcionalidades relacionadas a cada tipo. Por exemplo, para a persona 1, os principais elogios foram direcionados a funcionalidade "Escrever tweet". Vale ressaltar que nenhuma funcionalidade encontrada foi relacionada aos tipos "Dúvida", "Comparação", "Sugestão" e "Ajuda". Por isso, não foram geradas algumas relações, como as funcionalidades mais sugeridas pelo usuário.

#### 6.4 Apresentação das personas

As duas personas geradas neste experimento são apresentadas nas Figuras 24 e 25.

Figura 24 – Persona do Grupo 1.



Fonte: elaborado pela autora.

<sup>3</sup> UI Faces. Acessado em: 15 de novembro de 2018. Disponível em <<https://uifaces.co/>>.

Figura 25 – Persona do Grupo 2.



Fonte: elaborado pela autora.

## 6.5 Discussão

Nesta seção, apresenta-se uma discussão dos resultados obtidos no experimento e de fatores que poderiam, de alguma forma, enviesar ou melhorar os resultados.

**Quanto ao tamanho da base:** mesmo com um número pequeno de postagens e usuários (27) analisados de forma manual, foi possível perceber que é possível realizar a geração de personas, no contexto de avaliação textual.

**Quanto aos dados utilizados:** a API do Twitter não fornece, diretamente, alguns dados que são importantes para a criação das personas, como os dados demográficos. É possível que em outra API, tais dados possam ser extraídos. Além disso, novos dados podem ser encontrados em outras APIs, ou até mesmo não identificados. Por exemplo, se a avaliação textual for realizada com dados extraídos de uma loja de aplicativo, não se sabe quais informações podem ser extraídas sobre os usuários.

**Quanto a veracidade dos dados:** a análise dos dados demográficos realizada (Seção 6.2.1) mostrou que grande parte dos dados públicos dos usuários extraídos de SS são válidos. Porém, alguns dados precisam ser desconsiderados para criar uma persona apenas com informações válidas. Por exemplo, uma persona não pode possuir como localização, uma cidade

fictícia. Outros estudos devem ser realizados, a fim de averiguar validar as personas quanto a sua utilidade.

**Quanto aos dados gerados:** com a análise das postagens, pôde-se perceber a necessidade de uma nova variável a ser estudada: frequência de uso. Algumas PRUs retrataram que o usuário pode ter criado sua conta há um tempo, mas não costuma utilizar com frequência. Por exemplo, a postagem "*Vim pro twitter porque insta e face caiu mas descobri que não lembro como usa isso*" foi escrita por um usuário com 8 anos de uso. Apenas com a informação do tempo de uso, tal problema pode ser caracterizado como "Aprendizado", porém, se adicionada a informação de frequência de uso, essa postagem pode estar relacionada com a faceta de usabilidade "Memorabilidade". Outros estudos podem ser realizados, a fim de encontrar novas informações que possam enriquecer as personas geradas.

**Quanto as etapas realizadas de forma manual:** apenas a etapa de extração foi realizada de forma automática, as demais foram realizadas de forma manual neste experimento. Porém, com o decorrer do experimento, foi possível perceber que essas etapas podem ser realizadas de forma automática, por meio de funções simples, utilizações de APIs e utilização de algoritmos de mineração de dados.

## 6.6 Conclusão

Neste capítulo, foi apresentada uma validação da metodologia de geração automática de personas, proposta neste estudo. Para isso, foi utilizada uma amostra de 200 postagens que passaram por análise quanto a validade de seus dados e forma de modelagem.

Duas personas foram geradas, nelas foram inseridas informações sobre os usuários e sua relação com os dados da avaliação textual. O processo de criação dessas personas mostra que a metodologia utilizada pode ser útil e que pode ser realizada de forma automática.

## **7 CONCLUSÕES E TRABALHOS FUTUROS**

Neste capítulo são descritas as considerações finais deste trabalho quanto a sua contribuição e trabalhos futuros.

Este estudo teve como principal objetivo a criação de uma metodologia para geração automática de personas, a fim de aplicá-las em um contexto de avaliação textual de sistemas.

Com o estudo realizado, foi visto que a geração automática das personas é um assunto novo e com interesse crescente. Além disso, foi possível perceber que é possível gerar personas a partir dos dados públicos extraídos do twitter, além de realizar uma associação direta com os resultados da avaliação textual. Com isso, sua principal contribuição consiste em uma nova percepção sobre os usuários dentro da avaliação textual, com a utilização da metodologia proposta.

### **7.1 Contribuições do trabalho**

Como apresentado no decorrer deste estudo, sua principal contribuição foi uma metodologia para geração automática de personas a partir de dados públicos de usuários em SS. Com isso, os usuários poderão ser gerados juntamente com o decorrer da avaliação textual, não sendo necessário utilizar outras pesquisas para seu estudo.

Além disso, foi apresentado um estado da arte de geração automática de personas, por meio de um mapeamento sistemático. Este mapeamento, mostrou a relevância deste estudo e apresentou as principais etapas realizadas, os dados utilizados, requisitos para adicionar na ferramenta UUX-Posts, em quais contextos as personas são aplicadas e como validá-las. Com isso, interessados na criação de personas, poderão realizá-la de uma forma menos lenta e cansativa.

Como publicações, foi submetido um resumo do mapeamento sistemático deste estudo para a terceira edição dos Encontros Universitários na Universidade Federal do Ceará – Campus Russas, onde foi aceito e apresentado durante o evento, no período de 06 a 09 de novembro de 2018.

### **7.2 Trabalhos Futuros**

Este é um estudo inicial quanto o entendimento dos usuários em avaliações textuais, principalmente na metodologia MALTU. Por isso, este trabalho gera diversas possibilidades para

trabalhos futuros, tanto para a evolução da metodologia, quanto para a validade dessas personas.

Portanto, como trabalhos futuros devem ser melhor estudadas as formas para a geração automática de personas. Para isso, devem ser analisados quais os algoritmos de mineração de dados podem ser aplicados e a arquitetura que dará suporte para a geração automática das personas. Com essa definição, pode-se aplicar os requisitos, identificados no mapeamento sistemático, na ferramenta UUX-Posts.

Um outro estudo deve ser realizado a fim de analisar novas informações sobre os usuários e como identificá-las. Por exemplo, podem ser analisadas informações adquiridas com a própria postagem, ou até mesmo, com a descrição do usuário no SS.

Já quanto a validade dessas personas, devem ser realizados estudos que verifiquem o quanto elas impactam nos resultados da avaliação textual e como os dados podem ser interpretados a partir dessas novas informações.

Por fim, serão escritos artigos sobre o mapeamento sistemático e a metodologia de geração de personas.

## REFERÊNCIAS

- AN, J.; CHO, H.; KWAK, H.; HASSEN, M. Z.; JANSEN, B. J. Towards automatic persona generation using social media. In: **Future Internet of Things and Cloud Workshops (FiCloudW), IEEE International Conference on**. [S.l.]: IEEE, 2016. p. 206–211.
- AN, J.; KWAK, H.; JANSEN, B. J. Validating social media data for automatic persona generation. In: IEEE. **Computer Systems and Applications (AICCSA), 2016 IEEE/ACS 13th International Conference of**. [S.l.], 2016. p. 1–6.
- AN, J.; KWAK, H.; JANSEN, B. J. Personas for content creators via decomposed aggregate audience statistics. In: ACM. **Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017**. [S.l.], 2017. p. 632–635.
- AN, J.; KWAK, H.; JUNG, S.-g.; SALMINEN, J.; JANSEN, B. J. Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data. **Social Network Analysis and Mining**, Springer, v. 8, n. 1, p. 54, 2018.
- ANG, L. W. The design of presence and user testing on mentor-mentee consultation system (memoc). In: ACM. **Proceedings of the 2nd International Conference in HCI and UX Indonesia 2016**. [S.l.], 2016. p. 44–53.
- BUCHINGER, D.; CAVALCANTI, G. A. de S.; HOUNSELL, M. da S. Mecanismos de busca acadêmica: uma análise quantitativa. **Revista Brasileira de Computação Aplicada**, v. 6, n. 1, p. 108–120, 2014.
- COOPER, A. **The inmates are running the asylum:[Why high-tech products drive us crazy and how to restore the sanity]**. [S.l.]: Sams Publishers, 1999.
- COOPER, A.; REIMANN, R.; CRONIN, D. **About face 3: the essentials of interaction design**. [S.l.]: John Wiley & Sons, 2007.
- DODGE, J.; HILTON, M.; METOYER, R. A.; HUNTER, J.; SMELTZER, K.; VIJAY, C.; ATKINSON, A. Deriving age diverse personas from a participatory design study on home electricity feedback. In: ACM. **Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems**. [S.l.], 2017. p. 959–968.
- DRESCH, A.; LACERDA, D. P.; JÚNIOR, J. A. V. A. **Design science research: método de pesquisa para avanço da ciência e tecnologia**. [S.l.]: Bookman Editora, 2015.
- FREITAS, L. M.; SILVA, T. H. O. da; MENDES, M. S. Evaluation of spotify: an evaluation textual experience using the maltu methodology. In: ACM. **Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems**. [S.l.], 2016. p. 50.
- GOLDBERG, L. R. An alternative "description of personality": the big-five factor structure. **Journal of personality and social psychology**, American Psychological Association, v. 59, n. 6, p. 1216, 1990.
- GRUDIN, J.; PRUITT, J. Personas, participatory design and product development: An infrastructure for engagement. In: **PDC**. [S.l.: s.n.], 2002. v. 7, p. 144–152.

GUIMARÃES, B. **Uma Taxonomia Adaptativa por Persona Construída a partir da Participação da Comunidade Brasileira de Interação Homem-Computador**. Tese (Doutorado) — Dissertação (Mestrado em Informática)—Universidade de Fortaleza, Fortaleza, 2012.

GUO, A.; MA, J. An integrative and precise approach in personality computing based on ontic personae modeling. In: IEEE. **Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence & Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), 2017 IEEE 15th Intl.** [S.l.], 2017. p. 9–15.

GUO, A.; MA, J. Scenario-based modeling of ontic personae for automatic personality perception. In: IEEE. **2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCAL-COM/UIC/ATC/CBDCom/IOP/SCI).** [S.l.], 2017. p. 1–7.

GUO, A.; MA, J. Archetype-based modeling of persona for comprehensive personality computing from personal big data. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 18, n. 3, p. 684, 2018.

HOLANDA, K. d. **Um framework de elaboração de personas e sua aplicação para a eliciação de requisitos e para a análise das interações em sistemas sociais**. Tese (Doutorado) — Dissertação de Mestrado, 2010.

JANSEN, A.; MECHELEN, M. V.; SLEGGERS, K. Personas and behavioral theories: A case study using self-determination theory to construct overweight personas. In: ACM. **Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.** [S.l.], 2017. p. 2127–2136.

JANSEN, B. J.; JUNG, S.-G.; SALMINEN, J.; AN, J.; KWA, H. Leveraging social analytics data for identifying customer segments for online news media. In: IEEE. **Computer Systems and Applications (AICCSA), 2017 IEEE/ACS 14th International Conference on.** [S.l.], 2017. p. 463–468.

JANSEN, B. J.; JUNG, S.-G.; SALMINEN, J.; AN, J.; KWAK, H. Viewed by too many or viewed too little: Using information dissemination for audience segmentation. **Proceedings of the Association for Information Science and Technology**, Wiley Online Library, v. 54, n. 1, p. 189–196, 2017.

JUNG, S.-G.; AN, J.; KWAK, H.; SALMINEN, J.; JANSEN, B. J. Inferring social media users' demographics from profile pictures: A face++ analysis on twitter users. 2017.

JUNG, S.-G.; AN, J.; KWAK, H.; AHMAD, M.; NIELSEN, L.; JANSEN, B. J. Persona generation from aggregated social media data. In: ACM. **Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems.** [S.l.], 2017. p. 1748–1755.

JUNG, S. G.; SALMINEN, J.; AN, J.; KWAK, H.; JANSEN, B. J. Automatically conceptualizing social media analytics data via personas. In: AAAI PRESS. **12th International AAAI Conference on Web and Social Media, ICWSM 2018.** [S.l.], 2018. p. 715–716.



- JUNG, S.-g.; SALMINEN, J.; KWAK, H.; AN, J.; JANSEN, B. J. Automatic persona generation (apg): A rationale and demonstration. In: ACM. **Proceedings of the 2018 Conference on Human Information Interaction & Retrieval**. [S.l.], 2018. p. 321–324.
- KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. 2007.
- LEITAO, T.; MORGADO, C.; CUNHA, J. C. Measuring popularity in social network groups. In: IEEE. **Cloud and Green Computing (CGC), 2012 Second International Conference on**. [S.l.], 2012. p. 485–492.
- LIBERATI, A.; ALTMAN, D. G.; TETZLAFF, J.; MULROW, C.; GØTZSCHE, P. C.; IOANNIDIS, J. P.; CLARKE, M.; DEVEREAUX, P. J.; KLEIJNEN, J.; MOHER, D. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. **PLoS medicine**, Public Library of Science, v. 6, n. 7, p. e1000100, 2009.
- LIMA, A. C. E. S. *et al.* Mineração de mídias sociais como ferramenta para a análise de tríade da persona virtual. Universidade Presbiteriana Mackenzie, 2016.
- LIU, H.; MAES, P. What would they think?: a computational model of attitudes. In: ACM. **Proceedings of the 9th international conference on Intelligent user interfaces**. [S.l.], 2004. p. 38–45.
- MARCUS, A. User interface design's return on investment: Examples and statistics. In: **Cost-Justifying Usability (Second edition)**. [S.l.]: Elsevier, 2005. p. 17–39.
- MASIERO, A. A.; FERREIRA, L. A.; JR, P. T. A. Algoritmos de clusterização e python científico apoiando modelagem de usuário. In: BRAZILIAN COMPUTER SOCIETY. **Companion Proceedings of the 11th Brazilian Symposium on Human Factors in Computing Systems**. [S.l.], 2012. p. 47–50.
- MENDES, M. **MALTU—Um modelo para avaliação da interação em sistemas sociais a partir da linguagem textual do usuário**. Tese (Doutorado) — Tese (doutorado). Universidade Federal do Ceará, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2015.
- MENDES, M. S.; FURTADO, E. S. Uux-posts: a tool for extracting and classifying postings related to the use of a system. In: ACM. **Proceedings of the 8th Latin American Conference on Human-Computer Interaction**. [S.l.], 2017. p. 2.
- MIJAČ, T.; JADRIĆ, M.; ČUKUŠIĆ, M. The potential and issues in data-driven development of web personas. In: **MIPRO 2018-41st International Convention (Business Intelligence Systems/miproBIS)**. [S.l.: s.n.], 2018.
- NIELSEN, L. Acting as someone like me: personas in participatory innovation. **Proceedings of the PIN-C**, p. 1–7, 2012.
- NIELSEN, L.; HANSEN, K. S. Personas is applicable: a study on the use of personas in denmark. In: ACM. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. [S.l.], 2014. p. 1665–1674.
- NÓBREGA, C. L. d. O.; FURTADO, M. Um framework de elaboração de persona empresa para suporte na análise de valor de negócio na aplicação em sistemas de redes sociais. **Universidade de Fortaleza, Fortaleza**, 2011.

ORTBAL, K.; FRAZZETTE, N.; MEHTA, K. Constructed stakeholder personas: an educational tool for social entrepreneurs. **Procedia Engineering**, Elsevier, v. 159, p. 230–248, 2016.

PEREIRA, R.; BARANAUSKAS, M. C. C.; SILVA, S. R. P. da. Softwares sociais: uma visão orientada a valores. In: BRAZILIAN COMPUTER SOCIETY. **Proceedings of the IX Symposium on Human Factors in Computing Systems**. [S.l.], 2010. p. 149–158.

POHL, K.; RUPP, C. **Fundamentos da engenharia de requisitos**. [S.l.]: International Requirements Engineering Board (IREB), 2012.

PRATES, R. O.; BARBOSA, S. D. J. Avaliação de interfaces de usuário—conceitos e métodos. In: **Jornada de Atualização em Informática do Congresso da Sociedade Brasileira de Computação, Capítulo**. [S.l.: s.n.], 2003. v. 6, p. 28.

PRUITT, J.; ADLIN, T. **The persona lifecycle: keeping people in mind throughout product design**. [S.l.]: Elsevier, 2010.

PRUITT, J.; GRUDIN, J. Personas: practice and theory. In: ACM. **Proceedings of the 2003 conference on Designing for user experiences**. [S.l.], 2003. p. 1–15.

ROGERS, Y.; SHARP, H.; PREECE, J. **Design de Interação**. [S.l.]: Bookman Editora, 2013.

SALMINEN, J.; JANSEN, B. J.; AN, J.; JUNG, S.-G.; NIELSEN, L.; KWAK, H. Fixation and confusion: Investigating eye-tracking participants' exposure to information in personas. In: ACM. **Proceedings of the 2018 Conference on Human Information Interaction & Retrieval**. [S.l.], 2018. p. 110–119.

SALMINEN, J.; JUNG, S.-G.; AN, J.; KWAK, H.; JANSEN, B. J. Findings of a user study of automatically generated personas. In: ACM. **Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems**. [S.l.], 2018. p. LBW097.

SALMINEN, J.; KWAK, H.; SANTOS, J. M.; JUNG, S.-G.; AN, J.; JANSEN, B. J. Persona perception scale: Developing and validating an instrument for human-like representations of data. In: ACM. **Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems**. [S.l.], 2018. p. LBW075.

SALMINEN, J.; NIELSEN, L.; JUNG, S.-g.; AN, J.; KWAK, H.; JANSEN, B. J. “is more better?”: Impact of multiple photos on perception of persona profiles. In: ACM. **Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems**. [S.l.], 2018. p. 317.

SALMINEN, J.; SENGÜN, S.; KWAK, H.; JANSEN, B.; AN, J.; JUNG, S.-G.; VIEWEG, S.; HARRELL, D. F. Generating cultural personas from social data: A perspective of middle eastern users. In: IEEE. **Future Internet of Things and Cloud Workshops (FiCloudW), 2017 5th International Conference on**. [S.l.], 2017. p. 120–125.

SANTOS, M.; RABELO, J.; BARRETO, R.; CONTE, T. Persona security: A technique for supporting the elicitation of security requirements. In: **Proceedings of the International Conference on Software Engineering and Knowledge Engineering**. [S.l.: s.n.], 2014.

SILVA, T. H. O.; FREITAS, L. M.; MENDES, M. S. Beyond traditional evaluations: user's view in app stores. In: ACM. **Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems**. [S.l.], 2017. p. 15.

TERRY, M.; KAY, M.; VUGT, B. V.; SLACK, B.; PARK, T. Ingimp: introducing instrumentation to an end-user open source application. In: ACM. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. [S.l.], 2008. p. 607–616.

WÄRNESTÅL, P.; SVEDBERG, P.; NYGREN, J. Co-constructing child personas for health-promoting services with vulnerable children. In: ACM. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. [S.l.], 2014. p. 3767–3776.

WATANABE, Y.; WASHIZAKI, H.; HONDA, K.; NOYORI, Y.; FUKAZAWA, Y.; MORIZUKI, A.; SHIBATA, H.; OGAWA, K.; ISHIGAKI, M.; SHIIZAKI, S. *et al.* Id3p: Iterative data-driven development of persona based on quantitative evaluation and revision. In: IEEE PRESS. **Proceedings of the 10th International Workshop on Cooperative and Human Aspects of Software Engineering**. [S.l.], 2017. p. 49–55.

WATANABE, Y.; WASHIZAKI, H.; HONDA, K.; FUKAZAWA, Y.; TAGA, M.; MATSUZAKI, A.; SUZUKI, T. Retrospective based on data-driven persona significance in b-to-b software development. In: ACM. **Proceedings of the 40th International Conference on Software Engineering: New Ideas and Emerging Results**. [S.l.], 2018. p. 89–92.

WOHLIN, C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: ACM. **Proceedings of the 18th international conference on evaluation and assessment in software engineering**. [S.l.], 2014. p. 38.

ZHANG, X.; BROWN, H.-F.; SHANKAR, A. Data-driven personas: Constructing archetypal users with clickstreams and user telemetry. In: ACM. **Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems**. [S.l.], 2016. p. 5350–5359.

## **APÊNDICE A – FORMULÁRIO DE EXTRAÇÃO - INFORMAÇÕES SOBRE A TÉCNICA**

Este documento apresenta os dados extraídos, quanto à técnica, de acordo com o formulário 4 apresentado na Seção 3.2.4.

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a técnica			
						Nome	Autores	Etapas	Etapa(s) automática(s)
0	Findings of a User Study of Automatically Generated Personas	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	CHI 2018	2018	ACM	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Análise, Geração e Interação
1	Algoritmos De Clusterização e Python Científico Apoiando Modelagem De Usuário	Masiero, Andrey Araujo and Ferreira, Leonardo A. and Aquino, Jr., Plinio Thomaz	11th Brazilian Symposium on Human Factors in Computing Systems	2012	ACM	#	Andrey Araujo Masiero, Leonardo A. Ferreira, Plinio Thomaz Aquino Jr	Recebe os dados coletados e os normaliza, utilizando a equação de normalização. Depois, aplica um algoritmo de clusterização em Python. Ao executar o algoritmo ele retorna na variável labels o valor do índice que o usuário pertence. Depois separa os elementos em seus grupos para gerar as Personas. Em seguida, combina os valores das variáveis e encontra um valor em comum entre elas.	Análise dos dados (Clusterização)
2	Automatic Persona Generation (APG): A Rationale and Demonstration	Jung, Soon-gyo and Salminen, Joni and Kwak, Haewoon and An, Jisun and Jansen, Bernard J.	Conference on Human Information Interaction and Retrieval 2018	2018	ACM	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
3	Persona Generation from Aggregated Social Media Data	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeed and Nielsen, Lene and Jansen, Bernard J.	CHI 2017	2017	ACM	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
4	Ingimp: Introducing Instrumentation to an End-user Open Source Application	Terry, Michael and Kay, Matthew and Van Vugt, Brad and Slack, Brandon and Park, Terry	SIGCHI Conference on Human Factors in Computing Systems	2008	ACM	#	Michael Terry, Matthew Kay, Brad Van Vugt, Brandon Slack, Terry Park	Extração de dados da ferramenta. Depois disso, os dados passam por uma análise dos dados (clusterização) utilizando dois algoritmos, hand-coded rules e k-means.	Extração e análise

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a técnica			
						Nome	Autores	Etapas	Etapa(s) automática(s)
5	Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	CHI 2016	2016	ACM	Data-Driven Persona	#	Coleta de dados; Agregação dos cliques de usuários em cliques de vídeo; Estruturação em fluxos de trabalho por meio de clustering hierárquico. Análise estatística manual para seleção de 5 personas representativas	Extração e parte da análise
7	What Would They Think?: A Computational Model of Attitudes	Liu, Hugo and Maes, Pattie	9th International Conference on Intelligent User Interfaces	2004	ACM	WWTT	#	Coleta de dados, utilizando a ferramenta, pegando as informações sobre o blog da pessoa, assim pegando as postagens. Analisa os dados. Apresenta a persona digital da persona, gerada a partir de uma análise automatizada de textos pessoais usando processamento de linguagem natural e sensoriamento de texto com base no senso comum	Extração, Análise dos dados
9	Validating social media data for automatic persona generation	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	IEEE/ACS International Conference on Computer Systems and Applications	2017	Engineering Village	#	#	Extração dos dados diretamente do Youtuber do canal AJ+	Extração e análise
12	Archetype-based modeling of persona for comprehensive personality computing from personal big data	Guo, Ao and Ma, Jianhua	Journal Sensors (Switzerland)	2018	Engineering Village	SCENARIO-BASED MODELING OF ONTIC PERSONAE FOR PERSONALITY PECEPTION	#	Coleta de dados -> Detecção de cenários -> Modelagem da Ontic Personae -> Percepção da Personalidade	Análise dos dados e Geração
17	Scenario-based modeling of ontic personae for automatic personality perception	Guo, Ao and Ma, Jianhua	SmartWorld / SCALCOM/ UIC/ ATC/ CBDCOM/ IOP/ SCI 2017	2017	Engineering Village	SCENARIO-BASED MODELING OF ONTIC PERSONAE FOR PERSONALITY PECEPTION	#	Coleta de dados -> Detecção de cenários -> Modelagem da Ontic Personae -> Percepção da Personalidade	Análise dos dados e Geração
21	Automatically conceptualizing social media analytics data via personas	Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	ICWSM 2018	2018	Engineering Village	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
23	An Integrative and Precise Approach in Personality Computing Based on Ontic Personae Modeling	Guo, Ao and Ma, Jianhua	IEEE Cyber Science and Technology Congress, DASC-PICom-DataCom-CyberSciTec	2018	Engineering Village	SCENARIO-BASED MODELING OF ONTIC PERSONAE FOR PERSONALITY PECEPTION	#	Coleta de dados -> Detecção de cenários -> Modelagem da Ontic Personae -> Percepção da Personalidade	Análise dos dados e Geração

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a técnica			
						Nome	Autores	Etapas	Etapa(s) automática(s)
35	Towards Automatic Persona Generation Using Social Media	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohammed Z. and Jansen, Bernard J.	4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	2016	IEEE	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
36	Generating Cultural Personas from Social Data: A Perspective of Middle Eastern Users	Salminen, Joni and Sengün, Sercan and Kwak, Haewoon and Jansen, Bernard J. and An, Jisun and Jung, Soon-Gyo and Vieweg, Sarah and Harrell, D. Fox	5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	2017	IEEE	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
40	Leveraging Social Analytics Data for Identifying Customer Segments for Online News Media	Jansen, Bernard J. and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)	2017	IEEE	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
2272	Personas for content creators via decomposed aggregate audience statistics	An, Jisun and Kwak, Haewoon and Jansen, Bernard J	International Conference on Advances in Social Networks Analysis and Mining 2017	2017	Google Academic	#	#	Extração dos dados -> Cria a matriz W e H, identificando os grupos de usuários e suas interações -> Procura padrões com algoritmos de clusterização -> Adiciona personalidade e outros detalhes	Extração e Análise
2274	Viewed by too many or viewed too little: Using information dissemination for audience segmentation	Jansen, Bernard J and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	Association for Information Science and Technology	2017	Google Academic	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a técnica			
						Nome	Autores	Etapas	Etapa(s) automática(s)
2280	Inferring Social Media Users' Demographics from Profile Pictures: A Face++ Analysis on Twitter Users	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Salminen, Joni and Jansen, Bernard J	ICEB 2017 Proceedings	2017	Google Academic	Não apresenta	#	Extração das imagens -> Analisar imagens com o Face++ -> Obter dados demográficos -> Aplicar na geração das personas	Extração e Análise
2282	Fixation and Confusion: Investigating Eye-tracking Participants' Exposure to Information in Personas	Salminen, Joni and Jansen, Bernard J and An, Jisun and Jung, Soon-Gyo and Nielsen, Lene and Kwak, Haewoon	Conference on Human Information Interaction and Retrieval 2018	2018	Google Academic	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
2283	"Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles	Salminen, Joni and Nielsen, Lene and Jung, Soon-gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J	CHI 2018	2018	Google Academic	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
2285	Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data	An, Jisun and Kwak, Haewoon and Jung, Soon-gyo and Salminen, Joni and Jansen, Bernard J	Social Network Analysis and Mining	2018	Google Academic	Automatic Persona Generation (APG)	#	1) configuração; 2) coleção; 3) geração; e 4) interação	Extração, Geração e Interação
2326	ID3P: Iterative Data-Driven Development of Persona based on quantitative evaluation and revision	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Noyori, Yuki and Fukazawa, Yoshiaki and Morizuki, Aoi and Shibata, Hiroyuki and Ogawa, Kentaro and Ishigaki, Mikako and Shiizaki, Satiyo and others	10th International Workshop on Cooperative and Human Aspects of Software Engineering	2017	Google Academic	ID3P	#	1) Iniciar, aplicando GQM+S -> 2) Desenvolver personas por uma abordagem chamada Data-Driven Persona Construction -> 3) Deduzir as premissas para planejar estratégias -> 4) Planejar e executar estratégias -> 5) Revisar as personas -> 6) Verificar as premissas e avaliar as estratégias	Análise dos dados e parte da extração



ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a técnica			
						Nome	Autores	Etapas	Etapa(s) automática(s)
2329	Retrospective based on data-driven persona significance in B-to-B software development	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Fukazawa, Yoshiaki and Taga, Masahiro and Matsuzaki, Akira and Suzuki, Takayoshi	40th International Conference on Software Engineering : New Ideas and Emerging Results	2018	Google Academic	ReD2PS	#	1) Construa personas de usuário final. 2) Identifique os usuários alvo em cada iteração. 3) Adapte usuários-alvo a personas	Extração e Análise
2331	Persona Perception Scale: Developing and Validating an Instrument for Human-Like Representations of Data	Salminen, Joni and Kwak, Haewoon and Santos, Jo (Jo) M and Jung, Soon-Gyo and An, Jisun and Jansen, Bernard J	CHI 2018	2018	Google Academic		Não apresenta		
2334	The Potential and Issues in Data-Driven Development of Web Personas	Mijac, Tea and Jadric, Mario and Cukusic	MIPRO 2018	2018	Google Academic	Data-driven Personas	#	Extração dos dados, a partir da ferramenta Google Analytics -> Análise dos dados, mineração -> Geração das personas -> Enriquecimento com dados demográficos, extraídos por grupos de usuários ->	Extração e Análise

## **APÊNDICE B – FORMULÁRIO DE EXTRAÇÃO - INFORMAÇÕES SOBRE OS DADOS**

Este documento apresenta os dados extraídos, quanto aos dados das personas, de acordo com o formulário 4 apresentado na Seção 3.2.4.

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre as personas geradas	
						Dados extraídos	Características
0	Findings of a User Study of Automatically Generated Personas	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	CHI 2018	2018	ACM	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	<p>Perfil da persona: Imagem, nome, gênero, idade e país.            Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional.</p> <p>Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante.</p> <p>Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse.</p> <p>Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.</p>
1	Algoritmos De Clusterização e Python Científico Apoiando Modelagem De Usuário	Masiero, Andrey Araujo and Ferreira, Leonardo A. and Aquino, Jr., Plinio Thomaz	11th Brazilian Symposium on Human Factors in Computing Systems	2012	ACM	Dados de questionários	Não gera persona de forma visual, apenas apresenta um cluster dos dados encontrados.
2	Automatic Persona Generation (APG): A Rationale and Demonstration	Jung, Soon-gyo and Salminen, Joni and Kwak, Haewoon and An, Jisun and Jansen, Bernard J.	Conference on Human Information Interaction and Retrieval 2018	2018	ACM	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	<p>Perfil da persona: Imagem, nome, gênero, idade e país.            Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional.</p> <p>Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante.</p> <p>Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse.</p> <p>Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.</p>
3	Persona Generation from Aggregated Social Media Data	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeed and Nielsen, Lene and Jansen, Bernard J.	CHI 2017	2017	ACM	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	<p>Dados demográficos: idade, gênero e localização (clusterização dos dados);</p> <p>Nome da persona: dicionário de nomes feito por eles;</p> <p>Foto da persona: fotos tagadas com base nos dados demográficos;</p> <p>Outros detalhes da persona: tópicos de interesse, possíveis quotes.</p>
4	Ingimp: Introducing Instrumentation to an End-user Open Source Application	Terry, Michael and Kay, Matthew and Van Vugt, Brad and Slack, Brandon and Park, Terry	SIGCHI Conference on Human Factors in Computing Systems	2008	ACM	Dados demográficos, Logs de interação, Logs do cenário (Comandos do usuário; eventos principais da interface do usuário; principais informações demográficas; informações sobre o ambiente computacional; descrições dos usuários.)	Não apresentado

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre as personas geradas	
						Dados extraídos	Características
5	Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	CHI 2016	2016	ACM	Logs de interação (Cliques coletados automaticamente a partir de dados de telemetria relacionados ao uso real do produto no campo)	Não apresentado
7	What Would They Think?: A Computational Model of Attitudes	Liu, Hugo and Maes, Pattie	9th International Conference on Intelligent User Interfaces	2004	ACM	Imagem, Postagens (dados de uma URL para um weblog ou para um corpus de textos).	Memória reflexiva, memória episódica e memória imprimir
9	Validating social media data for automatic persona generation	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	IEEE/ACS International Conference on Computer Systems and Applications	2017	Engineering Village	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Não apresentado
12	Archetype-based modeling of persona for comprehensive personality computing from personal big data	Guo, Ao and Ma, Jianhua	Journal Sensors (Switzerland)	2018	Engineering Village	Dados demográficos, Informações do contexto (Localização, outras pessoas envolvidas no cenário e comportamento do usuário)	Gera uma Primary Ontic Persona, especificando o nome, gênero, nacionalidade, ocupação e tagline (frase de descrição). Gera uma Facet Ontic Persona, especificando a função, habilidades, atitudes e motivações. Personalidade da persona, apresentando suas características com base no modelo Big-Five.
17	Scenario-based modeling of ontic personae for automatic personality perception	Guo, Ao and Ma, Jianhua	SmartWorld / SCALCOM/ UIC/ ATC/ CBDCom/ IOP/ SCI 2017	2017	Engineering Village	Dados demográficos, Informações do contexto (Localização, outras pessoas envolvidas no cenário e comportamento do usuário)	Gera uma Primary Ontic Persona, especificando o nome, gênero, nacionalidade, ocupação e tagline (frase de descrição). Gera uma Facet Ontic Persona, especificando a função, habilidades, atitudes e motivações. Personalidade da persona, apresentando suas características com base no modelo Big-Five.
21	Automatically conceptualizing social media analytics data via personas	Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	ICWSM 2018	2018	Engineering Village	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Não apresenta
23	An Integrative and Precise Approach in Personality Computing Based on Ontic Personae Modeling	Guo, Ao and Ma, Jianhua	IEEE Cyber Science and Technology Congress, DASC-PICom-DataCom-CyberSciTec	2018	Engineering Village	Dados demográficos, Dados do Cenário (Localização, outras pessoas envolvidas no cenário e comportamento do usuário)	Gera uma Primary Ontic Persona, especificando o nome, gênero, nacionalidade, ocupação e tagline (frase de descrição). Gera uma Facet Ontic Persona, especificando a função, habilidades, atitudes e motivações. Personalidade da persona, apresentando suas características com base no modelo Big-Five.

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre as personas geradas	
						Dados extraídos	Características
35	Towards Automatic Persona Generation Using Social Media	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohammed Z. and Jansen, Bernard J.	4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	2016	IEEE	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Dados demográficos: idade, gênero e localização (clusterização dos dados); Nome da persona: dicionário de nomes feito por eles; Foto da persona: fotos tageadas com base nos dados demográficos; Outros detalhes da persona: tópicos de interesse, possíveis quotes.
36	Generating Cultural Personas from Social Data: A Perspective of Middle Eastern Users	Salminen, Joni and Sengün, Sercan and Kwak, Haewoon and Jansen, Bernard J. and An, Jisun and Jung, Soon-Gyo and Vieweg, Sarah and Harrell, D. Fox	5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	2017	IEEE	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Perfil da persona: Imagem, nome, gênero, idade e país. Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional. Pessoas como a persona: dados da clusterização. Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante. Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse. Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.
40	Leveraging Social Analytics Data for Identifying Customer Segments for Online News Media	Jansen, Bernard J. and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)	2017	IEEE	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Perfil da persona: Imagem, nome, gênero, idade e país. Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional. Pessoas como a persona: dados da clusterização. Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante. Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse. Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.
2272	Personas for content creators via decomposed aggregate audience statistics	An, Jisun and Kwak, Haewoon and Jansen, Bernard J	International Conference on Advances in Social Networks Analysis and Mining 2017	2017	Google Academic	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Objetivo da persona, dados demográficos, nome, foto e outros detalhes (descrição)
2274	Viewed by too many or viewed too little: Using information dissemination for audience segmentation	Jansen, Bernard J and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	Association for Information Science and Technology	2017	Google Academic	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Perfil da persona: Imagem, nome, gênero, idade e país. Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional. Pessoas como a persona: dados da clusterização. Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante. Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse. Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre as personas geradas	
						Dados extraídos	Características
2280	Inferring Social Media Users' Demographics from Profile Pictures: A Face++ Analysis on Twitter Users	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Salminen, Joni and Jansen, Bernard J	ICEB 2017 Proceedings	2017	Google Academic	Imagem do perfil	Dados demográficos
2282	Fixation and Confusion: Investigating Eye-tracking Participants' Exposure to Information in Personas	Salminen, Joni and Jansen, Bernard J and An, Jisun and Jung, Soon-Gyo and Nielsen, Lene and Kwak, Haewoon	Conference on Human Information Interaction and Retrieval 2018	2018	Google Academic	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Perfil da persona: Imagem, nome, gênero, idade e país. Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional. Pessoas como a persona: dados da clusterização. Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante. Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse. Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.
2283	"Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles	Salminen, Joni and Nielsen, Lene and Jung, Soon-gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J	CHI 2018	2018	Google Academic	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Perfil da persona: Imagem, nome, gênero, idade e país. Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional. Pessoas como a persona: dados da clusterização. Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante. Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse. Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.
2285	Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data	An, Jisun and Kwak, Haewoon and Jung, Soon-gyo and Salminen, Joni and Jansen, Bernard J	Social Network Analysis and Mining	2018	Google Academic	Dados demográficos (idade, gênero, localização) e Logs de interação (logs de interação do usuário com o conteúdo online em questão)	Perfil da persona: Imagem, nome, gênero, idade e país. Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional. Pessoas como a persona: dados da clusterização. Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante. Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse. Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.
2326	ID3P: Iterative Data-Driven Development of Persona based on quantitative evaluation and revision	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Noyori, Yuki and Fukazawa, Yoshiaki and Morizuki, Aoi and Shibata, Hiroyuki and Ogawa, Kentaro and Ishigaki, Mikako and Shiizaki, Satiyo and others	10th International Workshop on Cooperative and Human Aspects of Software Engineering	2017	Google Academic	Dados de questionários (Manual) e Logs de interação( Automático), sendo o id do usuário, categoria de artigos lidos pelo usuário, os artigos e número de compartilhamentos do artigo no Twitter e Facebook)	Nome, idade, foto, objetivo no sistema, preferências

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre as personas geradas	
						Dados extraídos	Características
2329	Retrospective based on data-driven persona significance in B-to-B software development	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Fukazawa, Yoshiaki and Taga, Masahiro and Matsuzaki, Akira and Suzuki, Takayoshi	40th International Conference on Software Engineering : New Ideas and Emerging Results	2018	Google Academic	Logs de interação (D1) Histórico de registro dos usuários finais para construir personas, D2) Problemas que incluem o recurso relacionado em versões anteriores e D3) Problemas que os desenvolvedores desejam identificar usuários-alvo.)	Não apresenta
2331	Persona Perception Scale: Developing and Validating an Instrument for Human-Like Representations of Data	Salminen, Joni and Kwak, Haewoon and Santos, Jo (a) M and Jung, Soon-Gyo and An, Jisun and Jansen, Bernard J	CHI 2018	2018	Google Academic	Não apresenta	Perfil da persona: Imagem, nome, gênero, idade e país. Descrição da persona: tópico recapitulando informações básicas, combinando com os tópicos de interesse e o dispositivo utilizado pelo usuário. Além disso, carreira, status de relacionamento, parentesco e nível educacional. Pessoas como a persona: dados da clusterização. Tópicos de interesse: tópicos com que a persona interage com o conteúdo. Tópicos mais interessante e menos interessante. Quotes (Perguntas): comentários em redes sociais associados aos tópicos de interesse. Vídeos mais assistidos: mostra os conteúdos com que a persona mais interagiu, ligados ao conteúdo online.
2334	The Potential and Issues in Data-Driven Development of Web Personas	Mijac, Tea and Jadric, Mario and Cukusic	MIPRO 2018	2018	Google Academic	Dados demográficos, Logs de interação (Para cada ClientID foram extraídos os seguintes dados: número de usuários, novos usuários, número de sessões, bounce rate, page/session, tempo médio de cada sessão, transações, revenue, país e categoria)	Nome, idade, gênero, dispositivo, tempo gasto

## **APÊNDICE C – FORMULÁRIO DE EXTRAÇÃO - INFORMAÇÕES SOBRE A FERRAMENTA**

Este documento apresenta os dados extraídos, quanto as ferramentas identificadas, de acordo com o formulário 4 apresentado na Seção 3.2.4.



ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a ferramenta					
						Nome	Link	Plataforma	Autores	Universidade, País	Sistema utilizado para extrair dados
0	Findings of a User Study of Automatically Generated Personas	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	CHI 2018	2018	ACM	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Facebook e Youtube
1	Algoritmos De Clusterização e Python Científico Apoiando Modelagem De Usuário	Masiero, Andrey Araujo and Ferreira, Leonardo A. and Aquino, Jr., Plinio Thomaz	11th Brazilian Symposium on Human Factors in Computing Systems	2012	ACM			Não apresenta			
2	Automatic Persona Generation (APG): A Rationale and Demonstration	Jung, Soon-gyo and Salminen, Joni and Kwak, Haewoon and An, Jisun and Jansen, Bernard J.	Conference on Human Information Interaction and Retrieval 2018	2018	ACM	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Facebook e Youtube
3	Persona Generation from Aggregated Social Media Data	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeed and Nielsen, Lene and Jansen, Bernard J.	CHI 2017	2017	ACM	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Youtube
4	Ingimp: Introducing Instrumentation to an End-user Open Source Application	Terry, Michael and Kay, Matthew and Van Vugt, Brad and Slack, Brandon and Park, Terry	SIGCHI Conference on Human Factors in Computing Systems	2008	ACM	Ingimp	<a href="http://www.ingimp.org">http://www.ingimp.org</a>	Desktop	Michael Terry, Matthew Kay, Brad Van Vugt, Brandon Slack, Terry Park	University of Waterloo, Canadá	Ingimp

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a ferramenta					
						Nome	Link	Plataforma	Autores	Universidade, País	Sistema utilizado para extrair dados
5	Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	CHI 2016	2016	ACM				Não apresenta		
7	What Would They Think?: A Computational Model of Attitudes	Liu, Hugo and Maes, Pattie	9th International Conference on Intelligent User Interfaces	2004	ACM	What Would They Think? (WWTT)	Não disponibilizado	Desktop	#	MIT Media Laboratory, Cambridge, MA	What Would They Think? (WWTT)
9	Validating social media data for automatic persona generation	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	IEEE/ACS International Conference on Computer Systems and Applications	2017	Engineering Village				Não apresenta		
12	Archetype-based modeling of persona for comprehensive personality computing from personal big data	Guo, Ao and Ma, Jianhua	Journal Sensors (Switzerland)	2018	Engineering Village	Archetype-based personality computing system (ABPC system)	Não disponibiliza	#	#	#	Wearable Web (serviços de IoT), dados do Facebook e histórico do navegador
17	Scenario-based modeling of ontic personae for automatic personality perception	Guo, Ao and Ma, Jianhua	SmartWorld / SCALCOM/ UIC/ ATC/ CBDCOM/ IOP/ SCI 2017	2017	Engineering Village				Não apresenta		
21	Automatically conceptualizing social media analytics data via personas	Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	ICWSM 2018	2018	Engineering Village	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Facebook e Youtube
23	An Integrative and Precise Approach in Personality Computing Based on Ontic Personae Modeling	Guo, Ao and Ma, Jianhua	IEEE Cyber Science and Technology Congress, DASC-PICom-DataCom-CyberSciTec	2018	Engineering Village	Archetype-based personality computing system (ABPC system)	Não disponibiliza	#	#	#	Wearable Web (serviços de IoT), dados do Facebook e histórico do navegador

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a ferramenta					
						Nome	Link	Plataforma	Autores	Universidade, País	Sistema utilizado para extrair dados
35	Towards Automatic Persona Generation Using Social Media	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohammed Z. and Jansen, Bernard J.	4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	2016	IEEE	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Facebook, Twitter e Youtube
36	Generating Cultural Personas from Social Data: A Perspective of Middle Eastern Users	Salminen, Joni and Sengün, Sercan and Kwak, Haewoon and Jansen, Bernard J. and An, Jisun and Jung, Soon-Gyo and Vieweg, Sarah and Harrell, D. Fox	5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	2017	IEEE	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Youtube
40	Leveraging Social Analytics Data for Identifying Customer Segments for Online News Media	Jansen, Bernard J. and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)	2017	IEEE	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Youtube
2272	Personas for content creators via decomposed aggregate audience statistics	An, Jisun and Kwak, Haewoon and Jansen, Bernard J	International Conference on Advances in Social Networks Analysis and Mining 2017	2017	Google Academic				Não apresenta		
2274	Viewed by too many or viewed too little: Using information dissemination for audience segmentation	Jansen, Bernard J and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	Association for Information Science and Technology	2017	Google Academic	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Youtube

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a ferramenta					
						Nome	Link	Plataforma	Autores	Universidade, País	Sistema utilizado para extrair dados
2280	Inferring Social Media Users' Demographics from Profile Pictures: A Face++ Analysis on Twitter Users	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Salminen, Joni and Jansen, Bernard J	ICEB 2017 Proceedings	2017	Google Academic			Não apresentado			Twitter e Face++ (https://www.faceplusplus.com/)
2282	Fixation and Confusion: Investigating Eye-tracking Participants' Exposure to Information in Personas	Salminen, Joni and Jansen, Bernard J and An, Jisun and Jung, Soon-Gyo and Nielsen, Lene and Kwak, Haewoon	Conference on Human Information Interaction and Retrieval 2018	2018	Google Academic	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Youtube
2283	"Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles	Salminen, Joni and Nielsen, Lene and Jung, Soon-gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J	CHI 2018	2018	Google Academic	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Youtube
2285	Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data	An, Jisun and Kwak, Haewoon and Jung, Soon-gyo and Salminen, Joni and Jansen, Bernard J	Social Network Analysis and Mining	2018	Google Academic	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	Youtube e Facebook
2326	ID3P: Iterative Data-Driven Development of Persona based on quantitative evaluation and revision	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Noyori, Yuki and Fukazawa, Yoshiaki and Morizuki, Aoi and Shibata, Hiroyuki and Ogawa, Kentaro and Ishigaki, Mikako and Shiizaki, Satiyo and others	10th International Workshop on Cooperative and Human Aspects of Software Engineering	2017	Google Academic			Não apresentado			

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a ferramenta					
						Nome	Link	Plataforma	Autores	Universidade, País	Sistema utilizado para extrair dados
2329	Retrospective based on data-driven persona significance in B-to-B software development	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Fukazawa, Yoshiaki and Taga, Masahiro and Matsuzaki, Akira and Suzuki, Takayoshi	40th International Conference on Software Engineering : New Ideas and Emerging Results	2018	Google Academic				Não apresenta		
2331	Persona Perception Scale: Developing and Validating an Instrument for Human-Like Representations of Data	Salminen, Joni and Kwak, Haewoon and Santos, Jo{\~a}o M and Jung, Soon-Gyo and An, Jisun and Jansen, Bernard J	CHI 2018	2018	Google Academic	APG	<a href="https://persona.qcri.org/">https://persona.qcri.org/</a>	WEB	#	Hamad Bin Khalifa University, Qatar	#
2334	The Potential and Issues in Data-Driven Development of Web Personas	Mijac, Tea and Jadric, Mario and Cukusic	MIPRO 2018	2018	Google Academic				Não apresenta		

## **APÊNDICE D – FORMULÁRIO DE EXTRAÇÃO - INFORMAÇÕES SOBRE O CONTEXTO**

Este documento apresenta os dados extraídos, quanto ao contexto de aplicação das personas, de acordo com o formulário 4 apresentado na Seção 3.2.4.

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre o contexto de aplicação da persona
						Contexto
0	Findings of a User Study of Automatically Generated Personas	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	CHI 2018	2018	ACM	Marketing digital
1	Algoritmos De Clustering e Python Científico Apoiando Modelagem De Usuário	Masiero, Andrey Araujo and Ferreira, Leonardo A. and Aquino, Jr., Flávio Thomaz	11th Brazilian Symposium on Human Factors in Computing Systems	2012	ACM	Desenvolvimento de softwares
2	Automatic Persona Generation (APG): A Rationale and Demonstration	Jung, Soon-gyo and Salminen, Joni and Kwak, Haewoon and An, Jisun and Jansen, Bernard J.	Conference on Human Information Interaction and Retrieval 2018	2018	ACM	Marketing digital
3	Persona Generation from Aggregated Social Media Data	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeed and Nielsen, Lene and Jansen, Bernard J.	CHI 2017	2017	ACM	Marketing digital
4	Ingimp: Introducing Instrumentation to an End-user Open Source Application	Terry, Michael and Kay, Matthew and Van Vugt, Brad and Slack, Brandon and Park, Terry	SIGCHI Conference on Human Factors in Computing Systems	2008	ACM	Marketing
5	Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	CHI 2016	2016	ACM	Evolução de Software
7	What Would They Think?: A Computational Model of Attitudes	Liu, Hugo and Maes, Pattie	9th International Conference on Intelligent User Interfaces	2004	ACM	Marketing digital
9	Validating social media data for automatic persona generation	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	IEEE/ACS International Conference on Computer Systems and Applications	2017	Engineering Village	Marketing digital
12	Archetype-based modeling of persona for comprehensive personality computing from personal big data	Guo, Ao and Ma, Jianhua	Journal Sensors (Switzerland)	2018	Engineering Village	Não definido
17	Scenario-based modeling of online personae for automatic personality perception	Guo, Ao and Ma, Jianhua	SmartWorld / SCALCOM/ UIC/ ATC/ CBDCOM/ IOP/ SCI 2017	2017	Engineering Village	Não definido
21	Automatically conceptualizing social media analytics data via personas	Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	ICWSM 2018	2018	Engineering Village	Marketing Digital
23	An Integrative and Precise Approach In Personality Computing Based on Ontic Personae Modeling	Guo, Ao and Ma, Jianhua	IEEE Cyber Science and Technology Congress, DASC-FICom-DataCom-CyberSciTe c	2018	Engineering Village	Não definido

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre o contexto de aplicação da persona
						Contexto
35	Towards Automatic Persona Generation Using Social Media	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohammed Z. and Jansen, Bernard J.	4th International Conference on Future Internet of Things and Cloud Workshops (FICloudW)	2016	IEEE	Marketing digital
36	Generating Cultural Personas from Social Data: A Perspective of Middle Eastern Users	Salminen, Joni and Sengün, Sercan and Kwak, Haewoon and Jansen, Bernard J. and An, Jisun and Jung, Soon-Gyo and Vieweg, Sarah and Harrell, D. Fox	5th International Conference on Future Internet of Things and Cloud Workshops (FICloudW)	2017	IEEE	Marketing digital
40	Leveraging Social Analytics Data for Identifying Customer Segments for Online News Media	Jansen, Bernard J. and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)	2017	IEEE	Marketing digital
2272	Personas for content creators via decomposed aggregate audience statistics	An, Jisun and Kwak, Haewoon and Jansen, Bernard J	International Conference on Advances in Social Networks Analysis and Mining 2017	2017	Google Academic	Marketing digital
2274	Viewed by too many or viewed too little: Using information dissemination for audience segmentation	Jansen, Bernard J and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	Association for Information Science and Technology	2017	Google Academic	Marketing digital
2280	Inferring Social Media Users' Demographics from Profile Pictures: A Face++ Analysis on Twitter Users	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Salminen, Joni and Jansen, Bernard J	ICEB 2017 Proceedings	2017	Google Academic	Marketing digital
2282	Fixation and Confusion: Investigating Eye-tracking Participants' Exposure to Information in Personas	Salminen, Joni and Jansen, Bernard J and An, Jisun and Jung, Soon-Gyo and Nielsen, Lene and Kwak, Haewoon	Conference on Human Information Interaction and Retrieval 2018	2018	Google Academic	Marketing digital
2283	"Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles	Salminen, Joni and Nielsen, Lene and Jung, Soon-gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J	CHI 2018	2018	Google Academic	Marketing digital
2285	Customer segmentation using online platforms: Isolating behavioral and demographic segments for persona creation via aggregated user data	An, Jisun and Kwak, Haewoon and Jung, Soon-gyo and Salminen, Joni and Jansen, Bernard J	Social Network Analysis and Mining	2018	Google Academic	Marketing digital
2326	ID3P: Iterative Data-Driven Development of Persona based on quantitative evaluation and revision	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Noyori, Yuki and Fukazawa, Yoshiaki and Morizuki, Aoi and Shibata, Hiroyuki and Ogawa, Kentaro and Ishigaki, Mikako and Shizaki, Satiyo and others	10th International Workshop on Cooperative and Human Aspects of Software Engineering	2017	Google Academic	Marketing
2329	Retrospective based on data-driven persona significance in B-to-B software development	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Fukazawa, Yoshiaki and Taga, Masahiro and Matsuzaki, Akira and Suzuki, Takayoshi	40th International Conference on Software Engineering : New Ideas and Emerging Results	2018	Google Academic	Desenvolvimento de software



ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre o contexto de aplicação da persona
						Contexto
2331	Persona Perception Scale: Developing and Validating an Instrument for Human-Like Representations of Data	Salminen, Joni and Kwak, Haewoon and Santos, Jo (a) M and Jung, Soon-Gyo and An, Jisun and Jansen, Bernard J	CHI 2018	2018	Google Academic	Marketing digital
2334	The Potential and Issues in Data-Driven Development of Web Personas	Mijac, Tea and Jadric, Mario and Cukusic	MIPRO 2018	2018	Google Academic	Marketing

## **APÊNDICE E – FORMULÁRIO DE EXTRAÇÃO - INFORMAÇÕES SOBRE A VALIDAÇÃO**

Este documento apresenta os dados extraídos, quanto a validação das personas, de acordo com o formulário 4 apresentado na Seção 3.2.4.

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a validação		
						Técnica utilizada	Aplicação da técnica	O que validou?
0	Findings of a User Study of Automatically Generated Personas	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	CHI 2018	2018	ACM	Teste com usuários, aplicação de surveys qualitativo e quantitativo	<p>Teste: foi dado um cenário e o participante deveria gerar as personas utilizando o sistema. Durante o teste, foi realizada uma análise dos logs de interação dos participantes com o sistema. Com ele, foram apontadas as características da persona que tiveram mais visibilidade (as que o participante mais interagiu).</p> <p>Survey quantitativo: os participantes poderiam classificar o quanto cada parte da persona havia gerado, variando entre 2 e 5. Para cada característica foi gerada uma média de avaliação. Survey qualitativo: os usuários puderam escrever o que gostariam de ter no sistema</p>	Uso da persona
1	Algoritmos De Clusterização e Python Científico Apoiando Modelagem De Usuário	Masiero, Andrey Araujo and Ferreira, Leonardo A. and Aquino, Jr., Plinio Thomaz	11th Brazilian Symposium on Human Factors in Computing Systems	2012	ACM		Não houve validação	
2	Automatic Persona Generation (APG): A Rationale and Demonstration	Jung, Soon-gyo and Salminen, Joni and Kwak, Haewoon and An, Jisun and Jansen, Bernard J.	Conference on Human Information Interaction and Retrieval 2018	2018	ACM		Não houve validação	
3	Persona Generation from Aggregated Social Media Data	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeed and Nielsen, Lene and Jansen, Bernard J.	CHI 2017	2017	ACM		Não houve validação	
4	Ingimp: Introducing Instrumentation to an End-user Open Source Application	Terry, Michael and Kay, Matthew and Van Vugt, Brad and Slack, Brandon and Park, Terry	SIGCHI Conference on Human Factors in Computing Systems	2008	ACM	Entrevista e Think-Aboud	<p>Seis sujeitos participaram do estudo inicial, sendo todos alunos de graduação. Cada usuário foi solicitado a navegar até o site, localizar o software e fazer o download e instalá-lo. Eles foram então solicitados a criar um logotipo para o software. Quando terminamos, pedimos aos participantes que visualizassem suas estatísticas no site. As sessões duraram 45 a 60 minutos cada. Quando as personas geradas foram apresentadas, foram solicitadas sugestões para apresentação das personas.</p>	Personas

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a validação		
						Técnica utilizada	Aplicação da técnica	O que validou?
5	Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	CHI 2016	2016	ACM	Entrevista	Entrevistas com especialistas nos produtos para garantir que os fluxos de trabalho e as metas de persona representem o uso real do produto.	Uso da persona
7	What Would They Think?: A Computational Model of Attitudes	Liu, Hugo and Maes, Pattie	9th International Conference on Intelligent User Interfaces	2004	ACM	Experimento controlado	4 participantes que possuíam entre 18 e 28 anos e mantinham um blog por pelo menos 2 anos	Uso da persona
9	Validating social media data for automatic persona generation	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	IEEE/ACS International Conference on Computer Systems and Applications	2017	Engineering Village		Não houve validação	
12	Archetype-based modeling of persona for comprehensive personality computing from personal big data	Guo, Ao and Ma, Jianhua	Journal Sensors (Switzerland)	2018	Engineering Village	Estudo de caso	Para criar diferentes cenários a partir de vários tipos de dados pessoais, o comportamento, a localização e as pessoas de contato de um testador são coletados de acordo. A localização é determinada pelos dados do GPS coletados em um smartphone. Enquanto o comportamento do testador é detectado em vários dispositivos vestíveis, por exemplo, relógio inteligente. E a pessoa envolvida em um cenário diferente é obtida a partir do feedback do testador. Depois, foram criadas 6 personas óticas para oito cenários diferentes. Aplicação do questionário e fazer um cálculo de comparação com os dados obtidos através dos dados extraídos pela ferramenta.	Persona
17	Scenario-based modeling of ontic personae for automatic personality perception	Guo, Ao and Ma, Jianhua	SmartWorld / SCALCOM/ UIC/ ATC/ CBDCom/ IOP/ SCI 2017	2017	Engineering Village	Estudo de caso	Aplicação das fases do modelo em um cenário, coletando dados dos participantes por 16 horas.	Uso da persona
21	Automatically conceptualizing social media analytics data via personas	Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	ICWSM 2018	2018	Engineering Village		Não foi realizada	
23	An Integrative and Precise Approach in Personality Computing Based on Ontic Personae Modeling	Guo, Ao and Ma, Jianhua	IEEE Cyber Science and Technology Congress, DASC-PICom-DataCom-CyberSciTec	2018	Engineering Village	Estudo de caso	Para criar diferentes cenários a partir de vários tipos de dados pessoais, o comportamento, a localização e as pessoas de contato de um testador são coletados de acordo. A localização é determinada pelos dados do GPS coletados em um smartphone. Enquanto o comportamento do testador é detectado em vários dispositivos vestíveis, por exemplo, relógio inteligente. E a pessoa envolvida em um cenário diferente é obtida a partir do feedback do testador. Depois, foram criadas 6 personas óticas para oito cenários diferentes. Aplicação do questionário e fazer um cálculo de comparação com os dados obtidos através dos dados extraídos pela ferramenta.	Persona

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a validação		
						Técnica utilizada	Aplicação da técnica	O que validou?
35	Towards Automatic Persona Generation Using Social Media	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohammed Z. and Jansen, Bernard J.	4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	2016	IEEE		Não houve validação	
36	Generating Cultural Personas from Social Data: A Perspective of Middle Eastern Users	Salminen, Joni and Sengün, Sercan and Kwak, Haewoon and Jansen, Bernard J. and An, Jisun and Jung, Soon-Gyo and Vieweg, Sarah and Harrell, D. Fox	5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)	2017	IEEE	Entrevista	Foram realizadas entrevistas com pessoas de perfis parecidos com os das personas, para entender melhor os comportamentos das personas geradas.	Uso da persona
40	Leveraging Social Analytics Data for Identifying Customer Segments for Online News Media	Jansen, Bernard J. and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)	2017	IEEE		Não realizado	
2272	Personas for content creators via decomposed aggregate audience statistics	An, Jisun and Kwak, Haewoon and Jansen, Bernard J	International Conference on Advances in Social Networks Analysis and Mining 2017	2017	Google Academic	Estudo de caso	Obtiveram dados dos usuários da conta AJ+, um canal online de notícias. Após a extração das informações, analisaram os dados e criaram 10 personas	Dados das personas
2274	Viewed by too many or viewed too little: Using information dissemination for audience segmentation	Jansen, Bernard J and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	Association for Information Science and Technology	2017	Google Academic	Estudo de caso	Obtiveram dados dos usuários da conta AJ+, um canal online de notícias. Após a extração das informações, analisaram os dados e criaram 10 personas	Dados das personas

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a validação		
						Técnica utilizada	Aplicação da técnica	O que validou?
2280	Inferring Social Media Users' Demographics from Profile Pictures: A Face++ Analysis on Twitter Users	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Salminen, Joni and Jansen, Bernard J	ICEB 2017 Proceedings	2017	Google Academic		Não apresenta	
2282	Fixation and Confusion: Investigating Eye-tracking Participants' Exposure to Information in Personas	Salminen, Joni and Jansen, Bernard J and An, Jisun and Jung, Soon-Gyo and Nielsen, Lene and Kwak, Haewoon	Conference on Human Information Interaction and Retrieval 2018	2018	Google Academic	Experimento Controlado	Foi analisado o <i>eye-tracking</i> dos participantes, para testar o melhor design para a apresentação das informações das personas na ferramenta	Uso das personas
2283	"Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles	Salminen, Joni and Nielsen, Lene and Jung, Soon-gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J	CHI 2018	2018	Google Academic	Experimento Controlado	Foram realizadas entrevistas e analisado o <i>eye-tracking</i> dos participantes, com o intuito de descobrir qual a melhor opção para apresentação da persona. Se: (a) várias fotos da mesma pessoa em diferentes situações contextuais, neste caso, a persona vista no trabalho e situações de lazer; e (b) várias fotos de pessoas diferentes, mas todas com as propriedades da persona que são particularmente visíveis em fotos, como sexo, etnia e idade aproximada	Uso das personas
2285	Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data	An, Jisun and Kwak, Haewoon and Jung, Soon-gyo and Salminen, Joni and Jansen, Bernard J	Social Network Analysis and Mining	2018	Google Academic	Análise quantitativa dos dados obtidos	Foram realizadas duas análises usando uma metodologia de dez vezes no conjunto de dados. As duas análises são (a) a previsão do cliente demográfico mais impactante para um determinado segmento comportamental do cliente e (b) a previsão de novas visualizações de vídeo por grupo demográfico.	Dados das personas
2326	ID3P: Iterative Data-Driven Development of Persona based on quantitative evaluation and revision	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Noyori, Yuki and Fukazawa, Yoshiaki and Morizuki, Aoi and Shibata, Hiroyuki and Ogawa, Kentaro and Ishigaki, Mikako and Shiizaki, Satiyo and others	10th International Workshop on Cooperative and Human Aspects of Software Engineering	2017	Google Academic	Estudo de caso	Aplicou o ID3P no serviço Netallica, serviço que provém artigos no web para usuários. Foram usados dois tipos de conjuntos de dados: D1) uma pesquisa por questionário implementada por uma empresa de pesquisa e D2) os dados de registro de usuários em um serviço. O questionário foi preenchido por 723 usuários da Netallica, incluindo 217 usuários em seus 20 anos, selecionados aleatoriamente por uma empresa de pesquisa.	Uso das personas

ID	Título	Autor(es)	Local de publicação	Ano	Base de dados	Sobre a validação		
						Técnica utilizada	Aplicação da técnica	O que validou?
2329	Retrospective based on data-driven persona significance in B-to-B software development	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Fukazawa, Yoshiaki and Taga, Masahiro and Matsuzaki, Akira and Suzuki, Takayoshi	40th International Conference on Software Engineering : New Ideas and Emerging Results	2018	Google Academic	Estudo de caso	Calcula o PerSil, cálculo que reflete o quanto os desenvolvedores consideram usuários finais reais, concentrando-se nos usuários-alvo. Para isso, foi realizada uma análise quantitativa	Uso das personas
2331	Persona Perception Scale: Developing and Validating an Instrument for Human-Like Representations of Data	Salminen, Joni and Kwak, Haewoon and Santos, Jo (Jo M) and Jung, Soon-Gyo and An, Jisun and Jansen, Bernard J	CHI 2018	2018	Google Academic	Entrevista com experts e Aplicação de Survey	Verificaram se as personas apresentadas na ferramenta estavam de acordo com as seguintes medidas, com base em uma revisão da literatura: Credibilidade, Clareza, Completude, Consistência, Consistência, Empatia, Familiaridade, Amigável, Interpersonal attraction, Liking, Similaridade e Utilidade. Com a aplicação do survey, fizeram os seguintes passos: Primeiro, uma Análise Fatorial Exploratória (EFA) inicial foi feita para identificar dimensões com o potencial de ter múltiplos fatores. Segundo, uma Análise Fatorial Confirmatória (AFC) foi conduzida separadamente para cada dimensão (devido ao baixo tamanho da amostra). Finalmente, os Alfas de Cronbach foram calculados para cada subescala para estimar a confiabilidade.	Uso das personas
2334	The Potential and Issues in Data-Driven Development of Web Personas	Mijac, Tea and Jadric, Mario and Cukusic	MIPRO 2018	2018	Google Academic		Não apresenta	

## **APÊNDICE F – CLASSIFICAÇÃO DOS ESTUDOS DE ACORDO COM OS CRITÉRIOS DE QUALIDADE**

Este documento apresenta a classificação dos estudos de acordo com os Critérios de Qualidade, de acordo com o formulário 3 apresentado na Seção 3.2.5.



ID	Título	Autor(es)	Ano	Base de dados	Critérios de Qualidade				Total
					Apresenta ferramenta?	Extrai dados automaticamente?	Descreve o processo de geração das personas?	Realiza validação das personas?	
0	Findings of a User Study of Automatically Generated Personas	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	2018	ACM	1	1	1	1	4
7	What Would They Think?: A Computational Model of Attitudes	Liu, Hugo and Maes, Pattie	2004	ACM	1	1	1	1	4
12	Archetype-based modeling of persona for comprehensive personality computing from personal big data	Guo, Ao and Ma, Jianhua	2018	Engineering Village	1	1	1	1	4
23	An Integrative and Precise Approach in Personality Computing Based on Ontic Personae Modeling	Guo, Ao and Ma, Jianhua	2018	Engineering Village	1	1	1	1	4
36	Generating Cultural Personas from Social Data: A Perspective of Middle Eastern Users	Salminen, Joni and Sengün, Sercan and Kwak, Haewoon and Jansen, Bernard J. and An, Jisun and Jung, Soon-Gyo and Vieweg, Sarah and Harrell, D. Fox	2017	IEEE	1	1	1	1	4
2274	Viewed by too many or viewed too little: Using information dissemination for audience segmentation	Jansen, Bernard J and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	2017	Google Academic	1	1	1	1	4
2283	"Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles	Salminen, Joni and Nielsen, Lene and Jung, Soon-gyo and An, Jisun and Kwak, Haewoon and Jansen, Bernard J	2018	Google Academic	1	1	1	1	4
2285	Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data	An, Jisun and Kwak, Haewoon and Jung, Soon-gyo and Salminen, Joni and Jansen, Bernard J	2018	Google Academic	1	1	1	1	4
2	Automatic Persona Generation (APG): A Rationale and Demonstration	Jung, Soon-gyo and Salminen, Joni and Kwak, Haewoon and An, Jisun and Jansen, Bernard J.	2018	ACM	1	1	1	0	3
3	Persona Generation from Aggregated Social Media Data	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeed and Nielsen, Lene and Jansen, Bernard J.	2017	ACM	1	1	1	0	3
35	Towards Automatic Persona Generation Using Social Media	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohammed Z. and Jansen, Bernard J.	2016	IEEE	1	1	1	0	3
40	Leveraging Social Analytics Data for Identifying Customer Segments for Online News Media	Jansen, Bernard J. and Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	2017	IEEE	1	1	1	0	3
2272	Personas for content creators via decomposed aggregate audience statistics	An, Jisun and Kwak, Haewoon and Jansen, Bernard J	2017	Google Academic	0	1	1	1	3
2282	Fixation and Confusion: Investigating Eye-tracking Participants' Exposure to Information in Personas	Salminen, Joni and Jansen, Bernard J and An, Jisun and Jung, Soon-Gyo and Nielsen, Lene and Kwak, Haewoon	2018	Google Academic	1	1	0	1	3
2326	ID3P: Iterative Data-Driven Development of Persona based on quantitative evaluation and revision	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Noyori, Yuki and Fukazawa, Yoshiaki and Morizuki, Aoi and Shibata, Hiroyuki and Ogawa, Kentaro and Ishigaki, Mikako and Shiizaki, Satiyo and others	2017	Google Academic	0	1	1	1	3

ID	Título	Autor(es)	Ano	Base de dados	Critérios de Qualidade				Total
					Apresenta ferramenta?	Extrai dados automaticamente?	Descreve o processo de geração das personas?	Realiza validação das personas?	
2329	Retrospective based on data-driven persona significance in B-to-B software development	Watanabe, Yasuhiro and Washizaki, Hironori and Honda, Kiyoshi and Fukazawa, Yoshiaki and Taga, Masahiro and Matsuzaki, Akira and Suzuki, Takayoshi	2018	Google Academic	0	1	1	1	3
4	Ingimp: Introducing Instrumentation to an End-user Open Source Application	Terry, Michael and Kay, Matthew and Van Vugt, Brad and Slack, Brandon and Park, Terry	2008	ACM	1	1	0	0	2
5	Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	2016	ACM	0	1	0	1	2
9	Validating social media data for automatic persona generation	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	2016	Engineering Village	0	1	1	0	2
17	Scenario-based modeling of ontic personae for automatic personality perception	Guo, Ao and Ma, Jianhua	2017	Engineering Village	0	0	1	1	2
21	Automatically conceptualizing social media analytics data via personas	Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	2018	Engineering Village	1	1	0	0	2
2280	Inferring Social Media Users' Demographics from Profile Pictures: A Face++ Analysis on Twitter Users	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Salminen, Joni and Jansen, Bernard J	2017	Google Academic	0	1	0	1	2
2331	Persona Perception Scale: Developing and Validating an Instrument for Human-Like Representations of Data	Salminen, Joni and Kwak, Haewoon and Santos, João M and Jung, Soon-Gyo and An, Jisun and Jansen, Bernard J	2018	Google Academic	1	0	0	1	2
2334	The Potential and Issues in Data-Driven Development of Web Personas	Mijac, Tea and Jadric, Mario and Cukusic	2018	Google Academic	0	1	1	0	2
1	Algoritmos De Clusterização e Python Científico Apoiando Modelagem De Usuário	Masiero, Andrey Araujo and Ferreira, Leonardo A. and Aquino, Jr., Plinio Thomaz	2012	ACM	0	0	1	0	1

## **APÊNDICE G – DADOS UTILIZADOS PARA A GERAÇÃO DE PERSONAS**

Este documento apresenta os dados utilizados para a geração das personas, apresentada na Seção 6.3.

Postagem	Classificação da postagem			Usuário					
	PRU?	Funcionalidade	Tipo	Tempo de uso	Popularidade	Localização	Dispositivo	Gênero	Idade
Teve uma época que eu tinha receio de falar qualquer coisa que fosse no Twitter porque um certo @ sempre reclamava... <a href="https://t.co/L0jYwYQ67q">https://t.co/L0jYwYQ67q</a>	Sim	Escrever tweet	Crítica	2 anos		0 Ceará, Brasil	iPhone	Feminino	25
essas horas eu to muito agradecida por ter twitter porque senão eu nem me sentiria livre pra escrever um oi em outra rede social como o fb	Sim	Escrever tweet	Elogio	2 anos		0 Duque de Caxias, Rio de Janeiro	Mobile Web	Feminino	20
Ei, você! Usuário do #facebook. Bem vindo ao @twitter. Aqui as coisas são um pouquinho diferentes. Não peça likes,... <a href="https://t.co/MhJQNX8Q22">https://t.co/MhJQNX8Q22</a>	Sim	Curtir tweet	Ajuda	4 anos	0.5	Brazil	Web	Masculino	22
Eu não posso falar o nome aqui no Twitter porque aparece na clipagem toda vez que alguém cita eles. En... <a href="https://t.co/SS7WHk36vU">https://t.co/SS7WHk36vU</a>	Sim	Mencionar usuário	Crítica	8 anos	1	Florianópolis, Brasil	iPhone	Feminino	24
Publiquei aqui no Twitter mais cedo. Obrigado por divulgar	Sim	Escrever tweet	Elogio	9 anos	3.5	Rio de Janeiro	Web	Masculino	24
Publiquei aqui no Twitter mais cedo. Obrigado!	Sim	Escrever tweet	Elogio	9 anos	0.5	Rio de Janeiro	Web	Masculino	24
#EthelineFreire Tetê, vim te visitar neste no twitter. Aqui é a deep web das redes sociais... Rs	Sim	Visitar perfil	Elogio	8 anos		0 Brasília, Distrito Federal	Web	Masculino	23
Bem eu instalei pela primeira vez o Twitter porque ya eu gosto de escrever	Sim	Escrever tweet	Elogio	0 dias	0.5	Portugal	iPhone	Feminino	18
enfim, na real eu vim no twitter porque eu queria falar de como eu sou louquíssima pra organizar meus horários de t... <a href="https://t.co/W3rO63y2uy">https://t.co/W3rO63y2uy</a>	Sim	Escrever tweet	Elogio	8 anos		0 São Paulo, Brasil	iPhone	Feminino	28
Num da nem pra reclamar no twitter mais	Sim	Reclamar no twitter	Crítica	3 anos		0 Curitiba, Paraná	Web	Masculino	19
Status: tretando no twitter porquê meu squad do Free fire tá offline	Sim	Escrever tweet	Comparação	8 anos	1	Manaus - AM	Android	Masculino	20
Nesse twitter aqui eu pego a visão de tudo muito rápido kkk	Sim	Feed	Elogio	2 anos	1.5	Belford Roxo, Brasi	Android	Feminino	27
Tava achando alguma coisa estranha usando o Twitter aqui no note. Me dei conta que não estava em modo noturno. Ne... <a href="https://t.co/yUbfB3K8qh">https://t.co/yUbfB3K8qh</a>	Sim	Modo noturno	Crítica	7 anos	1.5	Ijuí - RS	Web	Feminino	19
#dayum_nunes Por isso tenho uma relação amor/ódio com o Twitter, porque é que não posso editar os meus tweets?!!	Sim	Editar tweet	Crítica	3 anos	1.5	Angra do Heroísmc	Android	Masculino	31
Ate baixei o twitter aqui no celular dele pra poder reclamar rs	Sim	Escrever tweet	Elogio	6 anos		0 Florianópolis	iPhone	Feminino	19
Não gosto muito de entrar no Twitter, porque sempre tem tweets sobre o Vasco e analisando o atual momento do clube,... <a href="https://t.co/IN3XgrPVZD">https://t.co/IN3XgrPVZD</a>	Sim	Feed	Crítica	3 anos	0.5	Brasil	Android	Masculino	29

**PRUs sem funcionalidades**

Postagem	Classificação da postagem			Usuário					
	PRU?	Funcionalidade	Tipo	Tempo de uso	Popularidade	Localização	Dispositivo	Gênero	Idade
reinstalei o app do twitter porque desativei meu facebook e preciso manter um espaço para compartilhar as coisas qu	Sim	-	-	9 anos		0 Litoral - SP	Android	Masculino	23
Saí do face e pelo jeito vou ter que sair do twitter porque gente branca tem em todo lugar mesmo, pelamor!	Sim	-	Crítica	2 anos		0.5 Passo Fundo, Brasil	Web	Masculino	31
Eu amo meu Twitter porque aqui não tem nem ninguém da família nem o povo da Igreja, então a anarquia da minha mente... <a href="https://t.co/y4YlqwKZVd">https://t.co/y4YlqwKZVd</a>	Sim	-	Elogio / Crítica / Comparação	1 ano		0.5 Maringá, Brasil	Android	Masculino	25
#rafgabriel Acho Twitter mais perigoso que whats	Sim	-	-	9 anos		0 Brasil	Android	Feminino	29
Por isso eu uso twitterAqui é todo mundo fudido e ninguém tem vergonha de mostrar isso <a href="https://t.co/6SUFf8A4jl">https://t.co/6SUFf8A4jl</a>	Sim	-	Elogio	7 anos		1 Campo Limpo, São	Android	Masculino	28
twitter mais uma vez n deixou eu fazer meus trabalhos	Sim	-	Crítica	4 anos		1 São Miguel do Oeste	iPhone	Feminino	17
Twitter, mais um vicio pra coleção kk	Sim	-	Elogio	2 anos		1.5 Belo Horizonte, Br	Android	Masculino	20
o twitter mais parado do brasil é o meu	Sim	-	Crítica	2 meses		0 Porto Alegre, Brazil	iPhone	Feminino	21
eu quero amigos ô porra.Eu to muito triste.desabafo pro twitter porque psicologa ta caro, família ta estressada,... <a href="https://t.co/HMuHc3PlpW">https://t.co/HMuHc3PlpW</a>	Sim	-	-	5 anos		11.5 Lages, Santa Catari	iPhone	Feminino	21
não fico mt tempo no twitter mais pq tô ocupada demais dividindo meu tempo entre trabalho/faculdade/stardew valley kkkkk	Sim	-	-	2 anos		0 Recife, Pernambuc	iPhone	Feminino	17

## **ANEXO A – RELATÓRIO DE SELEÇÃO DOS ESTUDOS**

Todo o processo de seleção dos artigos pode ser visto no anexo a seguir. Este relatório foi gerado pela ferramenta StArt, utilizada para gerenciar a revisão sistemática apresentada no Capítulo 3.

ID Paper	Title	ID Search	Authors	Status/Selection	Status/Extraction	(E) CE1	(E) CE2	(E) CE3	(E) CE4	(E) CE5	(E) CE6	(E) CE7	(I) C11	(I) C12	(I) C13
0	Findings of a Us	0	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	N	Y	Y
1	Algoritmos De C	0	Masiero, Andrey Araujo and Ferreira, Leonardo A. and Aquino, Jr., Plinio	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	N	N
2	Automatic Perso	0	Jung, Soon-gyo and Salminen, Joni and Kwak, Haewoon and An, Jisun	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	Y	N
3	Persona Gener:0	0	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeec	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	Y	N
4	Ingimp: Introduc	0	Terry, Michael and Kay, Matthew and Van Vugt, Brad and Slack, Brand	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	N	Y
5	Data-driven Per	0	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	Y	Y
6	Automatic Disc	0	Chattopadhyay, T. and Maiti, Santa and Pal, Arindam and Ghose, Avik	ACCEPTED	REJECTED	N	N	N	N	N	N	N	N	N	N
7	What Would Th	0	Liu, Hugo and Maes, Pattie	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	N	N
8	Towards autom:1	1	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohamr	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
9	Validating socia	1	An, Jisun and Kwak, Haewoon and Jansen, Bernard J.	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	N	N
10	Automa-person:1	1	Masiero, Andrey Araujo and Destro, Ricardo de Carvalho and Curioni, C	ACCEPTED	REJECTED	N	Y	N	N	N	N	N	N	N	N
11	Persona generz:1	1	Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon and Ahmad, Moeec	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
12	Archetype-base	1	Guo, Ao and Ma, Jianhua	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	Y	Y
13	Child-persona:1	1	da Costa, Ana Claudia and Rebelo, Francisco and Teles, Julia	ACCEPTED	REJECTED	N	Y	N	N	N	N	N	N	N	N
14	Integrating hum	1	Kanno, Taro and Ooyabu, Tomohiko and Furuta, Kazuo	ACCEPTED	REJECTED	N	Y	N	N	N	N	N	N	N	N
15	PersonA: Persu	1	Ayubi, Soleh U. and Parmanto, Bambang	REJECTED	UNCLASSIFIED	N	N	N	N	Y	N	N	N	N	N
16	Findings of a us	1	Salminen, Joni and Jung, Soon-Gyo and An, Jisun and Kwak, Haewoon	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
17	Scenario-based	1	Guo, Ao and Ma, Jianhua	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	N	N
18	A study on auto	1	Iizuka, Shigeyoshi	REJECTED	UNCLASSIFIED	N	N	N	N	N	Y	N	N	N	N
19	Generating culti	1	Salminen, Joni and Sengun, Sercan and Kwak, Haewoon and Jansen, B	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
20	Data-driven per	1	Zhang, Xiang and Brown, Hans-Frederick and Shankar, Anil	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
21	Automatically cc	1	Jung, Soon-Gyo and Salminen, Joni and An, Jisun and Kwak, Haewoon	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	Y	N
22	Creating persor	1	Masiero, Andrey Araujo and Aquino, Plinio Thomaz	ACCEPTED	REJECTED	N	Y	N	N	N	N	N	N	N	N
23	An Integrative a	1	Guo, Ao and Ma, Jianhua	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	Y	Y
24	Neural personal	1	Zhang, Wei-Nan and Zhu, Qingfu and Wang, Yifa and Zhao, Yanyan an	REJECTED	UNCLASSIFIED	N	N	N	N	N	Y	N	N	N	N
25	Role transition r	1	Abidin, Siti Z. Z and Omar, Nasiroh and Idrus, Zainura	REJECTED	UNCLASSIFIED	N	N	N	N	N	Y	N	N	N	N
26	18th Internation	1		REJECTED	UNCLASSIFIED	N	N	N	Y	N	N	N	N	N	N
27	18th Internation	1		REJECTED	UNCLASSIFIED	N	N	N	Y	N	N	N	N	N	N
28	The RDFa cont:1	1	Khaili, Ali and Auer, Soren and Hladky, Daniel	REJECTED	UNCLASSIFIED	N	N	N	N	N	Y	N	N	N	N
29	24th Internation	1		REJECTED	UNCLASSIFIED	N	N	N	Y	N	N	N	N	N	N
30	18th European	1		REJECTED	UNCLASSIFIED	N	N	N	Y	N	N	N	N	N	N
31	Planning-based	1	Kurlander, David and Ling, Daniel T.	REJECTED	UNCLASSIFIED	N	N	N	N	N	Y	N	N	N	N
32	Requirements E	1		REJECTED	UNCLASSIFIED	N	N	N	Y	N	N	N	N	N	N
33	An Integrative a	2	A. Guo and J. Ma	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
34	Validating socia	2	J. An and H. Kwak and B. J. Jansen	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
35	Towards Autom	2	An, Jisun and Cho, Hoyoun and Kwak, Haewoon and Hassen, Mohamr	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	Y	N
36	Generating Cult	2	Salminen, Joni and Sengün, Sercan and Kwak, Haewoon and Jansen, B	ACCEPTED	ACCEPTED	N	N	N	N	N	N	N	Y	Y	Y
37	Automated Play	2	Holmgard, Christoffer and Green, Michael C. and Liapis, Antonios and T	ACCEPTED	REJECTED	N	N	N	N	N	N	N	N	N	N
38	An Expressive I	2	D. Fett and R. Küsters and G. Schmitz	REJECTED	UNCLASSIFIED	N	N	N	N	N	Y	N	N	N	N
39	The RDFa Cont	2	A. Khaili and S. Auer and D. Hladky	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
40	Leveraging Soc	2	Jansen, Bernard J. and Jung, Soon-Gyo and Salminen, Joni and An, Jis	ACCEPTED	ACCEPTED	N	N	N	N	Y	N	N	Y	N	N
41	Neural personal	3	Zhang, W.-N. and Zhu, Q. and Wang, Y. and Zhao, Y. and Liu, T.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
42	Findings of a us	3	Salminen, J. and Jung, S.-G. and An, J. and Kwak, H. and Jansen, B.J.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
43	An Integrative a	3	Guo, A. and Ma, J.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
44	Leveraging soci	3	Jansen, B.J. and Jung, S.-G. and Salminen, J. and An, J. and Kwa, H.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
45	Automatic Perso	3	Jung, S.-G. and Salminen, J. and Kwak, H. and An, J. and Jansen, B.J.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
46	Automatically cc	3	Jung, S.-G. and Salminen, J. and An, J. and Kwak, H. and Jansen, B.J.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
47	24th Internation	3		DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
48	Generating culti	3	Salminen, J. and ÅžengÅ¼n, S. and Kwak, H. and Jansen, B. and An, J	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
49	Validating socia	3	An, J. and Kwak, H. and Jansen, B.J.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
50	A study on auto	3	Iizuka, S.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
51	19th Internation	3		REJECTED	UNCLASSIFIED	N	N	N	Y	N	N	N	N	N	N
52	Towards autom:3	3	An, J. and Cho, H. and Kwak, H. and Hassen, M.Z. and Jansen, B.J.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
53	18th Internation	3		DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
54	18th Internation	3		DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
55	18th European	3		DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
56	An expressive r	3	Fett, D. and KÅ¼sters, R. and Schmitz, G.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
57	Requirements E	3		DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
58	Automa-person:3	3	Masiero, A.A. and Destro, R.C. and Curioni, O.A. and Aquino Jr., P.T.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
59	15th Internation	3		REJECTED	UNCLASSIFIED	N	N	N	Y	N	N	N	N	N	N
60	The RDFa cont:3	3	Khaili, A. and Auer, S. and Hladky, D.	DUPLICATED	UNCLASSIFIED	N	N	Y	N	N	N	N	N	N	N
61	SAGAS system:3	3	Villena, J. and Moreno, L. and MartÅ¼ez, P. and GonzÅ¼lez, J.C.	REJECTED	UNCLASSIFIED	N	N	N	N	N	Y	N	N	N	N









## ANEXO B – RESULTADO DE EXTRAÇÃO DE UMA POSTAGEM UTILIZANDO A API DO TWITTER

```

object(stdClass)#158 (2) { ["statuses"]=> array(15) { [0]=> object(stdClass)#2 (24) { ["created_at"]=>
string(30) "Wed Aug 22 12:52:03 +0000 2018" ["id"]=> float(1.0322490380397E+18) ["id_str"]=>
string(19) "1032249038039654406" ["text"]=> string(68) "Aguardando meu plano no dezzer acabar pra
eu voltar pro Spotify 😊" ["truncated"]=> bool(false) ["entities"]=> object(stdClass)#3 (4) {
["hashtags"]=> array(0) { } ["symbols"]=> array(0) { } ["user_mentions"]=> array(0) { } ["urls"]=> array(0) {
} } ["metadata"]=> object(stdClass)#4 (2) { ["iso_language_code"]=> string(2) "pt" ["result_type"]=>
string(6) "recent" } ["source"]=> string(82) "Twitter for iPhone" ["in_reply_to_status_id"]=> NULL
["in_reply_to_status_id_str"]=> NULL ["in_reply_to_user_id"]=> NULL ["in_reply_to_user_id_str"]=>
NULL ["in_reply_to_screen_name"]=> NULL ["user"]=> object(stdClass)#5 (42) { ["id"]=> int(114599842)
["id_str"]=> string(9) "114599842" ["name"]=> string(12) "Brunno Neves" ["screen_name"]=> string(11)
"NevesBrunno" ["location"]=> string(6) "Brasil" ["description"]=> string(49) "Muitos sabem quem eu
sou, mas poucos me conhecem." ["url"]=> string(23) "https://t.co/o1niivYDvt" ["entities"]=>
object(stdClass)#8 (2) { ["url"]=> object(stdClass)#7 (1) { ["urls"]=> array(1) { [0]=> object(stdClass)#6 (4)
{ ["url"]=> string(23) "https://t.co/o1niivYDvt" ["expanded_url"]=> string(32)
"http://Instagram.com/nevesbrunno" ["display_url"]=> string(25) "Instagram.com/nevesbrunno"
["indices"]=> array(2) { [0]=> int(0) [1]=> int(23) } } } } ["description"]=> object(stdClass)#9 (1) { ["urls"]=>
array(0) { } } } ["protected"]=> bool(false) ["followers_count"]=> int(9075) ["friends_count"]=> int(333)
["listed_count"]=> int(139) ["created_at"]=> string(30) "Tue Feb 16 00:35:39 +0000 2010"
["favourites_count"]=> int(3276) ["utc_offset"]=> NULL ["time_zone"]=> NULL ["geo_enabled"]=>
bool(true) ["verified"]=> bool(false) ["statuses_count"]=> int(6415) ["lang"]=> string(2) "pt"
["contributors_enabled"]=> bool(false) ["is_translator"]=> bool(false) ["is_translation_enabled"]=>
bool(false) ["profile_background_color"]=> string(6) "131516" ["profile_background_image_url"]=>
string(49) "http://abs.twimg.com/images/themes/theme14/bg.gif"
["profile_background_image_url_https"]=> string(50)
"https://abs.twimg.com/images/themes/theme14/bg.gif" ["profile_background_tile"]=> bool(true)
["profile_image_url"]=> string(75)
"http://pbs.twimg.com/profile_images/1031169648866287617/RIYLktCq_normal.jpg"
["profile_image_url_https"]=> string(76)
"https://pbs.twimg.com/profile_images/1031169648866287617/RIYLktCq_normal.jpg"
["profile_banner_url"]=> string(58) "https://pbs.twimg.com/profile_banners/114599842/1515692305"
["profile_link_color"]=> string(6) "1B95E0" ["profile_sidebar_border_color"]=> string(6) "FFFFFF"
["profile_sidebar_fill_color"]=> string(6) "EFEFEF" ["profile_text_color"]=> string(6) "333333"
["profile_use_background_image"]=> bool(true) ["has_extended_profile"]=> bool(true)
["default_profile"]=> bool(false) ["default_profile_image"]=> bool(false) ["following"]=> bool(false)
["follow_request_sent"]=> bool(false) ["notifications"]=> bool(false) ["translator_type"]=> string(7)
"regular" } ["geo"]=> NULL ["coordinates"]=> NULL ["place"]=> NULL ["contributors"]=> NULL
["is_quote_status"]=> bool(false) ["retweet_count"]=> int(0) ["favorite_count"]=> int(0) ["favorited"]=>
bool(false) ["retweeted"]=> bool(false) ["lang"]=> string(2) "pt" }

```