



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**DEPARTAMENTO DE COMPUTAÇÃO**  
**CURSO DE DOUTORADO EM CIÊNCIAS DA COMPUTAÇÃO**

**MARÍLIA SOARES MENDES**

**MALTU – UM MODELO PARA AVALIAÇÃO DA INTERAÇÃO EM  
SISTEMAS SOCIAIS A PARTIR DA LINGUAGEM TEXTUAL DO  
USUÁRIO**

**FORTALEZA — CE**

**2015**

**MARÍLIA SOARES MENDES**

**MALTU – UM MODELO PARA AVALIAÇÃO DA INTERAÇÃO EM  
SISTEMAS SOCIAIS A PARTIR DA LINGUAGEM TEXTUAL DO  
USUÁRIO**

Tese submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Doutor em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Prof. Dr. Miguel Franklin de Castro

Coorientadora: Profa. Dra. Maria Elizabeth Sucupira Furtado

**FORTALEZA — CE**

**2015**

Dados Internacionais de Catalogação na  
Publicação  
Universidade Federal do Ceará  
Biblioteca de Ciências e Tecnologia

---

M492m

Mendes, Marília Soares.

Maltu – um modelo para avaliação da interação em sistemas sociais a partir da linguagem textual do usuário / Marília Soares Mendes. – 2015.  
201 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Departamento de Computação, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2015.

Área de Concentração: Ciências da Computação.

Orientação: Prof. Dr. Miguel Franklin de Castro.

Coorientação: Profa. Dra. Maria Elizabeth Sucupira Furtado.

1. Interação homem-máquina. 2. Processamento de linguagem natural (Computação). 3. Computação. 4. Mineração de dados. I. Título.

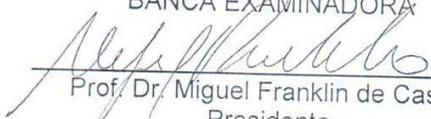
Marília Soares Mendes

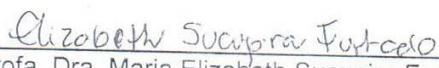
**MALTU - Modelo de Avaliação da Linguagem Textual do Usuário a partir de Postagens em Sistemas Sociais**

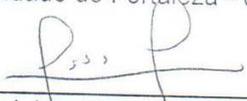
Tese submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação, da Universidade Federal do Ceará, como requisito para a obtenção do grau de Doutor em Ciência da Computação.

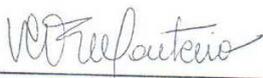
Aprovada em 25 de fevereiro de 2015

BANCA EXAMINADORA

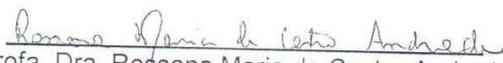
  
Prof. Dr. Miguel Franklin de Castro  
Presidente  
Universidade Federal do Ceará – UFC

  
Profa. Dra. Maria Elizabeth Sucupira Furtado  
Coorientadora  
Universidade de Fortaleza - UNIFOR

  
Prof. Dr. Jaime Hernán Sánchez Ilabaca  
Universidade do Chile

  
Profa. Dra. Vládia Célia Monteiro Pinheiro  
Universidade de Fortaleza - UNIFOR

  
Profa. Dra. Andreia Liborio Sampaio  
Universidade Federal do Ceará – UFC

  
Profa. Dra. Rossana Maria de Castro Andrade  
Universidade Federal do Ceará – UFC

Fortaleza, 25 de fevereiro de 2015

A Deus.

À minha família.

## AGRADECIMENTOS

A Deus! Tudo que o Senhor faz é perfeito! Senhor, obrigada por ter me dado toda a força necessária, ânimo, fé e perseverança para seguir em frente, não me permitindo desistir dos meus sonhos.

Aos meus queridos pais, Marcos e Nilce, por sempre me apoiarem, incentivarem e acreditarem em mim! Mamãe, pode contar para todo mundo agora!

Ao meu marido, Domingos. Casamo-nos em meio a este curso e já sabíamos que não seria fácil! Obrigada! Eu te amo!

Aos meus queridos sogros, Franzé e Edicinha, pelo apoio e compreensão nos momentos de ausência.

Aos irmãos de sangue: Marnilce, Vitoriano e Marcília, e aos irmãos adquiridos: Chiquinho, Eliezer e Nathalie. Tenho muita sorte por ter vocês!

Obrigada, também, a toda a minha família, pois são igualmente responsáveis diretos pela minha formação como pessoa.

Aos amigos compreensivos com minhas ausências, em especial minhas afilhadas queridas: Lyrê e Sâmia, que cresciam enquanto eu escrevia esta tese.

Aos meus orientadores, Miguel Franklin de Castro e Elizabeth Furtado. Prof. Miguel, obrigada pela a liberdade e confiança necessárias para produzir um trabalho que me fizesse sentir realizada por desenvolver. Um agradecimento especial à minha eterna e querida orientadora Profa. Beth, pela sua orientação acadêmica desde o tempo da graduação, que continuou no mestrado e agora no doutorado. Eu nunca terei palavras para agradecer tamanha dedicação, motivação, atenção e amizade. Espero continuar a cooperação, aprendendo e crescendo com seu conhecimento.

À UFC e à coordenação do MDCC, pelo apoio na publicação de alguns dos meus artigos relacionados a este trabalho.

Aos professores Vasco Furtado e Vlória Pinheiro, pelo apoio nas disciplinas que fiz como ouvinte na Unifor (PPGIA) a fim de aperfeiçoar este trabalho. Em especial, ao Prof. Vasco, muito obrigada pelos questionamentos, sugestões e valiosas contribuições!

À CAPES, pelo apoio financeiro com a manutenção da bolsa de auxílio.

Neste período de pesquisa, eu lecionei no Centro Universitário Estácio do Ceará, na Unifor e na UFC, e todos meus superiores foram compreensivos com minha situação de doutoranda. Eu agradeço à Janete Amaral (Coordenadora do curso Análise e

Desenvolvimento de Sistemas, da Estácio-CE), à Josyane de Souza (Coordenadora do Curso Sistemas de Informação, da Estácio-CE), ao Marcus Venicius (Coordenador do curso Análise e Desenvolvimentos de Sistemas, da Unifor) e ao Prof. Lindberg Lima Gonçalves (Diretor do campus da UFC, em Russas-CE).

Ao Laboratório de estudos do Usuário e da Qualidade de Uso de Sistemas (LUQS), do qual eu faço parte desde 2005. Lá, eu tive minha formação em IHC, suporte financeiro e o apoio de bolsistas de graduação. Agradeço, também, a todos seus participantes, pela ajuda, dicas, dúvidas, em especial aos amigos: Rafinha, Pati, Guido, Camila, Berg, Daniel, Ronaldo, Niedja, Dani e Natércia.

Aos alunos Fábio Theophilo e Diego Marino (bolsistas do LUQS), pela ajuda com a ferramenta de extração de postagens, e aos alunos da UECE, Unifor e Estácio, pela classificação manual das postagens. Aos especialistas de IHC convidados: Rafaela Lisboa, Patrícia Vasconcelos, Guido Militão, Camila Maia, Carlos Rosemberg, Ronaldo Mota, Kelma Madeira, Danielly Barboza, Samira Ribeiro e Albert Schilling, pela classificação e análise das postagens. Enfim, a todas as pessoas que, direta ou indiretamente, contribuíram para a realização deste trabalho.

Obrigada!

*“Use a linguagem que quiser, mas você nunca  
poderá dizer senão o que você é.”*

Ralph Waldo Emerson

## RESUMO

A área de Interação Humano-Computador (IHC) tem sugerido muitas formas para avaliar sistemas a fim de melhorar sua usabilidade e a eXperiência do Usuário (UX). O surgimento da web 2.0 permitiu o desenvolvimento de aplicações marcadas pela colaboração, comunicação e interatividade entre seus usuários de uma forma e em uma escala nunca antes observadas. Sistemas Sociais (SS) (e.g., Twitter, Facebook, MySpace, LinkedIn etc.) são exemplos dessas aplicações e possuem características como: frequente troca de mensagens e expressão de sentimentos de forma espontânea. As oportunidades e os desafios trazidos por esses tipos de aplicações exigem que os métodos tradicionais de avaliação sejam repensados, considerando essas novas características. Por exemplo, as postagens dos usuários em SS revelam suas opiniões sobre diversos assuntos, inclusive sobre o que eles pensam do sistema em uso. Esta tese procura testar a hipótese de que as postagens dos usuários em SS fornecem dados relevantes para avaliação da Usabilidade e da UX (UUX) em SS. Durante as pesquisas realizadas na literatura, não foi identificado nenhum modelo de avaliação que tenha direcionado seu foco na coleta e análise das postagens dos usuários a fim de avaliar a UUX de um sistema em uso. Sendo assim, este estudo propõe o MALTU – Modelo para Avaliação da interação em sistemas sociais a partir da Linguagem Textual do Usuário. A fim de fornecer bases para o desenvolvimento do modelo proposto, foram realizados estudos de como os usuários expressam suas opiniões sobre o sistema em língua natural. Foram extraídas postagens de usuários de quatro SS de contextos distintos. Tais postagens foram classificadas por especialistas de IHC, estudadas e processadas utilizando técnicas de Processamento da Linguagem Natural (PLN) e mineração de dados e, analisadas a fim da obtenção de um modelo genérico. O MALTU foi aplicado em dois SS: um de entretenimento e um SS educativo. Os resultados mostram que é possível avaliar um sistema a partir das postagens dos usuários em SS. Tais avaliações são auxiliadas por padrões de extração relacionados ao uso, aos tipos de postagens e às metas de IHC utilizadas na avaliação do sistema.

**Palavras-chave:** Interação Humano-Computador. Usabilidade. Experiência do usuário. Processamento da Linguagem Natural. Mineração de dados.

## **ABSTRACT**

The field of Human Computer Interaction (HCI) has suggested various methods for evaluating systems in order to improve their usability and User eXperience (UX). The advent of Web 2.0 has allowed the development of applications marked by collaboration, communication and interaction among their users in a way and on a scale never seen before. Social Systems (SS) (e.g. Twitter, Facebook, MySpace, LinkedIn etc.) are examples of such applications and have features such as: frequent exchange of messages, spontaneity and expression of feelings. The opportunities and challenges posed by these types of applications require the traditional evaluation methods to be reassessed, taking into consideration these new characteristics. For instance, the postings of users on SS reveal their opinions on various issues, including on what they think of the system. This work aims to test the hypothesis that the postings of users in SS provide relevant data for evaluation of the usability and of UX in SS. While researching through literature, we have not identified any evaluation model intending to collect and interpret texts from users in order to assess the user experience and system usability. Thus, this thesis proposes MALTU - Model for evaluation of interaction in social systems from the Users Textual Language. In order to provide a basis for the development of the proposed model, we conducted a study of how users express their opinions on the system in natural language. We extracted postings of users from four SS of different contexts. HCI experts classified, studied and processed such postings by using Natural Language Processing (PLN) techniques and data mining, and then analyzed them in order to obtain a generic model. The MALTU was applied in two SS: an entertainment and an educational SS. The results show that is possible to evaluate a system from the postings of users in SS. Such assessments are aided by extraction patterns related to the use, to the types of postings and to HCI factors used in system.

**Keywords:** Human Computer Interaction (HCI). Usability. User Experience. Natural Language Processing. Data Mining.

## LISTA DE FIGURAS

FIGURA 1 – TAXONOMIA DE MÉTODOS, TÉCNICAS E INSTRUMENTOS PARA A AVALIAÇÃO EMOCIONAL	37
FIGURA 2 – A DINÂMICA DO PROCESSO DE CONSTRUÇÃO DE UM SPLN	39
FIGURA 3 – VISÃO GERAL DOS DIFERENTES NÍVEIS DE PROCESSAMENTO LINGÜÍSTICO EM PLN	40
FIGURA 4 – ETAPAS DO KDD	43
FIGURA 5 – ABORDAGENS DA MINERAÇÃO NA WEB	46
FIGURA 6 – EXEMPLO DE UMA ÁRVORE DE DECISÃO	48
FIGURA 7 – EXEMPLO DE UMA CLASSIFICAÇÃO POR ÁRVORE DE DECISÃO	49
FIGURA 8 – FLUXO DO PROCESSO DE REVISÃO	56
FIGURA 9 – NUVEM DE PALAVRAS DOS PRINCIPAIS VEÍCULOS DE PUBLICAÇÃO DA REVISÃO	60
FIGURA 10 – NUVEM DE PALAVRAS DAS PALAVRAS-CHAVES DE ARTIGOS RELACIONADOS A SS DOS ANOS 2010, 2011 E 2012	63
FIGURA 11 – FERRAMENTA DE BUSCA DAS POSTAGENS	75
FIGURA 12 – POSTAGENS COM O PADRÃO [TWITTER ERRO]	76
FIGURA 13 – PRINCIPAIS SENTIMENTOS PERCEBIDOS NO SIGAA	103
FIGURA 14 – PRINCIPAIS CAUSAS/FUNCIONALIDADES PERCEBIDAS NO SIGAA	104
FIGURA 15 – PRINCIPAIS ELOGIOS PERCEBIDOS NO SIGAA	104
FIGURA 16 – CLASSIFICAÇÃO DAS PRUs POR ESPECIALISTAS	106
FIGURA 17 – QUADRO BRANCO RESULTANTE DO <i>BRAINSTORM</i> COM ESPECIALISTAS	106
FIGURA 18 – FERRAMENTA TEXTPROCESSINGTOOL	119
FIGURA 19 – TRECHO DO ARQUIVO ARFF GERADO PARA O TWITTER - SUBOBJETIVO 1	119
FIGURA 20 – RESULTADO MINERAÇÃO DO TWITTER – SUBOBJETIVO 1	121
FIGURA 21 - FREQUÊNCIA DOS ATRIBUTOS OBTIDOS PARA AS CLASSES PRU E NÃO-PRU DO TWITTER	122
FIGURA 22 – RESULTADO MINERAÇÃO DO SIGAA – SUBOBJETIVO 1	123
FIGURA 23 – RESULTADO MINERAÇÃO DO TWITTER – SUBOBJETIVO 2	124
FIGURA 24 – RESULTADO MINERAÇÃO DO SIGAA – SUBOBJETIVO 2	125
FIGURA 25 – RESULTADO MINERAÇÃO DO SUBOBJETIVO 3 – GRUPO 1	126
FIGURA 26 – RESULTADO MINERAÇÃO DO SUBOBJETIVO 3 – GRUPO 2	127
FIGURA 27 – RESULTADO MINERAÇÃO DO SUBOBJETIVO 3 – GRUPO 3	128
FIGURA 28 – RESULTADO DA MINERAÇÃO DO TWITTER, ALGORITMO J48, USANDO STEMMING	134
FIGURA 29 – TRECHO DA ÁRVORE GERADA PARA O TWITTER - SUBOBJETIVO 1, USANDO STEMMING	134
FIGURA 30 – PADRÕES DE UUX PARA O CONTEXTO DE JOGOS DE VIDEO GAME E SOFTWARES	136
FIGURA 31 – MODELO MALTU	138
FIGURA 32 – METODOLOGIA DE AVALIAÇÃO DO MODELO MALTU	139

FIGURA 33 – PRU VISCERAL DO TIPO CRÍTICA.....	143
FIGURA 34– PRU VISCERAL DO TIPO ELOGIO.....	143
FIGURA 35 – PRU COMPORTAMENTAL DO TIPO CRÍTICA .....	144
FIGURA 36 – PRU COMPORTAMENTAL DO TIPO DÚVIDA .....	144
FIGURA 37 – PRU REFLEXIVA DO TIPO ELOGIO .....	144
FIGURA 38– PRU REFLEXIVA DO TIPO CRÍTICA .....	145
FIGURA 39 – EXEMPLO DE CLASSIFICAÇÃO DE FUNCIONALIDADE EM UMA PRU.....	146
FIGURA 40 – CLASSIFICAÇÃO DE FUNCIONALIDADE EM UMA PRU .....	146
FIGURA 41– CLASSIFICAÇÃO DE FUNCIONALIDADE EM UMA PRU .....	147
FIGURA 42 – CLASSIFICAÇÃO DE UMA PRU A PARTIR DE METAS DE UUX .....	147
FIGURA 43 – CLASSIFICAÇÃO DE ARTEFATO EM UMA PRU .....	148
FIGURA 44 – CLASSIFICAÇÃO DE ARTEFATO EM UMA PRU .....	149
FIGURA 45 - ELEMENTOS CONSIDERADOS NO CONTEXTO DE AVALIAÇÃO NO MALTU.....	152
FIGURA 46 – TELA DE BUSCA DE POSTAGENS DO UUX-POST.....	153
FIGURA 47 – EXTRAÇÃO DE PRUS EM SS .....	153
FIGURA 48– EXTRAÇÃO DE PRUS A PARTIR DE UM BANCO DE POSTAGENS.....	153
FIGURA 49 – RESULTADO DE UMA BUSCA.....	154
FIGURA 50– RESULTADO DA AVALIAÇÃO DO TWITTER.....	156
FIGURA 51 – RESULTADO DA AVALIAÇÃO DO SIGAA.....	161
FIGURA 52 - ÁRVORE DE DECISÃO GERADA PELO ALGORITMO J48 PARA O TWITTER - SUBOBJETIVO 1 .....	192
FIGURA 53 - TRECHO 1 DA ÁRVORE DE DECISÃO PARA O SIGAA - SUBOBJETIVO 1.....	193
FIGURA 54 - TRECHO 2 DA ÁRVORE DE DECISÃO PARA O SIGAA - SUBOBJETIVO 1.....	194
FIGURA 55 - TRECHO 3 DA ÁRVORE DE DECISÃO PARA O SIGAA - SUBOBJETIVO 1.....	195
FIGURA 56 – ÁRVORE DE DECISÃO (J48) PARA O TWITTER - SUBOBJETIVO 2.....	196
FIGURA 57 – ÁRVORE DE DECISÃO (J48) PARA O SIGAA - SUBOBJETIVO 2 .....	197
FIGURA 58 – TRECHO 1 DA ÁRVORE DE DECISÃO (J48) PARA O SUBOBJETIVO 3 - GRUPO 1.....	198
FIGURA 59 – TRECHO 2 DA ÁRVORE DE DECISÃO (J48) PARA O SUBOBJETIVO 3 - GRUPO 1.....	199
FIGURA 60 – TRECHO DA ÁRVORE DE DECISÃO (J48) PARA O SUBOBJETIVO 3 - GRUPO 2.....	200
FIGURA 61 – TRECHO DA ÁRVORE DE DECISÃO (J48) PARA O SUBOBJETIVO 3 - GRUPO 3.....	200

## LISTA DE GRÁFICOS

GRÁFICO 1 – TEMAS ABORDADOS EM ESTUDOS DE SS .....	62
GRÁFICO 2 – ASPECTOS AVALIADOS EM TRABALHOS COMPLETOS DE AVALIAÇÃO EM SS DOS ANOS 2010, 2011 E 2012.....	63
GRÁFICO 3 – MÉTODOS UTILIZADOS EM TRABALHOS COMPLETOS DE AVALIAÇÃO EM SS DOS ANOS 2010, 2011 E 2012.....	64
GRÁFICO 4 – BOXPLOT DO PADRÃO DE EXTRAÇÃO [TWITTER] COM DEMAIS PADRÕES .....	78
GRÁFICO 5 – PORCENTAGEM DE TIPOS DE PRUS DO TWITTER .....	86
GRÁFICO 6 – PORCENTAGEM DE TIPOS DE PRUS DO SIGAA .....	89
GRÁFICO 7 – PORCENTAGEM DE PRUS OBTIDAS POR CLASSIFICAÇÃO POR INTENÇÃO DO USUÁRIO ...	97
GRÁFICO 8 – CLASSIFICAÇÃO POR INTENÇÃO DO USUÁRIO X TIPOS DE PRUS .....	98
GRÁFICO 9 – CRITÉRIO DE QUALIDADE DE USO = USABILIDADE X TIPO DE PRU = CRÍTICA.....	158
GRÁFICO 10 – CRITÉRIO DE QUALIDADE DE USO = UX X TIPO DE PRU = CRÍTICA.....	158
GRÁFICO 11 – CRITÉRIO DE QUALIDADE DE USO = UX X TIPO DE PRU = ELOGIO.....	158
GRÁFICO 12 – CRITÉRIO DE QUALIDADE DE USO = USABILIDADE X TIPO DE PRU = CRÍTICA.....	162
GRÁFICO 13 – CRITÉRIO DE QUALIDADE DE USO = UX X TIPO DE PRU = CRÍTICA .....	162
GRÁFICO 14 – CRITÉRIO DE QUALIDADE DE USO = UX X TIPO DE PRU = ELOGIO.....	162

## LISTA DE QUADROS

QUADRO 1 – METAS DA UUX .....	31
QUADRO 2– MÉTODOS DE AVALIAÇÃO DOS CRITÉRIOS DE QUALIDADES DE USO EM SiCOS.....	32
QUADRO 3 – MÉTODOS DE AVALIAÇÃO DA UX.....	34
QUADRO 4 – AS PRINCIPAIS DISCIPLINAS QUE ESTUDAM LÍNGUA .....	38
QUADRO 5 – EXEMPLO DE UM CONJUNTO DE DADOS: VERTEBRADOS.....	48
QUADRO 6 – CARACTERÍSTICAS APRESENTADAS POR NORMAN (2004) PARA OS TRÊS NÍVEIS DE PROCESSAMENTO.....	52
QUADRO 7 – EXEMPLOS DE SENTENÇAS OBJETIVAS E SUBJETIVAS COM E SEM A PRESENÇA DE SENTIMENTOS .....	54
QUADRO 8 – TRABALHOS DE AVALIAÇÃO EM SS DOS ANOS 2010, 2011 E 2012 .....	64
QUADRO 9– TRABALHOS CORRELATOS AO TEMA DA TESE.....	70
QUADRO 10 – INVESTIGAÇÕES REALIZADAS.....	72
QUADRO 11 – CLASSIFICAÇÃO DAS BUSCAS E PADRÕES DE EXTRAÇÃO.....	74
QUADRO 12 – EXEMPLOS DAS POSTAGENS OBTIDAS POR CADA PADRÃO.....	82
QUADRO 13 – EXEMPLOS DE TIPOS DE PRUs DO TWITTER .....	87
QUADRO 14 – POSTAGENS OBTIDAS NA ANÁLISE DO SIGAA.....	89
QUADRO 15 – EXEMPLOS DE TIPOS DE PRUs DO SIGAA.....	89
QUADRO 16 – COMENTÁRIO CARACTERIZANDO MAIS DE UM TIPO DE PRU.....	90
QUADRO 17 – DETALHES DO SISTEMA EM UMA PRU .....	91
QUADRO 18 – CLASSIFICAÇÃO DAS PRUs DE ACORDO COM O SENTIMENTO PERCEBIDO .....	94
QUADRO 19 – ASSOCIAÇÃO DAS PRUs COM O MODELO EMOCIONAL DE NORMAN .....	95
QUADRO 20 – EXEMPLO DE POSTAGENS CLASSIFICADAS POR METAS DE UUX .....	101
QUADRO 21 – EXEMPLO DE UM TRECHO DA PLANILHA COM POSTAGENS A SEREM CLASSIFICADAS .	102
QUADRO 22 – AGENDA DA INVESTIGAÇÃO COM ESPECIALISTAS DE IHC.....	105
QUADRO 23 – PRUs QUE CHAMARAM MAIS ATENÇÃO DOS ESPECIALISTAS.....	107
QUADRO 24 – RESULTADO DAS INVESTIGAÇÕES REALIZADAS.....	111
QUADRO 25 – CATEGORIA DE DADOS .....	113
QUADRO 26 – RESULTADO DA APLICAÇÃO COM O ANALISADOR SINTÁTICO PALAVRAS.....	116
QUADRO 27 – LEMMAS COLETADOS .....	117
QUADRO 28 – PADRÕES DE EXTRAÇÃO DE PRUs .....	130
QUADRO 29 – PADRÕES OBTIDO POR TIPOS DE PRUs.....	131
QUADRO 30 – PADRÕES OBTIDOS PARA AS METAS DE UUX.....	132
QUADRO 31 - OBJETIVOS DE AVALIAÇÃO USANDO O MALTU.....	141
QUADRO 32 – DESCRIÇÃO DE CARACTERÍSTICAS, CATEGORIA: TIPOS DE PRUs.....	142

QUADRO 33 – DESCRIÇÃO DE CARACTERÍSTICAS, CATEGORIA: CLASSIFICAÇÃO POR INTENÇÃO .....	143
QUADRO 34 - DESCRIÇÃO DE CARACTERÍSTICAS, CATEGORIA: CLASSIFICAÇÃO DE SENTIMENTOS POR POLARIDADE .....	145
QUADRO 35 – SUGESTÃO DE METAS PARA UUX NO MODELO MALTU .....	147
QUADRO 36– EXEMPLOS DE CLASSIFICAÇÃO DE PRUS USANDO O MODELO MALTU.....	149
QUADRO 37 - SUGESTÕES DE MEDIDAS PARA ANÁLISE DOS OBJETIVOS DE AVALIAÇÃO USANDO O MALTU.....	150
QUADRO 38 - ATIVIDADES DA METODOLOGIA DO MODELO MALTU.....	151
QUADRO 39 – CRITÉRIO DE QUALIDADE DE USO = USABILIDADE X TIPO DE PRU = CRÍTICAS X CAUSA .....	159
QUADRO 40 – CRITÉRIO DE QUALIDADE DE USO = USABILIDADE X TIPO DE PRU = ELOGIO X CAUSA .....	159
QUADRO 41 – CRITÉRIO DE QUALIDADE DE USO = UX X TIPO DE PRU = CRÍTICA X CAUSA .....	159
QUADRO 42 – PRINCIPAIS FUNCIONALIDADES NO CELULAR.....	160
QUADRO 43 – PRINCIPAIS FUNCIONALIDADES QUE OS USUÁRIOS TIVERAM DÚVIDAS .....	160
QUADRO 44 – SS = PRINCIPAIS SUGESTÕES DE FUNCIONALIDADES .....	160
QUADRO 45 – CRITÉRIO DE QUALIDADE DE USO = USABILIDADE X TIPO DE PRU = CRÍTICAS X CAUSA .....	163
QUADRO 46 – CRITÉRIO DE QUALIDADE DE USO = USABILIDADE X TIPO DE PRU = ELOGIO X CAUSA	163
QUADRO 47 – CRITÉRIO DE QUALIDADE DE USO = UX X TIPO DE PRU = CRÍTICA X CAUSA .....	163
QUADRO 48 – PRINCIPAIS FUNCIONALIDADES QUE OS USUÁRIOS TIVERAM DÚVIDAS .....	163
QUADRO 49– PRINCIPAIS SUGESTÕES DE FUNCIONALIDADES .....	164
QUADRO 50– PUBLICAÇÕES RELACIONADAS À TESE .....	167

## LISTA DE TABELAS

TABELA 1 – QUANTIDADE DE TRABALHOS SELECIONADOS EM CADA ETAPA DA SELEÇÃO.....	59
TABELA 2 – QUANTIDADE DE TRABALHOS PUBLICADOS POR ANO DE PUBLICAÇÃO .....	60
TABELA 3 – EXEMPLO DE UM ARQUIVO DE BUSCA DO DIA 02/10/2012 .....	77
TABELA 4 – CÁLCULO DA MÉDIA DAS 6 COLETAS [TWITTER ERRO].....	77
TABELA 5 – RESULTADO DAS BUSCAS INDIVIDUAIS .....	78
TABELA 6 – RESULTADO DAS BUSCAS ASSOCIADAS A SUBSTANTIVOS .....	79
TABELA 7 – RESULTADO DA BUSCA ASSOCIADA A ADJETIVOS .....	79
TABELA 8 – RESULTADO DAS BUSCAS ASSOCIADAS À INTERROGAÇÃO .....	79
TABELA 9 – RESULTADO DAS BUSCAS ASSOCIADAS A ADVÉRBIOS .....	80
TABELA 10 – APLICAÇÃO DA ANÁLISE DE POSTAGENS PARA ALUNOS DA GRADUAÇÃO .....	101
TABELA 11 – MÉDIA DE ACERTO DA TURMA POR CATEGORIA ANALISADA.....	102
TABELA 12 – PERFIL DOS ESPECIALISTAS DE IHC .....	105
TABELA 13 – QUANTIDADE DE PRUS OBTIDAS POR META DE UUX.....	114
TABELA 14 – QUANTIDADE DE PRUS CLASSIFICADAS POR METAS MAIS RELEVANTES DE UUX NOS SS TWITTER E SIGAA .....	115
TABELA 15– BALANCEAMENTO DAS METAS DE UUX .....	115
TABELA 16 – RESUMO DAS MINERAÇÕES REALIZADAS .....	129

## LISTA DE ABREVIATURAS E SIGLAS

EI	Extração de Informação
KDD	<i>Knowledge Discovery in Database</i> , descoberta de conhecimento em banco de dados
LUQS	Laboratório de estudos do Usuário e da Qualidade de Uso de Sistemas
MALTU	Modelo para Avaliação da interação em Sistemas Sociais a partir da Linguagem Textual do Usuário
Não-PRU	Postagem não relacionada ao uso do sistema
PLN	Processamento da Linguagem Natural
PRU	Postagem Relacionada ao Uso do sistema
SiCos	Sistemas Colaborativos
SIGAA	Sistema Integrado de Gestão de Atividades Acadêmicas, sistema acadêmico das Universidades Federais do Brasil
SS	Sistemas Sociais
TV	Torcida Virtual, sistema social interativo sobre times de futebol
UUX	Usabilidade e Experiência do Usuário
UUX-Posts	Buscador de postagens relacionadas a UUX
UX	<i>User eXperience</i> , Experiência do Usuário
WEKA	Software de mineração de dados

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>20</b>
1.1 CONTEXTUALIZAÇÃO E CARACTERIZAÇÃO DO PROBLEMA.....	20
1.2 HIPÓTESE E QUESTÕES DA PESQUISA.....	22
1.3 OBJETIVOS.....	23
1.4 METODOLOGIA.....	23
1.5 ESCOPO.....	23
1.6 CONTRIBUIÇÕES DA TESE.....	24
1.7 ORGANIZAÇÃO DA TESE .....	25
<b>2 FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>27</b>
2.1 SISTEMAS SOCIAIS.....	27
2.1.1 Postagens em Sistemas Sociais.....	28
2.2 USABILIDADE E EXPERIÊNCIA DO USUÁRIO.....	29
2.2.1 Métodos de avaliação da usabilidade .....	31
2.2.2 Métodos de avaliação da experiência do usuário .....	34
2.3 PROCESSAMENTO DA LINGUAGEM NATURAL.....	37
2.4 MINERAÇÃO DE DADOS.....	43
2.4.1 Algoritmos de mineração de dados .....	47
2.5 AVALIAÇÃO EMOCIONAL X ANÁLISE DE SENTIMENTOS.....	51
2.6 CONCLUSÃO DO CAPÍTULO.....	55
<b>3 TRABALHOS RELACIONADOS.....</b>	<b>56</b>
3.1 PROCESSO DE REVISÃO SISTEMÁTICA.....	56
3.1.1 Planejamento da revisão .....	57
3.1.2 Condução da revisão.....	58
3.1.3 Documentação.....	59
3.2 APRESENTAÇÃO DOS ESTUDOS.....	59
3.3 RESULTADO DAS QUESTÕES.....	60
3.3.1 QP: Qual o estado da arte sobre SS?.....	60
3.3.2 QS1: Quais são os principais aspectos avaliados em trabalhos sobre avaliação de SS?.....	63
3.3.3 QS2: Quais são os principais métodos de avaliação em trabalhos sobre SS?.....	64
3.4 TRABALHOS RELACIONADOS.....	68
3.5 CONCLUSÃO DO CAPÍTULO.....	70
<b>4 INVESTIGAÇÕES A PARTIR DAS POSTAGENS DOS USUÁRIOS EM SISTEMAS.....</b>	<b>71</b>
<b>SOCIAIS.....</b>	<b>71</b>
4.1 SISTEMAS SOCIAIS INVESTIGADOS .....	71
4.2 OS USUÁRIOS “FALAM” SOBRE O SISTEMA DURANTE O USO? .....	72
4.2.1 Sistema Social investigado .....	72
4.2.2 Metodologia.....	73

4.2.3 Resultados .....	78
4.2.4 Análise dos resultados .....	80
4.2.5 Conclusão: os usuários “falam” sobre o sistema durante o uso? .....	82
4.3 COMO OS USUÁRIOS “FALAM”? .....	84
4.3.1 Subobjetivo 1: identificar tipos de PRUs.....	84
4.3.2 Subobjetivo 2: investigar como os usuários expressam seus sentimentos nas PRUs .....	92
4.3.3 Discussão.....	99
4.3.4 Conclusão: como os usuários “falam”? .....	100
4.4 COMO AVALIAR? .....	100
4.4.1 Sistemas Sociais investigados.....	100
4.4.2 1º Estudo: classificação por alunos de IHC.....	100
4.4.3 2º Estudo: classificação por especialistas de IHC .....	105
4.4.4 Discussão.....	109
4.4.5 Conclusão: como avaliar?.....	109
4.5 CONCLUSÃO DO CAPÍTULO.....	110
<b>5 MINERAÇÃO DE DADOS .....</b>	<b>112</b>
5.1 OBJETIVOS.....	112
5.2 METODOLOGIA.....	112
5.2.1 Dados .....	113
5.2.2 Seleção: coleta de dados dos usuários .....	113
5.2.3 Pré-processamento .....	116
5.2.4 Transformação .....	118
5.2.5 Mineração de dados .....	119
5.2.6 Interpretação / avaliação .....	120
5.3 RESULTADOS.....	129
5.3.1 Padrões de extração de PRUs .....	129
5.3.2 Padrões para cada tipo de PRU.....	131
5.3.3 Padrões para cada meta de UUX.....	131
5.4 DISCUSSÃO.....	132
5.5 CONCLUSÃO DO CAPÍTULO.....	136
<b>6 MALTU - MODELO PARA AVALIAÇÃO DA INTERAÇÃO EM SISTEMAS SOCIAIS A PARTIR DA LINGUAGEM TEXTUAL DO USUÁRIO .....</b>	<b>137</b>
6.1 MODELO MALTU .....	137
6.2 METODOLOGIA.....	138
6.2.1 Contexto de avaliação .....	139
6.2.2 Extração de PRUs .....	141
6.2.3 Classificação das PRUs .....	142
6.2.4 Interpretação dos resultados .....	149
6.2.5 Relato dos resultados .....	151
6.3 FERRAMENTA UUX-POSTs .....	152

6.3.1 Estrutura da aplicação .....	152
6.4 LIMITAÇÕES DO MODELO E ALGUMAS ESTRATÉGIAS DE SOLUÇÃO.....	154
6.5 RESULTADO DE AVALIAÇÕES USANDO O MODELO MALTU .....	156
6.5.1 Resultado da avaliação do Twitter usando o modelo MALTU .....	156
6.5.2 Resultado da avaliação do SIGAA usando o modelo MALTU .....	160
6.6 CONCLUSÃO DO CAPÍTULO .....	164
<b>7 CONCLUSÃO .....</b>	<b>165</b>
7.1 RESULTADOS ALCANÇADOS .....	165
7.2 OPORTUNIDADES DE MELHORIA.....	168
7.2.1 Modelo MALTU .....	168
7.2.2 Ferramenta UUX-Post .....	170
7.3 TRABALHOS FUTUROS .....	170
<b>REFERÊNCIAS.....</b>	<b>172</b>
<b>APÊNDICES .....</b>	<b>188</b>

# 1 INTRODUÇÃO

Esta tese propõe um modelo para avaliação da interação do usuário a partir das postagens dos usuários em Sistemas Sociais. Nesta introdução, estão descritos os aspectos motivacionais, metodológicos e estruturais desta tese. Na seção 1.1 deste capítulo, o problema é contextualizado e caracterizado. Na seção 1.2, a hipótese e as questões da pesquisa que guiaram o desenvolvimento deste trabalho são apresentadas. Na seção 1.3, os objetivos e as principais metas deste trabalho são listados. Na seção 1.4, a metodologia utilizada para a elaboração desta tese é descrita. Na seção 1.5, é apresentado o escopo do trabalho, e na seção 1.6, são apresentadas as principais contribuições obtidas neste estudo. Por fim, a seção 1.7 encerra o capítulo apresentando a estrutura do restante desta tese.

## 1.1 CONTEXTUALIZAÇÃO E CARACTERIZAÇÃO DO PROBLEMA

Sistemas Sociais (SS) (e.g., Twitter, Facebook, MySpace, LinkedIn etc.) são ambientes interativos nos quais as pessoas se comunicam, interagem, colaboram e compartilham ideias e informações. Considere que os usuários, ao interagirem em SS, elogiam, tiram dúvidas ou reclamam do próprio SS. Imagine, agora, a quantidade de dados valiosos sobre a usabilidade e experiência do usuário (*User eXperience*, UX) que têm sido desperdiçados por falta de uma análise das mensagens dos usuários. Esta pesquisa foca na análise desses dados com o objetivo de apoiar a avaliação da Usabilidade e UX (UUX) nesses tipos de sistemas.

A primeira motivação para este trabalho foi o avanço tecnológico da web e de sistemas interativos. O advento da web 2.0 permitiu o desenvolvimento de novas aplicações marcadas pela colaboração, comunicação e interatividade entre seus usuários de uma forma e em uma escala nunca antes observadas (PEREIRA; BARANAUSKAS; SILVA, 2010). SS se encontram nessa categoria. Embora a área de Interação Humano-Computador (IHC) tenha sugerido muitas formas para avaliar sistemas a fim de melhorar sua UUX, as oportunidades e os desafios trazidos por esse novo tipo de *software* interativo exigem que os métodos tradicionais de avaliação sejam repensados, considerando os novos conceitos, como valores humanos e técnicos (PEREIRA; BARANAUSKAS; SILVA, 2010).

As características de SS foram uma motivação para este trabalho. A colaboração, por exemplo, tem incentivado a frequente troca de mensagens entre os usuários de SS. Segundo

Sharoda, Lichan e Chi (2011), recentes estudos têm descoberto que as pessoas estão usando mensagens em SS, especialmente, para fazer questões aos seus amigos. A espontaneidade em SS é outra característica motivadora para este trabalho. Em SS, os usuários postam mensagens espontaneamente sobre fatos cotidianos de seu dia. Diversos estudos (ARAMAKI; MASKAWA; MORITA, 2011; JAMISON-POWELL et al., 2012; NEWMAN et al., 2011; BISHOP et al., 2011; KAMAL et al., 2012) têm relatado, por exemplo, que os usuários compartilham seu estado de saúde entre seus contatos.

Além dessas, outra característica motivadora em SS é a expressão de sentimentos. Os usuários se voltam para as mídias sociais para expressarem seus sentimentos em torno de eventos importantes de suas vidas, como aniversários, propostas de casamento, nascimento de filhos etc. (BRUBAKER et al., 2012). Os usuários postam mensagens se estão felizes (ASIAEE et al., 2012) ou tristes (DE CHOUDHURY; DIAKOPOULOS; NAAMAN, 2012). Tal característica, inclusive, tem motivado a criação de novas funcionalidades em SS a fim de incentivar a expressão dos sentimentos por seus usuários. O Facebook, por exemplo, liberou recentemente uma nova ferramenta para inserir *emoticons*<sup>1</sup> nas atualizações de *status*. A ideia é que os usuários possam expressar, de modo visual, os seus sentimentos<sup>2</sup>. O Twitter, por sua vez, colocou novas opções de busca, incluindo a busca por postagens positivas ou negativas<sup>3</sup>.

No campo de avaliação de sistemas, perguntar aos usuários suas opiniões a respeito de um produto, por exemplo: como eles realizam o que querem, se eles o apreciam, se o produto é esteticamente atraente e se eles enfrentam problemas ao utilizá-lo, é uma forma de avaliar um sistema (PREECE; ROGERS; SHARP, 2005). As técnicas principais para coleta da opinião do usuário sobre o sistema são: pesquisa de campo, entrevista e questionários (PREECE; ROGERS; SHARP, 2005; CYBIS; BETIOL; FAUST, 2007; BARBOSA; SILVA, 2010; BECKER; TUMITAN, 2013). Segundo Becker e Tumitan (2013), tais técnicas têm apresentado as seguintes desvantagens: altos custos; restritas a uma amostra bem-definida; retorno demorado; pouco eficiente; e alta latência (devido ao longo tempo necessário entre a coleta de dados brutos e a disponibilização dos resultados). Além dessas, tais técnicas não consideram a espontaneidade do usuário no momento em que ele está usando o sistema. Acredita-se que a forma espontânea de descrever um problema do sistema a um amigo,

---

<sup>1</sup> *emoticon*, palavra derivada da junção dos seguintes termos em inglês: emotion (emoção) + icon (ícone) é uma sequência de caracteres tipográficos, tais como: :), ou ^-^ e :-); ou, também, uma imagem (usualmente, pequena), que traduz ou quer transmitir o estado psicológico, emotivo, de quem os emprega, por meio de ícones ilustrativos de uma expressão facial.

<sup>2</sup> Disponível em: <<http://www.engadget.com/2013/04/09/facebook-emotion-selection-tool/>>. Acesso em: 10 de abril de 2013

<sup>3</sup> Disponível em: <<https://twitter.com/search-advanced>>. Acesso em: 20 de fevereiro de 2014

durante o uso, seja diferente de uma descrição a um especialista. Preece, Rogers e Sharp levantaram a seguinte questão:

O que os usuários dizem nem sempre é o que fazem. As pessoas algumas vezes dão respostas que consideram mostrar-lhes em sua melhor forma, ou apenas podem esquecer o que aconteceu ou quanto tempo passaram realizando uma atividade em particular. Sendo assim, os avaliadores podem acreditar em todas as respostas que obtêm? Os respondentes estão dizendo “a verdade”, ou simplesmente fornecendo as respostas que supõem que o avaliador queira ouvir? (2005, p. 420).

Além disso, a linguagem predominante em SS consiste de textos escritos. Por que não aproveitar essa característica de comunicação em SS para obter dados relevantes sobre o uso do sistema?

A solução proposta nesta tese é considerar o processo de seleção e análise dos conteúdos postados por usuários em SS como uma estratégia para avaliação da UUX. Não se propõe automatizar totalmente a avaliação da UUX, pois é indiscutível a importância dos especialistas nesse processo. A proposta se refere a um modelo para apoiá-los na avaliação, fornecendo dados relevantes sobre o conteúdo proveniente da interação dos usuários em SS. Para tanto, serão realizados estudos em Processamento da Linguagem Natural (PLN) e mineração de dados. O PLN consiste no desenvolvimento de modelos computacionais para a realização de tarefas que dependem de informações expressas em alguma língua natural, como tradução e interpretação de textos, busca de informações em documentos etc. (RUSSELL; NORVIG, 1995). Mineração de dados é uma tecnologia que combina métodos tradicionais de análise de dados com algoritmos sofisticados para processar grandes volumes de dados (TAN; STEINBACH; KUMAR, 2009).

## 1.2 HIPÓTESE E QUESTÕES DA PESQUISA

Considerando que a coleta de opiniões dos usuários é um material valioso para avaliação de sistemas, esta tese de doutorado procura testar a seguinte hipótese: *as postagens dos usuários em SS fornecem dados relevantes para avaliação de aspectos da UUX.*

A partir dessa hipótese quatro Questões de Pesquisa (QP) foram formuladas:

- a) **QP1:** *Quais são os principais métodos de avaliação da UUX em SS?*
- b) **QP2:** *Os usuários “falam” sobre o sistema em uso?*
- c) **QP3:** *Como os usuários “falam”?*
- d) **QP4:** *Como avaliar?*

### 1.3 OBJETIVOS

Esta tese de doutorado tem como objetivo principal propor um modelo de avaliação de critérios de qualidade de uso em SS a partir das postagens dos usuários sobre o uso do sistema em interação. O modelo, denominado MALTU tem o intuito de apoiar avaliadores no entendimento das opiniões dos usuários sobre o sistema. Para o atendimento desse objetivo, os objetivos específicos são:

- a) realizar uma revisão bibliográfica sobre SS e métodos de avaliação (QP1);
- b) realizar investigações em postagens dos usuários em SS a fim de estudar suas características (QP2, QP3, QP4);
- c) propor um modelo de avaliação da UUX a partir da linguagem natural do usuário (QP4);

### 1.4 METODOLOGIA

A elaboração deste trabalho está orientada à seguinte metodologia:

- a) estudo da fundamentação teórica necessária para desenvolvimento desta tese (conceitos e tecnologias);
- b) realização de uma revisão bibliográfica a partir de uma revisão sistemática para identificação de trabalhos relacionados;
- c) investigações a partir da extração e análise de postagens dos usuários em SS de contextos distintos;
- d) mineração de dados das postagens dos usuários em SS a fim de obter dados relevantes sobre suas postagens;
- e) proposta de um modelo para avaliação textual usando as metas de UUX em SS;
- f) definir requisitos, especificação e implementação de uma ferramenta de suporte à avaliação textual da UUX em SS.

### 1.5 ESCOPO

Este trabalho tem como foco a proposição de um modelo de avaliação textual usando metas de UUX a partir das postagens dos usuários em SS. Os itens a seguir descrevem o escopo abordado neste estudo sobre alguns desses conceitos.

**Avaliação em SS:** existe uma diversidade de métodos de avaliação em sistemas, inclusive com o foco em SS. O escopo desta tese é focar na importância da opinião expressa no momento do uso do usuário para avaliação de SS, considerando suas características textuais de comunicação.

**Usabilidade e experiência do usuário:** o campo de avaliação de sistemas envolve uma série de critérios a serem avaliados em sistemas interativos, além de UUX, tais como: comunicabilidade, acessibilidade, privacidade, sociabilidade, entre outras. Este trabalho procura focar especificamente nas metas de UUX.

**Processamento da Linguagem Natural:** a área de PLN se dedica ao estudo, tratamento e compreensão da linguagem humana através de tecnologia computacional (DIAS; MALHEIROS, 2006). O escopo do PLN, nesta tese, consiste no apoio ao entendimento da linguagem do usuário ao se referir ao sistema durante seu uso.

**Mineração de dados:** os aspectos de mineração de dados abordados neste trabalho foram: 1) as etapas de descoberta de conhecimento em banco de dados (*Knowledge Discovery in Database*, KDD) (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996), a fim de guiar o processo de mineração dos textos obtidos pelos usuários em SS; e 2) algoritmos de classificação, com o objetivo de extrair padrões úteis de identificação desses textos. Os algoritmos utilizados nesta tese foram árvores de decisão e redes bayesianas.

**Avaliação emocional:** os aspectos de avaliação emocional abordados neste trabalho foram baseados nos estudos de Norman (2004), relacionando com o contexto textual das postagens.

**Análise de sentimentos:** este trabalho não utiliza algoritmos de análise de sentimentos nos experimentos. Mesmo assim, foram realizados estudos sobre sentenças objetivas, subjetivas, polaridade e intensidade de sentimentos a partir das postagens dos usuários. Temos ciência da importância da mesma para apoiar o modelo de avaliação proposto, de forma a identificar, automaticamente, opiniões positivas e negativas, e demais avaliações expressas em linguagem natural.

## 1.6 CONTRIBUIÇÕES DA TESE

Esta tese abre o caminho para uma nova linha de investigação sobre o estudo da expressão escrita dos usuários sobre o uso de sistemas. Ela contribui para IHC em geral e para o campo de avaliação de sistemas, especificamente com a produção de novos conhecimentos

e novas questões de pesquisa, além de ampliar o entendimento dos textos dos usuários no uso de SS, focando em sua linguagem, interação e motivação. Obtêm-se, em suas postagens, características específicas referentes à UUX, especificamente, ao domínio, ao sistema, ao usuário. Tais características devem motivar avaliadores, especialistas e designers a utilizar esse entendimento na avaliação e estudos em SS. Especificamente, as contribuições são:

- a) estudos teóricos em SS, de forma a apresentar características relevantes de avaliação em SS;
- b) experimentos práticos em SS, apresentando características das postagens dos usuários, classificadas em função de diversos fatores: tipo da postagem, intenção do usuário, sentimentos, metas de UUX etc.;
- c) disponibilização de uma base de dados (postagens dos usuários em SS) classificada, de forma a proporcionar estudos futuros de especialistas na área;
- d) padrões de extração para busca e classificação de postagens relacionadas ao uso de SS;
- e) um modelo e uma metodologia para avaliação das metas de UUX em SS a partir da linguagem textual do usuário;
- f) uma ferramenta para apoio à avaliação da UUX a partir das postagens dos usuários em SS.

## 1.7 ORGANIZAÇÃO DA TESE

Este capítulo apresentou a problemática que motiva esta tese de doutorado, bem como a hipótese, questões de pesquisa, objetivos e metas, metodologia, escopo, contribuições e publicações. O restante deste trabalho é composto por sete capítulos, organizados da seguinte forma:

- a) **Capítulo 2** — apresenta a fundamentação teórica, com alguns conceitos necessários para o melhor entendimento do trabalho, abordando SS, UUX, PLN, mineração de dados, avaliação emocional e análise de sentimentos;
- b) **Capítulo 3** — apresenta uma revisão sistemática e uma bibliográfica, com a identificação de trabalhos relacionados à SS, avaliação de SS e avaliação pela linguagem escrita;
- c) **Capítulo 4** — são apresentados experimentos realizados a partir das postagens dos usuários em SS. Cada experimento visou uma questão de pesquisa e, seguindo

uma metodologia de investigação, proporcionou resultados efetivos que contribuíram para o modelo proposto nesta tese;

- d) **Capítulo 5** — este capítulo apresenta o processo de mineração de dados realizado a partir das postagens dos usuários em SS;
- e) **Capítulo 6** — o Modelo MALTU é apresentado, com o detalhamento de suas partes. Uma ferramenta de extração de postagens relacionadas ao uso é apresentada;
- f) **Capítulo 7** — este capítulo é dedicado às considerações finais desta tese de doutorado. Nele, são descritos os resultados alcançados e uma discussão sobre as questões da pesquisa. Por fim, trabalhos futuros derivados desta pesquisa são apresentados.
- g) **Apêndices** — os apêndices apresentam material complementar deste trabalho. Questionários, metas de UUX e gráficos da mineração de dados são exibidos.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, é apresentada a fundamentação teórica sobre os temas a serem abordados nesta pesquisa. O objetivo é possibilitar o entendimento das colocações que serão feitas mais adiante. São apresentados conceitos sobre SS, UUX, métodos de avaliação, PLN, mineração de dados, algoritmos de mineração de dados, avaliação emocional e análise de sentimentos.

### 2.1 SISTEMAS SOCIAIS

Pereira, Baranauskas e Silva (2010) investigaram o conceito de *software* social do ponto de vista dos valores envolvidos (confiança, reputação, autonomia, colaboração etc.), chegando à seguinte definição:

Sistemas que permitem às pessoas, em suas particularidades e diversidades, se comunicarem (interagirem, colaborarem, compartilharem ideias e informações), mediando e facilitando qualquer forma de relacionamento social; sistemas cuja utilidade seja dependente e a estrutura moldada pela efetiva participação, interação e produção de conteúdo por parte dos usuários. (PEREIRA; BARANAUSKAS; SILVA, 2010, p. 151).

Este conceito de *software* social é amplo, abrangendo desde sites de colaboração de conteúdos, como: YouTube e Wikipédia, até sites de redes sociais, como o Facebook e o Twitter (PEREIRA; BARANAUSKAS; SILVA, 2010). Outros autores (PIMENTEL; FUKS, 2011; NICOLACI-DA-COSTA; PIMENTEL, 2011; PIMENTEL; GEROSA; FUKS, 2011; PRATES, 2011; CARVALHO, 2013) usam o termo Sistemas Colaborativos (SiCos) para designar ambos os termos: *groupware* e CSCW (*Computer Supported Cooperative Work*), sendo *groupware* sistemas computacionais usados para apoiar o trabalho em grupo e CSCW para designar tanto sistemas quanto os efeitos psicológicos, sociais e organizacionais do trabalho em grupo (NICOLACI-DA-COSTA; PIMENTEL, 2011). Para Gerosa e Fuks (2011), SS são sistemas de comunicação frequentemente usados na composição de SiCos:

Sistemas de comunicação são frequentemente usados na composição de sistemas colaborativos como: redes sociais, em que vários tipos de sistemas de comunicação são adaptados para possibilitar múltiplas formas de interação entre os usuários; ambientes de aprendizagem, em que vários sistemas de comunicação estão disponíveis para serem usados e configurados em cada curso; ou em ambientes virtuais, que geralmente contém um serviço de bate-papo e de audioconferência. (PIMENTEL; GEROSA; FUKS, 2011, p. 69).

Os SS concentram-se em capacitar seus usuários para se comunicar e interagir uns com os outros de diversas maneiras e para diversos fins (KARNIK et al., 2013; CHEN et al., 2010; ARAMAKI; MASKAWA; MORITA, 2011). Complementando esse conceito, Boyd e Ellison (2007) especificam SS como serviços baseados na web que permitem aos indivíduos (1) construir um perfil público ou semipúblico dentro de um sistema limitado, (2) articular uma lista de outros usuários com quem eles compartilham uma conexão, e (3) visualizar e percorrer sua lista de conexões e aquelas feitas por outras pessoas dentro do sistema. Este trabalho foca em SS, considerando a interação entre os usuários por meio de mensagens de texto, suas postagens.

### *2.1.1 Postagens em Sistemas Sociais*

Conforme dito na introdução, os SS trouxeram novas características de interação para os usuários. Usuários podem acessar SS por meio de diversos dispositivos de diferentes locais e a qualquer momento. A forma principal de interação nesses sistemas são as mensagens postadas, sejam elas de forma pública ou privada. Em suas postagens, os usuários tratam de diversos assuntos. O interesse deste trabalho está nas postagens públicas em linguagem natural nas quais o usuário se refere ao SS que ele está usando no momento (Postagens Relacionadas ao Uso, PRUs). Por exemplo: se o usuário estiver usando o Facebook, interessamos as PRUs do Facebook; se o SS em avaliação for o Twitter, as PRUs deverão ser sobre o Twitter.

Próximo a esse campo de análise de textos estão as pesquisas sobre revisões de produtos ou serviços na web. Nos últimos anos, o uso de sites para avaliação de produtos e serviços tornou-se cada vez mais comum. Sites como Booking<sup>4</sup>, Decolar<sup>5</sup>, Reclameaqui<sup>6</sup> e Tripadvisor<sup>7</sup> fornecem um espaço para que os clientes divulguem suas revisões sobre produtos e serviços.

Revisão é um pequeno texto detalhando prós e contras de um produto e, possivelmente uma avaliação dele e uma recomendação para potenciais compradores, escrito por um usuário do produto que esteve de posse do referido produto e usou-o durante algum tempo. Ela pode ser escrita por um revisor

---

<sup>4</sup> Disponível em: <booking.com>. Acesso em: 13 de maio de 2012

<sup>5</sup> Disponível em: <http://www.decolar.com/>. Acesso em: 13 de maio de 2012

<sup>6</sup> Disponível em: <http://www.reclameaqui.com.br/ >. Acesso em: 13 de maio de 2012

<sup>7</sup> Disponível em: <http://www.tripadvisor.com.br/>. Acesso em: 13 de maio de 2012

profissional ou por um usuário final comum. (HEDEGAARD; SIMONSEN, 2013, p. 2.090).

É importante destacar a reflexão entre os dois conceitos: revisões de produtos ou serviços e comentários dos usuários em SS. A partir de estudos empíricos, chegou-se à descrição das diferenças quanto aos seguintes aspectos:

- a) Forma — revisões são textos estruturados, ou seja, apresentam regularidade no formato de apresentação das informações, por exemplo, um formulário preenchido. Existem campos para pontuação, entrada do texto para avaliação e até um campo do aspecto a ser avaliado, enquanto que as postagens em SS são textos não estruturados, ou seja, não apresentam regularidade em seu formato. As postagens dos usuários podem apresentar imagens, diversos tipos de textos e caracteres e até links referenciando páginas;
- b) motivação — uma série de artigos on-line (RHODES, 2009; KARR, 2012; MERRITT, 2013; DITLEV, 2012) são escritos a fim de argumentar o porquê de as pessoas escreverem revisões. Entre os diversos motivos apresentados, em geral, os autores concordam que as revisões são feitas porque as pessoas se preocupam com seus colegas consumidores e querem ajudar aos outros na tomada de uma decisão (DITLEV, 2012). Em Mendes, Furtado e Castro (2014), foi realizada uma investigação sobre características das PRUs de SS e foi observado que os usuários elogiam, fazem críticas, comparações, tiram dúvidas e fornecem sugestões sobre o sistema, o que leva os autores a crer que tais comentários contêm relatos dos usuários sobre suas experiências de uso no sistema; e
- c) contexto — no momento da revisão, o revisor não está utilizando o sistema que ele quer analisar. O fato de o usuário comentar no próprio SS, que ele está usando, pode ser uma forma de solicitar ajuda a fim de resolver um problema no momento do uso.

## 2.2 USABILIDADE E EXPERIÊNCIA DO USUÁRIO

Usabilidade, segundo Preece, Rogers e Sharp (2005), é geralmente considerada como o fator que assegura que os produtos são fáceis de usar, eficientes e agradáveis – da perspectiva do usuário. Segundo a ISO 9241-11 (1998), usabilidade é uma medida em que um sistema, produto ou serviço pode ser usado por usuários específicos para se atingir objetivos

específicos com eficácia, eficiência e satisfação, em um determinado contexto de uso. A ISO 9241-11 (1998) define esses três objetivos da seguinte forma: eficácia consiste na exatidão e completude com que os usuários atingem objetivos específicos; eficiência consiste nos recursos gastos em relação à exatidão e completude com que os usuários atingem objetivos; e satisfação são atitudes positivas e ausência de desconforto em relação ao uso do produto. Preece, Rogers e Sharp (2005), por sua vez, apresentam seis metas de usabilidade: ser eficaz no uso (eficácia), eficiente no uso (eficiência), segura no uso (segurança), de boa utilidade (utilidade), capacidade de aprendizagem (aprendizado<sup>8</sup>) e capacidade de memorização (memorização<sup>9</sup>).

Preece, Rogers e Sharp (2005) também apresentam as metas decorrentes da experiência do usuário (*User eXperience*, UX), preocupando-se, principalmente, com a experiência que os sistemas proporcionarão aos usuários. Tais metas consistem em desenvolver sistemas que sejam: satisfatórios, agradáveis, divertidos, interessantes, úteis, motivadores, esteticamente apreciáveis, incentivadores de criatividade, compensadores e emocionalmente adequados. O conceito de UX é recente, definido pela primeira vez por Norman, Miller e Henderson (1995), e prontamente adotado por pesquisadores da comunidade de IHC. Tal conceito abrange áreas como psicologia, antropologia, sociologia, ciência da computação, design gráfico, entre outras. Segundo Forlizzi e Battarbee (2004), o termo está associado a diversos significados, que variam de usabilidade tradicional à beleza, aspectos hedônicos, afetivos ou experimentais do uso da tecnologia. De fato, existem várias definições para o termo. Nielsen e Norman<sup>10</sup> definem UX da seguinte forma: “A ‘experiência do usuário’ engloba todos os aspectos da interação do usuário final com a empresa, seus serviços e seus produtos”.

Law et al. (2009) realizaram uma pesquisa com 275 pesquisadores e profissionais da academia e da indústria com o objetivo de definir UX. Segundo eles, a maioria dos entrevistados concorda que UX é dinâmico, dependente do contexto e subjetivo. Eles concluíram que os pesquisadores definem UX como algo individual (em vez de social), que surge na interação com um produto, sistema, serviço ou um objeto. A proposta de definição de UX desses autores se aproxima do conceito da ISO 9241-210 (2008), que define UX em percepções e respostas das pessoas, resultantes do uso e/ou uso antecipado de um produto,

---

<sup>8</sup> As autoras usam o termo: *learnability*.

<sup>9</sup> As autoras usam o termo: *memorability*.

<sup>10</sup> Disponível em: < <http://www.nngroup.com/articles/definition-user-experience/>>. Acesso em: 12 de outubro de 2014

sistema ou serviço. A norma ressalta que a UX inclui todas as emoções, crenças, preferências, percepções, respostas físicas e psicológicas, comportamentos e realizações do usuário que ocorrem antes, durante e depois do uso.

Este trabalho adota essa última definição, assim como apresenta algumas diferenças entre os conceitos usabilidade e UX. Alguns trabalhos (BEVAN, 2009; ROTO; OBRIST; VÄÄNÄNEN-VAINIO-MATTILA, 2009; MAIA; FURTADO, 2014) ressaltam que as técnicas de usabilidade utilizam métricas que são difíceis de definir em avaliações da UX, conseqüentemente, levando a diferentes preocupações durante o desenvolvimento. Para Bevan (2009), as preocupações típicas da usabilidade incluem: avaliar a eficácia, eficiência, conforto, satisfação e capacidade de aprendizado dos usuários, a fim de identificar e corrigir problemas de usabilidade para tornar o produto fácil de usar. Enquanto isso, a UX preocupa-se em compreender e projetar a experiência do usuário com um produto e em identificar e evocar respostas emocionais dos usuários (BEVAN, 2009; FURTADO; FURTADO; VASCONCELOS, 2007).

Hedegaard e Simonsen (2013) fizeram um estudo das principais metas de UUX utilizados pelos seguintes autores: Folmer, Van Gurp e Bosch (2003); Seffah et al. (2006); Bevan (2009); Ketola e Roto (2008) e Bargas-Avila e Hornbaek (2011) e chegaram às seguintes metas (Quadro 1). A definição dos autores foi proposta com o objetivo de avaliar produtos e serviços a partir de revisões dos usuários. Tais metas são apresentadas de forma detalhada no APÊNDICE C desta tese.

Quadro 1 – Metas da UUX

<b>(Folmer, Van Gurp e Bosch, 2003; Seffah et al., 2006)</b>	<b>(Bevan, 2009)</b>	<b>(Ketola e Roto, 2008; Bevan, 2009)</b>	<b>(Bargas-Avila e Hornbaek, 2011)</b>
<b>Metas</b>			
Memorização Aprendizado Eficiência Erros / Eficácia Satisfação	Utilidade Prazer Conforto Confiança	Antecipação Usabilidade geral Hedônico Usabilidade detalhada Diferenças do usuário Suporte Impacto	Afeto e emoção Prazer, diversão Estética, atração Engajamento, fluidez Motivação Encantamento Frustração Hedônico

Fonte: Hedegaard e Simonsen (2013). Traduzido pela autora.

### 2.2.1 Métodos de avaliação da usabilidade

Em Prates (2011), no capítulo sobre a interação em sistemas colaborativos, são apresentados os seguintes critérios de qualidades de uso para SiCos: usabilidade, comunicabilidade, acessibilidade e sociabilidade. A autora apresentou os seguintes métodos de avaliação para cada critério de qualidade (Quadro 2).

Quadro 2– Métodos de avaliação dos critérios de qualidades de uso em SiCos

<b>Crítérios</b>	<b>Métodos</b>
Usabilidade	<ul style="list-style-type: none"> <li>▪ Avaliação heurística</li> <li>▪ Testes de usabilidade</li> </ul>
Comunicabilidade	<ul style="list-style-type: none"> <li>▪ Método de Inspeção Semiótica (MIS)</li> <li>▪ Método de Avaliação de Comunicabilidade (MAC)</li> </ul>
Acessibilidade	<ul style="list-style-type: none"> <li>▪ Avaliadores automáticos</li> <li>▪ Inspeção para verificação de conformidade com diretrizes de acessibilidade</li> <li>▪ Avaliação do uso de tecnologias assistivas</li> <li>▪ Avaliação com usuários</li> </ul>
Sociabilidade	<ul style="list-style-type: none"> <li>▪ Heurísticas para comunidades virtuais</li> <li>▪ Modelo Manas</li> </ul>

Fonte: Prates (2011).

Barbosa e Silva (2010), por sua vez, dividem os métodos de avaliação da usabilidade em: a) métodos de investigação, b) observação de uso e c) inspeção, descritos a seguir:

- a) métodos de investigação — permitem ao avaliador ter acesso, interpretar e analisar concepções, opiniões, expectativas e comportamentos do usuário relacionados com sistemas interativos. Tais métodos envolvem o uso de questionários, realização de entrevistas e grupos de foco;
- b) métodos de observação — fornecem dados sobre situações em que os usuários realizam suas atividades, com ou sem apoio de sistemas interativos. Esses métodos, por meio do registro dos dados observados, permitem identificar problemas reais que os usuários enfrentaram durante sua UX do sistema sendo avaliado. Nessa categoria, encontram-se: estudos de campo, teste de usabilidade, Método de Avaliação de Comunicabilidade (MAC) e prototipação em papel; e
- c) métodos de inspeção — não envolvem usuários e permitem ao avaliador identificar problemas que os usuários podem vir a ter quando interagirem com o sistema. Avaliação heurística, percurso cognitivo e a inspeção semiótica (MIS e MISI) fazem parte desta categoria.

A proposta desta tese consiste na investigação das opiniões e expectativas dos usuários por meio de suas postagens em SS. Levando em conta que os critérios de qualidade de uso a serem avaliados são UUX, os métodos apresentadas pelos autores citados nesse contexto são:

avaliação heurística, testes de usabilidade, questionários, entrevistas e grupos de foco, descritos a seguir.

O método de Avaliação Heurística é utilizado para encontrar problemas de usabilidade durante um processo de design interativo (NIELSEN, 1993 apud BARBOSA; SILVA, 2010). Os especialistas, orientados por um conjunto de princípios de usabilidade conhecidos como heurísticas, avaliam se os elementos da interface com o usuário estão de acordo com esses princípios (PREECE; ROGERS; SHARP, 2005). Segundo Cybis, Betiol e Faust (2007), as avaliações por especialistas podem produzir ótimos resultados no que diz respeito à rapidez de avaliação e à quantidade e importância de problemas diagnosticados. Entretanto, seus resultados dependem da competência dos avaliadores e das estratégias de avaliação empregadas (PREECE; ROGERS; SHARP, 2005; CYBIS; BETIOL; FAUST, 2007; ROCHA; BARANAUSKAS, 2003). Ademais, nesse método não são consideradas as opiniões dos usuários sobre o sistema.

O método Testes de Usabilidade permite perceber o usuário, seu ambiente e como suas tarefas são realizadas ao usar um sistema (MILLEN, 2000). Em tais testes, geralmente, mede-se o tempo que usuários típicos levam para completar tarefas claramente definidas e típicas, e o número e os tipos de erros que cometem são registrados (PREECE; ROGERS; SHARP, 2005). As desvantagens desse método são: o esforço em preparar visitas, conduzir e analisar os resultados e o risco do usuário sentir-se constrangido ao ser observado. O constrangimento do usuário é inerente a um teste na medida em que esse implica na observação de uma pessoa trabalhando com um sistema interativo (CYBIS; BETIOL; FAUST, 2007).

Os questionários têm como objetivo coletar rapidamente dados (principalmente quantitativos) de muitos usuários. Tal método tem mostrado muita eficiência como apoio aos demais métodos de avaliação, no entanto, se as perguntas não forem bem-formuladas ou forem tendenciosas, poderão induzir ou confundir o usuário. Além disso, ainda existe o risco de o questionário ser muito longo e os usuários responderem sem ler cuidadosamente as perguntas.

Entrevistas, tanto individuais como em grupo (grupo de foco, focal, *brainstorming*), são de fácil entendimento, baixo custo e têm o caráter mais qualitativo. No entanto, existe o esforço de recrutamento de usuários. Também como nos questionários, as perguntas devem ser bem-formuladas para se atingir o objetivo almejado e não confundir o usuário. Além disso, conforme mencionado neste trabalho, existem dúvidas sobre a veracidade das

informações dos usuários ao descrever um problema a um especialista (PREECE; ROGERS; SHARP, 2005).

### 2.2.2 Métodos de avaliação da experiência do usuário

No CHI 2009, na sessão de Grupos de Interesse Especial: “Avaliação de Experiência do Usuário – Você sabe qual método usar?” (BEVAN, 2009; OBRIST; ROTO; VÄÄNÄNEN-VAINIO-MATTILA, 2009), os participantes foram solicitados a descrever os métodos de avaliação de UX. Bevan (2009) coletou 36 métodos, que foram classificados pelos seguintes temas (Quadro 3): avaliação do contexto, avaliação de dados e avaliação por especialistas.

Segundo Bevan (2009), os métodos descritos no Quadro 3 se diferenciam da usabilidade tradicional, pois não visam apenas alcançar a eficácia, eficiência e satisfação, mas otimizar toda a UX por meio da interação real para a reflexão sobre a experiência.

Quadro 3 – Métodos de avaliação da UX

Evaluation methods	Evaluation context	Evaluation data	Expert evaluation
<b>Methods</b>	<b>Lab tests</b> <ul style="list-style-type: none"> <li>• Lab study with mind maps</li> <li>• Paper prototyping</li> </ul>	<b>User opinion/interview</b> <ul style="list-style-type: none"> <li>• Lab study with mind maps</li> <li>• Quick and dirty evaluation</li> <li>• Audio narrative</li> <li>• Retrospective interview</li> <li>• Contextual Inquiry</li> <li>• Focus groups evaluation</li> <li>• Observation \ Post Interview</li> <li>• Activity Experience Sampling</li> <li>• Sensual Evaluation Instrument</li> <li>• Contextual Laddering</li> <li>• Interview</li> <li>• ESM (Experience sampling Method)</li> </ul>	<ul style="list-style-type: none"> <li>• Expert evaluation</li> <li>• Heuristic matrix</li> <li>• Perspective-Based Inspection</li> </ul>
	<b>Field tests</b> <ul style="list-style-type: none"> <li>• Product / Tool Comparison</li> <li>• Competitive evaluation of prototypes in the wild</li> </ul>		
	<b>Field observation</b> <ul style="list-style-type: none"> <li>• Long term pilot study</li> <li>• Longitudinal comparison</li> <li>• Contextual Inquiry</li> <li>• Observation/Post Interview</li> <li>• Activity Experience Sampling</li> <li>• Longitudinal Evaluation</li> <li>• Ethnography</li> <li>• Field observations</li> <li>• Longitudinal Studies</li> </ul>		
	<b>Evaluation of groups</b> <ul style="list-style-type: none"> <li>• Evaluating collaborative user experiences</li> </ul>	<b>User questionnaire</b> <ul style="list-style-type: none"> <li>• Survey Questions-Emocards</li> <li>• Experience sampling triggered by events</li> <li>• SAM (Self Assessment Scale)</li> <li>• Magnitude Estimation</li> <li>• TRUE (Tracking Realtime User Experience)</li> <li>• Questionnaire (e.g. AttrakDiff)</li> </ul>	
	<b>Instrumented product</b> <ul style="list-style-type: none"> <li>• TRUE (Tracking Realtime User Experience)</li> </ul>		
	<b>Domain specific</b> <ul style="list-style-type: none"> <li>• Nintendo Wii</li> </ul>		

Evaluation methods	Evaluation context	Evaluation data	Expert evaluation
	<b>Children</b> <ul style="list-style-type: none"> <li>• OPOS-Outdoor Play Observation Scheme</li> <li>• This-or-that</li> </ul>	<b>Human responses</b> <ul style="list-style-type: none"> <li>• PURE (Preverbal User Reaction Evaluation)</li> <li>• Psycho-physiological measurements</li> </ul>	
	<b>Approaches</b> <ul style="list-style-type: none"> <li>• Evaluating UX jointly with usability</li> </ul>		

Fonte: CHI 2009 SIG *apud* Bevan (2009).

O site ALL ABOUT UX apresenta 35 métodos de avaliação<sup>11</sup> da UX durante a interação com um sistema. Desses, aqueles que consideram a opinião do usuário durante o uso do sistema são: ESM, *Activity Experience Sampling*, *Experience Sampling triggered by events*, TRUE e SEI, explicados nos próximos parágrafos.

ESM (*Experience Sampling Method*) é um método que pede aos participantes que parem em certos momentos e façam anotações de sua experiência em tempo real (OBRIST; ROTO; VÄÄNÄNEN-VAINIO-MATTILA, 2009). Uma de suas características consiste em contatar os usuários em vários momentos “desprevenidos” e não necessariamente com intervalo equivalente, gerando uma visão mais realista da interação com o sistema, impossibilitando o usuário de se “preparar para a realização de um teste”. Segundo ALL ABOUT UX (2015), um ponto fraco dessa técnica consiste na possibilidade de o ponto de parada ser impróprio para relatar a experiência do usuário. Nesse momento, se o usuário não pode usar o sistema e tem que interromper sua experiência, algumas emoções negativas podem acontecer.

O método *Activity Experience Sampling* recolhe amostras de experiências momentâneas no ambiente natural do indivíduo, como o uso de diários relatando sua experiência (RIEDIGER, 2009). A técnica *Experience sampling triggered by events* é uma variante dessas técnicas, porém captura os eventos dos usuários durante seu uso por meio de uma ferramenta (FETTER; GROSS; SCHIRMER, 2011; KHAN; MARKOPOULOS, 2015; MESCHTSCHERJAKOV; REITBERGER; TSCHELIGI, 2010). A partir dos eventos capturados, como o acesso a uma determinada tela do sistema, um questionário é apresentado ao usuário a fim de associar sua satisfação àquele determinado evento.

O TRUE (*Tracking Real Time User Experience*) auxilia o avaliador no rastreamento do comportamento dos usuários, monitorando digitalmente a interação entre o usuário e o

<sup>11</sup> Disponível em: <<http://www.allaboutux.org/snapshot-methods>>. Acesso em: 12 de janeiro de 2015.

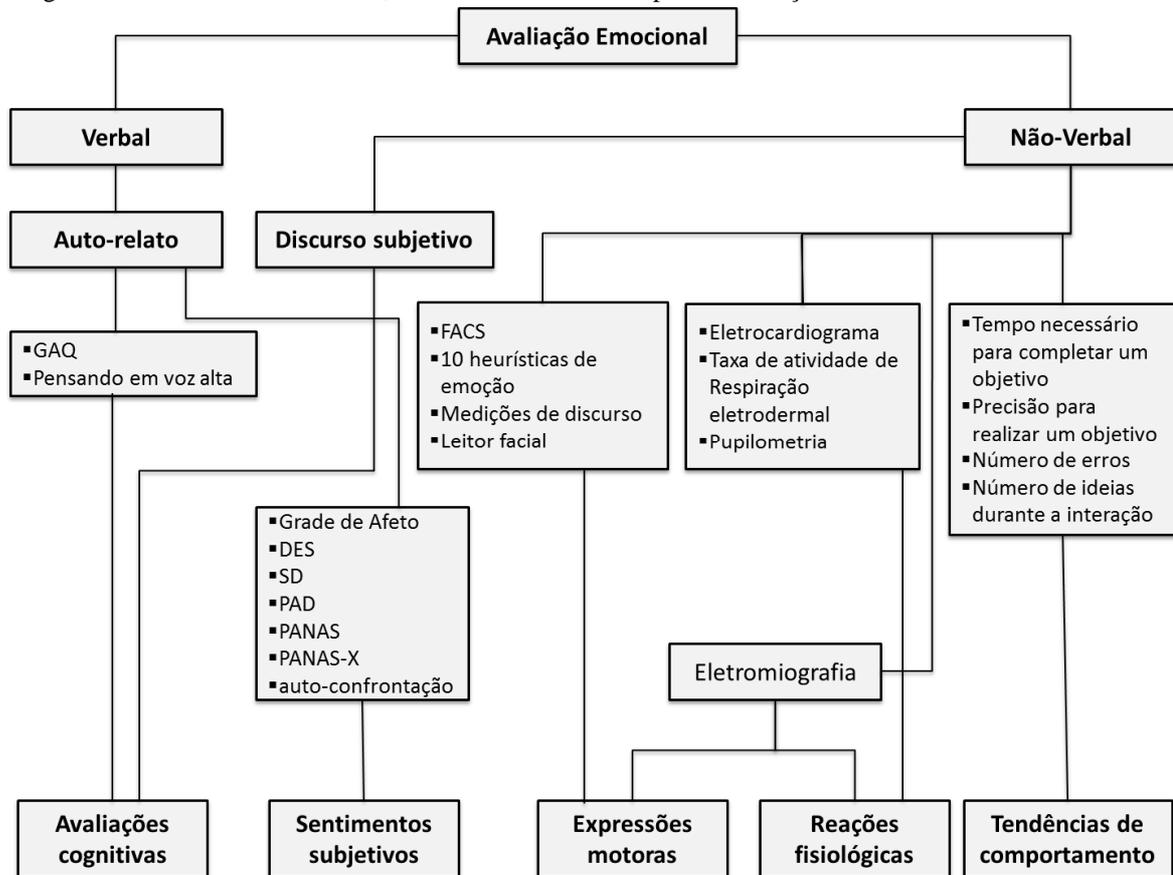
sistema, através da gravação em vídeo e podendo intercalar dados comportamentais e atitudes do usuário em relação ao sistema (KIM et al., 2008).

O SEI (*Sensual Evaluation Instrument*) consiste em um método destinado a fornecer um canal não verbal flexível de comunicação entre usuário e avaliador durante o desenvolvimento do sistema. Os usuários manifestam seus sentimentos por meio de objetos esculpidos à mão, indicando como eles estão se sentindo ao interagir com um sistema (ISBISTER et al., 2007).

Especificamente, este trabalho se baseia na avaliação da UUX por meio das postagens dos usuários durante o uso do sistema. No contexto de uso em SS, existe a frequente troca de mensagens entre os usuários e de forma espontânea. Na técnica ESM, os usuários irão relatar o uso aos especialistas. Conforme mencionado, considera-se que a forma espontânea de descrever um problema de uso a um amigo seja diferente de uma descrição a um especialista. Além disso, a linguagem predominante em SS consiste de textos escritos. Os eventos coletados dos usuários podem ser interessantes para análise de características como métricas de eficácia ou erros no sistema, mas os textos podem possibilitar a expressão das emoções dos usuários (BROOKE, 2009; ASIAEE et al., 2012).

Outra forma de avaliação da UX se dá por meio de avaliações emocionais. A Figura 1 ilustra uma taxonomia (XAVIER; NERIS, 2012), que classifica a avaliação emocional em verbal e não verbal, e apresenta algumas medidas para avaliação das emoções disponíveis na literatura. No contexto desta tese, interessam, apenas, as avaliações verbais, que são divididas em avaliações cognitivas: *Geneva Appraisal Questionnaire* (GAQ) e pensando em voz alta (*Think Aloud*); e sentimentos subjetivos: Grade de afeto, DES, SD, PAD, PANAS, PANAS-X e auto-confrontação. No entanto, nenhum dos métodos citados utiliza textos escritos por usuários durante o uso do sistema. Aqueles que não são orais utilizam questionários antes ou após o uso do sistema (GAQ, Grade de afeto, PANAS, PANAS-X e auto-confrontação).

Figura 1 –Taxonomia de métodos, técnicas e instrumentos para a avaliação emocional



Fonte: Xavier e Neris (2012). Traduzido pela autora.

Diversos autores (BEVAN, 2009; PREECE; ROGERS; SHARP, 2005; ROTO; OBRIST; VÄÄNÄNEN-VAINIO-MATTILA, 2009) acreditam que a triangulação de métodos facilita na avaliação da UX. A triangulação é uma estratégia de utilizar mais do que uma técnica de coleta ou análise de dados para obter diferentes perspectivas e confirmar as descobertas, permitindo obter resultados mais rigorosos e válidos (BARBOSA; SILVA, 2010). A análise de textos dos usuários irá fornecer mais uma contribuição na coleta de opiniões dos usuários.

### 2.3 PROCESSAMENTO DA LINGUAGEM NATURAL

A linguagem é um dos aspectos fundamentais do comportamento humano e um componente crucial de nossas vidas (ALLEN, 1994). Segundo o autor, a linguagem é estudada em várias disciplinas acadêmicas diferentes, com seus conjuntos de problemas e métodos de resolvê-los (Quadro 4).

Quadro 4 – As principais disciplinas que estudam língua

<b>Disciplina</b>	<b>Problemas típicos</b>	<b>Ferramentas</b>
<b>Linguística</b>	Como as palavras formam frases e sentenças? O que restringe os possíveis significados para uma frase?	Intuições sobre boa formação e significado; modelos matemáticos de estrutura (por exemplo, a teoria da linguagem formal, modelo semântico teórico).
<b>Psicolinguística</b>	Como as pessoas identificam a estrutura das sentenças? Como são os significados das palavras identificadas? Quando é que a compreensão ocorre?	Técnicas experimentais com base na medição do desempenho humano; análise estatística das observações.
<b>Filosofia</b>	Qual é o sentido, e como palavras e frases o adquirem? Como as palavras identificam objetos no mundo?	Argumentação da linguagem natural usando a intuição sobre contra-exemplos; modelos matemáticos (por exemplo, lógica e modelo de teoria).
<b>Linguística computacional</b>	Como a estrutura das frases é identificada? Como o conhecimento e raciocínio podem ser modelados? Como a linguagem pode ser usada para realizar tarefas específicas?	Algoritmos, estrutura de dados, modelos formais de representação e raciocínio. Técnicas de inteligência artificial (métodos de busca e representação).

Fonte: Allen (1994).

Este trabalho utiliza a linguística computacional. Segundo Vieira e Lima (2001), a linguística computacional pode ser entendida como a área de conhecimento que explora as relações entre linguística e informática, tornando possível a construção de sistemas com capacidade de reconhecer e produzir informação apresentada em linguagem natural. Tal área, segundo Othero e Menuzzi (2005), pode ser dividida em duas subáreas: linguística de *corpus* e PLN. Ainda segundo os autores, essa divisão nem sempre é nítida, uma vez que há muitos trabalhos que envolvem as duas áreas. A linguística de *corpus* preocupa-se basicamente com o trabalho a partir de “*corpora* eletrônico”, isto é, grandes bancos de dados que contenham amostras de linguagem natural (OTHERO; MENUZZI, 2005). Mais especificamente, a linguística de *corpus* se ocupa da coleta e exploração de conjuntos de dados linguísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística (SARDINHA, 2000).

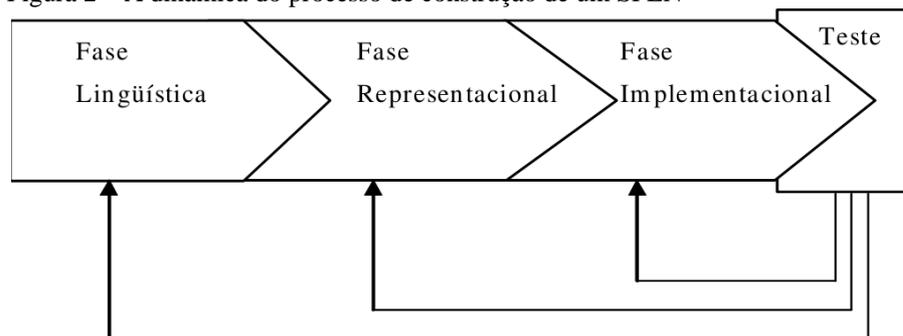
A área de PLN, por outro lado, preocupa-se diretamente com o estudo da linguagem voltado para a construção de *softwares* e sistemas computacionais específicos, como tradutores automáticos, reconhecedores automáticos de voz, geradores automáticos de resumos etc. (OTHERO; MENUZZI, 2005).

Segundo Vieira (2002), para o PLN, vários subsistemas são necessários para dar conta dos diferentes aspectos da língua: sons, palavras, sentenças e discurso nos níveis estruturais, de significado e de uso. De forma geral, o processo de desenvolvimento de Sistemas de PLN (SPLN) exige o trabalho de, no mínimo, três fases (DIAS DA SILVA et al., 2007):

- a) fase linguística: envolve a construção do corpo de conhecimentos sobre a própria linguagem, compreendendo os fenômenos linguísticos para o desenvolvimento do sistema;
- b) fase representacional: envolve a construção conceitual do sistema, projetando representações linguísticas e extralinguísticas em sistemas formais computacionalmente tratáveis;
- c) fase implementacional: envolve a codificação das representações elaboradas durante a fase anterior em termos de linguagens de programação e planejamento global do sistema.

Os autores consideram que essas três fases (Figura 2) devem ser desenvolvidas sucessiva, progressiva e ciclicamente, de forma que os testes possam contribuir para o aprimoramento dos resultados alcançados em cada fase.

Figura 2 – A dinâmica do processo de construção de um SPLN



Fonte: Dias da Silva et al., (2007).

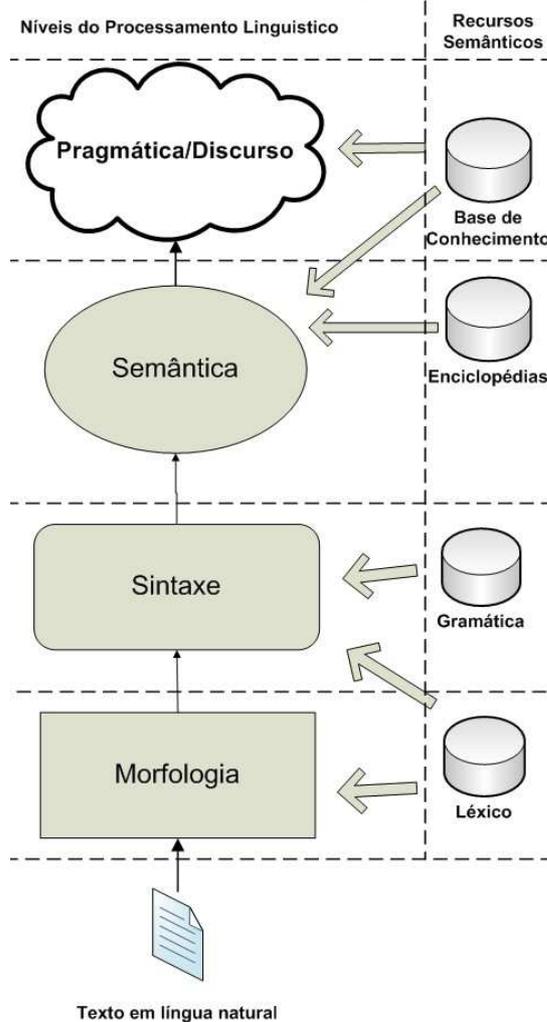
Projetar um SPLN envolve essencialmente (i) especificar, (ii) representar e (iii) codificar sistematicamente um volume considerável de informações (linguísticas e extralinguísticas), mecanismos de inferência e de controle dessas inferências, e, finalmente, projetar um sistema computacional (incluindo software e hardware) para o desenvolvimento e teste do próprio empreendimento. (DIAS DA SILVA et al., 2007, p. 16).

Dessa forma, Dias da Silva et al. (2007) consideram o estudo de PLN tanto importante para obtenção de subsídios para a implementação de programas computacionais como para desenvolver projetos de caráter acadêmico, utilizando-se do conhecimento linguístico para o tratamento das línguas naturais.

Em PLN, o material de entrada do processamento é um texto que deve ser analisado e recortado em unidades menores para a compreensão completa dos mecanismos de operação envolvidos em cada uma dessas unidades. A fim de compreender melhor como ocorre o

processamento de uma informação textual, deve-se analisar as fases do PLN. As quatro fases (ou níveis) que o programa deverá executar são: 1) Análise Morfológica; 2) Análise Sintática; 3) Análise Semântica; e 4) Análise Pragmática (Figura 3).

Figura 3 – Visão geral dos diferentes níveis de processamento linguístico em PLN



Fonte: Franco (2013) adaptada pela autora.

Segundo Silva et al. (2007) apud Franco (2013), os níveis de processamento linguístico em PLN são:

- 1) Análise morfológica — neste nível as unidades mínimas dotadas de significado são isoladas para a compreensão do processo de formação e flexão das palavras. Alguns exemplos de processadores nesse nível são: 1) Detector de Sentenças, 2) Tokenizador e 3) POS-Tagger. O Detector de Sentenças é responsável por dividir o texto em sentenças. O Tokenizador tem como principal objetivo separar as sentenças em

unidades menores chamadas de tokens, tais como palavras e símbolos de pontuação. No exemplo: “O Twitter está lento!”, tem-se os seguintes tokens: “O”, “Twitter”, “está”, “lento” e “!”. O POS-Tagger classifica cada token de acordo com sua classe gramatical, com base em consultas ao Léxico. Léxico é um recurso que guarda um conjunto de palavras e variações dessas palavras de acordo com a forma. Por exemplo: “O” -> Artigo, “Twitter”-> Substantivo, “está”-> Verbo, “lento”-> Adjetivo.

- 2) Análise sintática — neste nível considera-se que a organização das palavras e expressões resulta em determinadas funções que elas desempenham na sentença. Este nível é responsável por construir (ou recuperar) uma estrutura sintática válida para a sentença de entrada. Para tanto, é guiado por uma representação da gramática da língua em questão. Em se tratando de uma língua natural, em geral adota-se uma gramática “parcial” da língua natural, que, embora não abranja todas as construções da língua, contempla aquelas construções válidas de interesse para a aplicação. Alguns exemplos de processadores são: 1) *Lemmatizer*, 2) *Chunker* e 3) *Dependence Parser*. O *Lemmatizer* tem como objetivo retornar o lemma de uma palavra. O lemma é a forma canônica de uma palavra. Por exemplo, na palavra “é” o lemma é “ser”. O *Chunker* agrupa, identifica e classifica os sintagmas de uma sentença. Por exemplo, na sentença “O Twitter está travando demais hoje” tem-se os seguintes sintagmas: “O Twitter”-> Sintagma Nominal, “está travando demais hoje”-> Sintagma Verbal. O *Dependence Parser* retorna a árvore de dependência sintática de uma sentença. O parser realiza a análise sintática automática de frases em termos de suas funções gramaticais, ou seja, se colocarmos uma frase como “A menina joga”, o parser é capaz de processá-la e dar como saída a sinalização do que é artigo, sujeito e complemento da frase. A saída então será uma árvore onde os nós folha são as palavras da sentença. No *Dependence Parser* cada palavra da frase tem uma relação com seus dependentes. Por exemplo, saber a identidade do verbo ajuda a determinar qual é o sujeito e qual é o objeto na frase.
- 3) Análise semântica — esse nível é responsável pela interpretação de componentes da sentença ou da sentença como um todo e está presente sempre que a aplicação exigir algum tipo de interpretação. Nesse caso, é necessário conhecimento mais específico do domínio, presente no Modelo do Domínio. Por exemplo, para distinguir se a interpretação correta do termo “manga” é parte de um vestuário ou a fruta). Alguns exemplos de processadores nesse nível são: 1) *Word Sense Disambiguation* (WSD)

(STEVENSON; WILKS, 2003 apud FRANCO, 2013) que tem como função identificar o sentido correto de uma palavra ambígua, quando usada em uma sentença em particular. Por exemplo, na sentença “João gosta de manga” a palavra “manga” é usada no sentido de uma fruta; 2) *Semantic Role Labeling* (SRL) (GILDEA; JURAFSKY, 2002 apud FRANCO, 2013) - os papéis semânticos representam as relações lógicas entre um evento e seus participantes; 3) *Semantic Role Labeling* (SRL) é o processo de extrair automaticamente estruturas de papéis semânticos que permitem a análise do significado das sentenças. Por exemplo, na sentença “Marcos quebrou o vaso”, “Marcos” assume o papel de Agente, ou seja, causador voluntário de uma ação, e o “vaso” o papel de Paciente, ou seja, quem ou o quê sofre a ação; 4) *Named Entity Recognition* (NER) (NADEAU; SEKINE, 2007 apud FRANCO, 2013) tem como função identificar e classificar algumas entidades especiais do texto, tais como, localização, pessoas, tempo, organização, datas, quantidades e valores etc. Por exemplo, na sentença “Larissa estudou na UFC.”, o processador NER deve reconhecer “Larissa” como pessoa e “UFC” como instituição.

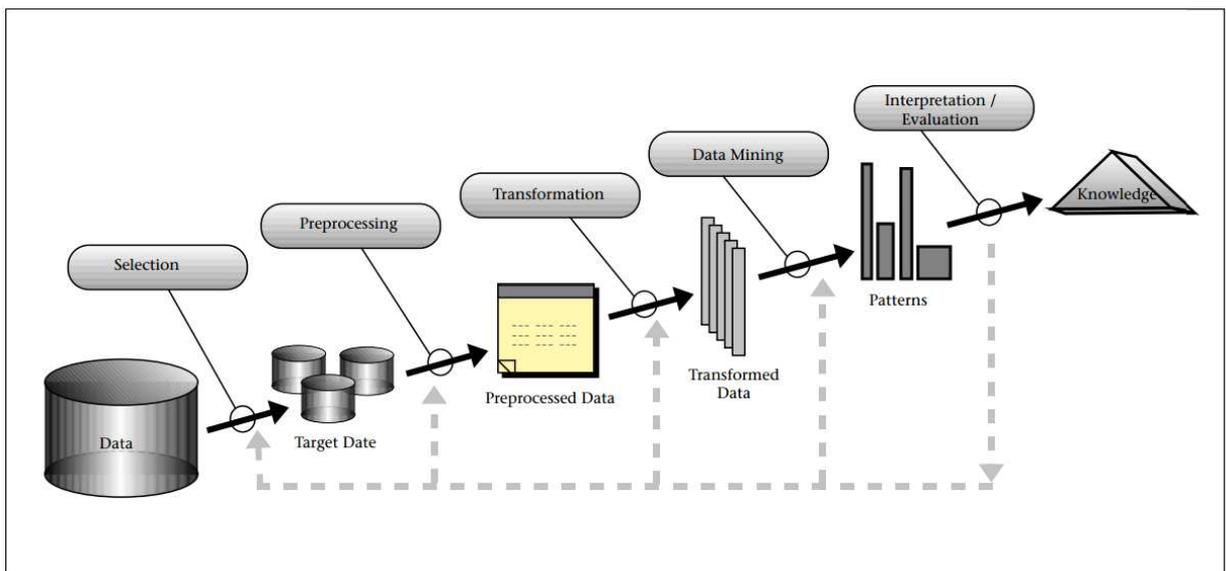
- 4) Análise discursiva (ou pragmática) — este nível permite a obtenção de uma representação do significado da mensagem original, levando em conta aspectos pragmáticos da comunicação. Por exemplo, nem sempre o caráter interrogativo de uma sentença expressa exatamente o caráter de solicitação de uma resposta. A sentença “Você sabe que horas são?” pode ser interpretada como uma solicitação para que as horas sejam informadas ou como uma repreensão por um atraso ocorrido. No primeiro caso, a pergunta informa ao ouvinte que o falante deseja obter uma informação e, portanto, expressa exatamente o caráter interrogativo. Entretanto, no segundo caso, o falante utiliza o artifício interrogativo como forma de impor sua autoridade. Estas diferentes interpretações são claramente relacionadas com a prática da língua no dia-a-dia e com as relações sociais presentes no contexto discursivo. Para apreender essa prática, os processadores recorrem a bases de conhecimento de mundo, especificamente, a bases de senso comum. Alguns processadores desse nível são: 1) *Coreference Resolution* (CR) (SOON; NG; LIM, 2001 apud FRANCO, 2013), que tem como principal função identificar cadeias de correferência, ou seja, um grupo de palavras ou expressões que se referem a uma mesma entidade. Por exemplo: “Nilce viajou. Ela estava em Israel.”, a cadeia {“Ela”, “Nilce”} refere-se a mesma entidade - uma pessoa chamada Nilce; 2) *Question & Answering* (LEE et al., 2001 apud

FRANCO, 2013), tem como principal função responder uma pergunta feita por um usuário. Por exemplo, “Onde nasceu o papa Francisco?” o sistema deve procurar uma fonte confiável e responder: “Em Flores, Buenos Aires, Argentina”; 3) *Sentiment Analyzer* (TABOADA; ANTHONY; VOLL, 2006; BONTICHEVA et al., 2013 apud FRANCO, 2013) refere-se ao uso de PLN, análise de texto e linguística computacional para identificar e extrair informações subjetivas de textos como tristeza, alegria, polaridade positiva ou negativa etc. (FRANCO, 2013).

## 2.4 MINERAÇÃO DE DADOS

Segundo Tan, Steinbach e Kumar (2009), a mineração de dados é o processo de descoberta automática de informações úteis em grandes depósitos de dados. Tal processo combina métodos tradicionais de análise de dados com algoritmos sofisticados para processar grandes volumes de dados e é uma parte integral da descoberta de conhecimento em banco de dados (KDD – *Knowledge Discovery in Databases*). Fayyad, Piatetsky-Shapiro e Smyth (1996) especificam o processo detalhadamente, com etapas como: 1) dados; 2) seleção; 3) pré-processamento; 4) transformação; 5) mineração de dados; e 6) interpretação/avaliação (Figura 4).

Figura 4 – Etapas do KDD



Fonte: Fayyad, Piatetsky-Shapiro; Smyth (1996).

Segundo Fayyad, Piatetsky-Shapiro e Smyth (1996) apud Corrêa (2003), as etapas pertencentes ao KDD são trabalhadas da seguinte forma:

- 1) dados — a descoberta de conhecimento em banco de dados se baseia nos dados armazenados de forma estruturada;
- 2) seleção de dados — nessa etapa, deve-se entender o domínio da aplicação e conhecer os objetivos finais a serem atingidos para que se possa selecionar um conjunto de dados, do qual será feita a descoberta do conhecimento;
- 3) pré-processamento — os dados são processados visando adequá-los aos algoritmos, ou seja, dados incompletos, problemas de definição de tipos, eliminação de tuplas repetidas etc. Algumas funções de pre-processamento de dados são:
  - a) limpeza de dados: é realizado um tratamento nos dados a fim de corrigir informações errôneas ou inconsistentes nas bases de dados de forma a não comprometer a conhecimento a ser extraído no final do processo de KDD;
  - b) *stemming*: aumenta a ocorrência do radical nos documentos, por exemplo: conectado, conectando, conectar, ficaria =>conect, conect, conect;
  - c) *thesaurus*: trata-se de uma compilação de palavras indexadas com significados similares (sinônimos), relacionados e opostos (antônimos). Palavras que possuem sinônimos podem ser expressas de uma só maneira. Por exemplo, as palavras “bonito”, “belo”, “lindo” são sinônimas. Com o objetivo de reduzir esses termos no texto, pode-se usar somente a palavra “belo” onde ocorrerem as palavras “bonito” ou “lindo” (MAZIERO et al., 2008);
  - d) *stopwords*: são palavras consideradas irrelevantes para a análise de textos ou para a busca. Na maioria das vezes, são palavras auxiliares ou conectivas, não fornecendo nenhuma informação discriminativa na expressão do conteúdo do texto. Palavras como pronomes, artigos, preposições e conjunções podem ser consideradas *stopwords*. A lista de *stopwords* é chamada de *Stoplist*.
- 4) transformação — nessa etapa, os dados deverão ser armazenados adequadamente para facilitar na utilização das técnicas de mineração de dados. Uma abordagem muito utilizada nesta etapa é a *bag-of-words*. Nessa abordagem, cada palavra encontrada na coleção de documentos torna-se um atributo no espaço vetorial (palavras que ocorrem no texto pelo menos uma vez).
- 5) mineração de dados — nessa etapa, são processados os algoritmos de aprendizado de máquina e de reconhecimento de padrões. O termo “padrões” significa alguma

abstração de um subconjunto de dados em alguma linguagem descritiva de conceitos (MORAIS; AMBRÓSIO, 2007). A maioria dos algoritmos é baseada em conceitos de aprendizagem de máquina, reconhecimento de padrões, estatística, classificação, clusterização e modelos gráficos;

- 6) interpretação/avaliação — a etapa final do processo consiste em interpretar os resultados obtidos e elaborar uma apresentação de modo que o usuário possa entender.

Os dados podem ser obtidos em um banco de dados ou pela utilização de técnicas de Recuperação de Informação (RI) ou de Extração de Informação (EI). Segundo Eikvil (1999), RI recupera documentos relevantes de uma coleção, enquanto EI extrai informações relevantes de documentos. A EI tem como finalidade extrair informações específicas do texto em língua natural (RILOFF; JONES, 1999). A informação extraída é determinada por um conjunto de padrões ou regras de extração específicas ao domínio de análise: padrões que podem ser escolhidos de maneira manual, por algum especialista, ou de forma automatizada (GAIZAUSKAS; WILKS, 1998).

O processo KDD também pode ser chamado de mineração de textos<sup>12</sup> (CORRÊA; 2003; BARION; LAGO, 2008; ARANHA; PASSOS, 2006; MOONEY; NAHM, 2005; SILBERCHATZ; KORTH; SUDARSHAN, 2006). No entanto, na mineração de dados, a extração dos dados é realizada em dados estruturados, e, em mineração de textos e mineração web<sup>13</sup>, os dados não são estruturados (MORAIS; AMBRÓSIO, 2007; BARION; LAGO, 2008; KAMBER, 2001 apud CORRÊA, 2003). Para tais processos, têm-se as seguintes definições:

A Mineração de Textos surgiu da intersecção de várias áreas de pesquisa, tais como Recuperação de Informação, Processamento da Linguagem Natural, Extração de Informação e Mineração de Dados, visando auxiliar na extração de informação útil em textos. (CORRÊA, 2003, p. 14).

Mineração Web pode ser conceituada como a descoberta e análise inteligente de informações úteis da Web. (COOLEY; MOBASHER; SRIVASTAVA, 1997, p. 558).

Segundo Marinho e Girardi (2003), um aspecto que diferencia a mineração de dados tradicional da mineração na web é que dados da web são muito mais sofisticados e dinâmicos

---

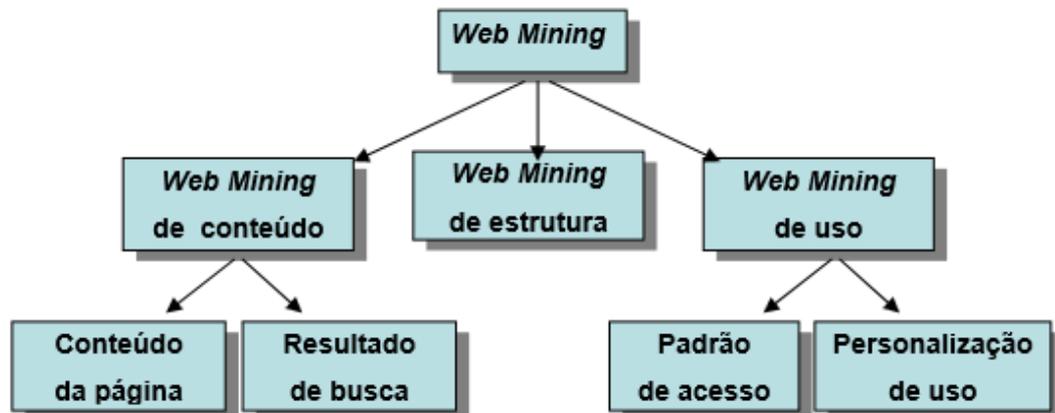
<sup>12</sup> *Text Mining*. Na literatura, é possível encontrar outras nomenclaturas sinônimas como Mineração de Dados em Textos (*Text Data Mining*) ou descoberta de conhecimento a partir de banco de dados textuais (*Knowledge Discovery from Textual Databases*) (LOPES, 2004).

<sup>13</sup> *Web Mining*

do que sistemas de armazenamento de bancos de dados tradicionais. O processo de mineração web pode ser dividido em três categorias que se subdividem em outras (Figura 5):

- a) mineração web de conteúdo, que consiste em minerar o conteúdo extraído de textos, imagens, áudios, vídeos e registros estruturados, como listas e tabelas;
- b) mineração web de estrutura, que consiste na tentativa de extração de informação baseada nas ligações entre páginas por meio dos hiperlinks contidos nelas. Pode-se, por exemplo, verificar a popularidade de uma página em um determinado contexto pela quantidade de referências a ela em outras páginas. Além desse, outras características podem ser extraídas a partir da análise da estrutura de grafos da web (SINGH; KUMAR, 2009); e
- c) mineração web de uso, que visa utilizar técnicas de mineração para descobrir padrões de uso das aplicações na tentativa de prever o comportamento do usuário ao interagir com a web a partir dos dados gerados pelas transações cliente-servidor.

Figura 5 – Abordagens da mineração na web



Fonte: Carrilho Junior (2007).

Esta tese se situa na área categorizada como mineração web de conteúdo, uma vez que será analisado o conteúdo encontrado em SS, como base de dados. Pretende-se extrair somente os textos – postagens dos usuários e extrair padrões úteis de identificação desses dados. Segundo Castanheira (2008), a mineração de dados difere de técnicas estatísticas porque, ao invés de verificar padrões hipotéticos, utiliza os próprios dados para descobrir tais padrões. Para esse fim, a área de mineração baseia-se na utilização de algoritmos capazes de vasculhar grandes bases de dados, de modo eficiente, e revelar padrões interessantes implícitos neles (GONÇALVES, 2015). Segundo Tan, Steinbach e Kumar (2009), os

algoritmos de mineração de dados são divididos em: algoritmos de classificação, de regressão, de associação, de agrupamento e de detecção de anomalias.

Neste trabalho, optou-se pela realização de experimentos utilizando algoritmos de classificação. A tarefa de classificação é uma função de aprendizado que mapeia dados de entrada, ou conjunto de dados de entrada, em um conjunto finito de classes (CASTANHEIRA, 2008). Segundo Castanheira (2008), o objetivo de um algoritmo de classificação é encontrar alguma correlação entre os atributos e uma classe, de modo que o processo de classificação possa usá-la para prever a classe de um exemplo novo e desconhecido. Os principais algoritmos de classificação são: árvores de decisão, redes neurais, redes bayesianas, máquinas de vetores de suporte, regras de decisão, método kNN (*k-nearest-neighbor*), algoritmos genéticos e conjuntos fuzzy (TAN; STEINBACH; KUMAR, 2009). Os algoritmos utilizados nesta tese são: árvores de decisão e redes bayesianas, explicados a seguir.

#### 2.4.1 Algoritmos de mineração de dados

Esta seção apresenta um resumo dos algoritmos de mineração de dados árvores de decisão e redes bayesianas, utilizados neste estudo.

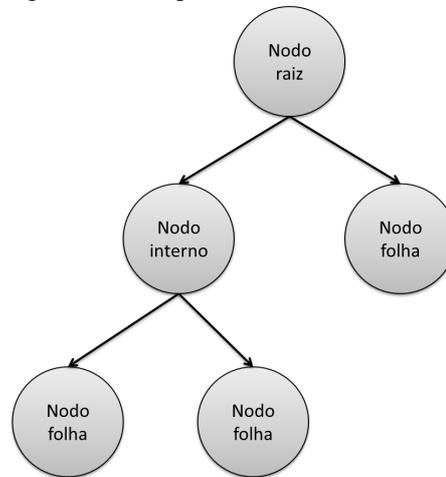
O algoritmo de classificação por árvores de decisão é uma técnica de classificação simples e muito usada (TAN; STEINBACH; KUMAR, 2009). Os classificadores baseados em árvore de decisão têm, como base, a estratégia dividir para conquistar<sup>14</sup>, em que os dados de um problema são divididos em vários subconjuntos, de forma a cada subconjunto ser formado de acordo com características semelhantes dos dados. Desta maneira, os classificadores baseados em árvores de decisão buscam meios de dividir um conjunto de dados em vários subconjuntos, conhecidos como nodos. A classificação, através de uma árvore de decisão, ocorre à medida que são percorridos os caminhos descritos pelos nodos, até ser encontrado um nodo que contém a característica determinante do caminho seguido, recebendo então o nome de folha.

A estrutura básica de uma árvore (Figura 6) pode ser formada por três tipos de nodos: o nodo raiz, que representa o início da árvore, os nodos internos, que dividem um determinado atributo e geram ramificações, e os nodos folha ou terminais, que contêm as informações de classificação do algoritmo (TAN; STEINBACH; KUMAR, 2009).

---

<sup>14</sup> *Divide and conquer* (SOUSA, 1998)

Figura 6– Exemplo de uma árvore de decisão



O algoritmo de classificação por árvores de decisão é um algoritmo supervisionado, ou seja, é necessário conhecer todos os registros e as informações da base de dados utilizadas no treinamento (RUSSELL; NORVIG, 1995). O Quadro 5 apresenta um exemplo de um conjunto de dados usado para classificar vertebrados nas seguintes classes: mamíferos, aves, peixes, répteis e anfíbios. O conjunto de atributos inclui propriedades de um vertebrado como sua temperatura corporal, cobertura de pele, método de reprodução, habilidade de voar e habilidade de viver na água. Cada registro é conhecido como uma instância ou exemplo.

Quadro 5 – Exemplo de um conjunto de dados: vertebrados

Nome	Temperatura corporal	Cobertura de pele	Dá cria	Ser Aquático	Ser Aéreo	Possui Pernas	Hiberna	Rótulo da classe
Humano	Sangue quente	Cabelo	Sim	Não	Não	Sim	Não	Mamífero
Pítton	Sangue frio	Escamas	Não	Não	Não	Não	Sim	Réptil
Salmão	Sangue frio	Escamas	Não	Sim	Não	Não	Não	Peixe
Baleia	Sangue quente	Cabelo	Sim	Sim	Não	Não	Não	Mamífero
Sapo	Sangue frio	Nenhuma	Não	Sim	Não	Sim	Sim	Anfíbio
Morcego	Sangue quente	Cabelo	Sim	Não	Sim	Sim	Sim	Mamífero
Pomba	Sangue quente	Penas	Não	Não	Sim	Sim	Não	Ave
Gato	Sangue quente	Pêlo	Sim	Não	Não	Sim	Não	Mamífero
Tubarão	Sangue frio	Escamas	Não	Sim	Não	Não	Não	Peixe
Tartaruga	Sangue frio	Escamas	Não	Semi	Não	Sim	Não	Réptil
Pinguim	Sangue quente	Penas	Não	Semi	Não	Sim	Não	Ave
Salamandra	Sangue frio	Nenhuma	Não	Semi	Não	Sim	Sim	Anfíbio

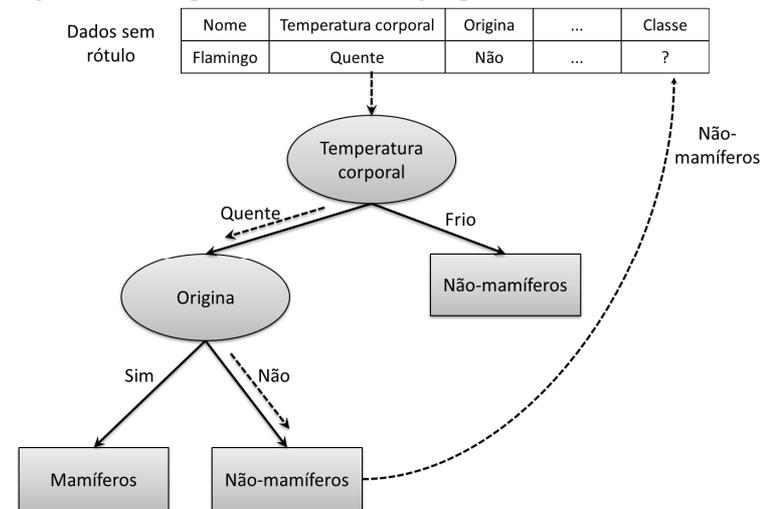
Fonte: Tan; Steinbach; Kumar (2009).

O processo de geração da árvore realizada pelo algoritmo inicia-se com a definição de quais são os elementos, ou seja, os valores dos nodos da árvore. Tan, Steinbach e Kumar (2009) apresentam um exemplo (Figura 7) de uma estrutura de uma árvore de decisão para o

problema de classificação de mamíferos, usando o conjunto de dados do Quadro 5 como treinamento:

O nodo raiz usa o atributo Temperatura Corporal para separar vertebrados de sangue quente dos de sangue frio. Já que todos os vertebrados de sangue frio não são mamíferos, um nodo folha com rótulo Não Mamífero é criado como o filho à direita do nodo raiz. Se o vertebrado tiver sangue quente, um atributo subsequente, Origina é usado para distinguir mamíferos de outras criaturas de sangue quente, que são na sua maioria pássaros (TAN; STEINBACH; KUMAR, 2009, p.178).

Figura 7 – Exemplo de uma classificação por árvore de decisão



Fonte: Tan; Steinbach; Kumar (2009)

Na Figura 7, para identificar se um novo registro (Flamingo) pertence à classe dos Mamíferos ou Não-mamíferos, o processo de classificação começa pelo nodo raiz, aplicando a condição de teste ao registro (Qual a temperatura corporal?) e seguindo a ramificação apropriada baseados nos resultados do teste (Quente). Isso levará a outro nodo interno, para o qual uma nova condição de teste é aplicada (Origina?), ou a um nodo folha (Frio). O rótulo da classe associada ao nodo folha é, então, atribuído ao registro (Não-mamífero). As linhas tracejadas representam o resultado da aplicação de testes sobre atributos para o vertebrado não rotulado.

Segundo Tan, Steinbach e Kumar (2009), há muitas árvores de decisão que podem ser construídas a partir de um determinado conjunto de atributos. Alguns dos algoritmos de árvores de decisão existentes são: CART, ID3 e C4.5.

Neste trabalho será usado o algoritmo C4.5 (QUINLAN, 1993 apud TAN; STEINBACH; KUMAR, 2009) no software WEKA<sup>15</sup> (*Waikato Environment for Knowledge*

<sup>15</sup> Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 15 de outubro de 2013

*Analysis*). O WEKA é um software de domínio público, da Universidade de Waikato, Nova Zelândia. Ele é constituído de uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados. Ele pode ser usado para aplicar métodos de aprendizado a um conjunto de dados e analisar a saída para extrair informações a partir dos dados de entrada (WITTEN; FRANK, 1999 apud ROSSI; REZENDE, 2011).

O algoritmo C4.5 é denominado de J48 no WEKA. Um dos aspectos para a grande utilização do algoritmo J48 pelos especialistas em mineração de dados é que o mesmo se mostra adequado para os procedimentos, envolvendo as variáveis (dados) qualitativas contínuas e discretas presentes nas bases de dados (LIBRELOTTO, 2013). Segundo Librelotto (2013), o algoritmo J48, proposto por QUINLAN (1993), é considerado o que apresenta o melhor resultado na montagem de árvores de decisão, a partir de um conjunto de dados de treinamento.

Segundo Fagundes da Silva (2004), esse algoritmo pode ser utilizado para a classificação de documentos descritos em termos dos mesmos atributos usados na sua representação. Neste caso, os atributos são os tokens e as classes são as categorias a serem classificadas. Isso é feito percorrendo-se a árvore recursivamente, até se chegar à folha que determina a classe a que o documento pertence ou, alternativamente, sua probabilidade de pertencer àquela classe. Por exemplo, para classificar textos de um jornal nas seguintes classes: Informática e Saúde, os atributos: “computador”, “impressora” e “software” têm uma maior probabilidade de pertencer à classe Informática, pois ocorrem com mais frequência nesta categoria, enquanto que os atributos: “dengue”, “hospital” e “emergência” têm uma maior probabilidade de pertencer a classe Saúde.

O uso desse algoritmo neste trabalho tem como objetivo aprender regras de classificação das PRUs e de suas características. Segundo Correa (2002) apud Fagundes da Silva (2004), uma maneira de se efetuar o aprendizado de regras consiste na construção inicial de uma árvore de decisão e na posterior tradução dessa para um conjunto equivalente de regras, onde a cada folha da árvore corresponde uma regra correspondente.

O outro algoritmo utilizado neste trabalho é o Redes bayesianas. Este algoritmo também é de aprendizado supervisionado e baseia-se no teorema de probabilidade de Bayes, também conhecido por classificador de Naïve Bayes (TAN; STEINBACH; KUMAR, 2009). Ele permite calcular as probabilidades de um novo documento pertencer a cada uma das categorias e atribuir a este às categorias com valores maiores de probabilidade (LEWIS; RINGUETTE, 1994 apud FAGUNDES DA SILVA, 2004). Segundo Grobelnik e Mladenic

(1998) apud Fagundes da Silva (2004), o classificador também pode informar a probabilidade daquele documento pertencer a uma determinada categoria, dentro de uma estrutura hierárquica. Ele supõe que todos os atributos dos exemplos são independentes uns dos outros, dado o contexto da categoria, isto é, a probabilidade de ocorrer uma palavra não depende da existência de outra(s). (DOMINGOS; PAZZANI, 1997 apud TAN; STEINBACH; KUMAR, 2009; FAGUNDES DA SILVA, 2004).

A técnica Naïve Bayes é uma das técnicas de aprendizagem bayesiana mais utilizadas em problemas de classificação de textos (MITCHELL, 1997; McCALLUM; NIGAM, 1998) apud (FAGUNDES DA SILVA, 2004) e, por esse motivo, ela também será utilizada neste trabalho. A ideia básica é usar a probabilidade na estimativa de uma dada categoria presente em um documento ou texto para identificação de padrões relevantes sobre o uso do sistema.

## 2.5 AVALIAÇÃO EMOCIONAL X ANÁLISE DE SENTIMENTOS

As emoções são um fator importante da vida e desempenham um papel essencial para entender o comportamento do usuário na interação com o computador (CRISTESCU, 2008). Hume (2015) fez uma distinção entre afeto, emoção e humor, da seguinte forma: afeto é um termo genérico que cobre uma ampla gama de sentimentos que as pessoas experimentam. É um conceito amplo que abrange ambas as emoções e estados de espírito; emoções são sentimentos intensos que são dirigidos a alguém ou alguma coisa; e humor é um sentimento que tende a ser menos intenso do que as emoções e que, muitas vezes (mas nem sempre), não possuem um estímulo contextual (HUME, 2015). O autor também afirma que existem dezenas de emoções, como raiva, desprezo, entusiasmo, inveja, medo, frustração, decepção, vergonha, nojo, felicidade, ódio, esperança, ciúme, alegria, amor, orgulho, surpresa e tristeza.

Ekman (1992) apud Thelwall, Wilkinson e Uppal (2010), por outro lado, limita a quantidade de emoções em seis emoções básicas: felicidade, tristeza, surpresa, medo, nojo e raiva, considerando que as demais são combinações dessas. Thelwall, Wilkinson e Uppal (2010), observaram que há duas dimensões fundamentais, em vez de uma série de tipos de emoções diferentes: positiva ou negativa. Voeffray (2011) considera que emoções positivas e negativas influenciam o consumo de produtos e a compreensão do processo de decisão, sendo cruciais para medir a expressão emocional.

Outra visão sobre a emoção é apresentada por Norman (2004). O autor, em seus estudos sobre emoção, caracterizou o comportamento afetivo dos seres humanos em três níveis de estrutura do cérebro:

*O nível visceral*, automático ou pré-programado, faz julgamentos rápidos — como o que é bom ou ruim, seguro ou perigoso. *O nível comportamental* refere-se aos processos cerebrais que controlam a maior parte de nossas ações — como andar de bicicleta, tocar um instrumento musical, dirigir um carro. *O nível reflexivo* refere-se à interpretação, compreensão e raciocínio e à parte contemplativa do cérebro. É nele que são processadas ações como apreciar uma obra de arte, sentir saudades de um amigo, torcer para um time de futebol (NORMAN, 2004, p. 14).

Norman (2004) associou estes três níveis do processamento humano a aspectos do design. O *design visceral* diz respeito às aparências — ao primeiro impacto causado por um produto. O *design comportamental* diz respeito ao uso sob o ponto de vista objetivo e refere-se à função que o produto desempenha, à eficácia com que cumpre sua função, à facilidade com que o usuário o compreende e o opera e demais aspectos relacionados ao modo como o produto “se comporta” junto ao usuário. O *design reflexivo*, por sua vez, considera a racionalização de um produto, a memória afetiva e os significados atribuídos aos produtos e a seu uso. Mais especificamente, o autor descreve as seguintes características para cada nível:

Os níveis *visceral* e *comportamental* se referem ao “*agora*”, seus sentimentos e experiências enquanto se está de fato vendo ou usando o produto. Mas o nível reflexivo se estende por muito mais tempo — por meio da reflexão você se lembra do passado e considera o futuro. O *design reflexivo*, portanto, tem a ver com relações de longo prazo, com os sentimentos de satisfação produzidos por ter, exibir, e usar um produto. O sentido de identidade própria de uma pessoa está situado no nível reflexivo, e é nele que a interação entre o produto e sua identidade é importante, conforme demonstra o orgulho (ou a vergonha) de ser dono ou de usar o produto. A interação entre cliente e serviço é importante nesse nível (NORMAN, 2004, p. 58).

O Quadro 6 ilustra, de forma resumida, as características apresentadas por Norman para os três níveis de processamento.

Quadro 6 – Características apresentadas por Norman (2004) para os três níveis de processamento

Nível	Características apresentadas por Norman
Visceral	aparência, “primeiro impacto causado por um produto”, “primeiras impressões”, “manifesta no tempo presente”, atração
Comportamental	“efetividade do uso”, “função que o produto desempenha”, “eficácia com que o produto cumpre sua função”, “facilidade com que o usuário compreende e opera o produto”, “como o produto se “comporta” com o usuário”, função, desempenho, usabilidade, uso, “manifesta no tempo presente”, utilidade
Reflexivo	“memória afetiva”, “particularidades culturais e individuais”, pessoal, lembranças, consciência, racionalização, inteligibilidade, “ponto de vista subjetivo”, “estende-se por muito tempo”, envolve sentimentos”, afeto

Segundo De Lera e Garreta-Domingo (2007), avaliar as emoções dos usuários não é uma tarefa fácil e de baixo custo. Profissionais de usabilidade até recentemente baseavam-se na avaliação da satisfação dos usuários, negligenciando as emoções por meio de testes de desempenho como número de erros, tempo e outras medidas objetivas e também de avaliações subjetivas da experiência do usuário como aplicações de questionários pós-teste.

Na literatura, existe uma variedade de métodos, técnicas, ferramentas e instrumentos para avaliação das emoções. Segundo Scherer (2005) apud Xavier (2013), para uma avaliação emocional<sup>16</sup> é preciso avaliar:

(1) as contínuas mudanças nos processos de avaliação em todos os níveis de processamento do sistema nervoso central; (2) os padrões de respostas gerados nos sistemas nervoso neuroendócrino, autonômico e somático; (3) as mudanças motivacionais produzidas pelo resultado da avaliação de tendências de ação particulares; (4) os padrões de expressões facial e vocal, bem como os movimentos do corpo; e (5) a natureza do estado dos sentimentos subjetivos experimentados que refletem todas essas mudanças dos componentes. (SCHERER, 2005 apud XAVIER, 2013, p. 44).

Xavier (2013) resume esse texto, sugerindo a avaliação de cinco componentes (subsistemas) da emoção: (1) avaliações cognitivas, (2) reações fisiológicas, (3) tendências comportamentais, (4) expressões motoras e (5) sentimentos subjetivos. Conforme mencionado, o autor classificou os métodos de avaliação emocional em medidas verbais e não verbais. Medidas verbais são aquelas em que o usuário explicitamente fala o que está sentindo ou utiliza palavras escritas para verbalizar as emoções. Medidas não verbais são consideradas discretas, independente de cultura e linguagem, elas podem ser subjetivas, como a utilização de símbolos universais (DESMET et al., 2003 apud XAVIER, 2013). Como exemplo de avaliação não verbal, tem-se a avaliação por expressões faciais (DE LERA; GARRETA-DOMINGO, 2007).

Xavier (2013) levanta uma discussão sobre a descrição verbal de um sentimento. Scherer (2005) apud Xavier (2013) define sentimento como um dos componentes da emoção, não representando o estado emocional de um indivíduo, mas sim uma atribuição verbal que uma pessoa faz sobre um episódio emocional.

---

<sup>16</sup> Muitos autores se referem à avaliação emocional como afetiva (XAVIER, 2013), tal como o termo emoção e afeto (BENTLY; JOHNSON; BAGGO, 2005 apud XAVIER, 2013; HUME, 2015).

O sentimento é a interpretação consciente que o indivíduo realiza sobre o que ele sente no momento, e que é transmitida de uma maneira verbal, por exemplo, quando uma pessoa diz que se sente “feliz” por ter conseguido fazer uma compra na Internet ou que está “frustrada” por não conseguir se conectar à internet. (XAVIER, 2013, p. 29).

Paralelo a essa linha, existe outro campo de pesquisa chamado Análise de sentimentos (ou mineração de opinião). A análise de sentimentos é o campo de estudo que analisa as opiniões das pessoas, sentimentos, avaliações, atitudes e emoções relacionados a produtos, serviços, organizações, pessoas, problemas, eventos etc., expressos em textos (revisões, blogs, discussões, notícias, comentários, *feedback*, ou qualquer outros documentos) (LIU, 2012). Para Wilson, Wiebe e Hoffmann (2009), a análise de sentimentos é um tipo de análise de subjetividade que foca em identificar opiniões positivas e negativas, emoções e avaliações expressas em linguagem natural. Subjetividade, segundo Stein e Wright (1995), diz respeito à expressão de si mesmo, à representação da perspectiva ou ponto de vista de um locutor (ou de um agente locutório) em um discurso.

Segundo Liu (2012), existem dois conceitos importantes que estão intimamente relacionados ao sentimento e opinião, que são: a subjetividade e a emoção. O autor exemplifica estes conceitos, apresentando a diferença entre sentenças objetivas e subjetivas, sendo a objetiva caracterizada por apresentar algumas informações factuais sobre o mundo, enquanto uma sentença subjetiva expressa alguns sentimentos pessoais, opiniões ou crenças. Segundo o autor, ambos os tipos de sentenças podem apresentar a presença ou a ausência de sentimentos. O quadro 7 apresenta alguns exemplos de sentenças objetivas e subjetivas, com e sem a presença de sentimentos.

Quadro 7 – Exemplos de sentenças objetivas e subjetivas com e sem a presença de sentimentos

Tipo	Sentimento	Sentença
Objetiva	Ausência	“Comprei uma nova televisão.”
Objetiva	Presença	“Então, ela parou de funcionar ontem.”
Subjetiva	Ausência	“Eu acho que ele já tem televisão.”
Subjetiva	Presença	“Odiei a nova televisão que comprei!”

Existem trabalhos nessa área que visam descobrir termos subjetivos ou neutros de frases ou sentenças (WIEBE; BRUCE; O’HARA, 1999; PANG; LEE; VAITHYANATHAN, 2002; KIM; HOVY, 2004), outros que focam no reconhecimento de polaridade a fim de classificar sentenças em positivas ou negativas (PANG; LEE; VAITHYANATHAN, 2002; TURNEY, 2002; ESULI; SEBASTIANI, 2006) e, ainda, aqueles que tentam identificar

diferentes graus de positividade e negatividade, como fortemente negativo, fracamente negativo, neutro, fracamente positivo e fortemente positivo (ESULI; SEBASTIANI, 2006; BROOKE, 2009; WILSON; WIEBE; HOFFMANN, 2009; CARRILLO DE ALBORNOZ et al., 2010).

Nesta tese, não são utilizados algoritmos de análise de sentimentos. No entanto, são realizados estudos de seus conceitos para avaliação de aspectos da UUX, como, por exemplo, identificar se uma opinião do usuário sobre o sistema é positiva ou negativa.

## 2.6 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou os conceitos de SS, UUX, PLN, mineração de dados, análise emocional e análise de sentimentos, a fim de possibilitar o entendimento do contexto e escopo abordados desta tese.

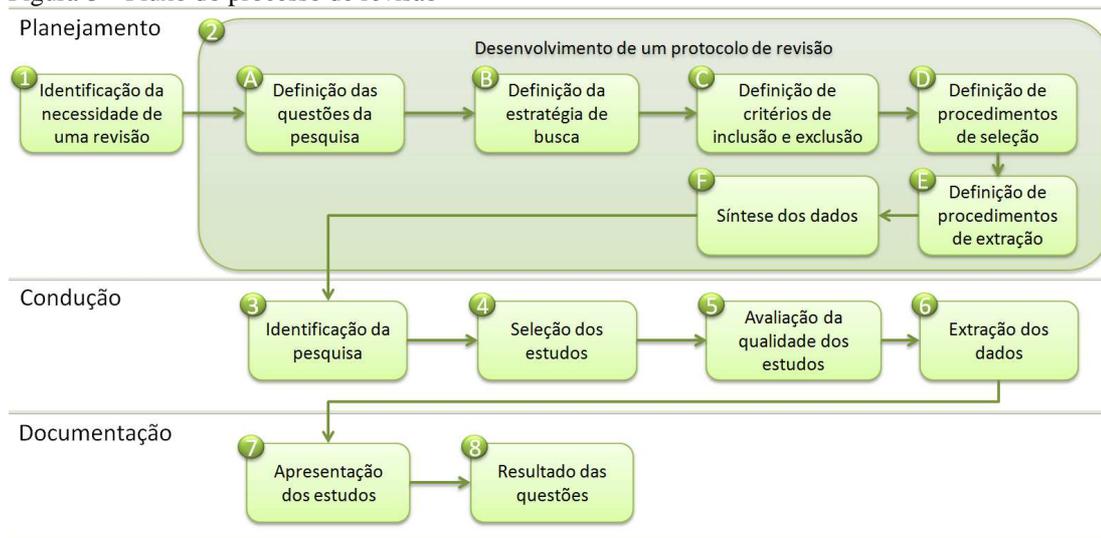
### 3 TRABALHOS RELACIONADOS

Neste capítulo, é apresentada a revisão sistemática realizada, bem como alguns trabalhos correlatos à tese. O objetivo da revisão sistemática é fornecer um estado da arte sobre SS, procurando responder os principais aspectos avaliados e os principais métodos utilizados em trabalhos sobre avaliação de SS. São apresentados trabalhos que têm focado em analisar textos de usuários a fim de avaliar ou estudar UUX.

#### 3.1 PROCESSO DE REVISÃO SISTEMÁTICA

Uma revisão sistemática é um meio de avaliar e interpretar todas as pesquisas disponíveis relevantes para uma determinada questão de pesquisa, área de tópico, ou fenômeno de interesse (KITCHENHAM, 2004). Segundo Kitchenham (2004), as revisões sistemáticas visam apresentar uma avaliação justa de um tema de pesquisa, utilizando uma metodologia confiável e rigorosa. Para a execução da revisão, foi utilizado como base o guia para revisão sistemática de Kitchenham (2004). A Figura 8 exibe o fluxo do processo de revisão utilizado neste trabalho a partir das etapas de Kitchenham (2004). Nas próximas seções, serão descritas as etapas necessárias para o planejamento, condução e documentação da revisão.

Figura 8 – Fluxo do processo de revisão



Fonte: Kitchenham (2004). Adaptado pela autora

### 3.1.1 Planejamento da revisão

**1 – Identificação da necessidade de uma revisão:** existem várias técnicas para avaliar diversos aspectos de SS, no entanto, não é fácil encontrar trabalhos que utilizem as postagens dos usuários no sistema para fazer avaliação dele.

**2 – Desenvolvimento de um protocolo de revisão:** o protocolo de revisão é composto de seis subetapas descritas a seguir.

**A – Definição das questões da pesquisa:** esta revisão sistemática teve como questão de pesquisa principal (QP): “Qual é o estado da arte sobre SS?”. Foram definidas duas subquestões de pesquisa (Questões Secundárias, QS) que ajudam a responder à questão principal:

- a) QS1: Quais são os principais aspectos avaliados em trabalhos sobre avaliação de SS?
- b) QS2: Quais são os principais métodos de avaliação em trabalhos sobre SS?

**B – Definição da estratégia de busca:** em todas as consultas foram utilizadas as seguintes palavras-chave: Facebook, Twitter, Tweet, *Tweets*, “Rede Social”, “Redes Sociais”, “*Social systems*”, “*Social system*”, “*Social Network*”, “*Social Networks*” ou “*Social networking*”<sup>17</sup>. Foram obtidas 999 publicações. Os artigos foram pesquisados no indexador ACM Digital Library<sup>18</sup>, por meio do mecanismo de busca avançada.

**C — Definição de critérios de inclusão e exclusão:** algumas restrições (R) foram utilizadas para limitar a busca, tais como:

- a) **R1:** trabalhos completos (*journal or proceeding*), em português ou inglês, e publicados entre os anos de 2010 a 2012.
- b) **R2:** trabalhos cujo tema é relacionado a avaliação.

**D – Definição de procedimentos de seleção:**

- a) etapa 1: a estratégia de busca é aplicada nas fontes.
- b) etapa 2: para selecionar um conjunto inicial de estudos, os títulos e resumos de todos os artigos obtidos foram lidos e confrontados com os critérios de inclusão e exclusão do item (C, R1), são extraídos: ano e veículo de publicação;

<sup>17</sup> String usada: Searching for: ((Title:Facebook or Title:Twitter or Title:Tweet or Title:*Tweets* or Title:"Social systems" or Title:"Social system" or Title:"Social Network" or Title:"Social Networks" or Title:"Social networking") and (PublishedAs:journal OR PublishedAs:proceeding) and (FtFlag:yes)) and (FtFlag:yes)

<sup>18</sup> ACM. Disponível em <http://portal.acm.org>

- c) etapa 3: todos os artigos selecionados na etapa 2 tiveram uma leitura diagonal e foram novamente confrontados com os critérios do item (C, R2), são extraídos: título, autores, tema e palavras-chave;
- d) etapa 4: todos os artigos selecionados na etapa 3 tiveram uma leitura completa. Os artigos incluídos são documentados e são extraídas as demais informações (observações, método(s) de avaliação e aspecto(s) avaliado(s)).

**E – Definição de procedimentos de extração:** a extração das informações dos artigos foi organizada em uma planilha com itens direcionados a responder as questões de pesquisa da revisão. A planilha continha os seguintes itens a serem preenchidos para cada trabalho: título, ano de publicação, veículo de publicação, autores, palavras-chave, tema, observações, método(s) de avaliação e aspecto(s) avaliado(s).

**F — Síntese de dados:** segundo Kitchenham e Charters (2007), a síntese de dados envolve resumir os resultados dos estudos incluídos, podendo ser descritiva e, quando possível, também quantitativa (envolvendo análise estatística). Nesta revisão, foi realizada uma análise quantitativa, a partir das publicações obtidas e descritiva, pela leitura completa dos trabalhos mais relevantes.

### *3.1.2 Condução da revisão*

A condução da revisão consistiu de quatro subetapas (numeradas de 3 a 6, segundo a Figura 8), descritas a seguir:

**3 – Identificação da pesquisa:** a coleta de informações desta revisão sistemática ocorreu no mês de fevereiro de 2013. A execução das demais atividades ocorreu no primeiro semestre de 2013. Como resultado obtido, foram encontrados 999 trabalhos na ACM Digital Library.

**4 – Seleção dos estudos:** para um refinamento dos resultados, foram seguidas as etapas do procedimento de seleção (E). A partir da etapa 1 foram extraídos 999 trabalhos, que depois de inspecionados, percebeu-se que 395 destes não se tratava de trabalhos completos. Tais trabalhos foram eliminados, resultando em 604 trabalhos para a etapa 2. Na etapa 3, por meio de uma leitura diagonal, foram selecionados trabalhos cujo tema fosse relacionado à avaliação, resultando em 42 trabalhos. A partir de uma leitura completa, 20 trabalhos foram obtidos como resultado da etapa 4 (Tabela 1).

Tabela 1 – Quantidade de trabalhos selecionados em cada etapa da seleção

<b>Etapas</b>	<b>Quantidade de trabalhos</b>
Etapa 1	999
Etapa 2	604
Etapa 3	42
Etapa 4	20

**5 – Avaliação da qualidade dos estudos:** a avaliação da qualidade dos artigos para a última etapa foi realizada de forma simplificada, verificando apenas a presença ou não de alguma forma de avaliação em SS.

**6 – Extração dos dados:** o processo de extração dos dados ocorreu em três situações:

- a) na primeira etapa, com os 999 trabalhos obtidos. Foram preenchidas as informações: ano e veículo de publicação;
- b) na segunda etapa, com os 604 trabalhos. Foram preenchidas as informações: título, autores, tema e palavras-chave;
- c) uma vez definido o conjunto dos trabalhos selecionados, a partir de uma leitura completa, foram preenchidas as informações restantes: observações, método(s) de avaliação e aspecto(s) avaliado(s) para os 20 trabalhos de avaliação.

Essa atividade durou aproximadamente quatro meses para ser concluída.

### 3.1.3 Documentação

A documentação da revisão consistiu de duas subetapas (numeradas de 7 e 8, segundo a Figura 8), descritas a seguir:

**7 – Apresentação dos estudos:** apresenta uma visão geral dos resultados obtidos pela revisão. Esses resultados mostram informações gerais, não diretamente associadas à avaliação, mas aos principais temas encontrados em trabalhos sobre SS, que servem para contextualizar a época da realização da revisão.

**8 – Resultado das questões:** apresenta o resultado das questões de pesquisa (questão principal e questões secundárias).

## 3.2 APRESENTAÇÃO DOS ESTUDOS



cada trabalho. Como resultados da análise, os principais temas abordados nesta revisão, foram:

- a) **estudo sobre conteúdos postados**, sua relevância, popularidade, credibilidade, interação, compartilhamento, recomendação e tipos de conteúdos. Nessa categoria, também estão os trabalhos de PLN (REITTER; LEBIERE, 2010; NADAMOTO et al., 2010; KINSELLA; MURDOCK; O'HARE, 2011), mineração de dados (CORREA; SUREKA, 2011; SILVA; AIZAWA, 2010; CRAMPES; PLANTIÉ, 2011; BHAT; ABULAISH, 2012) e análise de sentimentos (JIANG et al., 2011; TSOLMON; KWON; LEE, 2012; RABELO; PRUDÊNCIO; BARROS, 2012; BIFET; FRANK, 2010; QIU et al., 2012). Esta categoria obteve a maior porcentagem dos trabalhos analisados: 35% (209 trabalhos);
- b) **estudo sobre usuários**, envolvendo seus sentimentos, comportamentos, colaboração, relacionamentos, quantificação e categorização. Nessa categoria, estão trabalhos como o de Lipczac, Sigurbjornsson e Jaimes (2012), que realizaram análises explícitas (baseada em contatos, como seus relacionamentos) e implícitas (baseadas em ações, como comentários e seleção de fotos) dos usuários em SS a fim de encontrar semelhanças entre eles. Nessa linha, Lim e Datta (2012) fizeram um estudo de comunidades que partilham interesses comuns no Twitter. Chung, Koo e Park (2012) investigaram por que as pessoas compartilham informações em SS. Sibona e Walczak (2011) investigaram o que determina um usuário ser amigo ou deixar de ser amigo no Facebook. Prieto e Leahy (2012) fizeram um estudo de como os idosos usam SS e os fatores que influenciam seu uso. Esta categoria obteve a segunda maior porcentagem dos trabalhos analisados: 25% (148 trabalhos);
- c) **propostas de SS ou de funcionalidades para melhorar SS**: nessa categoria, encontram-se trabalhos que se empenham em melhorar ou propor novas funcionalidades para os SS existentes, ou até mesmo, propor novos SS. Carvalho e Furtado (2012) propuseram uma ferramenta denominada *Uxmarks*<sup>19</sup>, que utiliza o conteúdo (vídeos, blogs, notícias etc.) compartilhado em redes sociais para “preencher” um repositório de dados relacionados a UX. Seus usuários podem editar cada assunto colaborando com o conhecimento apresentado. Jehl, Hieber e Riezler (2012) propõem uma melhoria da tradução automática do Twitter. O

---

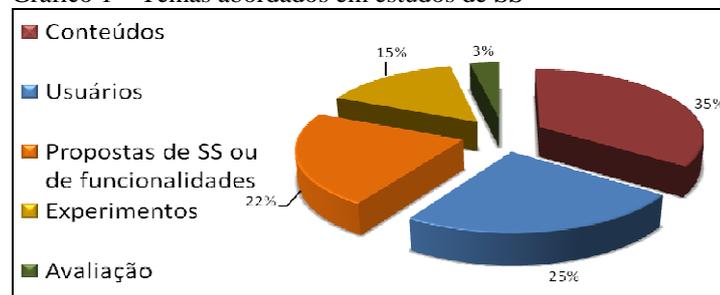
<sup>19</sup> Disponível em: <<http://www.uxmarks.com/>>. Acesso em: 12 de janeiro de 2015.

trabalho de Stein, Chen e Mangla (2011) apresenta o *Facebook Immune System*, um sistema com o objetivo de tornar o Facebook o lugar mais seguro na Internet para as pessoas e suas informações. Segurança e privacidade foram os principais temas para propostas de funcionalidades em SS (FAHL et al., 2012; BOSMA; MEIJ; WEERKAMP, 2012; KISILEVICH; MANSMANN, 2010; MASCETTI et al., 2011; MAZZIA; LEFEVRE; ADAR, 2012). Esta categoria obteve a terceira maior porcentagem dos trabalhos analisados: 22% (135 trabalhos);

- d) **experimentos**, estudos ou análise de redes sociais referentes a armazenamento de imagens, simulação de multidão (LI et al., 2012), propagação de influência (GOYAL; BONCHI; LAKSHMANAN, 2010; STIBE; KUKKONEN, 2012; SUN; NG, 2012; CULOTTA, 2010), propagação de vídeo (LI et al., 2012), estimação do tamanho do SS (KATZIR; LIBERTY; SOMEKH, 2011) etc. Esta categoria obteve 15% dos trabalhos analisados (92 trabalhos); e
- e) **avaliação** de aspectos como: usabilidade, UX, privacidade, segurança, comunicação, credibilidade, compartilhamento, acessibilidade etc. Esta categoria obteve apenas 20 trabalhos e 3% dos trabalhos analisados. Esta categoria, foco desta proposta de tese, será detalhada nas QS1 e QS2.

Alguns desses trabalhos pertenceram a mais de uma categoria, por exemplo, Agosto e Abbas (2010) avaliam a privacidade, segurança e comunicação e fazem um estudo das preferências dos usuários ou ainda o trabalho de Wang, Xu e Grossklags (2011) que avaliam a privacidade e apresentam uma proposta de um SS. O resultado (Gráfico 1) considera esse aspecto.

Gráfico 1 – Temas abordados em estudos de SS



A partir das palavras-chaves<sup>20</sup> dos 604 trabalhos, excluindo aquelas utilizadas na busca, por serem relacionadas a SS e serem comuns (Facebook, Twitter, Tweet, *Tweets*, “Rede Social”, “Redes Sociais”, “*Social systems*”, “*Social system*”, “*Social Network*”, “*Social*

<sup>20</sup> *Keywords*, disponibilizados pelos próprios autores do artigo.



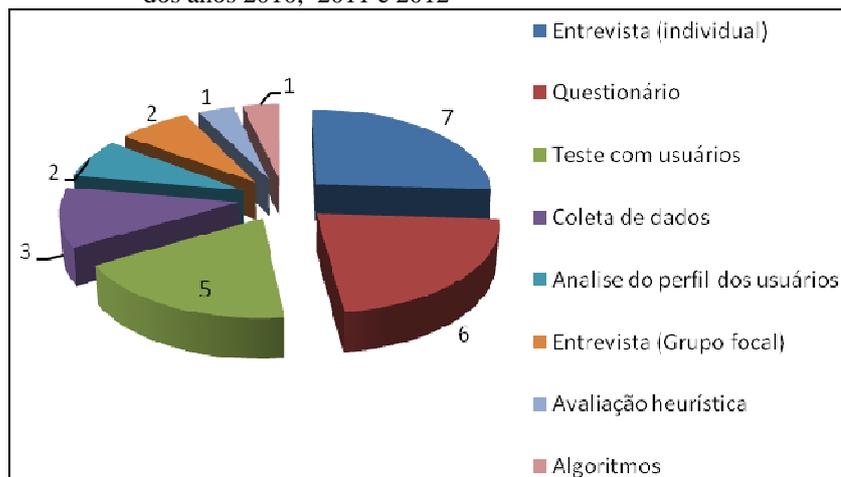
Alguns dos trabalhos avaliaram mais de um aspecto e utilizaram mais de um método de avaliação. Tais trabalhos serão descritos na próxima seção.

### 3.3.3 QS2: Quais são os principais métodos de avaliação em trabalhos sobre SS?

O objetivo desta QS foi fazer um levantamento dos métodos utilizados para avaliação. Foram contabilizados 7 métodos: entrevista (individual e grupo focal), questionário, teste com usuários, coleta de dados, análise do perfil dos usuários, avaliação heurística e algoritmos.

O método mais utilizado foi entrevista (Gráfico 3), responsável por 45% dos métodos utilizados, seguida de questionário, teste com usuários etc. O Quadro 8 apresenta os 20 trabalhos obtidos, seus aspectos avaliados e métodos utilizados.

Gráfico 3 – Métodos utilizados em trabalhos completos de avaliação em SS dos anos 2010, 2011 e 2012



Quadro 8 – Trabalhos de avaliação em SS dos anos 2010, 2011 e 2012

	Trabalho	Aspecto(s) avaliado(s)	Método(s) utilizado(s)
1	Bergmann e Silveira, 2012	<ul style="list-style-type: none"> <li>Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>Questionário</li> </ul>
2	Agosto e Abbas, 2010	<ul style="list-style-type: none"> <li>Privacidade</li> <li>Segurança</li> <li>Comunicação</li> </ul>	<ul style="list-style-type: none"> <li>Entrevistas</li> <li>Grupo focal</li> </ul>
3	Kelley et al., 2011	<ul style="list-style-type: none"> <li>Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>Entrevista</li> </ul>
4	Perotti e Hair, 2011	<ul style="list-style-type: none"> <li>UX</li> </ul>	<ul style="list-style-type: none"> <li>Entrevista</li> </ul>
5	Shi et al., 2011	<ul style="list-style-type: none"> <li>Satisfação do usuário</li> </ul>	<ul style="list-style-type: none"> <li>Questionário</li> </ul>
6	Staddon et al., 2012	<ul style="list-style-type: none"> <li>Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>Questionário</li> </ul>
7	Lampinen et al., 2011	<ul style="list-style-type: none"> <li>Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>Entrevistas</li> <li>Grupo focal</li> </ul>
8	Castillo, Mendoza e Poblete, 2011	<ul style="list-style-type: none"> <li>Credibilidade</li> </ul>	<ul style="list-style-type: none"> <li>Coleta de dados</li> <li>Algoritmos</li> </ul>
9	Mao, Shuai e Kapadia, 2011	<ul style="list-style-type: none"> <li>Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>Coleta de dados</li> </ul>

	<b>Trabalho</b>	<b>Aspecto(s) avaliado(s)</b>	<b>Método(s) utilizado(s)</b>
10	Egelman, Oates e Krishnamurthi, 2011	<ul style="list-style-type: none"> <li>• Privacidade</li> <li>• Usabilidade</li> </ul>	<ul style="list-style-type: none"> <li>• Entrevista</li> <li>• Teste com usuários</li> </ul>
11	Reynolds et al., 2011	<ul style="list-style-type: none"> <li>• Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>• Questionário</li> <li>• Análise do perfil dos usuários</li> </ul>
12	Newman et al., 2011	<ul style="list-style-type: none"> <li>• Compartilhamento</li> <li>• Comunicação</li> </ul>	<ul style="list-style-type: none"> <li>• Teste com usuários</li> </ul>
13	Wang, Xu e Grossklags, 2011	<ul style="list-style-type: none"> <li>• Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>• Teste com usuários</li> <li>• Entrevista</li> </ul>
14	Liu et al., 2011	<ul style="list-style-type: none"> <li>• Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>• Questionário</li> </ul>
15	Silva et al., 2012	<ul style="list-style-type: none"> <li>• Privacidade</li> <li>• Segurança</li> </ul>	<ul style="list-style-type: none"> <li>• Análise do perfil dos usuários</li> </ul>
16	Buzzi, Buzzi e Leporini, 2011	<ul style="list-style-type: none"> <li>• Usabilidade</li> <li>• Acessibilidade</li> </ul>	<ul style="list-style-type: none"> <li>• Avaliação heurística</li> </ul>
17	Staddon et al., 2012	<ul style="list-style-type: none"> <li>• Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>• Questionário</li> </ul>
18	Fahl et al., 2012	<ul style="list-style-type: none"> <li>• Privacidade</li> <li>• Usabilidade</li> </ul>	<ul style="list-style-type: none"> <li>• Entrevista</li> <li>• Teste com usuários</li> </ul>
19	Mazzia, LeFevre e Adar, 2012	<ul style="list-style-type: none"> <li>• Privacidade</li> </ul>	<ul style="list-style-type: none"> <li>• Coleta de dados</li> </ul>
20	Lányi, Nagy e Sik, 2012	<ul style="list-style-type: none"> <li>• Acessibilidade</li> </ul>	<ul style="list-style-type: none"> <li>• Teste com usuários</li> </ul>

Desses trabalhos, aqueles relacionados com o contexto de pesquisa abordado nesta tese são os que avaliam aspectos de usabilidade ou UX (trabalhos de números 4, 5, 10, 16 e 18, de acordo com o Quadro 8) e aqueles que utilizam como método de avaliação os dados dos usuários coletados de SS: métodos coleta de dados e análise do perfil dos usuários (trabalhos de números 8, 9, 11 e 15, de acordo com o Quadro 8). Tais trabalhos são descritos a seguir.

Perotti e Hair (2011), por meio do método de entrevistas, procuraram entender a UX em SS. Foram 86 usuários entrevistados que descreveram quais recursos são importantes para eles e o contexto para a sua utilização. O objetivo dos autores era entender como os usuários estão participando nesses sites e por quê. Os autores apresentaram como resultados os SS mais utilizados pelos usuários, com quem eles mais interagem (amigos, família, colegas de trabalho etc.), os recursos do sistema mais valiosos para eles e suas atividades mais praticadas.

O método de entrevista também foi utilizado por Egelman, Oates e Krishnamurthi (2011). Eles realizaram entrevistas com 33 usuários do Facebook para examinar como eles lidaram com as limitações de configurações de privacidade na interface do sistema. Os autores complementaram sua avaliação com testes com usuários e perceberam nos testes mais esclarecimento de quando e onde os usuários tinham dúvida sobre privacidade e usabilidade na interface.

Fahl et al. (2012) criaram um mecanismo de encriptação de dados para o Facebook e, a fim de avaliar a privacidade do Facebook e a usabilidade do mecanismo, realizaram testes com usuários, seguidos de entrevistas. O foco da avaliação, entretanto, não era o Facebook, e sim o mecanismo de encriptação proposto pelos autores. Foram 514 entrevistados e 96 usuários participantes dos testes. Os autores concluíram que, embora os usuários tivessem alguma dificuldade no processo de encriptação proposto por eles, sentiram-se mais protegidos com uma ferramenta de proteção de seus dados.

Com o objetivo de avaliar a usabilidade e a acessibilidade do Twitter, Buzzi, Buzzi e Leporini (2011) utilizaram o método avaliação heurística. A avaliação foi realizada pelos três autores (sendo um deles cego), que são especialistas em usabilidade e acessibilidade. Eles utilizaram um leitor de tela e apresentaram o resultado da avaliação fazendo uma comparação com um trabalho anterior de avaliação do Facebook. Suas conclusões foram que o Twitter, embora tenha apresentado melhores resultados que o Facebook, ainda apresenta problemas de usabilidade, além de não satisfazer aos princípios básicos de acessibilidade e design universal.

O trabalho de Shi et al. (2011) teve como objetivo avaliar a satisfação do usuário. Tal aspecto pertence à usabilidade (ISO 9241-11,1998; PREECE; ROGERS; SHARP, 2005) e à UX (ISO 9241-11,1998; BEVAN, 2008; PREECE; ROGERS; SHARP, 2005). Os autores tinham como objetivo investigar a motivação e os fatores que influenciam a intenção dos usuários a continuar a usar o Facebook. Eles aplicaram um questionário on-line e obtiveram 125 respondentes. Os resultados mostraram que a satisfação foi importante na determinação da intenção dos usuários em continuar a usar o Facebook.

Os trabalhos descritos acima avaliaram a usabilidade, UX ou metas de ambos (satisfação). Eles utilizaram os métodos: entrevistas, questionários, testes com usuários e avaliação heurística. Em nenhum deles, foram utilizados os dados dos usuários, coletados em SS. São descritos, a seguir, trabalhos cujo método utilizado foi a coleta de dados dos usuários a fim de avaliar o sistema.

Os trabalhos de Mao, Shuai e Kapadia (2011) e Castillo, Mendoza e Poblete (2011) utilizaram a coleta de dados como método para avaliação. Tal técnica se refere à extração de dados e tem sido muito utilizada em pesquisas em SS, principalmente na investigação e análise de dados e não em avaliações de sistemas. Nesse estudo, entretanto, essa técnica foi utilizada em dois trabalhos investigados, com o objetivo de avaliar a privacidade (MAO; SHUAI; KAPADIA, 2011) e a credibilidade das informações (CASTILLO; MENDOZA; POBLETE, 2011). Mao, Shuai e Kapadia (2011) fizeram uma avaliação a fim de destacar as

ameaças à privacidade enfrentadas pelos usuários. Segundo os autores, tais ameaças são os próprios conteúdos postados pelos usuários, como os seguintes temas sensíveis a privacidade: férias/viagens, postagens sob a influência de álcool e postagens de doenças. Eles utilizaram uma ferramenta de extração de postagem e coletaram 575.689 postagens sobre férias, 21.297 postagens sobre bebidas e 149.636 postagens sobre doença, concluindo o quão os usuários estão expondo sua privacidade. Nesse trabalho, apesar da análise nas postagens dos usuários, não houve uma interpretação a fim de obter alguma conclusão sobre a avaliação do sistema e sim sobre os conteúdos sensíveis à privacidade postados por usuários no Twitter.

Castillo, Mendoza e Poblete (2011) fizeram um estudo sobre a credibilidade das informações postadas no Twitter e observaram que ele também é muito usado para espalhar desinformação e boatos falsos, muitas vezes de forma não intencional. Então, os autores, com o objetivo de avaliar a credibilidade dessas informações, fizeram uma extração de dados do Twitter e, em seguida, com o apoio de uma ferramenta, fizeram uma avaliação automática da credibilidade das postagens, verificando se uma notícia postada era verdadeira ou não. Nesse caso, os autores utilizaram as postagens dos usuários para avaliar seu conteúdo e não aspectos do sistema.

A análise de perfis públicos de usuários, assim como coleta de dados, não é um método de avaliação comumente utilizado em IHC. No entanto, com o surgimento de SS, essa estratégia tem sido utilizada a fim de obter e validar os dados dispostos por usuários. Tal estratégia foi utilizada por Silva et al. (2012) e Reynolds et al. (2011). Os dois trabalhos avaliaram a privacidade e segurança em SS. No primeiro caso, os autores propõem um indicador capaz de estimar o nível de vulnerabilidade de um usuário no Facebook. O indicador proposto é gerado com base na observação das características de conexões de amigos do usuário e na quantidade de informação publicada, permitindo que o usuário seja alertado sobre o seu nível de exposição e de seus amigos. Os autores utilizaram uma aplicação responsável por navegar nas páginas do Facebook e realizar o armazenamento de perfis públicos, escolhidos de forma aleatória. Foram coletados 11 atributos (nome, gênero, educação, interesses, empresa, sobre mim – descrição sumária do perfil do usuário, e-mail, telefone, endereço, sítio web e configurações de privacidade) de 75.011 usuários do Facebook. Para cada um dos atributos coletados, era armazenado um valor lógico referente ao seu preenchimento, sendo “1” para preenchido, e “0” para não preenchido. Para configuração da privacidade, era armazenado “0” para o perfil que não alterou nenhuma configuração dos atributos coletados e “1” para o perfil que ocultou a exibição de algum atributo. Com isso, os

autores puderam analisar os níveis de vulnerabilidade dos usuários pelas informações coletadas e configuração de privacidade alterada, concluindo que 94,37% dos usuários possuíam baixo nível de vulnerabilidade. Nesse trabalho, as postagens dos usuários não foram coletadas a fim de fornecer alguma percepção sobre o sistema.

No segundo caso, Reynolds et al. (2011) objetivaram testar se as percepções de privacidade para o compartilhamento de informações dos usuários correspondia ao seu comportamento real. Os autores contrataram 103 usuários do Facebook para participarem da investigação. Na investigação, os usuários utilizavam a ferramenta proposta para coletar seus dados (mensagens do mural, atualizações de *status*, os amigos, as configurações de privacidade para mensagens e informações básicas). A coleta gerava um questionário com 40 questões, que eles respondiam, de forma a esclarecer o significado de cada elemento coletado. Postagens também eram coletadas, por exemplo, “a noite passada foi muito divertida, mas eu estou com uma enorme ressaca” ou “Eu estou atrasado para uma aula / reunião”. Os usuários eram questionados sobre quais as informações que eles quiseram compartilhar e às quais eles desejavam esconder. A grande maioria dos usuários, em suas respostas ao questionário, afirmara ter intenção de explicitamente esconder o conteúdo que eles criaram, mas poucos sabiam realmente usar os meios disponíveis para fazer. Os resultados revelaram que, apesar dos usuários demonstrarem preocupações com a privacidade, estas não eram refletidas em suas postagens e controle de privacidade. Este trabalho utilizou as postagens dos usuários a fim de checar a configuração de privacidade com os usuários. No entanto, não foi realizada uma análise deles a fim de apresentar alguma percepção sobre o sistema avaliado.

Observa-se que os dados coletados dos usuários são importantes em SS a fim de obter informações sobre os usuários ou o estudo de seu conteúdo postado. No entanto, não foram encontradas, nesta revisão sistemática, um modelo ou método que avaliem a UUX do sistema a partir dos textos dos usuários.

### 3.4 TRABALHOS RELACIONADOS

Esta seção discute alguns trabalhos correlatos à tese.

Alguns trabalhos que têm focado em narrativas de usuários a fim de estudar UUX ou UX são: Hassenzahl, Diefenbach e Göritz (2010), Korhonen, Arrasvuori e Väänänen-Vainio-Mattila (2010), Olsson e Salo (2012), Tuch, Trusell e Hornbæk (2013) e Hedegaard e Simonsen (2013).

Hassenzahl, Diefenbach e Göritz (2010), com o objetivo de estudar UX a partir de experiências positivas dos usuários, coletaram 500 textos escritos por usuários de produtos interativos (telefones celulares, computadores etc.) e apresentaram medidas de classificação de uma meta UX em positiva. Korhonen, Arrasvuori e Väänänen-Vainio-Mattila (2010) coletaram 116 relatos de experiência de usuários experientes sobre seus produtos pessoais (como smartphones e MP3 players). O objetivo era a avaliação da UX desses produtos e os usuários teriam de relatar seus sentimentos pessoais, valores e interesses quando os utilizam. Olsson e Salo (2012) coletaram 90 relatos escritos de usuários iniciantes de aplicações móveis de realidade aumentada. O foco também era avaliar a UX desses produtos e a análise consistia em determinar o assunto de cada texto e classificar, concentrando a atenção nas experiências mais satisfatórias e mais insatisfatórias. Seguindo essa linha, o trabalho de Tuch, Trusell e Hornbæk (2013) estudou 691 narrativas geradas por usuários com experiências positivas e negativas em tecnologias a fim de estudar a UX com elas.

Nos quatro trabalhos citados, as informações foram extraídas manualmente, a partir de textos gerados pelos usuários. Os usuários foram especificamente solicitados a escrever os textos ou responder um questionário (HASSENZAHN; DIEFENBACH; GÖRITZ, 2010; TUCH, TRUSELL; HORNBÆK, 2013), ao contrário da coleta espontânea do que eles postam no próprio sistema.

Hedegaard e Simonsen (2013) propuseram uma ferramenta para avaliar a UUX de produtos a partir de sites de comentários (*reviews*, revisões) sobre produtos. Os autores coletaram 3.492 revisões de produtos no site epinions.com, de forma automática (usando uma ferramenta de EI). Hedegaard e Simonsen (2013), também como os outros autores, não avaliaram o sistema em uso, e sim outros produtos utilizados pelos usuários. Nesse caso, os textos foram escritos por revisores de produtos, e, conforme já mencionado, acredita-se que a postura do usuário em um site de avaliação de produtos é diferente de quando um usuário está usando um sistema, encontra um problema e resolve relatar esse problema para desabafar ou mesmo para sugerir uma solução.

Em nenhum dos trabalhos encontrados é realizada a avaliação do sistema. Os dados são extraídos a fim de realizar um estudo sobre características da UUX. Em nenhum deles, houve a sugestão de um modelo para avaliar sistemas a partir dos textos dos usuários. O Quadro 9 apresenta um resumo dos trabalhos correlatos à tese.

Quadro 9– Trabalhos correlatos ao tema da tese

	<b>Hassenzahl; Diefenbach; Göriz (2010)</b>	<b>Korhonen; Arrasvuori; Väänänen- Vainio-Mattila (2010)</b>	<b>Olsson; Salo (2012)</b>	<b>Tuch; Trusell; Hornbæk (2013)</b>	<b>Hedegaard; Simonsen (2013)</b>
<b>Forma de extração / Técnica</b>	Manual / Questionário	Manual / Storytelling	Manual / Relatórios	Manual / Questionário	Automática / EI
<b>Local de extração</b>	Questionário online (Survey-Monkey)	Relatórios escritos	Relatórios escritos	Questionário online	Site (epinions.com)
<b>Quantidade de textos analisados</b>	500	116	90	691	3.492
<b>Produto ou sistema avaliado</b>	Produtos interativos (celulares, computadores etc.)	Produtos pessoais (smartphones, MP3 players etc.)	Aplicações móveis de realidade aumentada	Produtos interativos	Softwares e vídeo games
<b>Usuários</b>	Usuários dos produtos	Usuários dos produtos	Usuários dos produtos	Usuários dos produtos	Revisores de produtos
<b>Metas avaliadas</b>	UX (Experiências positivas)	UX	UX (Experiências mais satisfatórias e mais insatisfatórias)	UX (Experiências positivas e negativas)	UUX
<b>Existe um modelo de avaliação?</b>	Não	Não	Não	Não	Não

### 3.5 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou uma revisão sistemática a fim de fornecer um estado da arte sobre SS, bem como fazer um estudo de trabalhos com foco em avaliação em SS. Observou-se que os dados dos usuários têm sido coletados a fim de realizar diversas investigações sobre SS. O que torna ainda mais motivadora a proposta de avaliação para esses tipos de sistemas. No entanto, não foram encontrados métodos que avaliem o sistema a partir dos textos dos usuários.

Dessa forma, foram realizados estudos em trabalhos relacionados ao tema da tese. Nesses trabalhos, foi observado que, embora sejam realizadas avaliações de produtos a partir dos textos dos usuários, não existe um modelo que facilite esta forma de avaliação. Além disso, os textos investigados foram escritos por usuários mediante a solicitação dos pesquisadores, não considerando a espontaneidade do usuário no momento de uso do próprio sistema.

## 4 INVESTIGAÇÕES A PARTIR DAS POSTAGENS DOS USUÁRIOS EM SISTEMAS SOCIAIS

Neste capítulo, são apresentadas três investigações realizadas a partir das postagens dos usuários em quatro SS. O objetivo de cada investigação foi responder uma questão da pesquisa. São apresentados, para cada investigação, os SS analisados, os participantes, a metodologia e a quantidade de postagem analisada. Os SS investigados foram: Twitter<sup>21</sup>, Facebook<sup>22</sup>, Torcida Virtual<sup>23</sup> e SIGAA<sup>24</sup>, descritos a seguir.

### 4.1 SISTEMAS SOCIAIS INVESTIGADOS

O Twitter é um dos maiores serviços de *microblogging* na internet, com mais de 600 milhões de usuários ativos<sup>25</sup>. Microblogs são mensagens curtas de texto que os usuários usam para compartilhar todos os tipos de informações com o mundo. No Twitter, esses microblogs são chamadas de “*Tweets*”, e podem conter notícias, anúncios, assuntos pessoais, piadas, opiniões etc. Os contatos no Twitter são classificados como “*Followers*” e “*Following*” (“Seguidores” e “Seguidos”, tradução brasileira). “Seguidores” são outros usuários e visualizam suas postagens. “Seguidos” refletem a quantidade de usuários (contas) do Twitter que o usuário optou por seguir (visualizar suas postagens)<sup>26</sup>.

O Facebook é, atualmente, a maior rede social online, com 1,3 bilhão de usuários ativos<sup>27</sup>. O principal objetivo do Facebook é possibilitar o compartilhamento de conteúdos dos seguintes tipos: textos, links, imagens ou vídeos. O conteúdo é compartilhado na página principal do Facebook, chamada de mural ou “Feed de notícias”. Os contatos de um usuário do Facebook são denominados de “amigos”. As postagens de um usuário são visíveis para qualquer pessoa com permissão para ver o perfil completo.

O Torcida Virtual (TV) trata-se de um SS que congrega torcedores dos times de futebol do Brasil, permitindo aos seus usuários mapearem, interagirem e consultarem informações relevantes sobre seu time. Neste SS, existe uma página principal chamada de

<sup>21</sup> Disponível em: <<https://twitter.com/>>. Acesso em: 12 de janeiro de 2015

<sup>22</sup> Disponível em: <<https://www.facebook.com/>>. Acesso em: 12 de janeiro de 2015.

<sup>23</sup> Disponível em: <<http://torcidas.esporte.ig.com.br/>>. Acesso em: 12 de janeiro de 2015.

<sup>24</sup> Disponível em: <[www.si3.ufc.br/](http://www.si3.ufc.br/)>. Acesso em: 12 de janeiro de 2015.

<sup>25</sup> Disponível em: <<http://www.statisticbrain.com/twitter-statistics/>>. Acesso em: 12 de janeiro de 2015

<sup>26</sup> Disponível em: <<https://support.twitter.com/groups/50-welcome-to-twitter/topics/204-the-basics/articles/364620-o-abc-do-twitter>>. Acesso em: 12 de janeiro de 2015

<sup>27</sup> Disponível em: <<http://www.statisticbrain.com/facebook-statistics/>>. Acesso em: 12 de janeiro de 2015

“zona de zoação” na qual seus usuários postam mensagens (sobre qualquer assunto) e visualizam as mensagens postadas por seus contatos.

O SIGAA (Sistema Integrado de Gestão de Atividades Acadêmicas) é o sistema de controle acadêmico das Universidades Federais do Brasil, e, por meio dele, os estudantes podem ter acesso a várias funcionalidades, como: comprovante de Matrícula, histórico acadêmico, realização de matrícula etc. O sistema permite a troca de mensagens a partir de um fórum de discussões. Seus usuários são alunos e funcionários da universidade.

Tais investigações tiveram como objetivo responder às questões da pesquisa levantadas nesta tese. O Quadro 10 apresenta um resumo das investigações descritas neste capítulo.

Quadro 10 – Investigações realizadas

Fatores	Investigações		
	Os usuários “falam” sobre o sistema durante o uso?	Como os usuários “falam”?	Como avaliar?
<b>Sistemas analisados</b>	<ul style="list-style-type: none"> <li>• Twitter</li> </ul>	<ul style="list-style-type: none"> <li>• Twitter</li> <li>• SIGAA</li> <li>• Facebook</li> <li>• Torcida Virtual</li> </ul>	<ul style="list-style-type: none"> <li>• Twitter</li> <li>• SIGAA</li> </ul>
<b>Participantes</b>	<ul style="list-style-type: none"> <li>• 1 especialista e professora de IHC</li> </ul>	<ul style="list-style-type: none"> <li>• 7 bolsistas de graduação</li> <li>• 2 professores de IHC</li> <li>• 4 especialistas de IHC</li> </ul>	<ul style="list-style-type: none"> <li>• 17 alunos de graduação</li> <li>• 12 especialistas de IHC</li> </ul>
<b>Metodologia</b>	<ol style="list-style-type: none"> <li>1. Definição de padrões de extração;</li> <li>2. Extração das postagens dos usuários usando os padrões definidos; e</li> <li>3. Análise das postagens obtidas.</li> </ol>	<ul style="list-style-type: none"> <li>• Extração e Análise das postagens</li> </ul>	<ol style="list-style-type: none"> <li>1. Classificação das PRUs em metas de UUX; e</li> <li>2. Discussão sobre as classificações.</li> </ol>
<b>Quantidade de postagens analisadas</b>	<ul style="list-style-type: none"> <li>• Twitter: 295.797</li> </ul>	<ul style="list-style-type: none"> <li>• Twitter: 295.797</li> <li>• SIGAA: 24.743</li> <li>• Facebook: 100</li> <li>• TV: 100</li> </ul>	<ul style="list-style-type: none"> <li>• Twitter: 500</li> <li>• SIGAA: 650</li> </ul>

## 4.2 OS USUÁRIOS “FALAM” SOBRE O SISTEMA DURANTE O USO?

A investigação descrita nesta seção procurou responder a segunda questão da pesquisa: verificar se os usuários falam sobre o SS durante seu uso.

### 4.2.1 Sistema Social investigado

O foco desta primeira análise foi o SS Twitter, que foi escolhido tanto por seu caráter interativo e social quanto por sua interação se constituir principalmente de textos escritos.

#### 4.2.2 Metodologia

A fim de investigar se os usuários falam sobre o sistema durante o uso, decidiu-se procurar por PRUs no Twitter. A técnica utilizada para obter tais postagens foi a EI, mencionada no capítulo 2, que tem como finalidade extrair informações específicas por meio de um conjunto de padrões ou regras de extração. Dessa forma, as postagens dos usuários seriam extraídas do Twitter e depois analisadas a fim de verificar se eram PRUs ou não.

Nesse sentido, a metodologia dessa investigação foi estruturada em três etapas: 1) definição de padrões de extração; 2) extração das postagens dos usuários usando os padrões definidos; e 3) análise das postagens obtidas.

##### 4.2.2.1 Definição de padrões de extração

A fim de definir padrões de extração eficazes<sup>28</sup>, foram formuladas as seguintes suposições:

- a) os usuários, ao falarem do SS, utilizam o nome do sistema ou o nome de suas funcionalidades;
- b) os usuários, ao falarem do SS, utilizam substantivos relacionados à usabilidade do sistema, como: “usabilidade”, “interface”, “erro”, “problema” e “bug”;
- c) os usuários, ao falarem do SS, utilizam interrogações: 1) para fazer perguntas sobre o uso do sistema e; 2) para expressar seus sentimentos. Segundo Preece, Rogers e Sharp (2002), quando os usuários estão aborrecidos ou irritados por alguma novidade, geralmente utilizam símbolos do teclado para enfatizar sua fúria ou frustração, como pontos de exclamação (!), letras maiúsculas e pontos de interrogação (?);
- d) os usuários ao falarem do SS utilizam adjetivos para qualificá-los, como: “fácil”, “difícil”, “bom”, “ruim”. Segundo Taboada et al. (2011), grande parte de pesquisas relacionadas à análise de sentimentos tem-se centrado sobre o uso de adjetivos; e

---

<sup>28</sup> Um padrão de extração eficaz será aquele que extrair PRUs.

e) as PRUs possuem advérbios que caracterizam o comportamento dos usuários acerca do sistema, como: “bem”, “mal”, “muito”, “pouco”, “mais”, “menos”. Segundo Benamara et al. (2007), advérbios são muito utilizados em pesquisas relacionadas a análise de sentimentos.

A partir dessas suposições, os padrões de extração foram definidos e organizados em cinco tipos de buscas a serem efetuadas no Twitter. Para a suposição 1, foram definidos como padrões de extração o nome do sistema (“Twitter”) e de suas principais funcionalidades (“*Tweets*”, “*Followers*” e “*Followings*”). Por utilizar somente uma palavra como padrão, essas buscas foram chamadas de individuais (Quadro 11). As demais buscas são combinações desses padrões com os demais padrões relacionados às demais suposições, gerando 64 possíveis padrões de extração.

Quadro 11 – Classificação das buscas e padrões de extração

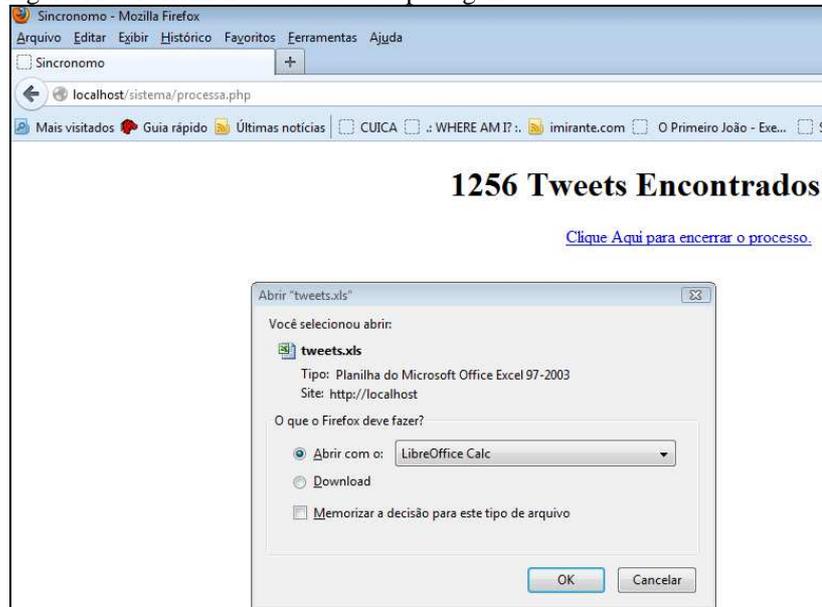
<b>Buscas</b>	<b>Padrões de extração</b>
<b>Individuais</b>	[Twitter], [ <i>Tweets</i> ], [Followers], [ <i>Followings</i> ]
<b>Associadas a substantivos</b>	[Twitter usabilidade], [Twitter interface], [Twitter erro], [Twitter problema], [Twitter <i>bug</i> ], [ <i>Tweets</i> usabilidade], [ <i>Tweets</i> interface], [ <i>Tweets</i> erro], [ <i>Tweets</i> problema], [ <i>Tweets</i> <i>bug</i> ], [Followers usabilidade], [Followers interface], [Followers erro], [Followers problema], [Followers <i>bug</i> ], [ <i>Followings</i> usabilidade], [ <i>Followings</i> interface], [ <i>Followings</i> erro], [ <i>Followings</i> problema], [ <i>Followings</i> <i>bug</i> ]
<b>Associadas a interrogação</b>	[Twitter “?”], [ <i>Tweets</i> “?”], [Followers “?”], [ <i>Followings</i> “?”]
<b>Associada a Adjetivos</b>	[Twitter fácil], [Twitter difícil], [Twitter bom], [Twitter ruim], [ <i>Tweets</i> fácil], [ <i>Tweets</i> difícil], [ <i>Tweets</i> bom], [ <i>Tweets</i> ruim], [Followers fácil], [Followers difícil], [Followers bom], [Followers ruim], [ <i>Followings</i> fácil], [ <i>Followings</i> difícil], [ <i>Followings</i> bom], [ <i>Followings</i> ruim]
<b>Associadas a Advérbios</b>	[Twitter bem], [Twitter mal], [Twitter muito], [Twitter pouco], [Twitter mais], [Twitter menos], [ <i>Tweets</i> bem], [ <i>Tweets</i> mal], [ <i>Tweets</i> muito], [ <i>Tweets</i> pouco], [ <i>Tweets</i> mais], [ <i>Tweets</i> menos], [Followers bem], [Followers mal], [Followers muito], [Followers pouco], [Followers mais], [Followers menos], [ <i>Followings</i> bem], [ <i>Followings</i> mal], [ <i>Followings</i> muito], [ <i>Followings</i> pouco], [ <i>Followings</i> mais], [ <i>Followings</i> menos]

#### 4.2.2.1.1 Ferramenta de extração das postagens

O Twitter, do momento desta pesquisa até o término da escrita deste trabalho, não possui uma forma de apresentar uma estimativa das buscas realizadas em seu sistema. Por esse motivo, foi desenvolvida uma ferramenta a fim de estimar cada busca realizada. A ferramenta usa as linguagens PHP, Javascript e AJAX (*Asynchronous JavaScript* e XML) para auxiliar nas múltiplas chamadas da API de busca do Twitter. Essa ferramenta foi útil para extração das postagens do Twitter, nessa investigação. Os padrões de extração foram utilizados na ferramenta para obtenção das postagens. Optou-se por extrair somente postagens

definidas com o perfil público<sup>29</sup>. A Figura 11 ilustra a ferramenta de busca das postagens no momento em que uma busca é realizada. Ela retorna a quantidade e as postagens obtidas (*Tweets*) armazenadas em um arquivo Excel (xls).

Figura 11 – Ferramenta de busca das postagens



#### 4.2.2.2 Extração das postagens dos usuários usando os padrões definidos

As coletas ocorreram duas vezes por mês, durante três meses. No início e no meio de cada mês, nas seguintes datas de 2012: 02/10, 17/10, 02/11, 17/11, 02/12 e 17/12, seguindo a ordem definida das buscas (Quadro 11). Em cada coleta eram efetuadas todas as cinco buscas. Os resultados de uma busca por padrão gerava um arquivo xls, contendo de 0 até 1500 postagens<sup>30</sup>. Foram obtidos 64 arquivos (1 para cada padrão) por coleta. Durante as 6 coletas, 384 arquivos foram obtidos. Dessa forma, foram extraídas 295.797 postagens de usuários com perfil público do universo mundial do Twitter. Cada arquivo de busca possuía postagens com até dois dias da data pesquisada.

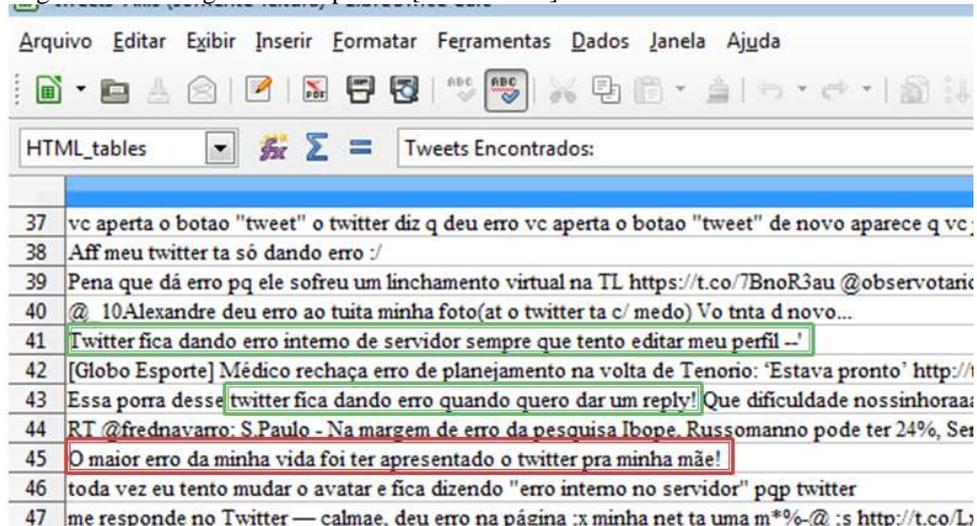
#### 4.2.2.3 Análise das postagens obtidas

<sup>29</sup> O perfil público do Twitter permite o acesso de qualquer pessoa à suas informações e postagens

<sup>30</sup> A API do Twitter, na época desta extração (outubro a dezembro de 2012) permitia o acesso a apenas as 15 primeiras páginas de uma busca. Cada página armazenava 100 postagens, de forma que o máximo de postagens obtidas era de até 1500 postagens em cada busca.

Cada arquivo de busca continha PRUs e não-PRUs, por exemplo, o padrão [Twitter erro] apresentou a seguinte PRU: “O **Twitter** fica dando **erro** quando eu quero dar uma resposta”, mas também apresentou a seguinte não-PRU: “O maior **erro** da minha vida foi ter apresentado o **Twitter** para minha mãe!”. As PRUs podem ser perguntas, reclamações ou mesmo elogios sobre o Twitter. A Figura 12 apresenta alguns exemplos de postagens obtidas, com o padrão [Twitter erro].

Figura 12 – Postagens com o padrão [Twitter erro]



A análise das postagens consistiu em avaliar a eficácia de cada padrão de extração utilizado. A eficácia ( $ef$ ) de um padrão consistiu na porcentagem de TPRUs (Total de PRUs) obtidas em relação ao TP (Total de Postagens) extraídas (Equação 4.1.1).

$$ef = \frac{(TPRUs * 100)}{TP} \quad (4.1.1)$$

Para exemplificar, tem-se a primeira coleta (Tabela 3), para o padrão de extração [Twitter usabilidade] foram obtidas 23 postagens, e, dentre essas postagens, 13 eram PRUs, totalizando 57% de PRUs, que é a eficácia da busca para esse padrão. Para o padrão de extração [Twitter interface], foram obtidas 1.496 postagens, com 1.065 PRUs, totalizando 71% de eficácia para o padrão [Twitter interface]. Esse resultado fornece uma porcentagem de eficácia de PRUs relacionadas a cada padrão, permitindo a avaliação da eficácia dos padrões utilizados. No entanto, tal resultado fornece apenas a eficácia de um dia de coleta (02/10/2012).

Tabela 3 – Exemplo de um arquivo de busca do dia 02/10/2012

Padrão de extração	TP	TPRUs	ef (em %)
[Twitter usabilidade]	23	13	57
[Twitter interface]	1.496	1.065	71

A fim de desconsiderar picos de problemas no sistema e assuntos mais comentados no momento, foi realizada uma média das 6 coletas para cada padrão (Tabela 4).

Tabela 4 – Cálculo da média das 6 coletas [Twitter erro]

	Coletas efetuadas						Média das 6 coletas
	02/10/2012	17/10/2012	02/11/2012	17/11/2012	02/12/2012	17/12/2012	
<b>TP</b>	1087	973	1475	541	1435	1382	1148
<b>TPRUs</b>	921	575	1342	456	1378	1052	954
<b>ef (em %)</b>	85	59	91	84	96	76	81

A análise foi realizada pela autora deste trabalho e consistia em: ler e contar as PRUs obtidas para cada padrão de extração nas seis coletas efetuadas. O cálculo da eficácia e das médias das coletas era feito automaticamente pelo excel. Na contagem das PRUs, também foram consideradas postagens retuitadas (*Retweet*<sup>31</sup>, RT), considerando como similar a opinião (interesse ou sentimento) do usuário sobre um assunto já postado por outro(s) usuário(s).

As seis coletas foram realizadas em datas distintas, conforme mencionado, a fim de evitar picos de assuntos muito comentados no momento. Ainda assim, a eficácia de um padrão pode variar consideravelmente se, em determinada data, o padrão de extração utilizado for mencionado em demasia por algum motivo. Considerando esse fato, foi utilizada a medida desvio padrão para medir a dispersão dos valores individuais da eficácia de um padrão. O desvio padrão é a média quadrática dos desvios em relação à média aritmética de uma distribuição de frequências.

Para o cálculo do desvio padrão (dp), foi utilizada a média da eficácia de um padrão ( $Mef$ ) (Tabela 4) e, a seguir, os desvios para mais e para menos a partir de cada eficácia obtida por busca (Equação 4.1.2).

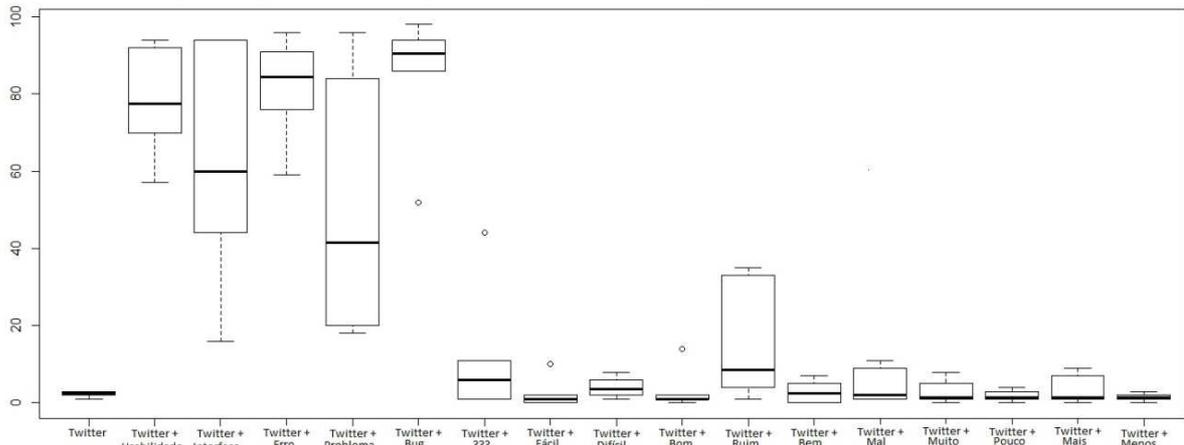
$$dp = \sqrt{\left( \frac{\sum (ef_i - ef)^2}{Mef - 1} \right)} \quad (4.1.2)$$

<sup>31</sup> O *retweet* é uma função do Twitter que consiste em replicar uma determinada mensagem de um usuário para a lista de seguidores, dando crédito a seu autor original.

Uma exemplificação da análise feita pode ser dada da seguinte forma: no padrão [Twitter erro] (Tabela 4), a média da eficácia das postagens obtidas nas 6 coletas foi  $Mef=81\%$ , resultantes das 6 coletas: 85%, 59%, 91%, 84%, 96% e 76%. Calculando o desvio padrão pela porcentagem das 6 coletas obtidas, deu um valor de  $dp = 13$ , significando que 66,6% dos valores de uma distribuição normal encontram-se dentro da faixa do desvio padrão, tanto para mais quanto para menos em relação à média. Apenas as porcentagens 59% e 96%, referentes à 2ª e à 5ª coleta, ficaram fora dessa média.

O Gráfico 4, do tipo boxplot, apresenta o resultado do padrão de extração Twitter com os demais padrões. O boxplot é um gráfico utilizado para avaliar a distribuição empírica do dados. Ele mostra graficamente a posição central dos dados (mediana), a tendência e os valores extremamente altos ou baixos. Por exemplo, o padrão [Twitter Bug] possui uma faixa de valores acima de 80%, mediana de 91% e valores extremos de 98% e 52%.

Gráfico 4 – Boxplot do padrão de extração [Twitter] com demais padrões



#### 4.2.3 Resultados

As Tabelas 5, 6, 7, 8 e 9 apresentam os resultados das buscas individuais, associadas a substantivos, a adjetivos, interrogação e advérbios, respectivamente.

Tabela 5 – Resultado das buscas individuais

Padrão de extração	TP	ef (em %)	dp (em %)
[Twitter]	1437	2	0,7
[Tweets]	1189	6	7,4
[Followings]	1292	2	0,6
[Followers]	1474	0	0

Tabela 6 – Resultado das buscas associadas a substantivos

Padrão de extração		TP	ef (em %)	dp (em %)
[Twitter]	Usabilidade]	22	78	14,1
	Interface]	609	61	30,7
	Erro]	1148	82	13,0
	Problema]	1041	50	34,0
	Bug]	978	85	16,8
[Tweets]	Usabilidade]	5	25	14,1
	Interface]	106	68	29,1
	Erro]	512	65	17,6
	Problema]	984	59	28,4
	Bug]	711	81	26,8
[Followings]	Usabilidade]	4	0	0
	Interface]	22	5	6,4
	Erro]	185	8	17,1
	Problema]	513	2	1,8
	Bug]	333	5	5,6
[Followers]	Usabilidade]	1	0	0
	Interface]	10	2	4,4
	Erro]	182	36	12,1
	Problema]	251	2	1,5
	Bug]	1049	12	2,9

Tabela 7 – Resultado da busca associada a adjetivos

Padrão de extração		TP	ef (em %)	dp (em %)
[Twitter]	Fácil]	1158	2	3,0
	Difícil]	951	4	2,3
	Bom]	1009	3	5,2
	Ruim]	881	15	15,2
[Tweets]	Fácil]	1019	4	2,1
	Difícil]	646	5	2,0
	Bom]	1149	2	0,9
	Ruim]	568	4	3,1
[Followings]	Fácil]	23	0	0
	Difícil]	25	1	1,8
	Bom]	331	0	0,5
	Ruim]	37	0	0
[Followers]	Fácil]	1486	0	0
	Difícil]	143	0	0
	Bom]	775	0	0
	Ruim]	34	2	2,2

Tabela 8 – Resultado das buscas associadas à interrogação

Padrão de extração	TP	ef (em %)	dp (em %)
[Twitter	1344	11	16,1
[Tweets	1487	24	18,2
[Followings	1025	1	0,8
[Followers	1481	0	0,5

Tabela 9 – Resultado das buscas associadas a advérbios

Padrão de extração	TP	ef (em %)	dp (em %)	
[Twitter]	Bem]	912	3	2,1
	Mal]	1077	4	4,1
	Muito]	1399	3	3,0
	Pouco]	1071	2	1,2
	Mais]	1324	3	3,9
	Menos]	1359	1	0,7
[Tweets]	Bem]	1086	1	0,9
	Mal]	854	1	0,4
	Muito]	1074	1	1,0
	Pouco]	904	2	1,0
	Mais]	1256	2	1,0
	Menos]	1167	1	0,6
[Followings]	Bem]	130	1	0,8
	Mal]	238	1	0,8
	Muito]	437	3	1,9
	Pouco]	66	1	0,8
	Mais]	735	3	1,9
	Menos]	111	4	4
[Followers]	Bem]	1248	0	0,5
	Mal]	766	0	0
	Muito]	581	0	0
	Pouco]	172	0	0
	Mais]	1432	0	0
	Menos]	1439	0	0

#### 4.2.4 Análise dos resultados

A média de *tweets* postados no Twitter por dia é de 58 milhões<sup>32</sup>, de forma que a amostra coletada de 295.797 *tweets* consiste em apenas 0,0005% dos dados do Twitter. Ainda assim, é possível tirar algumas conclusões investigativas. A primeira é que, embora os assuntos postados sejam diversos, os usuários falam sobre o sistema durante o uso e, dependendo de cada padrão de extração, falam mais ou menos.

Nessa análise, mais importante do que a média total de postagens obtida por cada padrão (TP) está a eficácia (*ef*) do padrão utilizado, pois podem ser obtidas muitas postagens com determinado padrão, mas nenhuma delas ser uma PRU, como o caso do padrão [Followers]: TP=1.474, *ef*=0 e *dp*=0 (Tabela 5) e, do contrário, poucas postagens obtidas, mas com uma grande proporção de PRUs, como no caso do padrão [Twitter usabilidade]: TP=22, *ef*=78 e *dp*=14,1 (Tabela 6).

Nesse sentido, os padrões com maior eficácia (acima de 50%), foram: [Twitter usabilidade]; [Twitter interface]; [Twitter erro]; [Twitter bug]; [Tweets interface]; [Tweets

<sup>32</sup> Disponível em: <<http://www.statisticbrain.com/twitter-statistics/>>. Acesso em: 01 de dezembro de 2014.

erro]; [*Tweets problema*]; e [*Tweets bug*]. Isso caracteriza os substantivos: “usabilidade”, “interface”, “erro”, “problema” e “bug” como os mais eficazes para encontrar PRUs. O nome do sistema “Twitter” e a funcionalidade “*Tweets*” também foram aquelas que mais retornaram PRUs. As demais funcionalidades “*Followings*” e “*Followers*” podem não ter sido consideradas como funcionalidades pelos usuários do Twitter, o que explicaria a baixa eficácia de PRUs obtida nesses padrões. Nos próximos parágrafos, os padrões serão analisados também de forma qualitativa, de acordo com as observações percebidas durante a análise.

Na busca individual, foram obtidas muitas postagens, com TP=1.025 para [*Followings*] a TP=1.474 para [*Tweets*]. No entanto, a eficácia de PRUs foi pouca, com  $ef=0\%$  para [*Followers*] a  $ef=6\%$  para [*Tweets*]. Observou-se que somente o nome do sistema ou de suas funcionalidades é insuficiente para recuperação de PRUs, pois o usuário também utiliza o nome do sistema para interagir com os demais, não somente para falar do sistema em uso.

As buscas associadas a substantivos retornaram com o mínimo TP=1, para [*Followers usabilidade*] ao máximo TP=1.148, para [Twitter erro] de postagens. Foram as buscas que retornaram mais PRUs (acima de 50%) (Tabela 6). As buscas associadas a adjetivos (Tabela 7) retornaram com o mínimo de TP=23 para [*Followings Fácil*] ao máximo TP=1.486 para [*Followers Fácil*], com baixas eficácias (0% a 15%) de PRUs. A maior eficácia das buscas com adjetivos refere-se ao padrão [Twitter ruim] ( $ef=15\%$ ).

A maior quantidade de postagens obtidas foi na busca com interrogações, com o mínimo de TP=1.025 para [*Followings “?”*] e o máximo TP=1.487 para [*Tweets “?”*] (Tabela 8). No entanto, a faixa das eficácias das PRUs ficou entre  $ef=0\%$  a  $ef=24\%$ , para [*Followers “?”*] e [*Tweets “?”*], respectivamente.

O resultado das buscas associadas a advérbios (Tabela 9) obteve quantidade mínima de TP=66 para [*Followings pouco*] e máxima de TP=1.439 para [*Followers menos*] e as mais baixas eficácias de PRUs, com  $ef=0\%$  para todos associados a “*Followers*” a  $ef=4\%$  para [Twitter mal]. Observou-se que os usuários utilizam advérbios para falarem de seus sentimentos em relação à sua vida pessoal. O Quadro 12 apresenta exemplos obtidos por cada padrão, PRUs e não-PRUs.

Quadro 12 – Exemplos das postagens obtidas por cada padrão

Buscas	Postagens	
	PRUs	Não-PRUs
<b>Individuais</b>	<p>“O <b>Twitter</b> deveria ter uma página de reclamações. Gostaria de saber por que deixam de me seguir.”</p> <p>“O ruim do <b>Twitter</b>, é que as <b>Tweets</b> não são ilimitadas :/”</p> <p>“Meu <b>twitter</b> ta com problema, não contabiliza os <b>Followers</b>”</p> <p>“Com essa atualização não consigo ver por onde meus <b>Followings</b> postam”</p>	<p>“Bom dia, galera do <b>Twitter!</b>”;</p> <p>“tem gente parando de me seguir no <b>Twitter</b>”;</p> <p>“quase chegando em 8.900 <b>Tweets!</b>!”</p> <p>“Quem quer mais <b>Followers</b>?”</p> <p>“Hoje eu vou assistir <b>The Following</b>”</p>
<b>Associadas a substantivos</b>	<p>“O <b>Twitter</b> não gosta de mim, só pode! ja tem 4 dias que toda hora tento mudar de capa ou de icon mas sempre dá <b>erro</b> :(”</p> <p>“App do <b>Twitter</b> dando <b>erro</b> em algumas mensagens... Só aqui???”</p> <p>“Demorou meio ano pra ler os <b>Tweets</b> grrr <b>bug</b> maldito”</p> <p>“Eu entro no seu perfil ta 1 <b>Followings</b> e 0 <b>Followers</b>, que <b>bug</b> nojento”</p> <p>“quantos <b>Followers</b> está aparecendo que eu tenho? pq aqui eh <b>bug</b>.”</p>	<p>“foi um <b>erro</b> te procurar no <b>Twitter</b>”</p> <p>“Eu não tenho nenhum <b>problema</b> em ter mais <b>Following</b> do que <b>Followers</b>”</p> <p>“Eu perdi 3 <b>Followers</b>, porque eu não entrei. O <b>problema</b> é que eu tenho que estudar”</p> <p>“Precisa de <b>Designer!</b> Revisão crítica de recursos visuais, detalhes da <b>interface</b> via <b>Twitter</b>”</p>
<b>Associadas a interrogações</b>	<p>“Por que <b>Twitter</b>? Por quê?? Por que você está repetindo meus <b>Tweets</b>?”</p>	<p>“Eu estou perdendo <b>Followers</b>. O que eu fiz para merecer isto?”.</p>
<b>Associadas a adjetivos</b>	<p>“<b>Twitter</b> e 140 caracteres: comunicação <b>difícil</b>.”</p> <p>“Eu sempre esqueço como é <b>bom</b> ter <b>Twitter</b> no celular”</p> <p>“A única coisa <b>ruim</b> sobre o <b>Twitter</b> que é difícil saber se alguém está on-line.”</p>	<p>“<b>Bom dia, Twitter</b>”.</p> <p>“Te amo Pai! Poha ele não tem <b>Twitter</b>, assim fica <b>difícil!</b>!”</p> <p>“Eu to tão <b>ruim</b> que nem consigo escrever <b>tweets</b>. ;s”</p>
<b>Associadas a advérbios</b>	<p>“O <b>Twitter</b> é <b>bem</b> mais rápido no celular do que o <b>Face o/ graças a Deus</b>”</p> <p>“Usando app do <b>Twitter</b> para Windows 8. Não está <b>mal!</b>”</p> <p>“porque o <b>Twitter</b> tem limite tão <b>pouco</b> de caracteres?”</p> <p>“pra que o <b>Twitter</b> foi inventar esse negocio de imagens recentes? <b>Muito</b> chato, quero tirar isso”</p>	<p>“Hoje não estou <b>muito</b> a fim de ficar no <b>Twitter</b>”</p> <p>“meu namorado não entra <b>mais</b> no <b>Twitter</b>”</p> <p>“Pelo <b>menos</b> do <b>Twitter</b> tenho que sair”.</p> <p>“Ainda <b>bem</b> que a professora de artes não tem <b>Twitter!</b>”</p>

#### 4.2.5 Conclusão: os usuários “falam” sobre o sistema durante o uso?

A realização dessa investigação respondeu à segunda questão da pesquisa: os usuários “falam” sobre o sistema durante o uso. Com a análise efetuada, também foi possível verificar a eficácia dos padrões de extração utilizados para extrair PRUs. No entanto, muitas PRUs podem ter ficado de fora, se os padrões definidos não tiverem sido abrangentes o suficiente. Por exemplo, palavras referentes às outras metas da UX, como satisfação, atratividade, não foram usadas. Optou-se, inicialmente, pelo critério usabilidade neste primeiro estudo, por considerá-la mais popular. Além disso, alguns padrões utilizados foram observados com

diferentes grafias: “Twitter” como “tt”, “muito” como “mt”, “problema” como “prob”, “pro”, “problem”, entre outras.

Outro desafio consistiu na linguagem analisada. Embora o objetivo da busca seja obter postagens na língua portuguesa, optou-se por usar o nome das funcionalidades em inglês nos padrões de extração, pelo seguinte fato: quando o Twitter surgiu, só existia na versão em inglês<sup>33;34;35;36</sup>, e os nomes de suas funcionalidades ficaram conhecidas em inglês, embora a quantidade de usuários dos outros países estivesse em constante crescimento<sup>37</sup>. Quando essa investigação foi realizada, o Twitter possuía pouco mais de 1 ano com a tradução para a língua portuguesa. Tal fato pode ter sido responsável pela quantidade de postagens em outras línguas, como o inglês, francês, espanhol e alemão, especialmente nas buscas individuais: [“Twitter”], [“Followings”], [“Followers”], [“Tweets”] e nas buscas associadas à interrogação e ao substantivo “bug”, que é um termo genérico. Todas as postagens, embora em línguas diferenciadas, foram traduzidas e contabilizadas, já que o foco era a eficácia de cada padrão de extração. Postagens com *links*, ruídos, ou sem clareza de entendimento, foram classificadas como não-PRUs.

Foram encontradas expressões como: “Grrr”, “rss”, “HAHAHA”, palavrões e símbolos expressivos de sentimentos. O Twitter não tinha *emoticons* na época dessa investigação e os usuários utilizavam parênteses e atalhos do sistema com a intenção de demonstrar sentimentos.

Depois que essa investigação foi realizada, o Twitter atualizou sua busca avançada, incluindo buscas por postagens positivas, negativas, perguntas, e *retweets*. A busca é simples e inclui somente os símbolos: “:)” para postagens positivas, “:(” para postagens negativas, “?” para perguntas e “RT” para *ReTweets*<sup>38</sup>. No entanto, ainda assim, segundo a análise efetuada, somente interrogação é insuficiente para buscas relacionadas a perguntas sobre o uso do sistema. Uma solução seria fazer uma seleção de advérbios apropriados para o domínio de

<sup>33</sup> Disponível em: < <http://tecnologia.uol.com.br/ultimas-noticias/redacao/2011/06/07/twitter-estreia-versao-em-portugues-do-brasil.jhtm> >. Acesso em: 24 de maio de 2013.

<sup>34</sup> Disponível em: <<http://tecnoblog.net/67403/twitter-em-portugues/>>. Acesso em: 24 de maio de 2013.

<sup>35</sup> Disponível em: <<http://g1.globo.com/tecnologia/noticia/2011/06/twitter-lanca-versao-do-site-em-portugues-para-brasileiros.html>>. Acesso em: 24 de maio de 2013.

<sup>36</sup> Disponível em: <<http://www.tecmundo.com.br/twitter/10606-twitter-em-portugues-a-espera-acabou.htm>>. Acesso em: 24 de maio de 2013.

<sup>37</sup> Disponível em: <<http://midiaboom.com.br/midia-social/twitter-de-cara-nova-e-sendo-traduzido-para-portugues/>>. Acesso em: 24 de maio de 2013.

<sup>38</sup> Disponível em: <<https://twitter.com/search-advanced>>. Acesso em: 20 de fevereiro de 2014

avaliação de sistema, considerando advérbios de interrogação para obtenção de postagens interrogativas, por exemplo: “como?”; “por quê?”.

Não foram considerados verbos nesta investigação. Os verbos representam as ações efetuadas no sistema e podem ser muito relevantes para a identificação de problemas. Verbos como: “conseguir”, “fazer”, “executar”, “parar” indicam ações e verbos como “irritar”, “odiar”, “adorar” representam sentimentos no uso do sistema.

### 4.3 COMO OS USUÁRIOS “FALAM”?

Uma vez validada a segunda questão da pesquisa, esta terceira tem como objetivo investigar como os usuários “falam” sobre o SS em uso, especificamente: 1) identificar tipos de PRUs; e 2) investigar como os usuários expressam seus sentimentos em relação ao sistema por meio de suas PRUs. A investigação descrita, a seguir, será dividida de acordo com estes dois subobjetivos.

#### 4.3.1 Subobjetivo 1: identificar tipos de PRUs

##### 4.3.1.1 Sistemas Sociais investigados: Twitter e SIGAA

O Twitter foi investigado usando as postagens extraídas na investigação anterior. A escolha do SIGAA tem a seguinte justificativa: obter uma maior variedade das características das PRUs em SS de diferentes contextos de uso. O SIGAA, no caso, é um sistema acadêmico com características sociais (comunidades, fóruns de discussão, chats etc.). As postagens extraídas são do momento da instalação do sistema (2º semestre de 2010) até janeiro de 2014. Existe, assim, uma maior possibilidade das postagens serem relacionadas às dificuldades dos usuários no começo do uso do sistema. Além disso, existe uma menor diversidade de usuários. São alunos e funcionários de uma universidade do estado do Ceará, com objetivos comuns de uso do sistema: visualizar notas, fazer matrícula, acompanhar aluno etc. Além dessas justificativas, na investigação anterior, com o Twitter, muitas PRUs podem ter ficado de fora, pois a extração foi realizada com padrões de extração, os quais, nem todos foram eficazes. Nessa investigação, foram coletadas todas as postagens de todos os fóruns do SIGAA, sem o uso de padrões de extração.

#### 4.3.1.2 Participantes

Esta análise foi realizada por sete bolsistas, alunos do curso de graduação em computação. Eles tiveram a ajuda de 2 pesquisadores, professores de IHC, que supervisionaram a atividade. Esta atividade foi remunerada com o valor de 1 bolsa de graduação e teve a duração média de 1 mês.

#### 4.3.1.3 Metodologia

A metodologia teve duas etapas: 1) Extração; e 2) Análise das postagens.

Como dito antes, para o Twitter, foram utilizadas as 295.797 postagens, extraídas na investigação anterior, e para o SIGAA, foram extraídas todas as postagens, sem a utilização de padrões de extração, de todos os fóruns<sup>39</sup> do SIGAA de todos os *campi* da UFC + UFCA (Fortaleza, Sobral, Quixadá e Cariri). Foram recuperadas 24.743 postagens a partir do momento em que o sistema foi instalado (2º semestre de 2010) até janeiro de 2014<sup>40</sup>.

A etapa de análise de postagens foi realizada em três subetapas: a primeira consistiu em descartar as postagens irrelevantes, como mensagens repetidas ou em outras línguas diferentes do português. A segunda subetapa consistiu em classificar cada postagem em PRU ou não-PRU e a terceira teve como objetivo analisar as PRUs, a fim de compreender como usuários “falam” sobre o sistema durante seu uso.

Cada bolsista realizou a atividade separadamente. Para facilitar o trabalho dos bolsistas, foram disponibilizados exemplos de PRUs e de não-PRUs classificadas na investigação anterior (Twitter) (Quadro 12). A classificação foi dividida da seguinte forma: 5 bolsistas para classificação do Twitter e 2 bolsistas para classificação do SIGAA. As postagens foram divididas de forma igual entre eles. Os bolsistas tinham as seguintes instruções:

- a) cada postagem deverá ficar somente com um determinado tipo (PRUs ou não-PRUs e tipos característicos de PRUs, como: dúvida, elogio etc.), não havendo a mesma postagem em mais de um tipo;
- b) não deverá ser excluída nenhuma postagem;

<sup>39</sup> Espaço destinado para os usuários postarem mensagens sobre qualquer assunto que lhes possam interessar e para os usuários do sistema cadastrados no mesmo

<sup>40</sup> As postagens foram cedidas pelo Prof. Miguel Franklin de Castro, Coordenador da Coordenadoria de Planejamento, Informação e Comunicação de Dados (COPIC), UFC

- c) se houver dúvida na classificação de alguma postagem, o bolsista deverá entrar em contato com os professores para esclarecimentos; e
- d) a postagem completa deverá ser copiada e colocada numa planilha, e não somente partes dela.

No final da análise, deverão ser entregues:

- a) atividade 1: duas planilhas (uma com as PRUs e uma com as não-PRUs);
- b) atividade 2: planilhas (uma para cada tipo de PRU).

Todas as mensagens selecionadas foram supervisionadas (analisadas) por dois pesquisadores, professores de IHC.

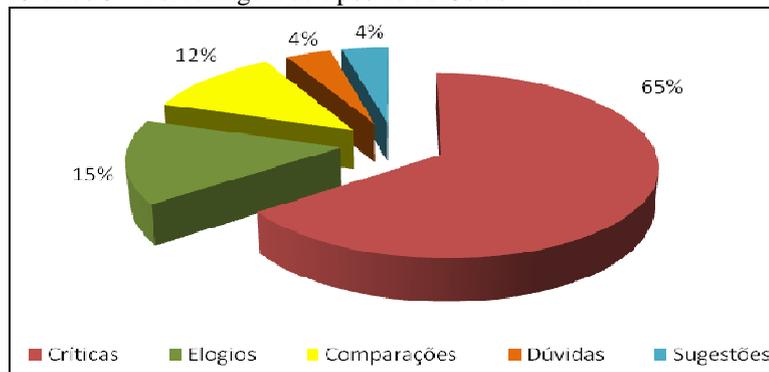
#### 4.3.1.4 Resultados

Os resultados a seguir serão descritos por SS analisado.

**Twitter:** conforme observado na investigação anterior, muitas das postagens extraídas do Twitter foram em outras línguas. Nesta presente investigação, essas postagens foram descartadas, assim como mensagens retuitadas e dados ruidosos, como *tags* HTML, totalizando 226.909 postagens descartadas e 68.888 postagens classificadas como válidas para análise. Destas, 1.452 foram PRUs.

Na terceira subetapa, com o objetivo de compreender como usuários “falam” sobre o sistema, bolsistas e professores observaram cinco principais tipos: críticas, elogios, dúvidas, sugestões e comparações. Então as PRUs foram classificadas nestes cinco tipos, resultando nas seguintes quantidades: 944, 218, 174, 58 e 58 para críticas, elogios, comparações, dúvidas e sugestões, respectivamente (Gráfico 5, dados em porcentagem). Alguns exemplos de tipos de PRUs são apresentados no Quadro 13.

Gráfico 5 – Porcentagem de tipos de PRUs do Twitter



Quadro 13 – Exemplos de tipos de PRUs do Twitter

<b>Tipos</b>	<b>Descrição</b>	<b>Exemplos</b>
Dúvidas	Questionamentos sobre o SS	<p>“Alguém está tendo problemas para fazer comentários na página do Twitter?”</p> <p>“Como eu coloco uma foto no Twitter? :(”</p> <p>“Alguém conhece um site que faça postagens no Twitter, google+ e Facebook?”</p> <p>“Eu bloqueei meu irmão no Twitter, agora como eu faço pra desbloquear??? Me esqueci”</p> <p>“Como faço pra mudar meu id do Twitter pelo celular?”</p> <p>“Alguém sabe como eu faço para uma pessoa não poder ver o meu Twitter?”</p>
Sugestões	Quando um usuário dá uma solução sobre a melhor forma de executar uma tarefa do sistema, ou quando ele sugere uma mudança no sistema.	<p>“O Twitter deveria mais de 140 caracteres.”</p> <p>“A interface do Twitter no android poderia mudar né?”</p> <p>“Bem que o Twitter poderia inventar uma opção de editar o tweet, assim quando escrevêssemos errado não precisaríamos apagar os Tweets KKKK”</p>
Críticas	Quando os usuários reclamam do SS ou reportam algum erro	<p>“Meu Twitter tem um problema, ele está repetindo meus Tweets”</p> <p>“Twitter por que você precisa me notificar quando os meus Followers seguirem outra pessoa?! Eu não me importo!”</p> <p>“Acho ‘triste’ esse site do Twitter. Usabilidade zero.”</p> <p>“O ruim do Twitter e que é difícil de saber se a pessoa esta On.”</p>
Elogios	Quando os usuários elogiam o SS	<p>“Ahhh eu estou muito feliz em usar o Twitter”;</p> <p>“Sabe o que eu adoro no Twitter? é que mesmo que eu clique sem querer em outra pag, o que tava digitando aqui continua aqui qnd volto :3”</p> <p>“Amo a nova interface do Twitter mesmo que tenha me custado 11.8MB”</p>
Comparações	Quando os usuários de SS fazem comparação com qualquer outro SS existente	<p>“Show de bola os App’s do Twitter, agora os dos Facebook são pesados e ruins de usabilidade.”</p> <p>“O aplicativo móvel do #Twitter dá um banho de usabilidade e desempenho no do #Facebook #semmais”</p> <p>“O Twitter é bem mais rápido no celular que o Face o/ graças a Deus!!”</p>

Após análise dos dados, foi aplicado um questionário (APÊNDICE A) e uma entrevista com os cinco alunos que analisaram o Twitter. Todos eles colocaram no questionário que não tiveram dificuldade em classificar postagens como PRUs e não-PRUs, no entanto, durante a entrevista, dois deles afirmaram sentir dificuldades apenas no começo da classificação; “depois era mais fácil”, segundo eles. Suas principais dificuldades eram relacionadas às postagens que tinham o nome do sistema mencionado, mas não estava claro o uso do sistema. Como exemplo, têm-se duas postagens nas quais eles tiveram dúvidas: 1) “Hoje eu não estou muito motivado para navegar no Twitter”. Apesar do nome “Twitter”, a mensagem não se refere ao uso do sistema e sim ao fato de o usuário não estar motivado a

navegar por motivo desconhecido. Tal postagem foi classificada como uma não-PRU; e 2) “Entrar no Twitter é fácil, difícil é sair! :)”. Já com esta postagem é possível inferir que o usuário gosta de estar no Twitter, tanto que relata que é difícil de parar de usá-lo. Tal postagem foi classificada como uma PRU.

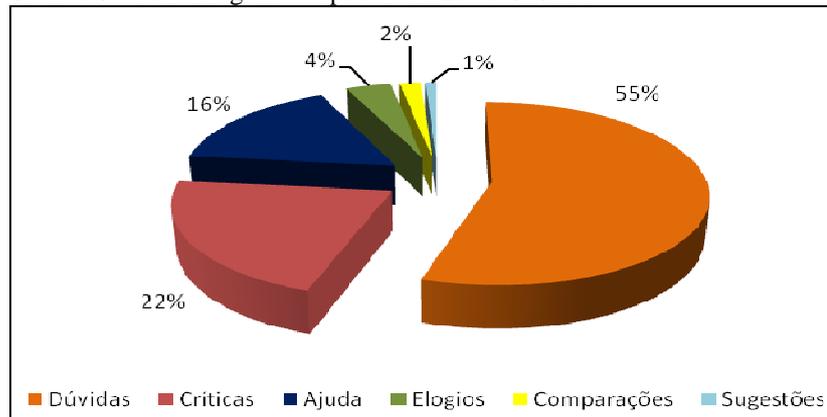
Todos os cinco bolsistas entregaram suas postagens em até 30 dias. As postagens que mais lhes chamaram atenção, segundo eles, foram postagens engraçadas, com críticas e palavrões.

Na classificação por tipos: críticas, sugestões, elogios, dúvidas e comparações, eles relataram que as críticas eram mais fáceis de classificar, mas se confundiam com as dúvidas dos usuários. No processo de análise, algumas postagens classificadas como dúvidas foram reclassificadas como críticas, como: “Por que o Twitter fica diminuindo meus Followers toda hora???”; e “Twitter, o que aconteceu com você? Por que esse erro? Por que não sou capaz de seguir de volta? Por quê?”. Mensagens como essas mostram claramente a intenção de criticar o sistema e não de fazer perguntas sobre ele.

**SIGAA:** nessa análise, não se teve problemas com a língua utilizada, pois todas as postagens extraídas foram em português, no entanto foram encontradas postagens repetidas, links e postagens sem clareza de entendimento. Essas foram descartadas, contabilizando um total de 6.077 postagens descartadas e 18.666 postagens válidas para análise. Tais postagens foram classificadas, totalizando 1.573 PRUs e 17.093 não-PRUs.

Na classificação por tipos de PRU para o SIGAA, observou-se a existência de postagens que não se encaixavam em nenhum tipo definido no Twitter (dúvida, crítica, sugestão, elogio ou comparação). Foi adicionado o tipo ajuda. Tais postagens tinham como principal característica ajudar os demais usuários do sistema, por exemplo: “Para conseguir visualizar as notas tem que clicar nas cadeiras matriculadas, depois na opção ALUNO no lado esquerdo, depois notas ou frequência”. A quantidade de postagens obtidas por cada tipo foi: 340 críticas, 867 dúvidas, 62 elogios, 29 comparações, 16 sugestões e 259 Ajuda (Gráfico 6, dados em porcentagem). O Quadro 14 apresenta alguns exemplos de PRUs e não-PRUs, e o Quadro 15 apresenta alguns exemplos de PRUs divididas por tipos. Os nomes dos usuários foram retirados.

Gráfico 6 – Porcentagem de tipos de PRUs do SIGAA



Quadro 14 – Postagens obtidas na análise do SIGAA

PRUs	Não-PRUs
<p>“Sabe como me dizer o que fazemos pra imprimir o histórico no SIGAA.. no módulo era autoexplicativo... aqui temos que clicar em cada disciplina pra ver a nota...”</p> <p>“Alguém pode explicar como olhar as notas, faltas e rendimento? Se souberem ajudem por favor! Obrigado”</p> <p>“só é comigo que falta o histórico? ou todos os outros já podem ver o histórico. Tenho medo com esta mudança de sistema ‘apagar’ algo que eu já tenha feito”</p>	<p>“Aula normal no dia 28/10”</p> <p>“o trancamento só pode ser feito na data que consta no calendário universitário.”</p> <p>“Alguém saberia informar se haverá aula normal na quinta-feira? E, se sim, onde será?”</p> <p>“Vamos fazer um abaixo assinado e enviar ao professor {nome} ou a quem seja necessário para que se abra uma nova turma?”</p>

Quadro 15 – Exemplos de tipos de PRUs do SIGAA

Categoria	Exemplos
Dúvidas	<p>“Estou com a seguinte dúvida: O novo histórico não apresenta o IRA, ou é o sistema que ainda não atualizou? Por que alguns dias atrás meu IRA estava na tela inicial do SIGAA e agora não está nem na tela, nem no histórico.”</p> <p>“ALGUÉM JÁ APRENDEU COMO TIRAR UM HISTÓRICO NESSE SISTEMA?”</p>
Sugestões	<p>“Esse sistema deveria funcionar no CHROME!!! O GOOGLE CHROME é muito melhor!!!!”</p> <p>“SERIA BOM SE PUDESSE FAZER AS PROVAS EM CASA, POR MEIO DESSE PORTAL...”</p> <p>“Este sistema seria mais eficiente caso as notas fossem expostas em um única pasta, e não ficar abrindo página por página tendo que ficar voltando sempre.”</p>
Críticas	<p>“Pelo amorde... Esse sistema tá ganhando de todos que eu já teve o desprazer de passar. Vou ter que ir para as aulas de duas cadeiras sem ter a menor ideia se estou matriculada.”</p> <p>“Também estou tendo um pouco de dificuldades com esse novo sistema.”</p>
Elogios	<p>“Parece ser bem interessante, gostei da interatividade.”</p> <p>“ai q chic, amei o novo sistema. ta parecendo orkut/facebook”</p>
Comparações	<p>“Se fosse no facebook, ia curtir o que a {nome} falou. :D”</p> <p>“heehh todo mundo ajudando, a interatividade aqui no sigaa ta igual o face”</p>
Ajuda	<p>“Vai na página da disciplina, no menu da turma vc seleciona Alunos &gt; Notas e Frequencia”</p> <p>“gente no site da UFC: www.ufc.br tem um link para um manual de uso desse sistema!”</p>

Os dois bolsistas que classificaram o SIGAA entregaram suas análises antes do prazo sugerido (20 e 22 dias) e, no momento da entrega, foi aplicado o mesmo questionário do Twitter (APÊNDICE A) e também realizada uma entrevista com cada um. Os dois colocaram no questionário que tiveram dificuldade em classificar postagens em PRUs e não-PRUs, e um deles expressou sua dificuldade da seguinte forma: “A minha maior dificuldade foi entender o que dizia cada postagem. Porque, além de ler cada uma, ainda tinha que interpretá-la e definir em qual situação ela se encaixava, não sei se isso aconteceu comigo por falta de experiência, mas é algo que vale a pena ser ressaltado. Em alguns casos, as postagens começam falando mal do sistema e terminam com um elogio.”

A observação desse bolsista se encaixa em um desafio percebido na análise dessas postagens: os dois tiveram muitas dúvidas em classificar postagens não somente pelo tamanho de cada postagem, mas pela complexidade do contexto e pela existência de mais de um tipo em uma mesma postagem. Para este fato, foi sugerido aos bolsistas que eles classificassem a postagem com o tipo mais determinante. No entanto, surgiram outras postagens, como a apresentada no Quadro 16, em que uma única postagem tem dois ou três tipos. Nesses casos, para uma melhor análise, é válido dividir uma postagem em sentenças.

Quadro 16 – Comentário caracterizando mais de um tipo de PRU

<b>Comentário caracterizando mais de um tipo de PRU</b>
<p><i>“Acho o novo portal com uma interface bem mais interativa <b>(Elogio)</b>, porém, como todo início de projeto, muitos ajustes terão que ser feitos <b>(Sugestão)</b>.</i></p> <p><i>Utilizar o Mozilla Firefox como <b>ÚNICO</b> navegador é um retrocesso, anula nossa democracia. <b>(Crítica)</b></i></p> <p><i>E a questão de precisarmos constantemente nos logar novamente é, no mínimo, chata. <b>(Crítica)</b></i></p> <p><i>Acho que tem muito a melhorar <b>(Crítica)</b>, mas gostei do novo estilo. <b>(Elogio)</b></i></p> <p><i>Att,</i> <i>{nome}”</i></p>

Outras postagens que os bolsistas tiveram dúvidas ao classificar foram: 1) “Conhecendo o sistema.”; 2) “Prezados alunos, as notas estão disponíveis no SIGAA”; 3) “Caros, acessem o SIGAA e vejam as suas presenças e faltas semanalmente.” Tais postagens fazem referência ao sistema, mas não estão falando do uso do sistema; logo, as três e as similares a estas foram classificadas como não-PRUs.

Como na análise do Twitter, algumas postagens que estavam classificadas como dúvidas foram reclassificadas como críticas, por exemplo: “Quando será que esse sistema será realmente funcional? Eis a pergunta.”, é uma crítica.

Outra observação desta análise foi que as postagens do SIGAA possuíam mais detalhes referentes a funcionalidades do sistema do que o Twitter. O Quadro 17 apresenta um exemplo de uma PRU com detalhes do sistema, como os passos para realizar uma determinada atividade (visualizar o histórico), elementos da interface (menu, botão, coluna, tópico) e ações (clicar, apertar).

Quadro 17 – Detalhes do sistema em uma PRU

<b>Detalhes do sistema em uma PRU</b>
<i>“Para ver o <b>histórico</b>, entre em ‘Ver turmas anteriores’. Esse <b>botão</b> aparece quando você está no ‘menu discente’ no tópico ‘Turmas do Semestre’. Depois, é só <b>clicar na cadeira</b> que você quer ver a nota, depois <b>apertar no botão ALUNOS da coluna da esquerda</b> e depois ser feliz! Abraço!”</i>

As postagens que os alunos acharam mais fáceis de classificar foram dos tipos: críticas e ajuda.

#### 4.3.1.5 Discussão

Esta seção tem como objetivo fazer uma análise quantitativa (a partir dos dados obtidos) e qualitativa (a partir das observações) entre os dois SS analisados.

A proporção de PRUs obtidas por postagens analisadas no SIGAA foi PRUs=8,42%, (TP=18.666), enquanto que no Twitter foi PRUs=2,1% (TP=68.888). Dada a diversidade de assuntos no Twitter (SHARODA; LICHAN; CHI, 2011), a porcentagem de PRUs é relativamente alta.

O tipo com maior quantidade de postagens no Twitter foi críticas, sendo responsável por 65% das postagens obtidas. A porcentagem de críticas obtidas no SIGAA foi de 22%. As características das críticas entre os dois SS variaram da seguinte forma: no SIGAA, em sua maioria, os usuários foram mais específicos em seus comentários, explicando detalhes do problema de uso, enquanto que no Twitter suas observações foram mais genéricas. Esse fato pode ser justificado: 1) pelo tamanho da postagem. No caso do Twitter, seus usuários só dispõem de 140 caracteres; 2) pela proximidade do relacionamento entre os usuários. É provável que no SIGAA, por ser um sistema de uma universidade, no qual seus usuários são universitários e funcionários, o relacionamento seja mais próximo, fazendo com que os usuários sejam mais detalhistas em suas mensagens, esperando obter uma resposta; e 3) pela própria intenção do usuário ao postar no sistema. É provável que alguns usuários desejem somente expressar sua satisfação ou indignação, sem exatamente esperar uma solução.

Em geral, um fator que vem endossar essa observação é que a categoria com maior postagem no SIGAA foi a de dúvida (55%), e o terceiro maior tipo foi de ajuda (16%), o que retrata a colaboração dos usuários na utilização desse sistema, para obter soluções para resolver um certo problema de uso. O Twitter obteve uma menor porcentagem de dúvidas (4%) e percebeu-se que seus usuários, quando falam do sistema, é com o principal objetivo de desabafar, elogiando ou criticando exageradamente o SS. Eles não têm tanto o objetivo de obter soluções. No entanto, as sugestões foram mais expressivas no Twitter (4%) do que no SIGAA (1%).

#### 4.3.1.6 Conclusão

Foram observadas, nessa investigação, características das postagens de dois SS de contextos diferentes: um SS popular (de entretenimento) – Twitter e um SS acadêmico de uma universidade – fóruns.

Foram observadas, também, postagens com maior e menor intensidade de sentimentos, surgindo a seguinte indagação: como os usuários expressam seus sentimentos em PRUs? Esta questão será investigada na próxima seção.

#### 4.3.2 Subobjetivo 2: *investigar como os usuários expressam seus sentimentos nas PRUs*

##### 4.3.2.1 Sistemas Sociais investigados

A investigação foi realizada em três SS: Twitter, Facebook e Torcida Virtual. Foram investigados três SS com o objetivo de obter uma maior diversidade de postagens analisadas.

##### 4.3.2.2 Procedimento

Esta investigação teve duas etapas: 1) extração e 2) análise das postagens dos usuários, essa última se subdividindo em duas subetapas: (a) classificação das PRUs pela presença ou ausência de sentimentos; e (b) associação das PRUs com o modelo emocional de Norman.

Investigar como os usuários expressam seus sentimentos nas PRUs relaciona-se com a motivação do usuário ao postar PRUs no sistema. Segundo Maitland e Chalmers (2011), a motivação dos usuários em postar mensagens sobre seus problemas está no apoio recebido

por seus contatos, como conselhos em momentos difíceis. Na interação com um sistema, Preece, Rogers e Sharp (2005) afirmam que, quando os usuários estão aborrecidos ou irritados por alguma novidade, reagem de forma exagerada, digitando coisas que nem sequer sonhariam em dizer pessoalmente. Norman (2004), por sua vez, afirma que uma pessoa reage emocionalmente a uma situação antes de avaliá-la cognitivamente. Nesse sentido, espera-se, nesta investigação, estudar a motivação e a expressão dos sentimentos dos usuários em PRUs e que características são relevantes para se avaliar um sistema.

Nos próximos parágrafos, são descritas as duas subetapas de classificação. A primeira classificação teve como objetivo diferenciar PRUs de acordo com seus sentimentos expressos, a fim de investigar se, para a avaliação da interação em sistemas, somente interessariam ser avaliadas as PRUs que expressam algum tipo de sentimento. Para a segunda classificação, fêz-se uma associação com o modelo emocional de Norman (2004) com o objetivo de identificar as intenções dos usuários sobre o sistema. Especificamente, verificar que comportamento motiva o usuário a postar uma PRU.

**Primeira investigação:** consistiu em analisar PRUs, a fim de separar aquelas que expressam algum sentimento (positivo ou negativo) e as que não expressam (neutro).

**Participantes:** 3 especialistas de IHC (pesquisadores ativos da área acadêmica) foram contatados para ler e classificar as PRUs dos usuários. Cada especialista realizou a atividade separadamente, que depois foi verificada pela autora deste trabalho.

**Metodologia:** a coleta dos dados foi realizada da seguinte forma: foram extraídas da página principal de cada SS, postagens de usuários públicos, usando a ferramenta de extração de dados (MENDES; FURTADO; CASTRO, 2013), descrita na investigação anterior, de janeiro a fevereiro de 2013. Os filtros utilizados foram o nome do sistema ou de suas funcionalidades, por exemplo: para extrair as postagens do Twitter, foram usados os padrões: [Twitter], [Tweets], [Followers] e [Following]; para o Facebook, os padrões foram: [Facebook], [Linha do tempo], [eventos] e [fotos]; e, para a TV, os padrões foram: [Torcida Virtual], [Torcida], [Virtual], [TV], [pontos], [site], [amigos]. Nem todas as postagens obtidas foram PRUs, então foram selecionadas as 100 primeiras PRUs de cada SS. Cada participante ficou com um SS e analisou 100 PRUs, classificando-as em dois tipos: com e sem sentimento. O participante lia a PRU, observando pistas indicativas de sentimento, como *emoticons*, interjeições, expressões como: “Grrrr”, “rss”, “hahaha”, palavrões, caixa alta e palavras que expressassem sentimentos, com ou sem ênfase (dobrando as letras, por exemplo).

**Resultados:** Os especialistas não relataram nenhuma dificuldade de classificação das PRUs nesses dois tipos. Ao contrário, relataram ser fácil classificar a presença ou ausência de sentimentos em uma PRU.

No Quadro 18 são apresentados exemplos e porcentagem das PRUs classificadas com a presença e ausência de sentimento nos três SS analisados. Observou-se que ambos os tipos de PRUs são interessantes para a obtenção de percepções dos usuários sobre o sistema: ou seus sentimentos ou suas características no uso do sistema.

Quadro 18 – Classificação das PRUs de acordo com o sentimento percebido

	<b>Com sentimento</b>	<b>Neutras</b>
<b>Twitter</b>	<p><b>71% das postagens</b></p> <p><i>“Adooooooooooooo quando eu quero fazer minhas postagens e o servidor dá erro!”</i></p> <p><i>“ODEIO a nova interface do Twitter!!!”</i></p>	<p><b>29% das postagens</b></p> <p><i>“Meu Twitter está com um problema. Ele não me deixa por gif no icon.”</i></p> <p><i>“Quando teremos a opção editar no Twitter? Sempre erro tudo e só vejo depois de publicado”</i></p>
<b>Facebook</b>	<p><b>59% das postagens</b></p> <p><i>“Isso é um tormento!!! Uma ferramenta maravilhosa pelo que nos poderia proporcionar, mas de uma usabilidade terrível!!!!!! Acorda Facebook!!”</i></p>	<p><b>41% das postagens</b></p> <p><i>“Como tira essa linha do tempo?”</i></p> <p><i>“Ei galera, tem como eu filtrar minhas postagens para que só meus amigos vejam?”</i></p>
<b>Torcida Virtual</b>	<p><b>37% das postagens</b></p> <p><i>“Já estou apaixonada por este site!”</i></p> <p><i>“{palavrão} o torcida era o site melhor do Brasil!!! O que fizeram com o site???”</i></p>	<p><b>63% das postagens</b></p> <p><i>“Alguém entende como se pontua ‘convidando amigos’?”</i></p> <p><i>“O TV não esta computando os meus pontos corretamente, alguém pode me explicar porque?”</i></p>

Outro fato observado foi que tanto sentenças objetivas, como subjetivas são importantes para obter dados sobre o sistema, por exemplo, a sentença subjetiva: “ODEIO a nova interface do Twitter!!!”, representa a opinião do usuário sobre o sistema, enquanto que a sentença objetiva: “O TV não esta computando os meus pontos corretamente, alguém pode me explicar porque?”, representa um problema ocorrido no sistema e uma dúvida do usuário.

Nessa análise, observou-se que os usuários manifestam seu sentimento em relação ao sistema ou em relação a algum erro do sistema, mas nem sempre especificam esse erro. Trata-se de um desabafo, de uma postagem impulsiva, com o objetivo principal de representar o que ele sentiu naquele momento. Também foi observado que, quanto mais o usuário expressa o sentimento em uma postagem, menos ele demonstra o interesse em obter uma solução; do contrário, quando ele foca em obter uma solução, ele não demonstra tanto sentimento.

Nesse ponto, faz-se uma associação com o modelo emocional proposto por Norman (2004), classificando as PRUs nas categorias: visceral, comportamental e reflexiva. Observou-se, nesta análise que, uma PRU visceral é aquela que possui uma maior intensidade de sentimentos, geralmente para criticar ou elogiar o sistema. Está relacionada principalmente à atração e às primeiras impressões. Não contém detalhes do uso e nem de funcionalidades do sistema.

PRUs com menor intensidade de sentimentos (neutras) e também PRUs caracterizadas por ser sentenças objetivas, que contem detalhes do uso, ações efetuadas, funcionalidades, são caracterizadas como PRUs comportamentais. As PRUs reflexivas são caracterizadas por serem subjetivas, apresentarem afeto ou uma situação de reflexão sobre o sistema. O Quadro 19 apresenta outros três tipos de PRUs propostos nesta tese a partir de uma associação das PRUs com o modelo emocional de Norman. A associação seguiu as características apresentadas por Norman para os três níveis de processamento: visceral, comportamental e reflexivo, apresentados no Quadro 6.

Quadro 19 – Associação das PRUs com o modelo emocional de Norman

PRU	Descrição de intenção	Exemplos das postagens analisadas
Visceral	PRU impulsiva, com o intuito principal de criticar / elogiar o sistema	(P1) “Tô xatiada com este Twitter que tá dando erro toda hora : (“ (P2) “Ahhh,eu gosto mais do Twitter!!! As outras redes sociais nem ligo +!!!” (P3) “Só aqui que o Twitter tá dando erro?” (P4) “O Twitter tá dando problema ou eu sou louca?” (P5) “Meu Twitter não funciona que {palavrão}” (P6) “Twitter anda mais insuportável esses dias..”
Comportamental	PRU objetiva, centrada no uso do sistema, com o objetivo de: relatar precisamente o problema ou questionar soluções.	(P7) “Twitter, qual o seu problema!!! Porque você não deixa eu usar o ponto de interrogação!!!” (P8) “Poucas coisas me irritam tanto quanto: <u>The Following download is not available</u> ” (P9) “muito legal isso <u>de poder dar enter</u> no Twitter” (P10) “A interface do Twitter já mudou faz um tempinho e eu ainda não acho as coisas onde clicar” (P11) “Ontem fui rever as pessoas que sigo e percebi que <u>deixei de seguir várias pessoas daqui</u> . Algum bug do Twitter?” (P12) “Meu Twitter está super ruim, as vezes <u>não dá para dar RT e toda hora tem que atualizar a página</u> ” (P13) “Fica dando erro ao mandar os Tweets no celular” (P14) “Meu Twitter tá com problema, <u>fica repetindo meus Tweets</u> ” (P15) “ <u>Quantos Followers está aparecendo que eu tenho? porque aqui é bug</u> ”
Reflexiva	PRU subjetiva, apresentando afeto ou alguma situação de reflexão sobre o sistema.	(P16) “Sabe o que eu sinto falta no Twitter? Uma maneira fácil de encontrar os primeiros Tweets de cada conta.” (P17) “A interface do Twitter no android poderia mudar, né?” (P18) “Alguém aí conhece um site ou app para fazer postagens ao mesmo tempo no Twitter, google+ e

PRU	Descrição de intenção	Exemplos das postagens analisadas
		<p><i>Facebook?"</i></p> <p>(P19) <i>"Sugestão para o Twitter: colocar um botão de edição. Assim não precisaremos apagar tudo porque erramos algo."</i></p> <p>(P20) <i>"Acho que eu to amando mais o Twitter que o face, lá eu nao me expresso tanto assimmm *_*"</i></p>

Baseada nessa associação, foi planejada uma segunda investigação.

**Segunda investigação:** teve como objetivo classificar as PRUs do sistema nos seguintes tipos: Visceral, Comportamental e Reflexiva, seguindo as definições explicitadas no Quadro 19.

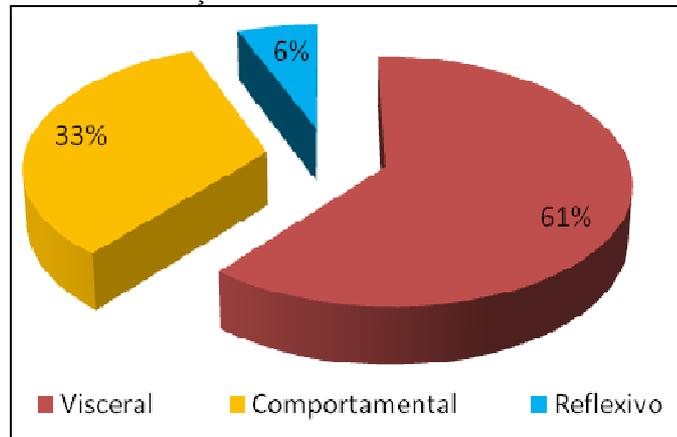
**Participantes:** o time de analisadores era composto por 4 especialistas em IHC. Três deles foram os mesmos da investigação anterior e o outro foi convidado para participar da metodologia descrita a seguir.

**Metodologia:** a investigação foi realizada em 1.452 PRUs de usuários públicos do Twitter, obtidas a partir da investigação anterior (seção 4.2, subobjetivo 1). As postagens tinham sido classificadas em 5 tipos pré-definidos: críticas, elogios, dúvidas, sugestões e comparações, e tinham os seguintes números de postagens: 944, 218, 58, 58 e 174, respectivamente.

O processo de classificação ocorreu da seguinte forma: 2 especialistas classificaram as postagens separadamente e 2 analisaram ambas as classificações. Nessa segunda classificação, no entanto, os especialistas relataram dificuldade em classificar as PRUs, principalmente aquelas do tipo comportamental e reflexiva. As dificuldades surgiram no início do processo e, com o decorrer das classificações, eles entenderam o conceito e seguiram classificando corretamente. Tanto as classificações foram realizadas de forma individual quanto as análises também. As análises consistiam em observar as classificações dos colegas e assinalar se concordavam ou discordavam das classificações realizadas. Foram apenas quatro discordâncias: duas entre postagens viscerais e comportamentais e duas entre postagens comportamentais e reflexivas. A divergência foi resolvida com uma conversa entre os quatro especialistas. O processo final de classificação das postagens foi tabulado pela autora desta tese.

**Resultados:** das 1.452 postagens analisadas, a maior proporção foi de postagens do tipo Visceral (61%), seguido pelo Comportamental (33%) e pela Reflexiva (6%) (Gráfico 7). Alguns exemplos de postagens de cada nível de processamento são apresentados no Quadro 19.

Gráfico 7 – Porcentagem de PRUs obtidas por classificação por intenção do usuário



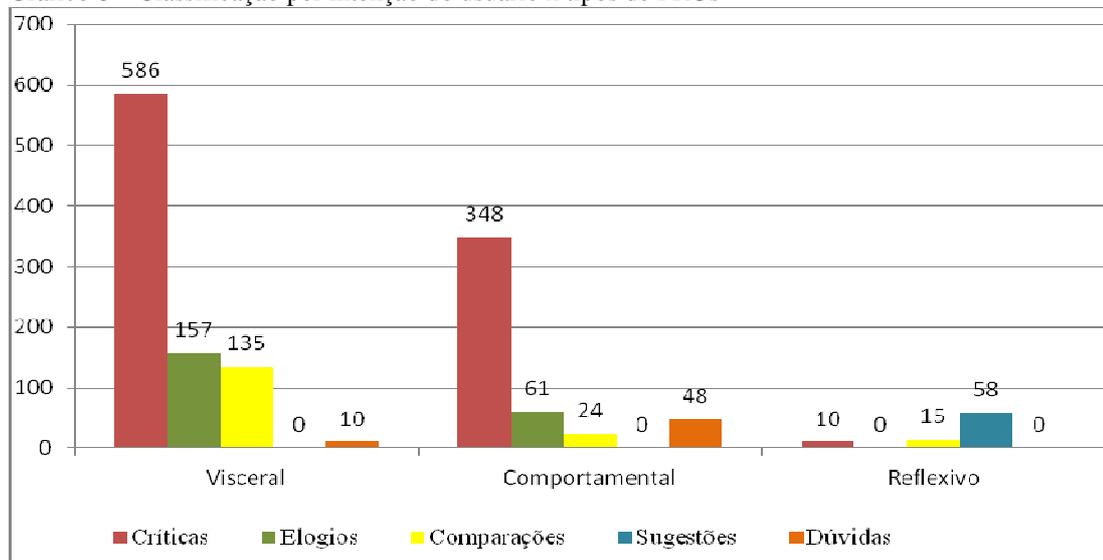
As 1.452 postagens foram classificadas por tipos de postagens em cada nível, conforme ilustra o Gráfico 8.

As postagens viscerais P1 – P6 do Quadro 19 tiveram um dos seguintes tipos (Gráfico 8): críticas (586), elogios (157), comparações (135) e dúvidas (10). Nenhuma sugestão se encaixou nesse nível. Os usuários reclamam, fazem elogios ou comparações no sistema, demonstrando sentimento, mas geralmente sem especificar qual é o erro (P1, P3) / qual é o problema (P4), porque gostam mais do Twitter (P2), porque o sistema não funciona (P5) ou o porque o sistema é insuportável (P6). No nível visceral, não existe nas postagens a necessidade expressa pelos usuários para encontrar uma solução / ou para obter uma explicação.

Como exemplos de postagens comportamentais, P7-P15 (Quadro 19), foram obtidos os seguintes tipos (Gráfico 8): críticas (348), elogios (61), comparações (24) e dúvidas (48). Novamente, nenhuma postagem do tipo sugestão foi classificada nesse nível. No entanto, nas postagens (P7, P8 e P12), os usuários demonstram algum sentimento sobre o uso do sistema e explicam o problema (P7, P14), o erro (P12, P13), o motivo de satisfação (P9) ou de insatisfação (P10) ou ainda expressam a necessidade de encontrar uma solução (P11, P15).

Como exemplo de postagens reflexivas, têm-se as postagens P16-P20 (Quadro 19). Nesse nível, foram encontradas postagens dos seguintes tipos (Gráfico 8): críticas (10), comparações (15) e sugestões (58). São sugestões baseadas na reflexão do uso do sistema. Na postagem P20, ainda pode-se perceber a motivação que o sistema fornece e que foi descrito pelo usuário.

Gráfico 8 – Classificação por intenção do usuário x tipos de PRUs



#### 4.3.2.3 Implicações para a avaliação da UUX a partir de postagens dos usuários

Descrevem-se, a seguir, algumas implicações resultantes dessa investigação que são relevantes para avaliação da UUX a partir de postagens dos usuários.

Uma possível investigação seria sobre o tipo de postagem mais relevante, em função dos objetivos dos avaliadores do sistema, que poderiam ser:

- para identificar o nível de satisfação/insatisfação do usuário com o sistema, as postagens com maior intensidade de sentimento, do tipo visceral, seriam mais relevantes;
- para identificar os problemas de uso do sistema em análise, as postagens do tipo comportamental seriam mais importantes; e
- para a análise de sugestões ou soluções, seriam mais importantes as postagens reflexivas.

Todos esses três tipos de postagens representando a intenção do usuário são importantes para avaliação da UUX do SS, especificamente os sentimentos que o sistema evoca de seus usuários. As mensagens que apresentam uma maior intensidade de sentimentos são relevantes para investigar a UX no sistema, enquanto que postagens com menor intensidade (neutras), as quais, geralmente, relatam problemas sobre o sistema, podem ser relevantes para investigar sua usabilidade.

Embora a maior intensidade de sentimentos observada seja em postagens do tipo visceral, elas podem não ter resultado eficiente para a identificação do problema.

Outros tipos de investigações podem ser identificados, mas ainda serão necessários mais estudos a partir das postagens em SS para classificação de cada tipo de postagem.

#### *4.3.3 Discussão*

Esta investigação teve como objetivo fazer um aprofundamento de como os usuários expressam seus sentimentos por meio das PRUs. Para tanto, foram realizados dois estudos com SS a fim de encontrar características relevantes para a UUX em SS. Observou-se que os usuários, ao postarem PRUs, têm objetivos diferenciados, como: desabafar, encontrar uma solução ou mesmo sugerir a melhoria do sistema. Cabe aos avaliadores decidirem o objetivo da avaliação no sistema.

Hedegaard e Simonsen (2013), em seu estudo sobre comentários de revisores sobre produtos, afirmaram que as revisões da Internet não contêm informações muito detalhadas sobre situações específicas de uso, ou de medições. Eles afirmaram que nenhum revisor escreve: “O número de clique do mouse para navegar da tela inicial para a funcionalidade que eu quero é 7, e que isso é irritante”. No entanto, nas investigações realizadas, foram encontradas postagens que relatam exatamente informações específicas sobre o uso, como: “Gente, desculpa, mas deu problema aqui no Twitter e eu não estou conseguindo colocar pontuação porque abre um menu do Twitter.”, o que caracteriza, também o fato de o usuário reportar o erro no próprio sistema em uso.

Outro ponto importante a ser considerado em avaliação da UUX pela coleta das postagens em SS é a possibilidade de um comentário de um usuário estimular aos demais usuários a comentarem também. Trata-se do “contágio emocional” (HATFIELD; CACIOPPO; RAPSON, 1993; BARSADE, 2002). Tal fator pode ter suas vantagens e desvantagens. Como desvantagem, existe a possibilidade de alguns comentários não representarem a real motivação do usuário ao postar, mas como vantagem, possibilitaria mais descrições dos problemas de uso, pelo estímulo dos demais.

#### *4.3.4 Conclusão: como os usuários “falam”?*

Essa investigação respondeu à terceira questão da pesquisa, como os usuários “falam” sobre o sistema. Eles reclamam do sistema, falam suas dúvidas, elogiam, dão sugestões, fazem comparações e ajudam os demais. Eles fazem isso com uma maior, menor ou nenhuma intensidade de sentimento; com mais, menos ou nenhum detalhe sobre o uso do sistema. Todas estas características são importantes para identificar dados relevantes sobre o uso do sistema.

#### 4.4 COMO AVALIAR?

Esta seção trata da quinta questão de pesquisa desta tese: investigar como avaliar um SS a partir de suas postagens.

Conforme contextualizado nesta tese e de largo conhecimento na literatura, para avaliar a UUX de um sistema, precisa-se analisar o sistema segundo metas de UUX, por exemplo: eficácia, eficiência, utilidade, satisfação do usuário, entre outros. Mais especificamente, espera-se chegar a algum resultado de avaliação da UUX a partir de uma análise das PRUs, especificamente: 1) classificar PRUs em metas de UUX; 2) descobrir, a partir de um conjunto de PRUs, quais são os principais problemas do sistema utilizado; e 3) identificar, a partir de um conjunto de PRUs, o contexto de uso do sistema.

##### *4.4.1 Sistemas Sociais investigados*

Foram utilizadas, nessa investigação, as PRUs do Twitter e do SIGAA obtidas a partir das investigações anteriores.

A fim de responder às questões levantadas nessa investigação, foram realizados dois estudos de classificação: 1) com alunos de graduação de duas disciplinas de IHC e 2) com especialistas de IHC, descritos nas próximas subseções.

##### *4.4.2 1º Estudo: classificação por alunos de IHC*

**Participantes:** participaram deste estudo alunos de duas turmas (disciplina de IHC) do curso de computação nos semestres 2014.1 e 2014.2. Na primeira turma, o SS avaliado foi o Twitter e na segunda, foi o SIGAA (Tabela 10).

Tabela 10 – Aplicação da análise de postagens para alunos da graduação

Semestre	SS	Total de alunos da disciplina	Total de postagens analisadas por aluno	Tempo para análise (em semanas)
2014.1	Twitter	7	50	2
2014.2	SIGAA	10	50	2

**Metodologia:** em ambas as turmas, a metodologia foi aplicada da seguinte forma: o conteúdo teórico referente às metas de UUX já havia sido previamente lecionado em aulas anteriores, bem como métodos de avaliação de UUX, então foi reservada uma aula para explicação do trabalho. Essa aula teve como tema: “Avaliação da usabilidade e experiência do Usuário em Sistemas Sociais a partir das postagens dos usuários”. Na primeira parte da aula, foram apresentadas as motivações em avaliar um SS a partir das postagens de seus usuários e os resultados das investigações anteriores. Na segunda parte da aula, foi realizada uma atividade com os alunos: classificação das postagens dos usuários em metas de UUX propostas por Hedegaard e Simonsen (2013), apresentadas no Quadro 1, bem como alguns exemplos de postagens classificadas pela autora deste trabalho (Quadro 20). O objetivo de aplicar essa atividade na classe com os alunos foi tirar suas dúvidas e prepará-los para o trabalho que seria aplicado para casa.

Quadro 20 – Exemplo de postagens classificadas por metas de UUX

Postagem	Meta
<i>“Esse Twitter está muito lerdo hoje”</i>	Eficiência
<i>“Adoooooro o Twitter”</i>	Satisfação
<i>“Fica dando erro ao mandar os Tweets no celular”</i>	Eficácia
<i>“Acho que eu to amando mais o Twitter que o face, lá eu não me expresso tanto assimmm *_*”</i>	Motivação
<i>“Já estou apaixonada por este site!!”</i>	Satisfação, encantamento
<i>“ODEIO a nova interface do Twitter!!!”</i>	Emoção, estética
<i>“Como tira essa linha do tempo?”</i>	Suporte
<i>“Demorou meio ano pra ler os Tweets grrr bug maldito”</i>	Eficiência, Emoção, frustração

Na atividade de classe, cada aluno recebeu uma folha com 10 postagens e levou em torno de 9 a 12 minutos para classificá-las. Após esse momento, suas postagens classificadas foram trocadas com os colegas, que tinham de corrigir, concordando ou discordando e, nesse último caso, apresentando a classificação correta. Após esse momento, as postagens foram

corrigidas e as dúvidas foram esclarecidas. Suas principais dúvidas foram referentes ao significado das metas e se podiam classificar uma postagem em mais de uma meta. Foi informado que uma postagem poderia ter mais de uma classificação. Em seguida, foi passado o trabalho para eles.

O trabalho consistia em: cada aluno classificar 50 postagens, dispostas em uma planilha, nas seguintes categorias: 1) tipo de postagem; 2) classificação por intenção do usuário; e 3) metas de UUX (Quadro 21). O prazo para classificação foi de 2 semanas e, ao final, o aluno deveria entregar a planilha com as 50 postagens classificadas e um questionário preenchido (APÊNDICE B). É importante mencionar que as postagens do SIGAA, com problemas de classificação na investigação 2 desta tese (Quadro 16), foram divididas em sentenças.

Quadro 21 – Exemplo de um trecho da planilha com postagens a serem classificadas

	Postagem	Tipo de postagem: crítica, elogio, dúvida, sugestão, comparação ou ajuda?	Classificação por intenção do usuário: visceral, comportamental ou reflexiva?	Classificação por metas de UUX: usabilidade, suporte, eficácia, eficiência etc.
1	Gostaria de saber como faz para adicionar optativas no SIGAA. Não consigo! Alguém poderia me ajudar?			
2	sério não consegui me matricular...esse sistema é muito confuso...			

**Resultados:** após duas semanas, os alunos entregaram suas postagens classificadas, que foram corrigidas, e tiveram uma nota equivalente a 10 pelos 150 itens avaliados (cada postagem tinha três classificações). Quem acertasse os 150, ficaria com 10 (foi aplicada uma regra de 3). Como resultado, em sua maioria, os alunos classificaram corretamente, sendo a média de notas da 1ª turma (Twitter) = 8,3 e da 2ª turma (SIGAA) = 8,0. A porcentagem média de acerto da turma por categoria analisada está disposta na Tabela 11. Os alunos, em sua maioria, colocaram mais de duas respostas por classificação de UUX. Ocorreu, durante a correção, casos em que apenas alguma classificação estava correta, mas o item era considerado correto se, pelo menos, uma das classificações estivesse correta.

Tabela 11 – Média de acerto da turma por categoria analisada

Semestre	SS	Tipo de postagem	Classificação por intenção do usuário	Classificação por meta de UUX
2014.1	Twitter	95%	77%	77%
2014.2	SIGAA	89%	76%	76%

Seus principais erros foram:

- a) em relação ao tipo de PRU: confusão entre os tipos dúvida e ajuda (para o SIGAA). Algumas PRUs eram dúvidas e eles classificaram como ajuda, por exemplo: “Alguém sabe como posso tirar um histórico por aqui?”, é uma dúvida, e “Vai em “ver disciplinas anteriores” que tem logo na tua página inicial.. clica na setinha azul que tem em cada disciplina, aí do lado esquerdo vai ter “alunos” e aí tem “ver notas”. ”, é uma ajuda;
- b) em relação à classificação por intenção do usuário: confusão entre os tipos comportamental e reflexiva (ocorreu em ambos SS). Algumas PRUs eram comportamentais (pois colocavam em detalhes, a atividade realizada no sistema) e foram classificadas como reflexiva; e
- c) em relação à classificação por metas de UUX: confusão entre as metas eficácia, eficiência e utilidade.

A classificação de 50 postagens possibilitou que cada aluno tivesse uma percepção do sistema, fornecendo um resultado de avaliação por meio de seus questionários (APÊNDICE B). Em suas respostas, eles identificaram os principais sentimentos dos usuários em relação ao sistema, suas reclamações e elogios. As Figuras 13, 14 e 15 ilustram a percepção da turma em relação ao conjunto de PRUs analisados do SIGAA.

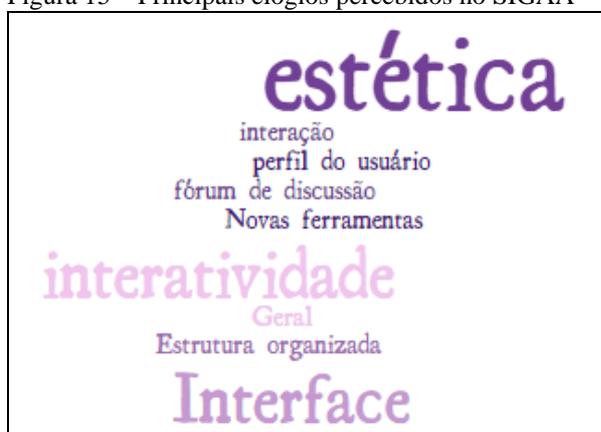
Figura 13 – Principais sentimentos percebidos no SIGAA



Figura 14 – Principais causas/funcionalidades percebidas no SIGAA



Figura 15 – Principais elogios percebidos no SIGAA



Com a aplicação nas duas turmas de alunos, foi concluído que é possível obter resultados de uma avaliação a partir de um conjunto de PRUs. Nesse estudo, foi possível fazer um relacionamento das categorias das PRUs classificadas pelos alunos. Por exemplo, 48% das críticas foram relacionados à meta eficácia, e 86% delas foi relacionada à meta frustração. No entanto, não foi solicitado ao aluno que, ao ler uma PRU, indicasse a funcionalidade a que o usuário se referia. No entanto, ao responder o questionário, o aluno descreveu as principais funcionalidades mencionadas pelos usuários (Figura 15). Percebe-se, que a funcionalidade é uma informação necessária a ser coletada. Com ela, poderiam ser descobertas, as principais funcionalidades com problemas de UUX no SS. Poderia ser realizado, também, um relacionamento das funcionalidades com as outras categorias de classificação sugeridas até então neste trabalho, por exemplo: “x% das **dúvidas** foi em relação à **funcionalidade y**”, ou “z% das PRUs do tipo **visceral** foi relacionada à **interface** do SS”, “w% das **críticas** foi relacionada à **funcionalidade y** e à meta **eficácia**”.

#### 4.4.3 2º Estudo: classificação por especialistas de IHC

**Participantes:** o segundo estudo foi realizado no dia 29 de novembro de 2014 e teve como participantes 12 especialistas de IHC, sendo 4 da academia e 8 da indústria. Foi considerado da academia quem leciona e da indústria quem trabalha com IHC (Tabela 12).

Tabela 12 – Perfil dos especialistas de IHC

Perfil dos participantes	
Gênero	Mulheres (8) Homens (4)
Idade	De 26 a 35 anos (7) De 36 a 45 anos (3) Acima de 45 anos (2)
Área	Academia (4) Indústria (8)
Tempo de experiência	De 4 a 8 anos (7) De 9 a 13 anos (3) Acima de 13 anos (2)

**Metodologia:** os participantes foram convidados com 2 semanas de antecedência, para participar de uma investigação de análise de postagens em metas de UUX. A agenda seguida foi a seguinte (Quadro 22).

Quadro 22 – Agenda da investigação com especialistas de IHC

Horário	Atividade	Descrição
08:00 às 08:30	Recepção dos especialistas	Cooffe break
08:30 às 09:00	Apresentação da investigação	Pequena apresentação explanando os principais conceitos do trabalho
09:00 às 09:30	Classificação	Cada especialista deverá classificar postagens dos usuários de um SS
09:30 às 10:00	Análise	Cada especialista deverá analisar as postagens classificadas por outro colega especialista
10:00 às 10:55	Brainstorm	Os participantes irão discutir sobre a experiência
10:55 às 11:00	Finalização e agradecimentos	Finalização da investigação

Os especialistas receberam: as postagens a serem classificadas, 22 metas para classificação de UUX (APÊNDICE C) e exemplos de PRUs classificadas por UUX (Quadro 20). Cada especialista ficou com 30 PRUs a serem classificadas. Metade dos especialistas ficou com PRUs do SIGAA e a outra metade com PRUs do Twitter. Na análise, os SS eram trocados; o especialista que tinha classificado o SIGAA analisava o Twitter, e vice-versa. Isso foi feito para que cada especialista tivesse uma visão de um SS diferente. A Figura 16 ilustra o momento de classificação das PRUs.

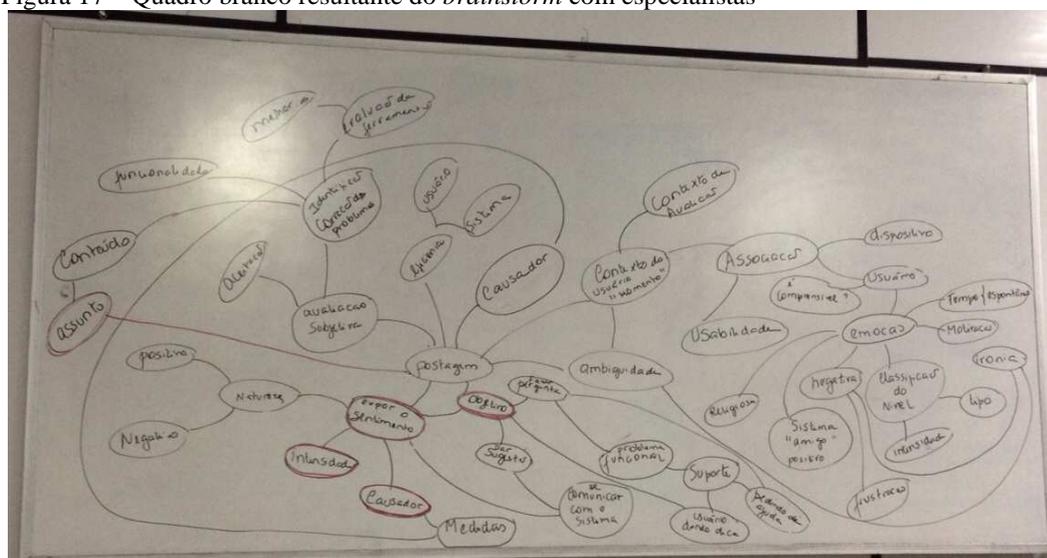
Figura 16 – Classificação das PRUs por especialistas



Após a classificação, cada especialista deveria comentar sobre uma PRU que lhe chamou mais atenção. Em seguida, foi realizado o *brainstorm*, momento no qual foi discutido sobre: os sentimentos do usuário, suas intenções de uso, tipos e a importância das PRUs (conclusões, ações ou como eles representariam).

Dois condutores da investigação fizeram anotações no quadro branco (Figura 17). Outras formas de coleta foram: filmagem e anotações em papel dos pontos mais importantes. Após o *brainstorm*, os participantes deveriam escrever, em uma folha em branco, suas principais dificuldades com a classificação das PRUs.

Figura 17 – Quadro branco resultante do *brainstorm* com especialistas



**Resultados:** os resultados serão apresentados de acordo com a geração de ideias do *brainstorm*, pelas informações escritas por eles ao final da investigação e pelas PRUs classificadas.

Antes do *brainstorm*, cada especialista deveria apresentar uma PRU que tivesse lhe chamado mais atenção e explicar o motivo. Nesse momento, cada especialista acabou apresentando mais de uma PRU. As justificativas mais interessantes estão apresentadas no Quadro 23. Outras justificativas foram relacionadas à dificuldade na classificação das PRUs, listadas a seguir.

Quadro 23 – PRUs que chamaram mais atenção dos especialistas

PRU escolhida	Justificativa
<p>“Maaass q <b>{palavrão}</b>! este twitter.. tá me dando problemas direto no celular... gostaria muito de manda estes fabricantes do Android ir <b>{palavrão}</b>”</p> <p>“Facebook só da Bug! Por isso eu amo o Twitter, porque haja o que hajar <b>{sic}</b>, ele sempre estará aqui pra eu poder reclamar da vida.”</p> <p>“PELO AMOR DE NOSSA SENHORA DAS LETRAS!!! como é que faz p/ visualizar os horários do próximo semestre AJUDA!! preciso ajustar minhas disciplinas c/ meu trabalho.”</p>	<p>A maioria das postagens citadas pelos especialistas foi relacionada à intensidade de sentimento percebida.</p>
<p>“<b>Querido twitter</b>, é <b>{palavrão}</b> ter que apagar o tweet quando eu erro alguma palavra, então coloque a opção “editar tweet”</p> <p>“<b>twitter</b> tira a quebra de linha por favor sem isso é bem melhor !”</p> <p>“<b>QUE {palavrão} TA FICANDO ISSO AQUI!!! TWITTER ME POUPE DE TANTO BUG!!!</b>”</p> <p>“<b>TWITTER AS VEZES EU PRECISO MUITO MAIS DO QUE 140 CARACTERES P PODER ESCREVER TUDO O QUE EU TÔ SENTINDO,SABIA?!</b>”,</p> <p>“<b>TWITTER QUAL O PROBLEMA EM MANDAR OS TWEETS NA ORDEM?!?!????????????</b>”</p>	<p>Tais PRUs foram escolhidas por causa da forma do usuário se expressar, como se estivesse falando com o próprio sistema.</p>
<p>“Dá uma raiva quando estou lendo um tweet e a tl atualiza :(”</p>	<p>Esta PRU foi escolhida porque caracteriza à seguinte situação: o usuário expressou seu sentimento aparentemente logo após a interrupção do Twitter pela atualização.</p>

### Classificação das PRUs:

#### a) Dificuldades de classificação:

As principais dificuldades citadas consistiram em: 1) quantidade alta (22 opções), gerando dúvidas no momento de classificação; e 2) algumas metas são interseções de outras,

como: usabilidade (eficácia, eficiência, satisfação, utilidade, aprendizado, segurança, memorização) e emoção (satisfação, frustração, afeto, prazer, encantamento).

**b) Classificações erradas:**

Durante a análise das PRUs classificadas pelos especialistas, foi percebido que algumas delas não tinham a classificação da usabilidade, somente a de UX, por exemplo, a seguinte PRU: “Síndrome da troca de tela do twitter. Toda vez que vou atualizar a tela, joga o dedo pra esquerda e vai pra aba ‘descobrir’ Grrrr”. O sentimento é tão perceptivo que eles classificaram em metas de UX: frustração, emoção e não classificaram a meta de usabilidade: segurança.

Em outro exemplo: “Odeio essa função do Twitter de só de encostar o dedo já favorita um tweet, dá nem pra stalkear sem medo”, os especialistas classificaram as metas de UX: frustração e emoção, mas não perceberam a meta de usabilidade, identificando um problema de segurança.

Um participante comentou que, fazendo a análise, percebeu que poderia ter feito uma classificação melhor. Ou que, à medida que vão se familiarizando com o método, vão tendo mais facilidade para fazer a classificação.

**c) Percepções gerais:**

Descrevem-se, a seguir, mais algumas percepções dos especialistas sobre a investigação realizada, são elas:

- a) “Quase sempre que houve um problema de eficácia/eficiência, acredito que houve frustração também.”;
- b) “Achei interessante os usuários fornecerem sugestões.”;
- c) “Percebi que no twitter, seus usuários não expõem tanto a parte funcional e no SIGAA, eles expõem mais detalhadamente o problema”;
- d) “Analisar uma postagem solta não diz tudo, mas a partir de um conjunto de PRUs é possível tecer informações sobre o contexto de uso...”.

Os especialistas tiveram um momento rápido de análise (10 minutos de classificação + 10 minutos de análise, com dois SS diferentes), mesmo assim foi possível eles relatarem algumas informações sobre o SS avaliado e seus usuários, como: “percebe-se que os usuários do SIGAA são usuários iniciantes”; “Os usuários do SIGAA ainda não estão acostumados com a plataforma”; “Algumas funcionalidades do SIGAA ainda não foram implementadas”; “O Twitter tem usuários satisfeitos, motivados e engajados! Não imaginava...”. Outras

discussões interessantes surgiram sobre o usuário: “qual o objetivo dele ao postar?”, “o que ele quer?”, “Qual sua prioridade? expressar o sentimento?, solicitar ajuda?”.

Sobre o sentimento expresso pelos usuários nas PRUs, um especialista comentou que em todos os casos existem sentimentos, diferenciando apenas a polaridade e a intensidade, por exemplo, muito alegre. Por outro lado, outros especialistas perceberam que o sentimento nem sempre estava expresso, embora problemas do sistema viessem relatados. Eles comentaram que seria interessante classificar, também, o que causou esse sentimento.

Os especialistas discutiram o que eles fariam com o resultado. Destacam-se as seguintes respostas: análise da aceitação do produto, evolução da aplicação, melhoria na interface ou correção de bugs do sistema.

#### *4.4.4 Discussão*

No estudo com os alunos, cada aluno analisou, em casa, 50 PRUs. E, com isso, foi possível que eles percebessem o contexto de uso do SS. Os especialistas também perceberam que é possível obter informações do contexto de uso a partir de um conjunto de PRUs. Perceberam, por exemplo, que o SIGAA era constituído principalmente de usuários iniciantes que estavam com dificuldades de adaptação à nova ferramenta. Perceberam, também, que ainda existiam funcionalidades não implementadas.

A aplicação deste estudo com especialistas também serviu para reforçar as demais investigações realizadas e descritas no decorrer deste capítulo.

#### *4.4.5 Conclusão: como avaliar?*

Conclui-se aqui essa investigação, abordando às seguintes atividades realizadas: 1) classificar PRUs em metas de UUX. Os alunos e especialistas fizeram classificações das PRUs em metas de UUX. Os alunos, em um teste-piloto antes da avaliação com os especialistas. A partir desta experiência, foram observadas algumas medidas a serem tomadas para facilitar o processo de classificação. Por exemplo: deve-se simplificar a lista de classificação das metas, retirando interseções entre elas para não confundir o especialista. Deve-se, também, separar as metas por critério de qualidade de uso, como: metas de usabilidade e metas de UX, evitando que eles se esqueçam de um dos critérios.

A segunda atividade realizada: descobrir, a partir de um conjunto de PRUs, quais são os principais problemas do sistema utilizado. Em algumas PRUs, o problema é claramente apresentado. Nessas, poderia ser feita uma classificação do problema (ou a causa da postagem) a fim de fazer um relacionamento com as demais classificações efetuadas. A classificação seria a causa ou a funcionalidade mencionada pelo usuário, ou seja, o assunto a que ele se refere em sua PRU.

A terceira atividade foi: identificar, a partir de um conjunto de PRUs, o contexto de uso do sistema. É possível perceber, o contexto de uso do sistema, somente a partir de um conjunto de PRUs classificadas. Assim como outros métodos de avaliação, como avaliação heurística, por exemplo, em que determinados itens do sistema devem ser avaliados a fim de chegar a uma conclusão sobre o mesmo. Para a avaliação a partir de PRUs, deve-se ter uma quantidade de itens (PRUs) avaliados a fim de chegar a alguma conclusão relevante.

Conclui-se, então, que é possível chegar a resultados de avaliação da UUX a partir da análise de PRUs. Essa investigação resultou em 1.150 PRUs analisadas, sendo 850 classificadas por alunos (350 do Twitter + 500 do SIGAA) e 300 classificadas por especialistas de IHC (150 para cada SS: Twitter e SIGAA).

#### 4.5 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou três investigações a partir das postagens dos usuários em quatro SS. Alguns desafios foram enfrentados durante estas investigações. Um deles consistiu no grande volume de dados a serem analisados. Outro desafio consistiu na análise e correção das postagens. A categorização das postagens em PRUs, por tipos, por intenção do usuário e por metas de UUX não é um processo trivial. No entanto, durante a classificação, foi possível perceber que tal processo é facilitado com o tempo, na medida em que os analisadores vão ganhando experiência com esse tipo de classificação. Também foi percebido que são necessários um modelo e uma metodologia (de como usar o modelo) para apoiar os avaliadores no entendimento dos conceitos e na classificação das PRUs. Analisar a linguagem escrita pelos usuários, usada para expressar seus sentimentos em relação ao sistema, é de fundamental importância, não somente para avaliação da UUX ou outros critérios de qualidade de uso, mas para o estudo de IHC.

O Quadro 24 apresenta as investigações realizadas neste capítulo, bem como os resultados obtidos.

Quadro 24 – Resultado das investigações realizadas

Fatores	Investigações		
	Os usuários “falam” sobre o sistema durante o uso?	Como os usuários “falam”?	Como avaliar?
<b>Sistemas analisados</b>	<ul style="list-style-type: none"> <li>• Twitter</li> </ul>	<ul style="list-style-type: none"> <li>• Twitter</li> <li>• SIGAA</li> <li>• Facebook</li> <li>• Torcida Virtual</li> </ul>	<ul style="list-style-type: none"> <li>• Twitter</li> <li>• SIGAA</li> </ul>
<b>Participantes</b>	<ul style="list-style-type: none"> <li>• 1 especialista e professora de IHC</li> </ul>	<ul style="list-style-type: none"> <li>• 7 bolsistas de graduação</li> <li>• 2 professores de IHC</li> <li>• 4 especialistas de IHC</li> </ul>	<ul style="list-style-type: none"> <li>• 17 alunos de graduação</li> <li>• 12 especialistas de IHC</li> </ul>
<b>Metodologia</b>	<ol style="list-style-type: none"> <li>1. Definição de padrões de extração;</li> <li>2. Extração das postagens dos usuários usando os padrões definidos; e</li> <li>3. Análise das postagens obtidas.</li> </ol>	<ul style="list-style-type: none"> <li>• Extração e Análise das postagens</li> </ul>	<ol style="list-style-type: none"> <li>1. Classificação das PRUs em metas de UUX; e</li> <li>2. Discussão sobre as classificações.</li> </ol>
<b>Quantidade de postagens analisadas</b>	<ul style="list-style-type: none"> <li>• Twitter: 295.797</li> </ul>	<ul style="list-style-type: none"> <li>• Twitter: 295.797</li> <li>• SIGAA: 24.743</li> <li>• Facebook: 100</li> <li>• TV: 100</li> </ul>	<ul style="list-style-type: none"> <li>• Twitter: 500</li> <li>• SIGAA: 650</li> </ul>
<b>Resultados</b>	<ul style="list-style-type: none"> <li>• Os usuários falam sobre o sistema durante seu uso;</li> <li>• Padrões de extração de PRUs.</li> </ul>	<ul style="list-style-type: none"> <li>• Classificação por Tipos de PRUs;</li> <li>• Classificação por intenção do usuário nas PRUs.</li> </ul>	<ul style="list-style-type: none"> <li>• Classificação das PRUs em metas de UUX;</li> <li>• Percepções dos especialistas sobre como avaliar um SS usando PRUs.</li> </ul>

## 5 MINERAÇÃO DE DADOS

Neste capítulo, é apresentado o processo de mineração de dados realizado com as postagens dos usuários, obtidas a partir das investigações anteriores. São apresentados os objetivos e a metodologia, bem como as técnicas, ferramentas e algoritmos utilizados. Por fim, são apresentados como resultados os padrões obtidos, seguidos de uma discussão sobre as abordagens adotadas.

### 5.1 OBJETIVOS

O objetivo em aplicar mineração de dados, nesta tese, foi descobrir dados relevantes sobre as PRUs classificadas no Twitter e SIGAA, ilustradas no Quadro 25, e com os seguintes subobjetivos:

- a) validar (e obter novos) padrões de extração de PRUs;
- b) obter padrões para cada tipo de PRU;
- c) obter padrões para cada meta de UUX;

O primeiro subobjetivo foi parcialmente trabalhado na primeira investigação, com o Twitter. Foram definidos padrões de extração a partir de algumas suposições e, em seguida, tais padrões foram analisados estatisticamente, de acordo com a quantidade de PRUs obtidas. No entanto, foi observado que nem todos os padrões definidos foram eficazes. Para atingir esse subobjetivo, será realizada a mineração de dados a partir das postagens de dois SS: Twitter e SIGAA (subseção 4.2).

O segundo subobjetivo visa identificar padrões nos tipos de PRUs: críticas, elogios, sugestões, dúvidas, comparações e ajuda (subseção 4.3).

O terceiro subobjetivo visa extrair um conjunto de padrões relacionados às metas de UUX. Para esse subobjetivo, serão utilizadas as PRUs classificadas por metas pelos alunos e especialistas de IHC (subseção 4.4).

### 5.2 METODOLOGIA

A metodologia desta mineração é explicada utilizando as etapas do processo de KDD, descrito no capítulo 2 desta tese (Figura 4). O KDD é um processo não trivial de identificação

de padrões novos, válidos e potencialmente úteis (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996).

### 5.2.1 Dados

Nesse processo, foram utilizadas as bases de dados classificadas do Twitter e do SIGAA, conforme o Quadro 25. Os tipos de classificação de cada subobjetivo são chamados de classe no processo de mineração de dados.

Quadro 25 – Categoria de dados

SS	Subobjetivo 1 (Tipos das postagens)	Subobjetivo 2 (Tipos de PRUs)	Subobjetivo 3 (Metas de UUX)
Twitter	1 - PRUs 2 - Não-PRUs	1 - Dúvida 2 - Crítica 3 - Elogio 4 - Sugestão 5 - Comparação	1. Afeto 2. Aprendizado 3. Confiança 4. Eficácia 5. Eficiência 6. Estética
SIGAA		1 - Dúvida 2 - Crítica 3 - Elogio 4 - Sugestão 5 - Comparação 6 - Ajuda	7. Frustração 8. Motivação 9. Satisfação 10. Utilidade 11. Suporte 12. Segurança

### 5.2.2 Seleção: coleta de dados dos usuários

Com a classificação do Twitter, o número de PRUs e de não-PRUs obtidas foi 1.452 e 17.093, respectivamente. Enquanto que para o SIGAA foram obtidas 1.573 PRUs e 67.436 não-PRUs. Considerando estas duas classes: PRUs e não-PRUs, tais dados são desbalanceados. Quando, em uma base de dados, uma classe aparece em maior quantidade que outra, a base é dita desbalanceada. Embora, em muitos dos problemas reais, a grande desproporção no número de casos pertencentes a cada uma das classes seja comum (PRATI, 2006; CASTANHEIRA, 2008), quando essa diferença é grande, os sistemas de aprendizado podem encontrar dificuldades em induzir o conceito relacionado à classe minoritária (PRATI, 2006).

Existem diversos métodos para balanceamento de classes. Prati (2006) cita três principais tipos: 1) atribuição de custos de classificação incorreta, que consiste em minimizar a taxa de erros da classificação incorreta; 2) *under-sampling*, que visa o balanceamento dos dados por meio da remoção de exemplos das classes majoritárias; e 3) *over-sampling*, que consiste na replicação de exemplos da classe minoritária.

Nesse processo, optou-se pelos métodos *under-sampling* e *over-sampling*, que consistem em métodos para balancear artificialmente um conjunto de dados (PRATI, 2006). Segundo Prati (2006), esses são dois métodos básicos mais usados para balancear a distribuição das classes: remover exemplos das classes mais populosas e inserir exemplos nas classes menos populosas. Em suas versões mais simples, essa adição/remoção é feita de maneira aleatória (PRATI, 2006). Desta forma, foi implementado um algoritmo em C para estes dois métodos. O algoritmo recebe a base de dados, um valor, e a opção (*under-sampling* ou *over-sampling*) e, desta forma, elimina ou replica um valor determinado de tuplas, aleatoriamente.

Para o subobjetivo 1 foi realizado o balanceamento *under-sampling* aleatório, com as classes Não-PRUs, cujo desbalanceamento é grande em relação às PRUs, em ambos os SS.

Para o subobjetivo 2, também foi aplicado o balanceamento *under-sampling* aleatório e as postagens foram balanceadas pela quantidade do menor tipo. No Twitter foram obtidas as seguintes quantidades de PRUs para os tipos “dúvidas” = 58, “sugestões” = 58, “comparações” = 174, “elogios” = 218 e “críticas” = 944, após o balanceamento, ficaram 58 PRUs para cada tipo. No SIGAA foram obtidas as seguintes quantidades de PRUs para os tipos: “comparação” = 27, “elogio” = 47, “crítica” = 288, “dúvida” = 853, “sugestão” = 102 e “comentários” = 256, após o balanceamento, totaliza 27 PRUs para cada tipo.

Na classificação por meta de UUX, foram utilizadas, para o Twitter, as 350 postagens classificadas pelos alunos e as 150 postagens classificadas pelos especialistas de IHC, ambas corrigidas por dois professores, totalizando 500 PRUs. Para o SIGAA, foram 500 PRUs classificadas pelos alunos e 150 dos especialistas, totalizando 650 PRUs, também corrigidas. A Tabela 13 apresenta a quantidade de PRUs obtidas por metas de UUX no Twitter e no SIGAA.

Tabela 13 – Quantidade de PRUs obtidas por meta de UUX nos SS Twitter e SIGAA

Metas de UUX	Twitter	SIGAA
Afeto	19	10
Antecipação	0	1
Aprendizado	10	141
Confiança	1	32
Eficácia	102	243
Eficiência	37	13
Encantamento	7	1
Engajamento	3	0
Estética	13	6
Frustração	308	415
Impacto	1	3
Memorização	4	7
Motivação	30	9

Metas de UUX	Twitter	SIGAA
Satisfação	115	72
Prazer	5	0
Utilidade	73	32
Suporte	0	26
Segurança	29	14

Algumas das metas obtiveram pouquíssimas ou nenhuma PRU. Logo, consideraram-se as 10 mais relevantes metas (com maior quantidade de PRUs e com, no mínimo, 10 PRUs obtidas em um dos SS). Ficaram doze metas (Tabela 14), que foram balanceadas em três grupos (Tabela 15), usando os dois métodos de balanceamento citados. Os grupos foram divididos pela quantidade de postagens, por exemplo, no grupo 1 a quantidades de postagens de cada grupo é até 50. O método de balanceamento foi escolhido com a intenção de aproximar a quantidade de postagens à média das postagens de cada grupo.

Tabela 14 – Quantidade de PRUs classificadas por metas mais relevantes de UUX nos SS Twitter e SIGAA

Metas UUX	Twitter	SIGAA	Total
Afeto	19	10	29
Aprendizado	10	141	151
Confiança	1	32	33
Eficácia	102	243	345
Eficiência	37	13	50
Estética	13	6	19
Frustração	308	415	723
Motivação	30	9	39
Satisfação	115	72	187
Utilidade	73	32	105
Suporte	0	26	26
Segurança	29	14	43

Tabela 15– Balanceamento das metas de UUX

Método de balanceamento utilizado	Metas UUX	Total
Grupo 1 (aplicar <i>over-sampling</i> , replicando dados para obtenção de classes com 50 exemplos)	Estética	19
	Suporte	26
	Afeto	29
	Confiança	33
	Motivação	39
	Segurança	43
	Eficiência	50
Grupo 2 (aplicar <i>over-sampling</i> , replicando dados da meta utilidade e <i>under-sampling</i> eliminando dados da meta satisfação para obtenção de classes com 151 exemplos)	Utilidade	105
	Aprendizado	151
	Satisfação	187
Grupo 3 (aplicar <i>under-sampling</i> , eliminando dados da meta Frustração para obtenção de classes com 345 exemplos)	Eficácia	345
	Frustração	723

### 5.2.3 Pré-processamento

Usualmente, o pré-processamento é composto de etapas baseadas na abordagem *bag-of-words*. Como essa abordagem transforma em atributo cada token presente no texto, é recomendada realizar uma “limpeza” do texto a fim de descartar palavras irrelevantes, otimizando os termos mais frequentes (preditivos). Tal processo é chamado de redução de dimensionalidade em textos. A precisão de muitos algoritmos de aprendizagem pode ser otimizada pela seleção dos termos mais preditivos (MARTINS, 2003; CAMILO; DA SILVA, 2009). O algoritmo Naïve Bayes, por exemplo, tende a ter um baixo desempenho sem a seleção dos termos mais preditivos (CAMILO; DA SILVA, 2009).

No pré-processamento foi utilizado o *Lemmatizer*, abordado no capítulo 2 desta tese, pelo seguinte motivo: a abordagem *stemming*, apesar de muito usada para padronização dos termos, pode apresentar problemas em palavras com acentuação no radical, como: representação, representações, com este processo ficaria assim: => representaçã, representaçõ (FAGUNDES DA SILVA, 2004). Para o processo do *lemmatizer* aplicado neste trabalho, foram realizados os seguintes passos:

a) análise sintática: foi utilizado o analisador sintático (*tagger-parser*) PALAVRAS<sup>41</sup>.

Segundo Fagundes da Silva (2004), esse analisador é robusto, pois possibilita a análise sintática de sentenças incorretas e até mesmo incompletas, destacando-se também que apresenta baixas taxas de erro (menos de 1% para classe de palavras e 3-4% na análise sintática). Como exemplo do processamento utilizado, considere a sentença **S1: o Twitter fica dando erro, quando eu quero dar uma resposta**, após passar pelo analisador PALAVRAS, ficaria conforme apresentado no Quadro 26 (os lemmas das palavras estão colocados entre colchetes).

Quadro 26 – Resultado da aplicação com o analisador sintático PALAVRAS

[c] <artd> DET M S @>NTwitterSOURCE:live3.runningtextA1	[Twitter] <Inst> PROP M S
@SUBJ>ficaSOURCE:live4.runningtextA1	[ficar] <fmc> V PR 3S IND VFIN @FAUXdandoSOURCE:
live5.runningtextA1	[dar] <vt> V GER @IMV @#ICL-AUX<erroSOURCE:live6.runningtextA1
[erro] <act-d> <ac> N M S @<ACC.SOURCE:live7.runningtextA1	[.] <co-fmc> PU
@COquandoSOURCE:live8.runningtextA1	[quando] <interr> ADV @ADVL>euSOURCE:live9.running
textA1	[eu] PERS M/F 1S NOM @SUBJ>queroSOURCE:live10.runningtextA1
V PR 1S IND VFIN @FMVdarSOURCE:live11.runningtextA1	[dar] <vt> V INF @IMV @#ICL-
<ACCumaSOURCE:live12.runningtextA1	[um] <arti> DET F S @>NrespostaSOURCE:live13.running
textA1	[resposta] <act-s> N F S @<ACCAguémSOURCE:live14.runningtextA1

<sup>41</sup> Disponível em: <<http://beta.visl.sdu.dk/visl/pt/parsing/automatic/upload.php>>. Acesso em: 16 de setembro de 2013.

Quadro 27 – Lemmas coletados

[o]	<artd>	DET M S @>	NTwitterSOURCE:live3.runningtextA1	[Twitter]	<inst>	PROPM S
@SUBJ>	ficaSOURCE:live4.runningtextA1	[ficar]	<fmc>	V PR 3S IND VFIN @FAUXdandoSOURCE:live5.runningtextA1	[dar]	<vt>
[erro]	<act-d>	<ac>	N M S @>	ACC.SOURCE:live7.runningtextA1	[.]	<co-fmc>
PU	@COquandoSOURCE:live8.runningtextA1	[quando]	<interr>	ADV @ADVL>	euSOURCE:live9.runningtextA1	[eu]
PERS M/F 1S NOM @SUBJ>	queroSOURCE:live10.runningtextA1	[querer]	<fmc>	V PR 1S IND VFIN @FMVdarSOURCE:live11.runningtextA1	[dar]	<vt>
V INF @IMV @#ICL->	ACCumaSOURCE:live12.runningtextA1	[um]	<arti>	DET F S @>	NrespostaSOURCE:live13.runningtextA1	[resposta]
<act-s>	N F S @>	ACCAlguemSOURCE:live14.runningtextA1				

- b) parser: foi desenvolvido, para essa mineração, um algoritmo usando a linguagem de programação C, a fim de extrair os lemmas no sintático (Quadro 27). Após a extração do sintático, a sentença ficaria da seguinte forma: **S2: o Twitter ficar dar erro, quando eu querer dar um resposta.** Com esse passo, substantivos, artigos, advérbios ficam no masculino e singular e os verbos ficam no infinitivo, aumentando a contagem de elementos. Similar ao processo de *stemming*, um verbo será indicado como o mesmo, independente do tempo verbal, por exemplo, têm-se as seguintes sentenças:

S3: “Não **consigo** clicar na ajuda!”;

S4: “Não estou **conseguindo** mudar meu perfil.”; e

S5: “Como faço para **conseguir** alterar minha senha?”;

Com estes passos, ficariam:

S6: “Não **conseguir** clicar na ajuda”;

S7: “Não estar **conseguir** mudar meu perfil”; e

S8: “Como fazer para **conseguir** alterar minha senha?”.

**2) Limpeza do texto:** foram realizados os seguintes passos: a) conversão para minúsculas; b) remoção de acentuações; c) remoção de dígitos decimais; e d) remoção de pontuações e caracteres especiais (@, #, \$, %, &, \*, +, ^, °, “, ¬, =, -, /, |, \ etc.). Para tais passos, foi utilizada a ferramenta wReplace<sup>42</sup>. O wReplace é um programa gratuito, que permite alterar/substituir múltiplas letras e palavras de cada vez. É possível criar as próprias regras e salvar para que possa ser reutilizado depois. Dessa forma, foram criadas regras para os seguintes passos listados acima.

<sup>42</sup> Disponível em: <[http://www.sharktime.com/us\\_wReplace.html](http://www.sharktime.com/us_wReplace.html)>. Acesso em: 13 de outubro de 2013.

#### 5.2.4 Transformação

A transformação consiste na criação do vetor de palavras. Conforme mencionado, a abordagem utilizada para criação do vetor foi a *bag-of-words* (BEKKERMAN; ALLAN, 2004 apud ROSSI; REZENDE, 2011). Cada palavra presente nas postagens torna-se um atributo e cada postagem classificada é chamada de exemplo ou instância.

O processo de transformação dos dados em vetor é muito trabalhoso, e, dada a quantidade de dados, a realização manual torna-se impossível. Para tanto, foram utilizadas duas ferramentas: a *Text Master Split & Join*<sup>43</sup>, para dividir as sentenças em arquivos, e a ferramenta *TPT – Text Preprocessing Tool*<sup>44</sup>, que recebe os arquivos em txt com as postagens e cria um arquivo padrão para o processo de mineração de dados, cuja extensão é arff (WITTEN, 2000 apud FAGUNDES DA SILVA, 2004).

A Figura 18 ilustra a ferramenta TPT com as opções utilizadas. A lista de *stopwords* utilizada neste trabalho foi a disponibilizada pela própria ferramenta, com 438 termos, contendo interjeições, advérbios, pronomes pessoais, artigos definidos e indefinidos e numerais (SOARES; PRATI; MONARD, 2008). Com a opção *Document Frequency cut*, as palavras que ocorrem em menor frequência são removidas. Foi utilizada o valor *default 2* para este parâmetro. A opção *Weight* permite escolher o peso dos termos nos documentos, sendo binário (0 – ausência e 1 – ocorrência) ou a frequência do termo. A Figura 19 ilustra um trecho de um arquivo ARFF gerado por esta ferramenta. Neste arquivo, cada palavra presente no texto é um atributo do tipo real. O arquivo apresenta duas classes: não-PRU e PRU e abaixo do @DATA estão as instancias classificadas. A ausência ou presença de um atributo é representada por valores 0s e 1s e a frequência é representada numericamente pela quantidade de vezes em que ele aparece.

---

<sup>43</sup> Disponível em: <<http://www.textmaster.ca/filesplit/download.html>>. Acesso em: 13 de outubro de 2013

<sup>44</sup> Disponível em: <<http://sites.labic.icmc.usp.br/tpt/>>. Acesso em: 13 de outubro de 2013

Figura 18 – Ferramenta TextProcessingTool

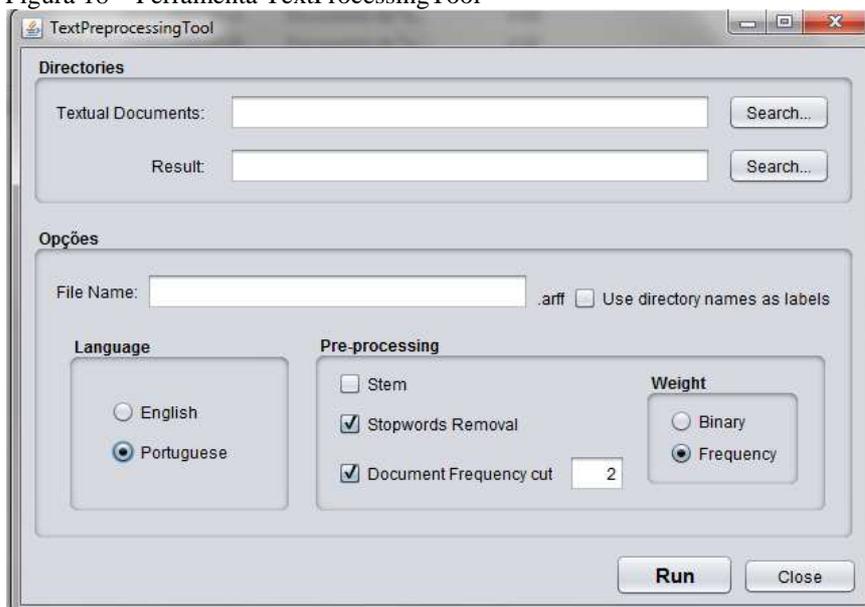
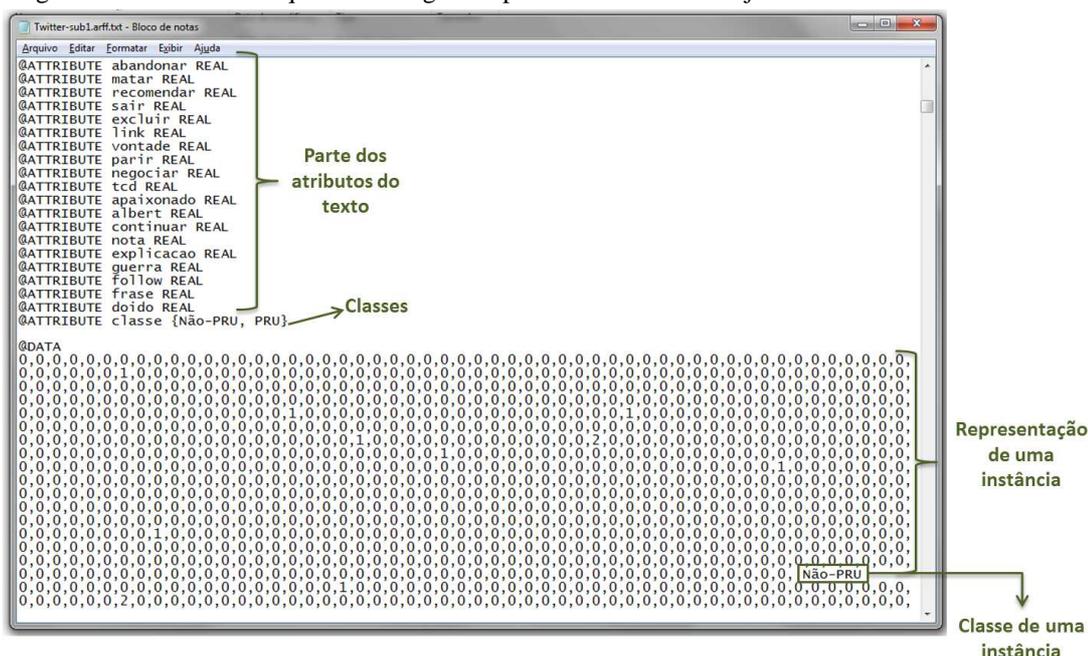


Figura 19 – Trecho do arquivo ARFF gerado para o Twitter - Subobjetivo 1



### 5.2.5 Mineração de dados

Para aplicar a técnica de mineração de dados, foram utilizados dois algoritmos no software WEKA, sendo eles: o J48 e o Naïve Bayes, apresentados no capítulo 2 desta tese. Cada algoritmo foi executado 2 vezes, utilizando 2 técnicas diferentes de teste. A primeira técnica foi a “Cross Validation”, com 10 folds e a segunda técnica foi o “Percentage Split”, com porcentagem 70. A “Cross Validation” é uma técnica amplamente utilizada para

validação cruzada de dados gerados por algoritmos de classificação. É necessário especificar em quantos subconjuntos será dividido o conjunto de dados, e, do total de subconjuntos, um deles é usado para testar as regras geradas por todos os outros (TAN; STEINBACH; KUMAR, 2009). No “Cross Validation 10 *folds*”, a amostra é dividida em k (10) partes de igual tamanho, e de preferência as partes (ou *folds*) devem possuir a mesma quantidade de padrões, garantindo a mesma proporção de classes para cada subconjunto. O algoritmo é treinado sob k-1 “*folds*” (subconjuntos) gerando as regras, posteriormente é validado sob o *fold* (subconjunto) que sobrou. O conjunto de treino é formado pelas 9 (nove) partes e testado calculando a taxa de acerto sob os dados da parte não utilizada, a parte que sobrou. No final, a taxa de acerto é uma média das taxas de acerto nas k iterações realizadas. A técnica de fragmentação “*Percentage Split*” separa uma porcentagem dos dados para testar as regras do classificador. Neste caso, 70% dos exemplos para treinamento e 30% para testes.

Em todos os grupos, foram utilizadas as configurações padrões para o algoritmo J48 e Naïve Bayes.

### 5.2.6 Interpretação / avaliação

Essa etapa será explicada de acordo com cada subobjetivo da mineração.

#### **Subobjetivo 1: padrões de extração de PRUs**

##### ➤ **Twitter**

A mineração do Twitter para o subobjetivo 1 apresentou 895 atributos e 952 instâncias (divididos em duas classes balanceadas: PRU e não-PRU). A Figura 20 ilustra os resultados das aplicações dos dois algoritmos, usando os dois tipos de testes. O resultado da aplicação de um algoritmo pode ser avaliado por um conjunto de métricas disponíveis no WEKA. As métricas mais usadas para avaliação dos dados são:

- a) *Correctly Classsified Instances* — porcentagem de registros classificados corretamente durante a construção do modelo de classificação;
- b) *Incorrectly Classsified Instances* — porcentagem de registros classificados incorretamente durante a construção do modelo de classificação;
- c) *Confusion Matrix* (matriz de confusão) — mostra como as instâncias do conjunto de teste foram classificadas. A classe que está sendo analisada aparece na linha, as classificações encontradas aparecem nas colunas e a diagonal da matriz

corresponde às classificações corretas. Por exemplo, na Figura 20-a, a matriz de confusão apresenta que para a classe não-PRU foram classificadas 435 sentenças corretamente, enquanto que 41 foram classificadas incorretamente. Para a classe PRU foram 447 sentenças classificadas corretamente e 29 classificadas incorretamente.

As demais métricas são utilizadas para avaliação do desempenho do algoritmo na mineração dos dados. Elas são geralmente utilizadas na escolha do melhor algoritmo para a implementação de um sistema, são elas: estatística de *Kappa*, métricas de avaliação de erros e detalhes de acurácia da classificação. Como o objetivo dessa mineração é o estudo dos dados, somente serão analisadas as métricas destacadas na Figura 20.

Figura 20 – Resultado mineração do Twitter – subobjetivo 1

a) J48 - Cross Validation(10 folds)	b) J48 - Percentage Split (70%)
<pre> === Stratified cross-validation === === Summary ===  Correctly Classified Instances      882      92.6471 % Incorrectly Classified Instances     70      7.3529 % Kappa statistic                    0.8529 Mean absolute error                 0.1024 Root mean squared error             0.252 Relative absolute error             20.4868 % Root relative squared error         50.3985 % Total Number of Instances          952  === Detailed Accuracy By Class ===            TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class           0.914   0.061   0.938     0.914   0.926     0.951   Não-PRU           0.939   0.086   0.916     0.939   0.927     0.951   PRU Weighted Avg.   0.926   0.074   0.927     0.926   0.926     0.951  === Confusion Matrix ===   a  b  &lt;-- classified as 435 41   a = Não-PRU  29 447   b = PRU </pre>	<pre> === Evaluation on test split === === Summary ===  Correctly Classified Instances      265      92.6573 % Incorrectly Classified Instances     21      7.3427 % Kappa statistic                    0.8525 Mean absolute error                 0.1063 Root mean squared error             0.2411 Relative absolute error             21.2201 % Root relative squared error         48.1051 % Total Number of Instances          286  === Detailed Accuracy By Class ===            TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class           0.925   0.072   0.918     0.925   0.921     0.96    Não-PRU           0.928   0.075   0.934     0.928   0.931     0.96    PRU Weighted Avg.   0.927   0.074   0.927     0.927   0.927     0.96  === Confusion Matrix ===   a  b  &lt;-- classified as 123 10   a = Não-PRU  11 142   b = PRU </pre>
c) Naïve Bayes - Cross Validation(10 folds)	d) Naïve Bayes - Percentage Split (70%)
<pre> === Stratified cross-validation === === Summary ===  Correctly Classified Instances      882      92.6471 % Incorrectly Classified Instances     70      7.3529 % Kappa statistic                    0.8529 Mean absolute error                 0.076 Root mean squared error             0.2586 Relative absolute error             15.1899 % Root relative squared error         51.7124 % Total Number of Instances          952  === Detailed Accuracy By Class ===            TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class           0.901   0.048   0.949     0.901   0.925     0.975   Não-PRU           0.952   0.099   0.906     0.952   0.928     0.977   PRU Weighted Avg.   0.926   0.074   0.928     0.926   0.926     0.976  === Confusion Matrix ===   a  b  &lt;-- classified as 429 47   a = Não-PRU  23 453   b = PRU </pre>	<pre> === Evaluation on test split === === Summary ===  Correctly Classified Instances      268      93.7063 % Incorrectly Classified Instances     18      6.2937 % Kappa statistic                    0.8731 Mean absolute error                 0.0641 Root mean squared error             0.2279 Relative absolute error             12.7906 % Root relative squared error         45.4691 % Total Number of Instances          286  === Detailed Accuracy By Class ===            TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class           0.91   0.039   0.953     0.91   0.931     0.974   Não-PRU           0.961   0.09   0.925     0.961   0.942     0.98    PRU Weighted Avg.   0.937   0.067   0.938     0.937   0.937     0.977  === Confusion Matrix ===   a  b  &lt;-- classified as 121 12   a = Não-PRU  6 147   b = PRU </pre>

O resultado obtido com a aplicação dos dois algoritmos e dois tipos de testes foram similares, com o melhor resultado de sentenças classificadas corretamente para o Naïve Bayes - percentage split = 93.7063% (Figura 20 - d). Nos quatro resultados, de acordo com as

matrizes de confusão obtidas, o maior número de sentenças classificadas corretamente foi da classe PRU.

O WEKA, com a aplicação do algoritmo J48, gerou uma árvore de tamanho 47, com 24 folhas (APÊNDICE D - Figura 52). Segundo Russel e Norvig (1995), os dados mais importantes na análise de uma árvore de decisão se encontram nos nodos-folhas mais próximos do nodo raiz. O atributo mais significativo foi o “twitter”, seguido de “tweet” e “facebook”.

O algoritmo Naïve Bayes apresenta a frequência obtida pelos atributos em cada classe. A Figura 21 apresenta um trecho da frequência obtida para as classes PRU e não-PRU. Desta forma, identificam-se os atributos mais relevantes para cada classe: os atributos: “ruim”, “difícil”, “orkut”, “facebook”, “problema” e “erro”, foram os mais relevantes para a classe PRU, enquanto que os atributos: “tweet”, “noite”, “dia” e “saber” foram mais relevantes para a classe não-PRU.

Figura 21 - Frequência dos atributos obtidos para as classes PRU e não-PRU do Twitter

```

Attributes: 895
[list of attributes omitted]
Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Naive Bayes Classifier

Attribute                Class
                          Não-PRU  PRU
                          (0.5)    (0.5)
=====
motivo
  mean                    0  0.0042
  std. dev.               0.1667 0.1667
  weight sum              476   476
  precision                1     1

horrivel
  mean                    0  0.0042
  std. dev.               0.1667 0.1667
  weight sum              476   476
  precision                1     1

engracado
  mean                    0.0063 0.0021
  std. dev.               0.1667 0.1667
  weight sum              476   476
  precision                1     1

```

## ➤ SIGAA

A mineração do SIGAA para o subobjetivo 1 apresentou 4.666 atributos e 1.292 instâncias balanceadas e divididas em duas classes: PRU e não-PRU.

O resultado obtido com os dois algoritmos e dois tipos de testes também foram similares, com melhor resultado de sentenças classificadas corretamente para o Naïve Bayes - Cross Validation (10 folds) = 91.7957% (Figura 22 - c). Nos quatro resultados, segundo as matrizes de confusão, o maior número de sentenças classificadas corretamente foi da classe PRU.

Figura 22 – Resultado mineração do SIGAA – subobjetivo 1

a) J48 - Cross Validation(10 folds)	b) J48 - Percentage Split (70%)
<pre> === Stratified cross-validation === === Summary === Correctly Classified Instances 1175 90.9443 % Incorrectly Classified Instances 117 9.0557 % Kappa statistic 0.8189 Mean absolute error 0.1093 Root mean squared error 0.2878 Relative absolute error 21.8654 % Root relative squared error 57.555 % Total Number of Instances 1292  === Detailed Accuracy By Class === TP Rate FP Rate Precision Recall F-Measure ROC Area Class 0.893 0.074 0.923 0.893 0.908 0.933 nao-PRU 0.926 0.107 0.897 0.926 0.911 0.933 PRU Weighted Avg. 0.909 0.091 0.91 0.909 0.909 0.933  === Confusion Matrix ===  a b &lt;-- classified as 577 69   a = nao-PRU 48 598   b = PRU </pre>	<pre> === Evaluation on test split === === Summary === Correctly Classified Instances 348 89.6907 % Incorrectly Classified Instances 40 10.3093 % Kappa statistic 0.7929 Mean absolute error 0.122 Root mean squared error 0.3086 Relative absolute error 24.3652 % Root relative squared error 61.6268 % Total Number of Instances 388  === Detailed Accuracy By Class === TP Rate FP Rate Precision Recall F-Measure ROC Area Class 0.88 0.088 0.899 0.88 0.89 0.924 nao-PRU 0.912 0.12 0.895 0.912 0.903 0.924 PRU Weighted Avg. 0.897 0.105 0.897 0.897 0.897 0.924  === Confusion Matrix ===  a b &lt;-- classified as 161 22   a = nao-PRU 18 187   b = PRU </pre>
c) Naïve Bayes - Cross Validation(10 folds)	d) Naïve Bayes - Percentage Split (70%)
<pre> === Stratified cross-validation === === Summary === Correctly Classified Instances 1186 91.7957 % Incorrectly Classified Instances 106 8.2043 % Kappa statistic 0.8359 Mean absolute error 0.0814 Root mean squared error 0.2845 Relative absolute error 16.2863 % Root relative squared error 56.9072 % Total Number of Instances 1292  === Detailed Accuracy By Class === TP Rate FP Rate Precision Recall F-Measure ROC Area Class 0.907 0.071 0.927 0.907 0.917 0.955 nao-PRU 0.929 0.093 0.909 0.929 0.919 0.972 PRU Weighted Avg. 0.918 0.082 0.918 0.918 0.918 0.964  === Confusion Matrix ===  a b &lt;-- classified as 586 60   a = nao-PRU 46 600   b = PRU </pre>	<pre> === Evaluation on test split === === Summary === Correctly Classified Instances 354 91.2371 % Incorrectly Classified Instances 34 8.7629 % Kappa statistic 0.8239 Mean absolute error 0.0875 Root mean squared error 0.2952 Relative absolute error 17.4807 % Root relative squared error 58.9443 % Total Number of Instances 388  === Detailed Accuracy By Class === TP Rate FP Rate Precision Recall F-Measure ROC Area Class 0.891 0.068 0.921 0.891 0.906 0.952 nao-PRU 0.932 0.109 0.905 0.932 0.918 0.965 PRU Weighted Avg. 0.912 0.09 0.913 0.912 0.912 0.959  === Confusion Matrix ===  a b &lt;-- classified as 163 20   a = nao-PRU 14 191   b = PRU </pre>

O WEKA, com a aplicação do algoritmo J48, gerou uma árvore de tamanho 119, com 60 folhas. Com isso, é possível fazer algumas observações:

- Com a visualização da primeira divisão da árvore de decisão (APÊNDICE D - Figura 53), pode-se constatar que: os atributos: “escrever”, “hífen”, “artigo”, “entregar”, “empresa” e “público”, são determinantes para classificação de não-

PRUs. É possível, também, observar em cada folha os casos corretamente classificados e os incorretamente classificados por cada relação. Nesta primeira divisão, são obtidos dados corretamente classificados em sua totalidade;

b) A Figura 54 (APÊNDICE D) ilustra outro trecho da árvore com alguns atributos relevantes para a determinação da classe PRU, como: “matricular”, “sigaa”, “sistema”, “aparecer” e “módulo”. Para determinação da classe não-PRU, tem-se os atributos: “manhã” e “encaminhar”;

c) Por fim, a árvore finaliza (Figura 55 - APÊNDICE D), apresentando mais alguns atributos relevantes para classificação de PRUs e não-PRUs.

## Subobjetivo 2: extração de padrões para cada tipo de PRU

### ➤ Twitter

A mineração do Twitter para o subobjetivo 2 apresentou 250 atributos e 290 instâncias (balanceadas e divididas em cinco classes: dúvida, crítica, elogio, sugestão e comparação). A Figura 23 ilustra os resultados das aplicações dos dois algoritmos, usando os dois tipos de testes.

Figura 23 – Resultado mineração do Twitter – subobjetivo 2

a) J48 - Cross Validation (10 folds)	b) J48 - Percentage Split (70%)
<pre> ==== Stratified cross-validation ==== ==== Summary ==== Correctly Classified Instances 172      59.3103 % Incorrectly Classified Instances 118    40.6897 % Kappa statistic 0.4914 Mean absolute error 0.1849 Root mean squared error 0.3526 Relative absolute error 57.7883 % Root relative squared error 88.1348 % Total Number of Instances 290  ==== Detailed Accuracy By Class ==== TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class 0.776    0.086    0.692    0.776    0.732    0.847    comparacao 0.448    0.086    0.565    0.448    0.5    0.764    elogio 0.586    0.216    0.405    0.586    0.479    0.738    critica 0.672    0.03    0.848    0.672    0.75    0.809    sugestao 0.483    0.091    0.571    0.483    0.523    0.72    duvida Weighted Avg. 0.593    0.102    0.616    0.593    0.597    0.775  ==== Confusion Matrix ==== a  b  c  d  e  &lt;-- classified as 45  3  7  2  1  a = comparacao 7  26  20  0  5  b = elogio 6  11  34  2  5  c = critica 2  2  3  39  10  d = sugestao 5  4  18  3  28  e = duvida </pre>	<pre> ==== Evaluation on test split ==== ==== Summary ==== Correctly Classified Instances 46      52.8736 % Incorrectly Classified Instances 41    47.1264 % Kappa statistic 0.4174 Mean absolute error 0.2057 Root mean squared error 0.3712 Relative absolute error 63.952 % Root relative squared error 92.2104 % Total Number of Instances 87  ==== Detailed Accuracy By Class ==== TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class 0.826    0.063    0.826    0.826    0.826    0.888    comparacao 0.105    0.044    0.4    0.105    0.167    0.738    elogio 0.545    0.342    0.188    0.545    0.279    0.653    critica 0.438    0.014    0.875    0.438    0.583    0.774    sugestao 0.667    0.101    0.632    0.667    0.649    0.876    duvida Weighted Avg. 0.529    0.093    0.621    0.529    0.532    0.802  ==== Confusion Matrix ==== a  b  c  d  e  &lt;-- classified as 19  2  0  1  1  a = comparacao 2  2  13  0  2  b = elogio 1  2  6  0  2  c = critica 0  0  7  7  2  d = sugestao 1  0  4  1  12  e = duvida </pre>
<pre> ==== Stratified cross-validation ==== ==== Summary ==== Correctly Classified Instances 184      63.4483 % Incorrectly Classified Instances 106    36.5517 % Kappa statistic 0.5431 Mean absolute error 0.1628 Root mean squared error 0.349 Relative absolute error 50.869 % Root relative squared error 87.2419 % Total Number of Instances 290  ==== Detailed Accuracy By Class ==== TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class 0.603    0.069    0.686    0.603    0.642    0.823    comparacao 0.483    0.103    0.538    0.483    0.509    0.798    elogio 0.639    0.138    0.536    0.638    0.583    0.81    critica 0.724    0.03    0.857    0.724    0.785    0.904    sugestao 0.724    0.116    0.609    0.724    0.661    0.876    duvida Weighted Avg. 0.634    0.091    0.645    0.634    0.636    0.842  ==== Confusion Matrix ==== a  b  c  d  e  &lt;-- classified as 35  7  8  2  6  a = comparacao 6  28  16  3  5  b = elogio 4  7  37  0  10  c = critica 4  4  2  42  6  d = sugestao 2  6  6  2  42  e = duvida </pre>	<pre> ==== Evaluation on test split ==== ==== Summary ==== Correctly Classified Instances 50      57.4713 % Incorrectly Classified Instances 37    42.5287 % Kappa statistic 0.4694 Mean absolute error 0.1738 Root mean squared error 0.3665 Relative absolute error 54.033 % Root relative squared error 91.0461 % Total Number of Instances 87  ==== Detailed Accuracy By Class ==== TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class 0.696    0.063    0.444    0.696    0.744    0.844    comparacao 0.211    0.074    0.444    0.211    0.286    0.796    elogio 0.545    0.197    0.286    0.545    0.375    0.763    critica 0.623    0.042    0.769    0.623    0.69    0.923    sugestao 0.778    0.145    0.583    0.778    0.667    0.891    duvida Weighted Avg. 0.575    0.095    0.607    0.575    0.571    0.848  ==== Confusion Matrix ==== a  b  c  d  e  &lt;-- classified as 16  2  2  1  2  a = comparacao 3  4  8  1  3  b = elogio 0  2  6  0  3  c = critica 1  1  2  10  2  d = sugestao 0  0  3  1  14  e = duvida </pre>

Este subobjetivo obteve uma menor porcentagem de postagens classificadas corretamente em comparação com o subobjetivo 1. Seu melhor resultado foi obtido pelo algoritmo Naïve Bayes - teste cross validation de 63.44% (Figura 23 - c). A categoria que

obteve uma melhor classificação com a aplicação do algoritmo J48 e Naïve Bayes – Percentage Split (70%) foi comparação.

O algoritmo J48 gerou uma árvore de tamanho 77 e número de folhas 39 (Figura 56 - APÊNDICE D). A partir dela, é possível chegar a algumas considerações, como por exemplo: se tem o atributo "dever", pertence a categoria sugestão.

### ➤ SIGAA

A mineração do SIGAA para o subobjetivo 2 apresentou 339 atributos e 162 instâncias (balanceadas e divididas em seis classes: dúvida, crítica, elogio, sugestão, comparação e ajuda) (Figura 24). Seu melhor resultado foi obtido pelo algoritmo Naïve Bayes - teste cross validation de 50.61% (Figura 24 - c). A categoria que obteve uma melhor classificação foi comparação (Figura 24-c).

A Figura 57 (APÊNDICE D) ilustra um trecho da árvore gerada com a aplicação do algoritmo J48. Com esta árvore, de tamanho 75 e número de folhas 38, foram levantadas as seguintes observações: a presença dos atributos: “Facebook”, “orkut” ou “rede” é determinante para classificar a classe comparação, enquanto que “parecer”, “criar” e “ferramenta” categoriza a classe elogio.

Figura 24 – Resultado mineração do SIGAA – subobjetivo 2

a) J48 - Cross Validation (10 folds)	b) J48 - Percentage Split (70%)
<pre> ==== Stratified cross-validation ==== ==== Summary ==== Correctly Classified Instances      59      36.4198 % Incorrectly Classified Instances    103     63.5802 % Kappa statistic                    0.237 Mean absolute error                 0.2187 Root mean squared error             0.4107 Relative absolute error              78.6718 % Root relative squared error         110.1282 % Total Number of Instances          162  ==== Detailed Accuracy By Class ==== TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class 0.333    0.03     0.692     0.333   0.45       0.7       elogio 0.296    0.141    0.296     0.296   0.296     0.66     critica 0.259    0.252    0.171     0.259   0.206     0.461    sugestao 0.222    0.067    0.4       0.222   0.286     0.606    ajuda 0.333    0.207    0.243    0.333   0.281     0.59     duvida Weighted Avg. 0.364    0.127    0.415   0.364   0.372     0.652  ==== Confusion Matrix ==== a b c d e f &lt;-- classified as 20 0 4 3 0 0   a = comparacao 3 9 3 7 1 4   b = elogio 2 1 8 8 2 6   c = critica 1 2 7 7 3 7   d = sugestao 2 0 1 7 6 11   e = ajuda 1 1 4 9 3 9   f = duvida </pre>	<pre> ==== Evaluation on test split ==== ==== Summary ==== Correctly Classified Instances      16      32.6531 % Incorrectly Classified Instances    33     67.3469 % Kappa statistic                    0.2042 Mean absolute error                 0.2339 Root mean squared error             0.40664 Relative absolute error              83.5669 % Root relative squared error         108.0704 % Total Number of Instances          49  ==== Detailed Accuracy By Class ==== TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class 0.625    0.024    0.833    0.625  0.714     0.806    comparacao 0.3    0.154   0.333    0.3    0.316     0.482    elogio 0.1     0.026   0.5      0.1    0.167     0.697    critica 0.5     0.488   0.125    0.5    0.2       0.5      sugestao 0.273   0.053   0.6      0.273  0.375     0.673    ajuda 0.25    0.044   0.333    0.25  0.286     0.569    duvida Weighted Avg. 0.327    0.116    0.483   0.327  0.347     0.631  ==== Confusion Matrix ==== a b c d e f &lt;-- classified as 3 0 0 3 0 0   a = comparacao 1 1 3 1 1 1   b = elogio 0 3 1 6 0 0   c = critica 0 2 0 3 1 0   d = sugestao 0 0 0 7 3 1   e = ajuda 0 1 0 2 0 1   f = duvida </pre>
c) Naive Bayes - Cross Validation (10 folds)	d) Naive Bayes - Percentage Split (70%)
<pre> ==== Stratified cross-validation ==== ==== Summary ==== Correctly Classified Instances      82      50.6173 % Incorrectly Classified Instances    80     49.3827 % Kappa statistic                    0.4074 Mean absolute error                 0.3912 Root mean squared error             0.62827 % Relative absolute error              104.9088 % Root relative squared error         162 Total Number of Instances          162  ==== Detailed Accuracy By Class ==== TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class 0.778    0.141    0.525    0.778   0.627     0.855    comparacao 0.519    0.089    0.538    0.519   0.528     0.796    elogio 0.481    0.104    0.481    0.481   0.481     0.761    critica 0.259    0.089    0.368    0.259   0.304     0.621    sugestao 0.444    0.052    0.632    0.444   0.522     0.716    ajuda 0.536    0.119    0.484    0.536   0.517     0.728    duvida Weighted Avg. 0.506    0.099    0.505   0.506   0.497     0.746  ==== Confusion Matrix ==== a b c d e f &lt;-- classified as 21 3 1 0 1 1   a = comparacao 10 14 0 0 1 2   b = elogio 5 1 13 4 0 4   c = critica 2 2 8 7 4 4   d = sugestao 2 4 1 3 12 5   e = ajuda 0 2 4 5 1 15   f = duvida </pre>	<pre> ==== Evaluation on test split ==== ==== Summary ==== Correctly Classified Instances      23      46.9388 % Incorrectly Classified Instances    26     53.0612 % Kappa statistic                    0.3611 Mean absolute error                 0.3987 Root mean squared error             63.5845 % Relative absolute error              106.0045 % Root relative squared error         49 Total Number of Instances          49  ==== Detailed Accuracy By Class ==== TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class 0.625    0.146    0.455    0.625  0.526     0.793    comparacao 0.4     0.128   0.444    0.4    0.421     0.723    elogio 0.5     0.077   0.625    0.5    0.556     0.895    critica 0.333   0.14    0.25    0.333  0.286     0.671    sugestao 0.455   0.053   0.714    0.455  0.536     0.634    ajuda 0.5     0.089   0.333    0.5    0.4       0.767    duvida Weighted Avg. 0.469    0.102    0.511   0.469  0.478     0.747  ==== Confusion Matrix ==== a b c d e f &lt;-- classified as 3 1 1 0 1 0   a = comparacao 3 4 0 0 1 1   b = elogio 0 0 5 4 0 1   c = critica 0 0 1 2 5 1   d = sugestao 0 3 0 2 5 1   e = ajuda 0 1 1 0 0 2   f = duvida </pre>

### Subobjetivo 3: extração de padrões para cada meta de UUX:

Com o objetivo de aproveitar melhor os exemplos classificados dos dois SS, essa mineração somou os exemplos do Twitter e do SIGAA e dividiu em três grupos, de forma a balancear as metas (Tabela 15). Os resultados serão apresentados a seguir pela mineração de cada grupo.

A mineração do grupo 1 apresentou 520 atributos e 170 instâncias (balanceadas e divididas em sete classes: estética, suporte, afeto, confiança, motivação, segurança e eficiência). A Figura 25 ilustra os resultados das aplicações dos dois algoritmos, usando os dois tipos de testes.

Figura 25– Resultado mineração do subobjetivo 3 – Grupo 1

a) J48 - Cross Validation(10 folds)	b) J48 - Percentage Split (70%)
Correctly Classified Instances 83 48.8235 %	Correctly Classified Instances 22 43.1373 %
Incorrectly Classified Instances 87 51.1765 %	Incorrectly Classified Instances 29 56.8627 %
Kappa statistic 0.3984	Kappa statistic 0.3274
Mean absolute error 0.1448	Mean absolute error 0.1735
Root mean squared error 0.3303	Root mean squared error 0.3616
Relative absolute error 59.2542 %	Relative absolute error 70.8324 %
Root relative squared error 94.4598 %	Root relative squared error 103.0645 %
Total Number of Instances 170	Total Number of Instances 51
=== Detailed Accuracy By Class ===	
TP Rate FP Rate Precision Recall F-Measure ROC Area Class	TP Rate FP Rate Precision Recall F-Measure ROC Area Class
0.615 0.229 0.327 0.615 0.427 0.759 afeto	0.667 0.286 0.333 0.667 0.444 0.69 afeto
0.68 0.076 0.607 0.68 0.642 0.837 confiança	0.286 0.068 0.4 0.286 0.333 0.659 confiança
0.44 0.097 0.44 0.44 0.44 0.713 eficiencia	0.375 0 1 0.375 0.545 0.765 eficiencia
0.25 0.006 0.8 0.25 0.381 0.883 estetica	0 0 0 0 0 0.867 estetica
0.385 0.111 0.385 0.385 0.385 0.778 motivacao	0.444 0.095 0.5 0.444 0.471 0.675 motivacao
0.385 0.076 0.476 0.385 0.426 0.736 segurancia	0.571 0.182 0.333 0.571 0.421 0.7 segurancia
0.577 0.007 0.938 0.577 0.714 0.943 suporte	0.6 0.043 0.6 0.6 0.6 0.909 suporte
Weighted Avg. 0.488 0.091 0.554 0.488 0.493 0.803	Weighted Avg. 0.431 0.106 0.463 0.431 0.409 0.738
=== Confusion Matrix ===	
a b c d e f g <-- classified as	a b c d e f g <-- classified as
16 2 0 0 7 1 0   a = afeto	6 2 0 0 1 0 0   a = afeto
1 17 1 0 1 4 1   b = confiança	0 2 0 0 0 3 2   b = confiança
6 0 11 1 4 3 0   c = eficiencia	3 0 3 0 1 1 0   c = eficiencia
7 1 4 4 0 0 0   d = estetica	6 0 0 0 0 0 0   d = estetica
13 1 1 0 10 1 0   e = motivacao	2 0 0 0 4 3 0   e = motivacao
4 4 4 0 4 10 0   f = segurancia	1 0 0 0 2 4 0   f = segurancia
2 3 4 0 0 2 15   g = suporte	
c) Naive Bayes - Cross Validation (10 folds)	d) Naive Bayes - Percentage Split (70%)
Correctly Classified Instances 132 77.6471 %	Correctly Classified Instances 33 64.7059 %
Incorrectly Classified Instances 38 22.3529 %	Incorrectly Classified Instances 18 35.2941 %
Kappa statistic 0.738	Kappa statistic 0.5876
Mean absolute error 0.0693	Mean absolute error 0.1002
Root mean squared error 0.2441	Root mean squared error 0.2941
Relative absolute error 27.9383 %	Relative absolute error 40.9148 %
Root relative squared error 69.8095 %	Root relative squared error 83.8124 %
Total Number of Instances 170	Total Number of Instances 51
=== Detailed Accuracy By Class ===	
TP Rate FP Rate Precision Recall F-Measure ROC Area Class	TP Rate FP Rate Precision Recall F-Measure ROC Area Class
0.962 0.069 0.714 0.962 0.82 0.975 afeto	0.444 0.167 0.364 0.444 0.4 0.815 afeto
0.76 0.021 0.864 0.76 0.809 0.91 confiança	0.714 0 1 0.714 0.833 0.984 confiança
0.64 0.069 0.615 0.64 0.627 0.858 eficiencia	0.75 0.07 0.667 0.75 0.706 0.916 eficiencia
0.75 0.006 0.923 0.75 0.828 0.968 estetica	0.5 0.133 0.333 0.5 0.4 0.902 estetica
0.615 0.028 0.8 0.615 0.696 0.882 motivacao	0.444 0 1 0.444 0.615 0.87 motivacao
0.731 0.028 0.826 0.731 0.776 0.925 segurancia	0.857 0.023 0.857 0.857 0.857 0.963 segurancia
0.962 0.042 0.806 0.962 0.877 0.984 suporte	1 0.022 0.833 1 0.909 0.991 suporte
Weighted Avg. 0.776 0.039 0.786 0.776 0.774 0.927	Weighted Avg. 0.647 0.061 0.721 0.647 0.658 0.911
=== Confusion Matrix ===	
a b c d e f g <-- classified as	a b c d e f g <-- classified as
25 0 0 0 1 0 0   a = afeto	4 0 1 4 0 0 0   a = afeto
0 19 0 0 1 3 2   b = confiança	1 5 0 0 0 1 0   b = confiança
1 0 16 1 2 1 4   c = eficiencia	1 0 6 0 0 0 1   c = eficiencia
2 0 2 12 0 0 0   d = estetica	1 0 2 3 0 0 0   d = estetica
5 2 3 0 16 0 0   e = motivacao	3 0 0 2 4 0 0   e = motivacao
1 1 5 0 0 19 0   f = segurancia	1 0 0 0 0 6 0   f = segurancia
1 0 0 0 0 0 25   g = suporte	0 0 0 0 0 0 5   g = suporte

O melhor resultado foi obtido com a aplicação do algoritmo Naive Bayes, usando o teste Cross Validation (77%), que é o mais indicado quando se minera textos com grande

quantidade de atributos (520), inclusive mais que as instâncias. Suporte e Afeto foram as metas mais acertadas. Ainda assim, é possível tirar algumas conclusões com a árvore gerada pelo algoritmo J48 (Figura 58 e 59 - APÊNDICE D). A árvore, de tamanho 85 e número de folhas 43 decidiu que com o atributo “lento”, é determinante para a classe eficiência. Os atributos: “design”, “bonito”, “estilo”, são determinantes para a classe estética. Os atributos: “preencher”, “solicitar”, “providenciar”, são determinantes para a classe suporte.

A mineração do grupo 2 apresentou 531 atributos e 226 instâncias (balanceadas e divididas em três classes: utilidade, aprendizado e satisfação). A Figura 26 ilustra os resultados das aplicações dos dois algoritmos, usando os dois tipos de testes.

Figura 26 – Resultado mineração do subobjetivo 3 – Grupo 2

a) J48 - Cross Validation (10 folds)	b) J48 - Percentage Split (70%)
<pre> === Stratified cross-validation === === Summary === Correctly Classified Instances      166      73.4513 % Incorrectly Classified Instances    60      26.5487 % Kappa statistic                    0.6017 Mean absolute error                 0.1986 Root mean squared error             0.3925 Relative absolute error             44.6698 % Root relative squared error         83.2542 % Total Number of Instances          226  === Detailed Accuracy By Class ===           TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class           0.72    0.106    0.771    0.72    0.745    0.837    aprendizado           0.763    0.16    0.707    0.763    0.734    0.861    satisfacao           0.72    0.132    0.73    0.72    0.725    0.855    utilidade Weighted Avg.  0.735    0.133    0.736    0.735    0.735    0.851  === Confusion Matrix ===  a  b  c  &lt;-- classified as 54 12  9   a = aprendizado  7 58 11   b = satisfacao  9 12 54   c = utilidade </pre>	<pre> === Evaluation on test split === === Summary === Correctly Classified Instances      45      66.1765 % Incorrectly Classified Instances    23      33.8235 % Kappa statistic                    0.5005 Mean absolute error                 0.2507 Root mean squared error             0.4476 Relative absolute error             56.279 % Root relative squared error         94.6973 % Total Number of Instances          68  === Detailed Accuracy By Class ===           TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class           0.826    0.156    0.731    0.826    0.776    0.825    aprendizado           0.85    0.292    0.548    0.85    0.667    0.797    satisfacao           0.36    0.047    0.818    0.36    0.5    0.641    utilidade Weighted Avg.  0.662    0.155    0.709    0.662    0.642    0.749  === Confusion Matrix ===  a  b  c  &lt;-- classified as 19  3  1   a = aprendizado  2 17  1   b = satisfacao  5 11  9   c = utilidade </pre>
c) Naïve Bayes - Cross Validation (10 folds)	d) Naïve Bayes - Percentage Split (70%)
<pre> === Stratified cross-validation === === Summary === Correctly Classified Instances      196      86.7257 % Incorrectly Classified Instances    30      13.2743 % Kappa statistic                    0.8008 Mean absolute error                 0.0899 Root mean squared error             0.288 Relative absolute error             20.2296 % Root relative squared error         61.0717 % Total Number of Instances          226  === Detailed Accuracy By Class ===           TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class           0.867    0.026    0.942    0.867    0.903    0.95    aprendizado           0.908    0.093    0.831    0.908    0.868    0.927    satisfacao           0.827    0.079    0.838    0.827    0.832    0.904    utilidade Weighted Avg.  0.867    0.067    0.87    0.867    0.868    0.927  === Confusion Matrix ===  a  b  c  &lt;-- classified as 65  4  6   a = aprendizado  1 69  6   b = satisfacao  3 10 62   c = utilidade </pre>	<pre> === Evaluation on test split === === Summary === Correctly Classified Instances      59      86.7647 % Incorrectly Classified Instances    9      13.2353 % Kappa statistic                    0.8033 Mean absolute error                 0.0828 Root mean squared error             0.27 Relative absolute error             18.5944 % Root relative squared error         57.1153 % Total Number of Instances          68  === Detailed Accuracy By Class ===           TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class           0.913    0    1    0.913    0.955    0.997    aprendizado           1    0.188    0.69    1    0.816    0.955    satisfacao           0.72    0    1    0.72    0.837    0.935    utilidade Weighted Avg.  0.868    0.055    0.909    0.868    0.871    0.962  === Confusion Matrix ===  a  b  c  &lt;-- classified as 21  2  0   a = aprendizado  0 20  0   b = satisfacao  0  7 18   c = utilidade </pre>

Novamente o algoritmo Naïve Bayes apresentou um melhor resultado de classificação (Figura 26– c), com 86% de acerto. A classe com maior taxa de acerto foi satisfação.

A mineração do grupo 3 apresentou 886 atributos e 447 instâncias (balanceadas e divididas em duas classes: eficácia e frustração). A Figura 27 ilustra os resultados das aplicações dos dois algoritmos, usando os dois tipos de testes.

Figura 27 – Resultado mineração do subobjetivo 3 – Grupo 3

a) J48 - Cross Validation (10 folds)	b) J48 - Percentage Split (70%)
<pre> === Stratified cross-validation === === Summary === Correctly Classified Instances      358      80.0895 % Incorrectly Classified Instances    89       19.9105 % Kappa statistic                    0.5931 Mean absolute error                 0.2271 Root mean squared error             0.3873 Relative absolute error             45.8825 % Root relative squared error         77.8617 % Total Number of Instances          447  === Detailed Accuracy By Class ===                 TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class                 0.716   0.13    0.818     0.716   0.764     0.862    eficacia                 0.87    0.284  0.79      0.87    0.828     0.862    frustracao Weighted Avg.   0.801   0.215  0.802     0.801   0.799     0.862  === Confusion Matrix ===   a  b  &lt;-- classified as 144 57   a = eficacia  32 214   b = frustracao </pre>	<pre> === Evaluation on test split === === Summary === Correctly Classified Instances      101      75.3731 % Incorrectly Classified Instances    33       24.6269 % Kappa statistic                    0.496 Mean absolute error                 0.2589 Root mean squared error             0.4385 Relative absolute error             52.2789 % Root relative squared error         88.0425 % Total Number of Instances          134  === Detailed Accuracy By Class ===                 TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class                 0.639   0.151  0.78      0.639   0.703     0.809    eficacia                 0.849   0.361  0.738     0.849   0.79      0.809    frustracao Weighted Avg.   0.754   0.265  0.757     0.754   0.75      0.809  === Confusion Matrix ===   a  b  &lt;-- classified as  39 22   a = eficacia  11 62   b = frustracao </pre>
c) Naïve Bayes - Cross Validation (10 folds)	d) Naïve Bayes - Percentage Split (70%)
<pre> === Stratified cross-validation === === Summary === Correctly Classified Instances      324      72.4832 % Incorrectly Classified Instances    123      27.5168 % Kappa statistic                    0.4483 Mean absolute error                 0.2795 Root mean squared error             0.5058 Relative absolute error             56.4797 % Root relative squared error         101.6754 % Total Number of Instances          447  === Detailed Accuracy By Class ===                 TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class                 0.736   0.285  0.679     0.736   0.706     0.768    eficacia                 0.715   0.264  0.769     0.715   0.741     0.771    frustracao Weighted Avg.   0.725   0.273  0.728     0.725   0.725     0.77  === Confusion Matrix ===   a  b  &lt;-- classified as 148 53   a = eficacia  70 176   b = frustracao </pre>	<pre> === Evaluation on test split === === Summary === Correctly Classified Instances      105      78.3582 % Incorrectly Classified Instances    29       21.6418 % Kappa statistic                    0.5595 Mean absolute error                 0.2179 Root mean squared error             0.4383 Relative absolute error             44.0018 % Root relative squared error         88.012 % Total Number of Instances          134  === Detailed Accuracy By Class ===                 TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class                 0.705   0.151  0.796     0.705   0.748     0.827    eficacia                 0.849   0.295  0.775     0.849   0.81      0.828    frustracao Weighted Avg.   0.784   0.229  0.785     0.784   0.782     0.828  === Confusion Matrix ===   a  b  &lt;-- classified as  43 18   a = eficacia  11 62   b = frustracao </pre>

Nessa mineração, o algoritmo J48 obteve um melhor resultado (Figura 27-a), com 80% de instancias classificadas corretamente. A classe com maior taxa de acertos foi frustração.

Foram realizados sete processos de mineração de dados, executados com dois tipos de algoritmos e dois tipos de testes, resultando em 28 resultados, dispostos na Tabela 16. As demais árvores obtidas com esta mineração estão dispostas no APÊNDICE D deste trabalho.

Tabela 16 – Resumo das minerações realizadas

Subobjetivos	SS	Algoritmo	Teste	Instancias classificadas corretamente (%)
Subobjetivo 1	Twitter	J48	Cross Validation (10 folds)	92.64
			Percentage Split (70%)	92.65
	Naïve Bayes	J48	Cross Validation (10 folds)	92.64
			Percentage Split (70%)	93.70
	SIGAA	J48	Cross Validation (10 folds)	90.94
			Percentage Split (70%)	89.69
Naïve Bayes	J48	Cross Validation (10 folds)	91.79	
		Percentage Split (70%)	91.23	
Subobjetivo 2	Twitter	J48	Cross Validation (10 folds)	59.31
			Percentage Split (70%)	52.87
	Naïve Bayes	J48	Cross Validation (10 folds)	63.44
			Percentage Split (70%)	57.87
	SIGAA	J48	Cross Validation (10 folds)	36.41
			Percentage Split (70%)	32.65
Naïve Bayes	J48	Cross Validation (10 folds)	50.61	
		Percentage Split (70%)	46.93	
Subobjetivo 3	Grupo 1 (Twitter + SIGAA)	J48	Cross Validation (10 folds)	48.82
			Percentage Split (70%)	43.13
	Naïve Bayes	J48	Cross Validation (10 folds)	77.64
			Percentage Split (70%)	64.70
	Grupo 2 (Twitter + SIGAA)	J48	Cross Validation (10 folds)	73.45
			Percentage Split (70%)	66.17
Naïve Bayes	J48	Cross Validation (10 folds)	86.72	
		Percentage Split (70%)	86.76	
Grupo 3 (Twitter + SIGAA)	J48	Cross Validation (10 folds)	80.08	
		Percentage Split (70%)	75.37	
Naïve Bayes	J48	Cross Validation (10 folds)	72.48	
		Percentage Split (70%)	78.35	

### 5.3 RESULTADOS

Os resultados são apresentados a seguir, por subobjetivos.

#### 5.3.1 Padrões de extração de PRUs

O Quadro 28 apresenta os termos mais frequentes das PRUs obtido com a mineração no Twitter e SIGAA, de acordo com suas categorias gramaticais.

Quadro 28 – Padrões de extração de PRUs<sup>45</sup>

<b>Categorias gramaticais</b>	<b>Twitter</b>	<b>SIGAA</b>
Verbos	Ter, dar, ficar, poder, fazer, dever, seguir, usar, querer, achar, saber, ver, entrar, existir, mudar, editar, deixar, falar, apagar, amar, olhar, gostar, dizer, vir, passar, começar, parecer, perder, indicar, responder, curtir, carregar	ter, ir, poder, fazer, dever, ver, parecer, dar, gostar, saber, usar, ajustar, compreender, seguir, clicar, conseguir, ajudar, solicitar
Substantivos <sup>46</sup>	Twitter, facebook, <b>tweet</b> , face, orkut, erro, problema, sugestao, pessoa, <b>botao</b> , <b>celular</b> , <b>app</b> , gente, <b>pagina</b> , comentario, <b>opcao</b> , caractere, conta, rede, <b>aplicativo</b> , <b>following</b> , <b>follower</b> , obrigado, tempo, vida, galera, <b>seguidor</b>	<b>Sistema</b> , matricula, disciplina, <b>sigaa</b> , aluno, nota, turma, <b>modulo</b> , <b>academico</b> , <b>ufc</b> , facebook, historico, horario, problema, semestre, prof, <b>plataforma</b> , <b>pagina</b> , coordenação, dado
Adjetivos	Bom, ruim, fácil, novo, direto, difícil, rápido, louco, maldito, triste, lento, confuso, perdido, lindo, legal, chato, feliz, querido, louco	Novo, bom, optativo, simples, antigo, esquerdo, fácil, interativo, obrigado, primeiro, absurdo, boa, igual, institucional, legal, útil
Advérbios	Não, Mais, bem, porque, como, quando, nem, assim, mal, hoje, entao, demais, algo, longe, onde, quando	Nao, como, bem, mais, aqui, nem, entao, quanto, porque, quando, tambem, hoje, melhor, menos, onde

A mineração de dados apresentada possibilitou uma análise das suposições utilizadas na primeira investigação desta tese:

- a) os usuários, ao falarem do SS, utilizam o nome do sistema ou o nome de suas funcionalidades. Pela análise estatística, tal padrão revelou-se ineficaz quando utilizado sozinho. Com a mineração de dados, observou-se que, realmente, os usuários utilizam o nome do sistema e de suas funcionalidades ao se referirem a ele. Foram encontradas outros nome de sistemas e funcionalidades, como: sistema, plataforma, página etc.;
- b) os usuários, ao falarem do SS, utilizam substantivos relacionados à usabilidade do sistema, como: “usabilidade”, “interface”, “erro”, “problema” e “bug”. Pela análise estatística, tais substantivos foram eficazes para a extração de PRUs. Pela mineração de dados, quase toda a totalidade dos substantivos obtidos se referia a funcionalidades do SS;
- c) os usuários, ao falarem do SS, utilizam interrogações: 1) para fazer perguntas sobre o uso do sistema; e 2) para expressar seus sentimentos. Pela análise estatística, as interrogações não foram eficazes quando utilizadas apenas com o padrão nome do sistema ou de suas funcionalidades. No processo de mineração, os símbolos de

<sup>45</sup> As palavras mais frequentes estão no topo. Apenas os resultados mais significativos são listados.

<sup>46</sup> Em vermelho, os nomes utilizados para referência ao sistema e às suas funcionalidades, artefatos e elementos da interface.

- pontuação foram eliminados, no entanto advérbios interrogativos, a exemplo de “como”, “porque”, “quando”, “quanto” e “onde”, foram obtidos;
- d) os usuários ao falarem do SS utilizam adjetivos para qualificá-los, como “fácil”, “difícil”, “bom”, “ruim”. Pela análise estatística, o padrão “ruim” foi o mais eficaz. Na mineração de dados, percebeu-se outros adjetivos, não somente para caracterizar o uso do sistema, mas também o próprio usuário, como “triste”, “feliz” e “confuso”;
- e) as PRUs possuem advérbios que caracterizam o comportamento dos usuários acerca do sistema, como: “bem”, “mal”, “muito”, “pouco”, “mais”, “menos”. Pela análise estatística, essa suposição não foi eficaz. Com a mineração, apenas “bem”, “mal”, e “mais” surgiram como relevantes. Foram obtidos outros advérbios que são relacionados ao contexto de uso do sistema, como: “quando”, “onde” e demais advérbios de interrogação e de negação.

### 5.3.2 Padrões para cada tipo de PRU

O Quadro 29 apresenta os termos mais frequente dos tipos de PRUs obtidos com a mineração no Twitter e SIGAA.

Quadro 29 – Padrões obtido por tipos de PRUs<sup>47</sup>

Tipos de PRUs	Twitter	SIGAA
Comparação	Facebook, mais, face, orkut, bem, ir, bom, ter, usar, achar, ver, dia, ruim	facebook, sistema, ir, aqui, parece, ver, sigaa, ufc, novo, social, rede, orkut, modulo
Elogio	Twitter, facebook, bem, bom, mais, ter, face, ficar, dar, celular, legal, adorar, gente	sistema, parece, novo, plataforma, gostar, bem, modulo, bom, ter, ir, mais, dar, acadêmico
Dúvida	Twitter, fazer, ruim, erro, dar, problema, tweet, saber, pagina, como, comentario, mais, aqui	aluno, matricula, sistema, fazer, ter, como, saber, ver, horário, disciplina, sigaa, bom, prof
Crítica	Twitter, mais, tweet, ficar, entrar, aqui, querer, porque, ir, gostar, tao, mal, chato	sistema, ter, matricula, bom, quando, nem, chrome, senha, usuário, como, disciplina, sigaa, nota
Sugestão	Twitter, ter, dever, tweet, sugestao, poder, mais, botão, seguir, dar, editar, ficar, querer, porque	sistema, matricula, aluno, ter, como, disciplina, sigaa, ir, novo, problema, letra, semestre, dia
Ajuda	-	matricula, ter, aluno, turma, sistema, nota, sigaa, disciplina, você, ver, clicar, dia, passar

### 5.3.3 Padrões para cada meta de UUX

<sup>47</sup> As palavras mais frequentes estão no topo. Apenas os resultados mais significativos são listados.

O Quadro 30 apresenta os padrões obtidos com a mineração no Twitter e SIGAA para as metas de UUX. Os nomes dos sistemas e de suas funcionalidades foram retirados do quadro.

Quadro 30 – Padrões obtidos para as metas de UUX<sup>48</sup>

<b>Categorias</b>	<b>Padrões</b>
Afeto	Gostar, rede, sistema, legal, adorar, amar, facebook, poder, comunicar, facilitar, criativo, eficiente, moderno, interessante, postar, existir, praticar, aproveitar, usufruir
Aprendizado	Saber, fazer, ver, ter, conseguir, poder, encontrar, aparecer, proximo, visualizar, clicar, achar, dizer, entender, entrar, informar, ajudar, novo, pagina, perder, campo, anterior, caminho, dificil, email, facil, menu, mexer, querer, aprender
Confiança	Sistema, parecer, fazer, gostar, praticar, interatividade, usar, poder, interessante, esperar, achar, acreditar, agradar, ajudar, apoio, resolver, melhorar, melhoria, começar
Eficácia	Sistema, ter, fazer, problema, saber, conseguir, aparecer, poder, dar, ver, erro, dizer, ficar, querer, tentar, dever, ajuste, gostar, colocar, realizar, continuar, mensagem, achar, mudar, esperar, abrir, pagina, entrar, opcao, permitir, faltar, excluir, acessar, mandar, jeito
Eficiência	ruim, ficar, lento, ver, dar, sistema, esperar, querer, carregar, erro, face, legal, fazer, problema, usar, achar, passar, aparecer, disciplina, pau, app, espaço, ler, tentar, abrir
Estética	Gostar, parecer, ficar, bonito, plataforma, design, atualizar, android, limpo, perfeito, acostumar, pagina, melhoria, adorar, adaptar, add, campo, curtir, escrever, esquerdo, estilo, mexer, simples, visual, windows, entrar, interface, diferente
Frustração	Ter, sistema, problema, fazer, ruim, saber, ficar, dar, poder, erro, ver, conseguir, querer, dizer, bug, aparecer, tempo, achar, mau, dever, entender, mudanca, mudar, passar, usar, entrar, funcionar, odiar, começar, tentar, abrir, acessar, colocar, complicar, falar, precisar
Motivação	Poder, adorar, desabafar, achar, criar, falar, ver, plataforma, social, saber, super, fazer, interatividade, usar, deixar, existir, amar, sistema, interessante, ajudar, igual, facebook, ficar, interativo, otimo, legal
Satisfação	Sistema, gostar, legal, achar, adorar, ter, face, facebook, demorar, parecer, poder, social, existir, ficar, entrar, interatividade, proximo, criar, amar, desabafar, facil, ver, interessante, bonito, ferramenta, novo, fazer, querer, falar, deus, usar, deixar, perfeito
Segurança	Saber, sistema, começar, mau, fazer, dar, aparecer, errar, apertar, poder, ruim, erro, digitar, constar, faltar, lista, odiar, ver, ficar, querer, mexer, perceber, problema, clicar, medo, enter, raiva, teclar, espaço, criar, remover, complicar
Suporte	Fazer, sistema, saber, preencher, ver, esperar, problema, solicitar, clicar, clique, consultar, providenciar, pedir, aparecer, dar, errar, poder, erro, pergunta, vir, acreditar, continuar, dever, resolver, funcionar, tentar, conseguir, cadastrar, complexo, procurar, selecionar
Utilidade	Poder, saber, usar, ver, excluir, dar, ficar, ruim, problema, dever, mozilla, querer, caractere, demorar, sistema, tirar, objetivo, limite, escrever, entrar, precisar, abrir, linha, face, deixar, utilizar, editar, palavra, gostar, achar, novo, fazer, otimo, ter, parecer, proximo

## 5.4 DISCUSSÃO

Nesta subseção, são discutidos alguns fatores que podem influenciar os dados apresentados.

**Contexto:** os resultados apresentados partiram da classificação de postagens de dois SS: Twitter e SIGAA. Estudos futuros, com outros sistemas com características diferentes, poderão produzir resultados diferentes.

<sup>48</sup> As palavras mais frequentes estão no topo. Apenas os resultados mais significativos são listados.

**Abordagem *bag-of-words*:** como mencionado, cada palavra encontrada torna-se um atributo para classificação em mineração. Porém, são ignoradas a ordem das palavras ou as palavras ao redor, informação de pontuação ou estrutural, e assume-se que a ocorrência das palavras em um documento é independente (ROSSI; REZENDE, 2011). Rossi e Rezende observam que um problema encontrado ao utilizar palavras simples como atributos dos documentos textuais é que conceitos do mundo real muitas vezes são compostos por duas ou mais palavras, e gerar atributos compostos por mais de uma palavra pode ser útil para o processo de mineração de textos. Por exemplo, observou-se que grande parte das não-PRUs continha a expressão: “‘*Bom dia*’ *Twitter*”, que tem um significado diferente de “*O Twitter é bom*”. A expressão “*Bom dia*”, se geralmente ocorre conjuntamente, seria colocada como um atributo. Por outro lado, alguns autores afirmam que resultados experimentais mostraram que representações mais sofisticadas perdem em desempenho com relação à abordagem *bag-of-words* (APTÉ; DAMERAU; WEISS, 1994; DUMAIS; PLATT; HECKERMAN; SAHAMI, 1998; LEWIS, 1992 apud MARTINS, 2003).

#### **Técnicas de pré-processamento utilizadas:**

- a) ***Lemmatizer x stemming*:** nesse experimento, optou-se por utilizar a técnica de *Lemmatizer* em vez de *stemming*. No entanto, outros experimentos foram realizados, com os mesmos dados, porém usando *stemming*. Os resultados de aprendizado pelos algoritmos foram melhores (91%) (Figura 28). No entanto, era impossível distinguir algumas categorias gramaticais importantes para os padrões de extração. Como exemplo, tem-se o seguinte termo “*segu*” (Figura 29), trata-se de um verbo? (“seguir”), substantivo (“seguidores”) para se referir à funcionalidade do Twitter? ou o adjetivo (“seguinte”)? Como o objetivo foi encontrar padrões relacionados às categorias, inclusive de categorias verbais, foi utilizado o lemma, pois ele conserva a classe gramatical.

Figura 28 – Resultado da mineração do Twitter, algoritmo J48, usando stemming.  
Cross Validation (10 folds)

```

Time taken to build model: 2.09 seconds

=== Stratified cross-validation ===
=== Summary ===

```

Correctly Classified Instances	1328	91.4601 %
Incorrectly Classified Instances	124	8.5399 %

```

Kappa statistic          0.8292
Mean absolute error      0.118
Root mean squared error  0.2736
Relative absolute error  23.5987 %
Root relative squared error 54.7287 %
Total Number of Instances 1452

=== Detailed Accuracy By Class ===

```

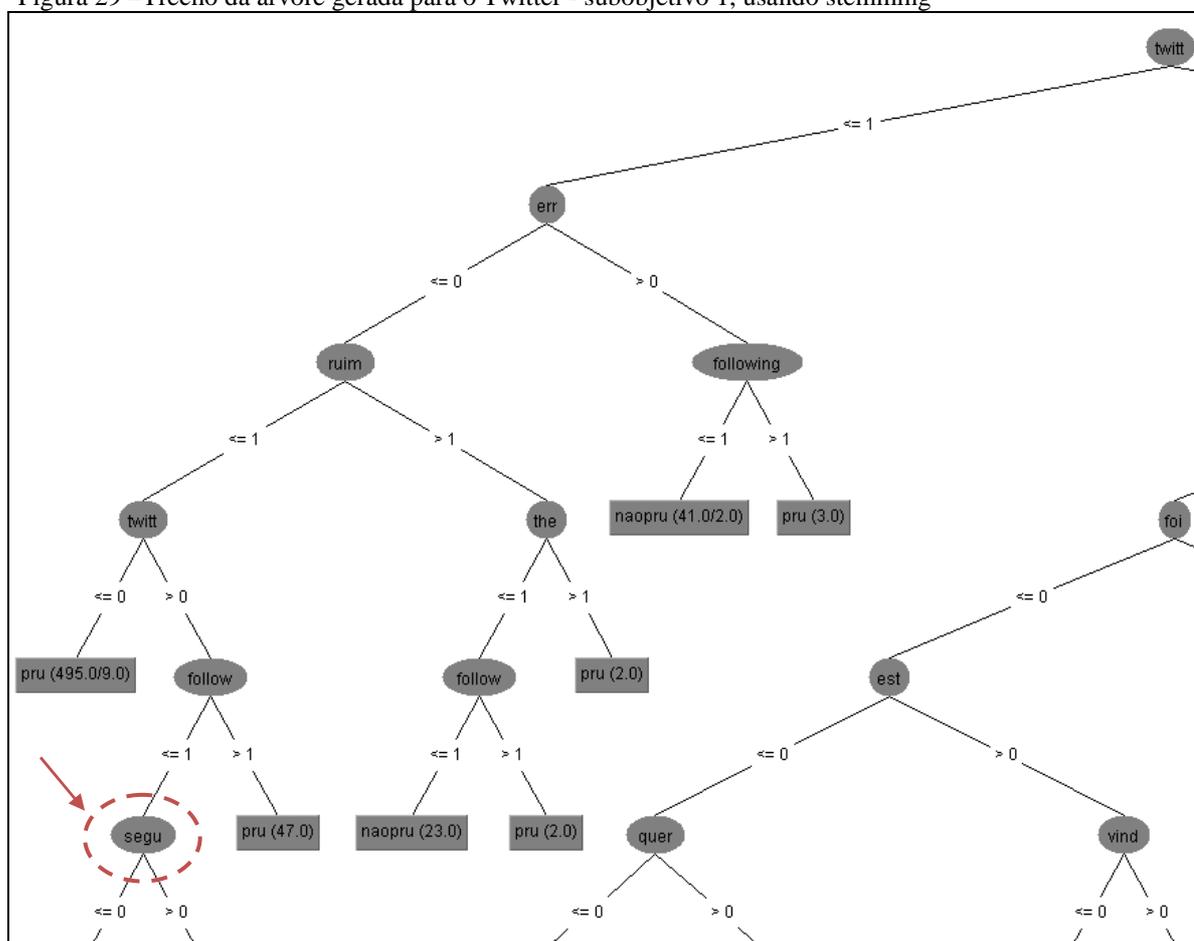
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.904	0.074	0.924	0.904	0.914	0.927	pru
	0.926	0.096	0.906	0.926	0.916	0.927	naopru
Weighted Avg.	0.915	0.085	0.915	0.915	0.915	0.927	

```

=== Confusion Matrix ===
 a  b  <-- classified as
656 70 | a = pru
 54 672 | b = naopru

```

Figura 29 – Trecho da árvore gerada para o Twitter - subobjetivo 1, usando stemming



- a) ***Não uso da técnica Thesaurus***: como o objetivo foi extrair novos padrões a partir dos textos classificados, decidiu-se não padronizar termos com vocabulário de sinônimos. No entanto, percebe-se uma possibilidade de como essa técnica poderia ter sido útil. Na troca de termos, como: “hahaha”, “rsrsrs”, “hehehe”, “kkkkk”, “huahuhau” por “risadas”, por exemplo. Ou outros termos representativos de emoções como “grrr”, “aaahh” e até mesmo uma lista de palavras pela palavra: “palavrão”. Tais elementos foram excluídos nessa análise e poderiam ter sido úteis na verificação de que em que categorias e metas eles seriam mais frequentes;
- b) **grafia incorreta**: embora tenha sido utilizado tratamento para letras duplicadas no processo de limpeza dos textos, ainda foram encontradas palavras grafadas de forma diferente, como: “liiindo”, “boom”, “facio”, “mecher”, “kerer”, “klicar”, “face”, “facebok” etc. Tais palavras diminuem sua frequência, pois concorrem com as palavras de grafia correta. “facio” e “fácil” são dois atributos diferentes;
- c) **metas de UUX**: o volume de dados para essa categoria foi obtido a partir de 1.150 PRUs classificadas. No entanto, foi insuficiente para estabelecer padrões para todas as metas de UUX avaliadas. Metas como antecipação, encantamento, engajamento, impacto, memorização e prazer, não foram mineradas.

**Padrões relacionados:** O trabalho de Hedegaard e Simonsen (2013) apresentou os seguintes padrões relacionados às metas de UUX (Figura 30). Os dados foram extraídos a partir de revisões de *video games* e *softwares*, minerados com a abordagem *bag-of-words* e o algoritmo *Support Vector Machine (SVM)*. Apesar do contexto distinto desta tese, pode-se observar algumas similaridades, como por exemplo, na meta aprendizado, padrões como: “fácil”, “aprender” e elementos da interface, como “menu”. Na meta satisfação, padrões como: “gostar”, “legal”, “amar”, “bonito”. Na meta de estética, foram encontrados padrões relacionados a elementos da tela e adjetivos caracterizando tais elementos. Os padrões fornecidos nesta mineração também podem contribuir para o contexto do estudo das metas de UUX.

Figura 30 – Padrões de UUX para o contexto de jogos de video game e softwares

Learnability	Errors/effectiveness	Satisfaction	Hedonic	Detailed usability	Pleasure	Affect and emotion	Enjoyment, Fun	Aesthetics, Appeal	Engagement, Flow	Frustration
intuit	glitch	great	fun	realli	fun	fun	fun	graphic	challeng	frustrat
easier	issu	love	enjoy	great	enjoy	enjoy	enjoy	sound	addict	incompatibilit
learn	lag	realli	frustrat	best	bore	bore	bore	music	replay	hardest
figur	camera	worth	annoy	nice	love	frustrat	entertain	realist	difficulti	perpetu
easiest	imprecis	nice	bore	worth	entertain	annoy	funni	voic	hour	insult
eas	bump	best	funni	problem	annoy	entertain	love	soundtrack	difficult	dissadvantag
straightforward	bug	graphic	love	love	felt	love	amus	effect	depth	needless
easi	configur	sound	entertain	overall	sooth	amus	humor	beauti	harder	nerv
sc	suspect	fun	humor	graphic	lighter	laugh	laugh	anim	moment	fuel
foreword	error	overall	hate	easi	workout	scari	excit	look	complex	afterward
practic	flaw	problem	sooth	recommend	grin	excit	sooth	environ	nonstop	grin
simpl	crowd	felt	issu	frustrat	addict	addict	grin	vivid	tough	needless
angl	dodg	disappoint	hum	pretti	humor	humor	lighter	audio	easi	la
menus	biowar	definit	lighter	bad	chore	sooth	kinda	atmosph	valu	unfair
steam	ai	favorit	cute	disappoint	incompatibilit	felt	fell	color	deep	plain
sacr	troubl	good	grin	fun	fell	fell	hilari	visual	therefor	flat
plasmid	semblanc	bad	excit	sound	rooftop	grin	workout	sceneri	keep	grow
experiment	mater	price	catchi	good	nevertheless	lighter	nevertheless	sprite	imposs	fusion
password	inconsist	interest	addict	learn	laugh	chore	zero	impress	engag	cheat
straight	suffer	cool	workout	favorit	afterward	truli	intrigu	render	tire	melodramat
thus	data	fan	chore	definit	told	engag	rooftop	model	hard	habit
minut	confus	improv	tens	better	intrigu	tens	shatter	cute	intens	gasp
master	technic	fantast	incompatibilit	feel	nostalgia	kinda	told	detail	sc	heck
smooth	prompt	better	fell	price	shatter	incompatibilit	scare	appeal	painstak	annoy
nunchuk	resolnt	perfect	afterward	improv	regardless	hilari	younger	realism	lenient	babysit
incred	respons	unfortun	nostalgia	interest	perpetu	workout	im	pixel	hardest	insan
sensor	load	lack	nevertheless	cool	moneybag	createur	cute	bright	gripper	vito
gps	primit	pretti	stagger	would	everytim	cute	queue	stun	becom	slog
casual	precis	feel	intrigu	perfect	adict	nostalgia	tedious	creepi	interest	scaletta
applet	delay	qualiti	regret	qualiti	countless	nevertheless	jeremi	hear	most	flavor

Fonte: Hedegaard; Simonsen (2013)

## 5.5 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou o processo de mineração de dados realizado com as postagens dos usuários. O objetivo da de mineração de dados foi obter padrões úteis relacionados às PRUs classificadas. Especificamente, com esse experimento, foi possível:

- validar (e obter novos) padrões de extração de PRUs;
- obter padrões para cada tipo de PRU; e
- obter padrões para cada meta de UUX.

Os padrões fornecidos nesta mineração irão auxiliar a classificação de sentenças de forma manual e possibilitar a extração e/ou classificação automática de sentenças.

## **6 MALTU - MODELO PARA AVALIAÇÃO DA INTERAÇÃO EM SISTEMAS SOCIAIS A PARTIR DA LINGUAGEM TEXTUAL DO USUÁRIO**

Neste capítulo, é apresentado o modelo MALTU e uma metodologia para a aplicação do modelo, com exemplos de como classificar as postagens dos usuários a fim de obter um resultado de avaliação. Para avaliação automática é apresentada a ferramenta UUX-POST. Em seguida, é realizada uma análise sobre limitações do modelo, discutindo possíveis problemas e soluções. Por fim, são apresentados os resultados de avaliação do Twitter e SIGAA, seguindo o modelo apresentado.

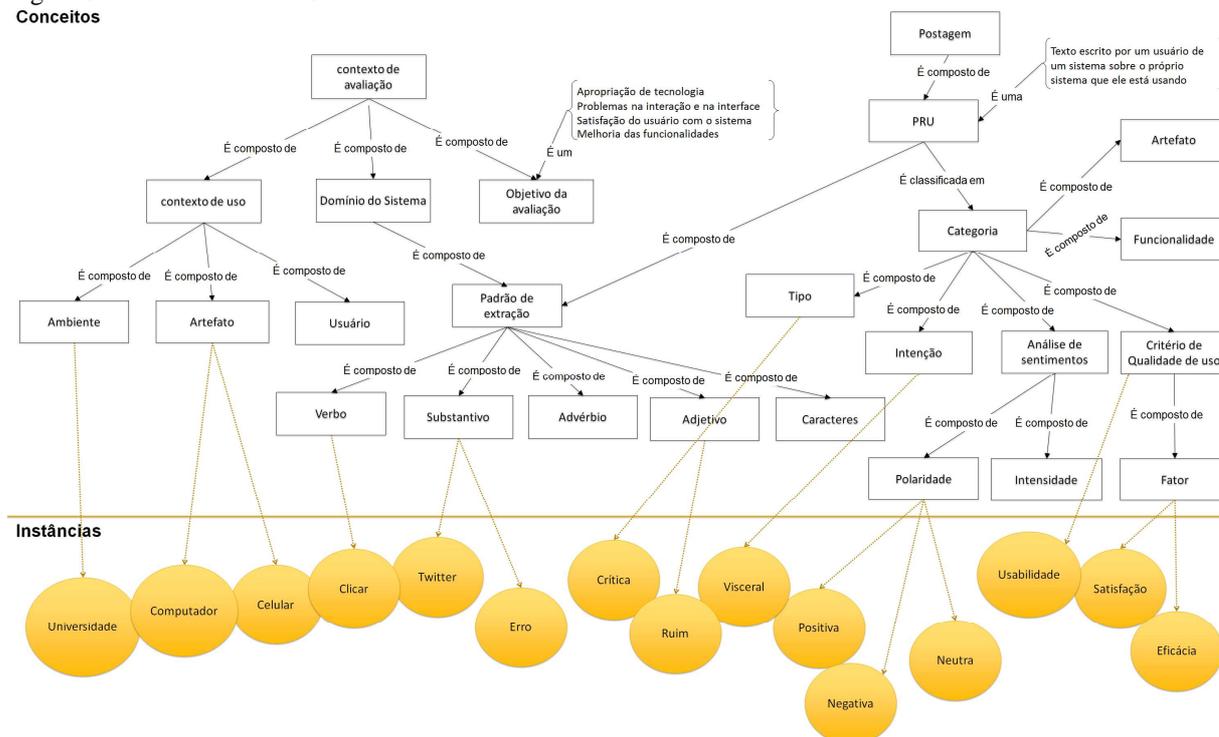
### **6.1 MODELO MALTU**

A questão que o modelo procura responder por meio de seus componentes é: como avaliar? O MALTU – Modelo para Avaliação da interação em SS a partir da Linguagem Textual do Usuário apresenta uma visão geral da avaliação a partir de textos espontâneos relacionados ao uso do sistema postados por usuários em SS.

A Figura 31 ilustra o modelo MALTU. Ele é dividido em duas camadas. A primeira ilustra os conceitos referentes à avaliação, representados com a notação de ontologia. Uma ontologia descreve os conceitos de um domínio e, também, as relações que existem entre estes conceitos (GRUBER, 2005 apud MORAES; AMBRÓSIO, 2007). A segunda camada apresenta algumas instâncias dos conceitos apresentados.

As atividades do avaliador compõem uma metodologia de avaliação de um SS a partir de um conjunto de PRUs (Figura 32). O modelo será explicado, conjuntamente com a metodologia, na próxima seção.

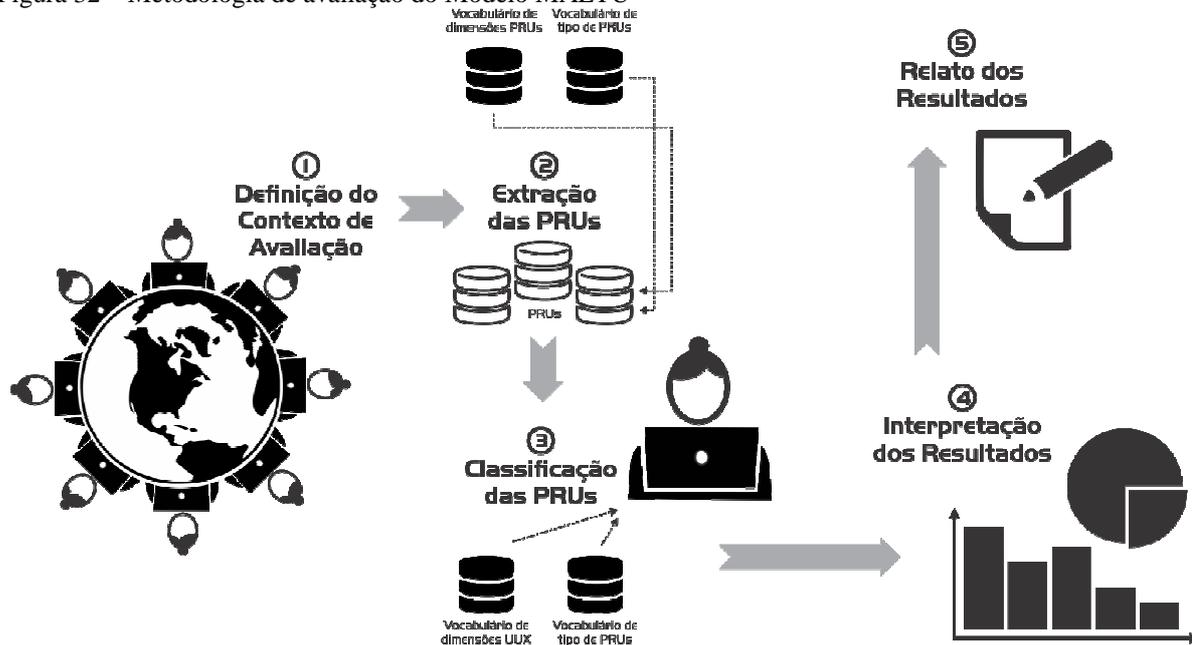
Figura 31 – Modelo MALTU  
Conceitos



## 6.2 METODOLOGIA

A metodologia de avaliação do modelo MALTU (Figura 32) tem como objetivo guiar um profissional de IHC na avaliação de um SS a partir de um conjunto de PRUs. A metodologia explica as cinco atividades de um avaliador, representadas em etapas sequenciadas para avaliação, são elas: (1) definição do contexto de avaliação; (2) extração das PRUs; (3) classificação das PRUs; (4) interpretação dos resultados e (5) relato dos resultados. Tais etapas serão explicadas, detalhadamente, nas próximas seções.

Figura 32 – Metodologia de avaliação do Modelo MALTU



### 6.2.1 Contexto de avaliação

A definição do contexto de avaliação no MALTU envolve definir: 1) o contexto de uso do sistema; 2) o domínio do sistema; e 3) os objetivos de avaliação. Na definição do contexto de uso, devem ser definidos os seguintes elementos: usuários, plataforma e ambiente.

Por se tratar de uma nova forma de avaliação, o modelo foi constituído a partir dos estudos apresentados nesta tese, com o acréscimo da definição de contexto de uso de Furtado (2012), que caracteriza o contexto de uso de um sistema interativo com a presença de três componentes: usuários, plataforma e ambiente:

O usuário é o indivíduo que interage com o sistema interativo realizando tarefas para atingir objetivos variados; a plataforma corresponde ao ambiente computacional, composto por dispositivos disponíveis (como computador, celular, TV digital) e de recursos de hardware (como monitor, mouse, controle remoto); e o ambiente se refere às propriedades do ambiente real, sendo relativas às condições do ambiente físico, como iluminação, tipo de barulho etc. (FURTADO, 2012, p 14).

Esta definição assemelha-se à proposta pela ISO 9241-11 (2011), que acrescenta, também ao ambiente, o contexto social. Para Dey (2001) apud Barbosa e Silva (2010), o contexto de uso é caracterizado por toda situação do usuário relevante a sua interação com o sistema, incluindo o momento de utilização do sistema (quando) e o ambiente físico, social e cultural em que ocorre a interação (onde).

Com o modelo MALTU não é possível identificar o ambiente físico do usuário, tendo em vista que SS são acessados de diversos dispositivos e a partir de lugares distintos. No entanto, é possível definir o ambiente social e cultural do usuário.

A definição do usuário envolve coletar características dos usuários do SS a ser avaliado, como idade, sexo, formação acadêmica, tempo de uso no sistema etc. Uma técnica recomendada para representação de um grupo de usuários reais é *personas* (PRUITT e GRUDIN, 2003). Uma série de trabalhos (MADEIRA et al., 2008; MADEIRA, 2010; NÓBREGA, 2011; FURTADO, MADEIRA e NÓBREGA, 2012; FURTADO, MILITAO e NÓBREGA, 2012) tem aplicado esta técnica para a modelagem de usuários em SS.

A definição da plataforma (artefato<sup>49</sup>) envolve definir a partir de qual(is) dispositivo(s) o usuário pode acessar o SS. Se é um sistema que permite ser acessado por vários dispositivos, eles deverão ser listados. Para a definição do Ambiente, devem ser considerados os seus aspectos sociais e culturais. Segundo a ISO 9241-210 (2008), os aspectos sociais e culturais do ambiente incluem fatores como práticas de trabalho, estrutura organizacional e atitudes.

Para definição do domínio do sistema a ser avaliado, devem-se definir prováveis termos, os quais seus usuários utilizam para falar do sistema, assim como das suas principais funcionalidades. Tais informações serão utilizadas na composição dos padrões de extração de PRUs. Segundo Barbosa e Silva (2010), raramente o sistema inteiro é avaliado, sendo necessário definir o escopo da avaliação, delimitando quais partes da interface devem fazer parte da avaliação. Desta forma, informações sobre o domínio do sistema também devem ser definidas a fim de extrair PRUs relacionadas ao escopo da avaliação a ser realizada.

Segundo (HIX e HARTSON, 1993; RUBIN, 1994; MACK e NIELSEN, 1994; SHARP et al., 2007) apud (BARBOSA e SILVA, 2010), os principais aspectos avaliados em sistemas são: avaliar a apropriação de tecnologia pelos usuários, ideias e alternativas de design, conformidade com um padrão e problemas na interação e na interface. Com o modelo MALTU é possível avaliar a apropriação de tecnologia e problemas na interação e na interface. Além destes, outros objetivos podem ser avaliados: satisfação dos usuários com o sistema e sugestão de novas funcionalidades. O Quadro 31, adaptado de Barbosa e Silva (2010), apresenta os objetivos de avaliação e perguntas sugeridas para cada objetivo de avaliação usando o modelo MALTU.

---

<sup>49</sup> Alguns autores utilizam o termo artefato para representar o meio de utilização do sistema, como plataforma computacional, internet, computador, celular etc. (FUKS et al., 2011). O MALTU adota este termo.

Quadro 31 - Objetivos de avaliação usando o MALTU

Objetivos	Perguntas
Avaliar a apropriação de tecnologia	<p>O quanto eles estão motivados para explorar novas funcionalidades?            Quais são os pontos fortes de fracos do sistema, na opinião dos usuários?            Quais os objetivos dos usuários podem ser alcançados através do sistema? E quais não podem? Quais necessidades e desejos foram ou não atendidos?            A tecnologia disponível pode oferecer maneiras mais interessantes ou eficientes de os usuários atingirem seus objetivos?            O que é possível modificar no sistema interativo para adequá-lo melhor ao ambiente de trabalho?            Por que os usuários não incorporaram o sistema no seu cotidiano?</p>
Identificar problemas na interação e na interface	<p>Em sua maioria, os usuários conseguem operar o sistema? Eles atingem seu objetivo com frequência? Com quanta eficiência? Quantos erros?            Que parte da interface e da interação deixam os usuários insatisfeitos?            Que parte da interface os desmotiva a explorar novas funcionalidades?            Que problemas de IHC dificultam ou impedem o usuário de alcançar seus objetivos?            Onde esses problemas se manifestam? Com que frequência tendem a ocorrer? Qual é a gravidade desses problemas?            Quais barreiras o usuário encontra para atingir seus objetivos?            Ele tem acesso a todas as informações oferecidas pelo sistema?</p>
Avaliar a satisfação dos usuários com o sistema	<p>Qual o nível de satisfação dos usuários com o sistema? O quanto eles estão satisfeitos? Qual seu motivo de satisfação/insatisfação?</p>
Identificar a sugestão de novas funcionalidades	<p>O que é possível modificar no sistema interativo para adequá-lo melhor ao ambiente de trabalho? Quais são os desejos e necessidades dos usuários?</p>

Fonte : Barbosa; Silva (2010). Adaptado pela autora desta tese

### 6.2.2 Extração de PRUs

Essa etapa consiste na obtenção de PRUs do sistema a ser avaliado. Essa obtenção pode ser realizada de duas formas: 1) extração manual ou 2) extração automática. Tais opções têm como apoio os padrões de extração fornecidos neste trabalho (Quadro 28).

A extração manual, tanto pode ser realizada com a busca de PRUs a partir dos padrões fornecidos em campos de busca do sistema a ser avaliado, como a partir de uma base de dados com postagens dos usuários. A extração automática consiste na extração de PRUs com apoio de ferramentas. Este trabalho apresenta uma ferramenta, denominada UUX-Post, que realiza a extração de PRUs a partir de postagens públicas dos usuários, tanto em SS, quanto a partir de uma base de dados. Essa ferramenta será apresentada na seção 6.3.

Embora os padrões de extração sugeridos neste trabalho sejam resultados de um processo bem trabalhado de classificação, ainda assim, pode ser necessária uma classificação manual de PRUs. Alguns exemplos de postagens classificadas nesta categoria (PRUs e não-PRUs) estão apresentados nos Quadros 12 e 14.

Após a extração, deve-se seguir para a próxima etapa: classificação das PRUs.

### 6.2.3 Classificação das PRUs

A atividade de classificação das PRUs envolve classificar uma PRU em diferentes categorias de classificação sugeridas pelo MALTU, que são: a) tipo; b) intenção; c) análise de sentimentos; d) funcionalidade; e) critérios de qualidade de uso; e f) artefato. Essas categorias são descritas a seguir:

#### a) Classificação por tipo:

Uma PRU pode ser classificada nos seguintes tipos: crítica, elogio, ajuda, dúvida, comparação e sugestão, conforme características apresentadas no Quadro 32. Exemplos de postagens classificadas nessa categoria estão apresentados nos Quadros 13 e 15.

Quadro 32 – Descrição de características, categoria: tipos de PRUs

<b>Tipos de PRUs</b>	<b>Características</b>
Crítica	Contém reclamação, erro, problema ou comentário negativo em relação ao sistema.
Elogio	Contém elogio ou comentário positivo sobre o sistema.
Dúvida	Contém dúvida ou questionamento sobre o sistema ou suas funcionalidades.
Comparação	Contém comparação com outro sistema.
Sugestão	Contém sugestão sobre uma mudança no sistema.
Ajuda	Contém passos para realizar uma atividade no sistema.

Para essa categoria, este trabalho forneceu padrões para tipos de PRUs, que auxiliam a classificação manual e possibilitam a extração e/ou a classificação automática a partir de ferramentas. Os padrões estão disponíveis no Quadro 29.

#### b) Classificação por intenção do usuário:

A classificação por intenção tem como objetivo classificar as PRUs de acordo com a intenção do usuário sobre o sistema, que são: visceral, comportamental e reflexiva, conforme características apresentadas no Quadro 33. Exemplos de postagens classificadas nesta categoria estão exemplificados no Quadro 19. Alguns detalhes desta classificação são apresentados nos próximos parágrafos.

Quadro 33 – Descrição de características, categoria: classificação por intenção

Classificação por intenção	Características
Visceral	Não apresenta detalhes, nem referências a funcionalidades do sistema. Alta intensidade de sentimento. Demais características: atração, primeiro impacto, exagero.
Comportamental	Contém detalhes de uso, problemas ou funcionalidades do sistema. Pouca ou nenhuma intensidade de sentimento. Demais características: ações, desempenho, função, uso.
Reflexiva	Faz referência a alguma funcionalidade do sistema, mas sem apresentar muitos detalhes do uso. Média intensidade de sentimentos. Demais características: afeto, lembrança, valor agregado, particularidades culturais ou individuais.

Uma PRU do tipo visceral é caracterizada por ser geralmente impulsiva, emotiva, com o intuito principal de criticar ou elogiar o sistema sem, no entanto, apresentar detalhes sobre ele. Uma PRU visceral é mais indicada para a classificação de metas UX. As Figuras 33 e 34 apresentam exemplos de PRUs viscerais, sendo que a primeira se trata de uma crítica e a segunda de um elogio. Em ambas não foi possível classificar a funcionalidade nem a meta de usabilidade. Quando não houver definição da causa, ela é caracterizada como visceral.

Figura 33 – PRU visceral do tipo crítica

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
Crítica	Visceral	-	-	Frustração

REALMENTE, PENSE NUM SISTEMA RUIM,  
AGORA ELEVE AO QUADRADO,  
MULTIPLIQUE POR MIL, E ASSIM AINDA FICA  
MUITO RUIM, O PIOR SISTEMA

Emoção

Figura 34– PRU visceral do tipo elogio

Emoção

Ahhh eu gosto mais do Twitter!!! As outras  
redes sociais nem ligo +!!!

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
Elogio	Visceral	-	-	Satisfação, encantamento

As PRUs comportamentais são mais centradas no uso do sistema. Possuem uma menor intensidade de sentimento e têm como objetivo principal: relatar precisamente o problema ou questionar soluções. Geralmente, seus tipos principais são: crítica, dúvida e ajuda. As Figuras 35 e 36 apresentam exemplos de PRUs comportamentais, sendo que a primeira se trata de

uma crítica e a segunda de uma dúvida. Uma PRU comportamental é a mais indicada para a classificação de metas de usabilidade.

Figura 35 – PRU comportamental do tipo crítica

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
Crítica	Comportamental	dar unfollow nos outros do nada	Segurança	Frustração

Causa  
 Síndrome da troca de tela do twitter.  
 Toda vez que vou atualizar a tela,  
 jogo o dedo pra esquerda e vai pra  
 aba 'descobrir' Grrrr  
 Emoção

Detalhes de uso

Figura 36 – PRU comportamental do tipo dúvida

Como fazer a rematrícula se o portal do discente o SIGAA não está aparecendo o menu discente?

Causa

Detalhes de uso

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
Dúvida	Comportamental	Fazer rematrícula, não está aparecendo o menu discente	Eficácia	-

As PRUs do tipo reflexiva possuem uma intensidade média de sentimento e são mais racionais, indicam valor agregado, motivação ou mesmo uma reflexão sobre o uso do sistema. Geralmente, seus tipos principais são: elogio, sugestão e comparação. As Figuras 37 e 38 apresentam exemplos de PRUs reflexivas, sendo que a primeira se trata de um elogio e a segunda de uma crítica. PRUs reflexivas, assim como as viscerais, são mais indicadas para a classificação de metas UX.

Figura 37 – PRU reflexiva do tipo elogio

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
Elogio	Reflexivo	interativo	-	Satisfação

Causa  
 Parece que o sistema está melhor agora do que no semestre passado, quando foi adotado pela primeira vez. Estou otimista quanto ao Sigaa, é bem interativo e tem tudo para ser, cada vez mais, uma boa ferramenta para toda a comunidade acadêmica.  
 Reflexão

Figura 38– PRU reflexiva do tipo crítica

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
crítica	Reflexiva	Fazer amigos	-	Frustração

Já tô achando que é mais facil, fazer amigos no } Reflexão  
 #Facebook do que no #Twitter =/ Emoção

Causa

### c) Análise de sentimentos:

Nesta categoria, apresentam-se duas formas de classificação que podem ser realizadas para analisar os sentimentos nas PRUs: por polaridade e por intensidade.

**Polaridade:** esta classificação tem como objetivo classificar uma PRU em positiva, neutra ou negativa, conforme as características apresentadas no Quadro 34.

Quadro 34 - Descrição de características, categoria: classificação de sentimentos por polaridade

Classificação de sentimentos por polaridade	Características
Positiva	Demonstra sentimento positivo
Neutra	Não demonstra nenhum tipo de sentimento
Negativa	Demonstra sentimento negativo

Tal tipo de classificação foi realizada, de forma manual, na subseção 4.3.2 e exemplos desta classificação são apresentados no Quadro 18. Apesar deste tipo de classificação não ser comumente realizada em avaliação de sistemas, sobretudo, de forma manual, tal forma de classificação de sentenças é largamente utilizada na área de análise de sentimentos, com o objetivo de obter a opinião positiva ou negativa de uma sentença. Diversos autores (BROOKE, 2009; KIM; HOVY, 2004; PANG; LEE; VAITHYANATHAN, 2002; TABOADA et al., 2011) têm focado na análise de produtos e serviços por meio dos comentários sobre estes. Taboada et al. (2011) apresentaram uma abordagem para extração de sentimentos em textos que classifica tanto a polaridade, quanto a intensidade de uma sentença. A utilização de abordagens similares a esta, para este tipo de classificação de PRUs, seria de grande auxílio.

Alguns autores (TUCH; TRUSELL; HORNBÆK, 2013; OLSSON; SALO, 2012; PARTALA; KALLINEN, 2012) fizeram uso de algoritmos para classificação de polaridade em sentenças para estudar a UX com produtos tecnológicos.

**Intensidade:** esta classificação tem como objetivo classificar o quanto de sentimento (positivo ou negativo) é expresso em uma PRU. Nos exemplos: “Estou gostando deste

sistema...” e “*Eu amo muito usar este sistema*”. O sentimento positivo observado é mais intenso na segunda PRU.

Neste trabalho não foram realizadas investigações com a intensidade de sentimentos em uma PRU. Na classificação por intenção, a intensidade do sentimento é observada em conjunto com outros fatores, como a descrição detalhada de um problema, reflexão etc. Desta forma, para a classificação individual de PRUs somente pela intensidade do sentimento é recomendada a classificação automática por meio de algoritmos de análise de sentimentos. O trabalho de Taboada et al. (2011), por exemplo, utiliza a seguinte abordagem: um dicionário de termos, no qual cada termo possui um valor positivo, negativo ou zero (se for neutro). Por exemplo, o termo “*monstruosidade*” possui valor “-5”, os termos “*ódio*” ou “*odiar*”, possuem valor “-4”, enquanto que “*satisfatório*” possui valor “1” e “*delicioso*”, valor “4”. Desta forma, o cálculo é realizado automaticamente nos termos das sentenças, calculando a intensidade de uma sentença.

#### d) Classificação por funcionalidade:

Conforme mencionado, as PRUs do tipo comportamental são aquelas que mais detalham o uso do sistema. Nestas, é possível classificar a funcionalidade referida pelo usuário. As Figuras 39, 40 e 41 apresentam exemplos de PRUs em que a funcionalidade está indicada.

Figura 39 – Exemplo de classificação de funcionalidade em uma PRU

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
Crítica	Comportamental	Função do twitter de só de encostar o dedo já favorita	Segurança	Frustração

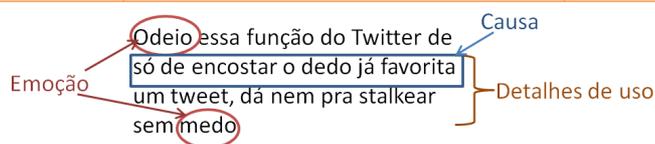


Figura 40 – Classificação de funcionalidade em uma PRU

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
Crítica	Comportamental	dar unfollow nos outros do nada	Segurança	Frustração

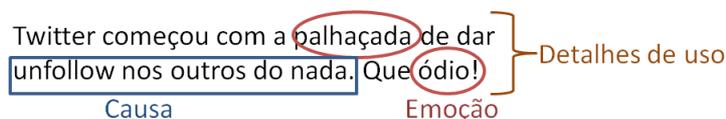
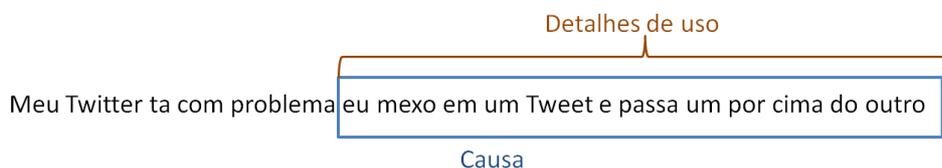


Figura 41– Classificação de funcionalidade em uma PRU

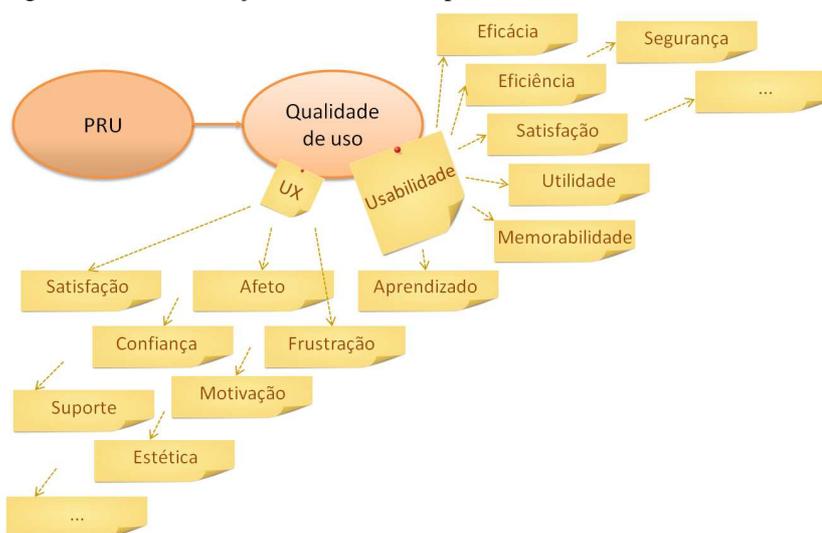
tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX
Crítica	Comportamental	Tweet	Eficácia	-



### e) Classificação por critérios de qualidade de uso:

Essa classificação envolve determinar os critérios de qualidade de uso usabilidade e/ou UX e suas metas a serem avaliadas. A Figura 42 ilustra a seleção de metas por critérios de qualidades de uso que são avaliadas no MALTU. As metas de cada critério de qualidade de uso escolhido também dependem do avaliador e do objetivo de avaliação.

Figura 42 – Classificação de uma PRU a partir de metas de UUX



Devido à dificuldade de classificação dos avaliadores ao utilizar muitas metas e metas relacionadas, este trabalho optou por sugerir metas mais distintas (Quadro 35). Outras sugestões são apresentadas no APÊNDICE C deste trabalho.

Quadro 35 – Sugestão de metas para UUX no modelo MALTU

Qualidades de uso	Metas	Referências
Usabilidade	Eficácia	ISO 9241-210 (2008); ISO 9241-11 (1998); PREECE, ROGERS; SHARP (2005); SHACKEL (1991); SHNEIDERMAN et al. (1998); NIELSEN (1993);
	Eficiência	ISO 9241-210 (2008); ISO 9241-11 (1998); PREECE; ROGERS; SHARP (2005); SHNEIDERMAN et al.

Qualidades de uso	Metas	Referências
		(1998); NIELSEN (1993); SHACKEL (1991);
	Satisfação	ISO 9241-210 (2008); ISO 9241-11 (1998); SHNEIDERMAN et al. (1998); NIELSEN (1993); SHACKEL (1991);
	Segurança	PREECE; ROGERS; SHARP (2005);
	Utilidade	PREECE; ROGERS; SHARP (2005);
	Memorabilidade	SHNEIDERMAN et al. (1998); PREECE; ROGERS; SHARP (2005); NIELSEN (1993); SHACKEL (1991);
	Aprendizado	SHNEIDERMAN et al. (1998); PREECE; ROGERS; SHARP (2005); SHACKEL (1991); NIELSEN (1993);
UX	Satisfação	PREECE; ROGERS; SHARP (2005);
	Afeto	BARGAS-AVILA; HORNBÆK (2011);
	Confiança	BEVAN (2008);
	Estética	BARGAS-AVILA; HORNBÆK (2011);
	Frustração	BARGAS-AVILA; HORNBÆK (2011);
	Motivação	BARGAS-AVILA; HORNBÆK (2011); KETOLA; ROTO (2008)
	Suporte	KETOLA; ROTO (2008)

Para essa categoria, este estudo forneceu padrões para classificação de metas de UUX (Quadro 30). Os padrões auxiliam a classificação manual e possibilitam a classificação automática, a partir de ferramentas de classificação. Alguns exemplos de PRUs classificadas nessa categoria são apresentados no Quadro 20.

#### e) Classificação por artefato:

Essa classificação consiste em identificar o dispositivo que o usuário está utilizando no momento. Geralmente, isso é feito na etapa de extração das postagens, quando é decidido a partir de que dispositivo as PRUs serão extraídas. No entanto, existem SS, como o Twitter e Facebook, por exemplo, em que as PRUs extraídas do sistema podem ser provenientes de dispositivos diversos. Nesse caso, é válido classificar o dispositivo mencionado pelo usuário nessa etapa de classificação. As Figuras 43 e 44 apresentam exemplos de classificação de artefato em uma PRU. Outros exemplos de classificação de PRUs são apresentados no Quadro 36.

Figura 43 – Classificação de artefato em uma PRU

tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX	Artefato
Elogio	Reflexivo	tweetcaster	Eficiência	Satisfação, Motivação	Celular

Emoção Gostei bem mais desse tweetcaster do q do twitter para Android, acho muito lerdo no meu celular o aplicativo oficial, esse é bem mais rápido

Dispositivo } Reflexão  
Causa

Figura 44 – Classificação de artefato em uma PRU

Essa <sup>Emoção</sup> <palavrão> que o @twitter inventou de } Detalhes de uso  
 Causa quebrar a linha é ruim pra ver via web, mas via }  
 celular é pior ainda. (MUITO PIOR) Parem com isso.  
 Dispositivo

Tipo	Classificação por intenção	Funcionalidade (causa)	Usabilidade	UX	Artefato
Crítica	Comportamental	Quebrar a linha	-	Frustração	Celular

Quadro 36– Exemplos de classificação de PRUs usando o modelo MALTU

PRU	Tipo	Classificação por intenção do usuário	Funcionalidade (causa)	Dispositivo	Usabilidade	UX
Gostei do novo formato da timeline do @twitter. Valoriza usabilidade: dá pra agir rápido c/ interações e isso valoriza meu tempo.	Elogio	Reflexiva	Novo formato da timeline do twitter	-	Utilidade	Satisfação
Gnt alguém sabe como compartilhar as fotos do insta no twitter? Já tentei de todo jeito e ta dando erro :c	Dúvida	Comportamental	Compartilhar as fotos do insta no twitter	-	Eficácia	Frustração
Ainda tenho esperança de algum dia o twitter atualizar e aparecer um botão de gravar áudio.. minha preguiça ta {palavrão}	Sugestão	Reflexiva	Botão de gravar áudio	-	Utilidade	Antecipação
Não consigo excluir disciplinas. Alguém sabe desvendar esse segredo? O sistema, aparentemente, exclui e quando imprimo o comprovante aparece a disciplina novamente.	Crítica	Comportamental	Excluir disciplinas	-	Eficácia	Frustração
Única coisa ruim do twitter pelo celular e que não da pra ver o trend topics. :(	Crítica	Comportamental	não da pra ver o trend topics.	Celular	Eficácia	Frustração

Este trabalho forneceu padrões para classificação de PRUs, por tipos de PRU, e metas de UUX. Além dos padrões possibilitarem a classificação automática, a partir de ferramentas de classificação, eles podem ser utilizados para extração das PRUs para cada classificação apresentada.

#### 6.2.4 Interpretação dos resultados

Segundo Barbosa e Silva (2010), em etapas de interpretação e consolidação de avaliações, os dados dos participantes devem ser organizados de modo a evidenciar as relações entre eles. Segundo os autores, para testes de usabilidade, por exemplo, consiste em testar hipóteses, descobrir tendências, comparar soluções alternativas e verificar se o sistema

atingiu as metas de usabilidade definidas no início do projeto. Em tais etapas, são geralmente utilizadas tabelas, gráficos, cálculos de média ou outro indicador relevante.

Kuniavsky (2003) apud Barbosa e Silva (2010) também recomenda que o avaliador categorize os problemas encontrados durante a interação, descrevendo categorias e o impacto imediato na usabilidade do sistema avaliado.

Seguindo essas recomendações, no MALTU, as medidas utilizadas pelo avaliador consistem na frequência dos resultados de classificação das PRUs e nos relacionamentos entre eles, como:

- a) a frequência de cada tipo de PRU;
- b) a frequência da classificação por intenção do usuário;
- c) a frequência por análise da polaridade e intensidade dos sentimentos;
- d) a frequência da causa (funcionalidade);
- e) a frequência de cada meta de qualidade de uso avaliada;
- f) o relacionamento entre a frequência da causa (funcionalidade) com o artefato;
- g) o relacionamento entre a frequência do tipo de PRU e a frequência da meta de qualidade de uso;
- h) relacionamento de todos os resultados ao objetivo da avaliação.

O Quadro 37 apresenta algumas sugestões de medidas para análise dos objetivos de avaliação usando o MALTU. Por exemplo, o avaliador, com o objetivo de realizar uma melhoria das funcionalidades poderá utilizar como medidas PRUs do tipo sugestão, como os resultados apresentados nos Quadros 46 e 51.

Quadro 37 - Sugestões de medidas para análise dos objetivos de avaliação usando o MALTU

<b>Objetivos</b>	<b>Medidas</b>
Avaliar a apropriação de tecnologia	Frequência da meta Motivação Frequência da meta Satisfação Metas de UX x Funcionalidade Tipo Crítica x Funcionalidade Tipo Elogio x Funcionalidade Tipo Comparação x Funcionalidade
Problemas na interação e na interface	Metas de Usabilidade x Funcionalidade Metas de UX x Funcionalidade Tipo Crítica x Funcionalidades Tipo Dúvida x Funcionalidade Crítica x Funcionalidade
Satisfação do usuário com o sistema	Metas de UX x funcionalidade Frequência da intenção do usuário Frequência da polaridade e intensidade de sentimentos do usuário
Melhoria das funcionalidades	Tipo Sugestão Tipo Sugestão x Funcionalidades Meta Utilidade x Tipo Elogio x Funcionalidades

### 6.2.5 Relato dos resultados

A atividade de relatar os resultados no MALTU consiste em gerar um relatório com as seguintes informações:

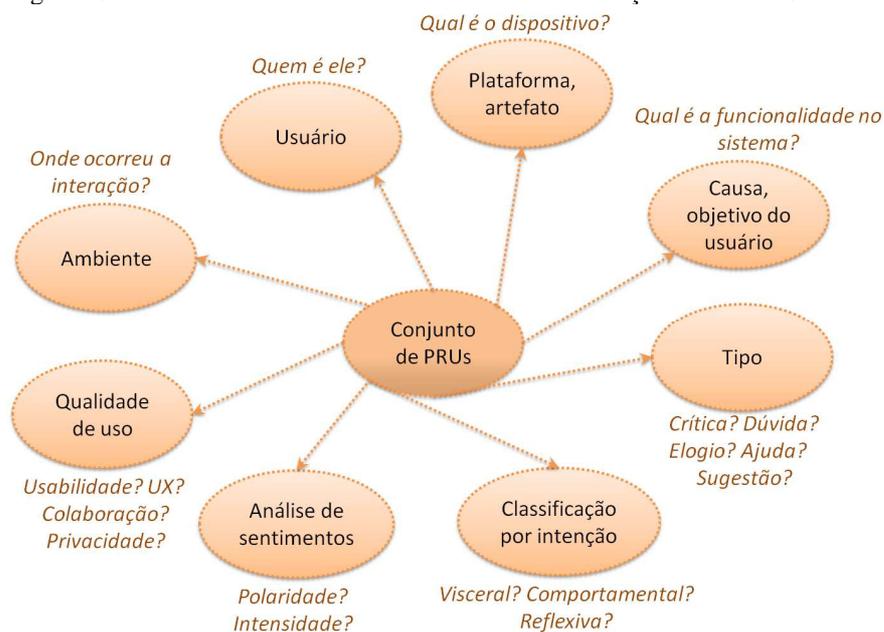
- a) contexto de uso do sistema;
- b) os objetivos e escopo da avaliação;
- c) informação da forma de extração e classificação (manual ou automática);
- d) informações sobre os participantes da avaliação (avaliadores: quantidade, experiência etc.);
- e) informação do artefato utilizado;
- f) tabelas e gráficos com o resultado das classificações obtidas;
- g) uma lista das causas (funcionalidades), relacionadas às metas avaliadas.

O Quadro 38 faz um resumo da metodologia apresentada para o modelo MALTU e a Figura 45 apresenta os elementos considerados para avaliação de um SS usando o MALTU.

Quadro 38 - Atividades da metodologia do modelo MALTU

<b>Avaliação da linguagem textual do usuário</b>	
<b>Atividade</b>	<b>Tarefa</b>
Definir contexto de avaliação	<ul style="list-style-type: none"> <li>▪ Identificar contexto de uso (ambiente, usuário e artefato)</li> <li>▪ Identificar domínio do sistema</li> <li>▪ Identificar objetivos da avaliação</li> </ul>
Extração das PRUs	<ul style="list-style-type: none"> <li>▪ Extrair as PRUs</li> </ul>
Classificação das PRUs	<ul style="list-style-type: none"> <li>▪ Classificar as PRUs</li> </ul>
Interpretação dos Resultados	<ul style="list-style-type: none"> <li>▪ Listar os dados resultantes da classificação: tabelas, gráficos etc.</li> <li>▪ Inspeccionar resultados</li> </ul>
Relato dos resultados	<ul style="list-style-type: none"> <li>▪ Gerar um relatório com os resultados da avaliação</li> </ul>

Figura 45 - Elementos considerados no contexto de avaliação no MALTU



### 6.3 FERRAMENTA UUX-POSTs

Para auxiliar na extração e classificação das postagens em SS, uma ferramenta foi construída. Esta ferramenta, denominada UUX-Posts<sup>50</sup>, é destinada a estudiosos e profissionais da área de IHC e tem como objetivo apoiar a avaliação de SS, fornecendo opiniões do usuário a respeito do sistema em uso.

As postagens são coletadas de perfis públicos de SS, como o Twitter, Facebook ou outro similar. A ferramenta ainda permite o *upload* de um banco de postagens em formato csv. A busca (*default*) é realizada usando um conjunto de padrões de identificação de PRUs fornecidos neste trabalho. Futuramente, irá possibilitar uma busca avançada, no qual o próprio usuário poderá alterar e definir seus próprios padrões de busca.

#### 6.3.1 Estrutura da aplicação

A UUX-Posts foi desenvolvida com as linguagens PHP, Javascript e AJAX (*Asynchronous JavaScript* e XML). A ferramenta coleta apenas postagens de usuários cujos perfis são públicos. Os dados coletados do usuário são referentes à informação disponibilizada

<sup>50</sup> Disponível em: <<http://www.uxmarks.com/luqs.unifor.br/uuxposts/>>. Acesso em: 12 de janeiro de 2015

pelo usuário em SS, como idade, sexo e localização. A UUX-Posts não coleta imagens e, em postagens com nomes de usuários, os nomes são omitidos, adquirindo a tag {user}.

A Figura 46 apresenta a tela de extração de PRUs do UUX-Posts, e a Figura 47 apresenta busca no Twitter. A Figura 48 apresenta a busca em um banco de postagens csv e a Figura 49 apresenta o resultado de uma busca no Twitter.

Figura 46 – Tela de busca de postagens do UUX-Post

The screenshot shows the 'Consultar Posts' interface. On the left, there is a sidebar with the text 'UUX Post', 'Analisador de postagens de Sites de Redes Sociais', and a menu with 'O que é', 'Consultar Posts', and 'Publicações'. The main area has a large green rounded rectangle containing a search bar. Below the search bar, it says 'Ou enviar arquivo:' followed by a button 'Escolher arquivo' and the text 'Nenhum arquivo selecionado'. At the bottom of the green area are two buttons: 'Verificar' and 'Avançado'.

Figura 47 – Extração de PRUs em SS

This screenshot is identical to Figure 46, but the search bar now contains the text 'twitter.com'.

Figura 48– Extração de PRUs a partir de um banco de postagens

This screenshot is identical to Figure 46, but the search bar now contains the text 'BancoSIGAA3.csv'.

Figura 49 – Resultado de uma busca

**UUX Post** Total de postagens: 30

*Analizador de postagens de Sites de Redes Sociais*

[Voltar](#)

O que é  
Consultar Posts  
Publicações

**Postagens obtidas**

Esse botão de seguir aqui na timeline do Twitter é um inferno pra min  
botão mute eu te amo nessas horas <http://t.co/RYLJw9FTO5>

{User} É tão facil mas eu acho no meu FC N EXISTE ESSE BOTÃO  <http://t.co/FOhRMt75R5>

Sério, Twitter, coloquem uma merda de botão editar!

RT {User}: {User} É tão facil mas eu acho no meu FC N EXISTE ESSE BOTÃO  <http://t.co/FOhRMt75R5>

Outras funcionalidades deverão ser implementadas a partir do resultado deste trabalho, permitindo a classificação manual e classificação automática de tipos de PRUs, por intenção, análise de sentimentos e classificação por metas de UUX. Em tais classificações, serão utilizados os padrões obtidos na mineração de dados. Para a interpretação dos resultados, serão gerados gráficos e tabelas com o cruzamento das categorias de classificação. Tais funcionalidades propostas poderão ser desenvolvidas em trabalhos futuros provenientes destas tese (seção 7.2.2).

## 6.4 LIMITAÇÕES DO MODELO E ALGUMAS ESTRATÉGIAS DE SOLUÇÃO

Esta seção apresenta algumas limitações do modelo e formas de lidar com elas. Tais limitações se encaixam nas etapas de extração e classificação das postagens de forma automática, com a utilização de ferramentas computacionais.

Uma pessoa pode ler uma postagem de um usuário e rapidamente classificar se o usuário está satisfeito ou não com o sistema. Se é uma dúvida, crítica, sugestão ou qualquer outra classificação apresentada neste trabalho. Esta tese provou isto. No entanto, automatizar esta atividade envolve estudos mais aprofundados nas áreas de PLN e mineração de dados. Embora esta tese tenha apresentado trabalhos correlatos a estas áreas, bem como a área de análise de sentimentos, algumas estratégias devem ser adotadas para construção ou escolha de sistemas para este fim. Desta forma, nos próximos parágrafos são apresentados possíveis problemas e algumas estratégias de solução.

- a) Problema: erros de ortografia, ou de digitação diminuem qualidade da extração e classificação automática de PRUs;

Solução: aplicar ferramentas de correção ortográfica;

- b) Problema: variações de gênero, número e grau são comuns. As variações dificultam as análises. Por exemplo, as tags "sociais" e "social" são variações de número e, embora escritas de forma diferente, são equivalentes;
- Soluções: 1) aplicação da técnica Stemming, explicada no capítulo 2 desta tese; 2) aplicação da técnica de Orengo e Huyck (2001), que desenvolveram um algoritmo e uma ferramenta (RSLP) para extração de radicais de palavras da língua portuguesa. Uma grande vantagem da utilização desta ferramenta é a utilização de um dicionário externo e editável, contendo cerca de 32 mil palavras, possibilitando remanejar seu conteúdo ou mesmo aperfeiçoar a extração através das regras;
- c) Problema: sinônimos, termos distintos com o mesmo valor semântico, como bonito e belo pode reduzir a classificação automática por categorias;
- Solução: aplicação da técnica Thesaurus, explicada no capítulo 2 desta tese;
- d) Problema: termos irrelevantes com baixa frequência que não agregam valor na classificação;
- Soluções: 1) aplicação da técnica StopWords, explicada no capítulo 2 desta tese; 2) aplicação da matriz de frequência. A matriz de frequência é uma técnica de pré-processamento no qual é construída uma matriz de representação a partir de um conjunto de textos, com as linhas correspondendo às palavras selecionadas e as colunas correspondendo aos textos da coleção. Segundo Morais e Ambrósio (2007), à cada entrada desta matriz é atribuído o valor da frequência absoluta de cada palavra em cada texto. A frequência absoluta de cada palavra, em cada uma de suas entradas na matriz, é transformada em seu logaritmo. Isto é feito baseando-se no fato de que um documento com, por exemplo, três ocorrências de uma mesma palavra, tende a ser mais importante do que um documento com apenas uma ocorrência, porém não três vezes mais importante. Em seguida, cada um dos novos valores de entrada é dividido pelo somatório do produto destes valores pelo logaritmo dos mesmos, para salientar a sua importância. (MORAIS E AMBRÓSIO, 2007).
- e) Problema: Polissemias, palavras que expressam vários significados dificultam a compreensão, porque desambiguação é dependente do contexto. O termo “show”, por exemplo, pode se referir tanto a “apresentação” ou “espetáculo” em “O show do Vitoriano foi ótimo!” e a um adjetivo em outro, como: “A interface está show!”.

Solução: alguns autores (GOMES, 2009; FURTADO et al., 2012; PINHEIRO et al., 2009; PINHEIRO, 2010) têm apresentado soluções para resolver este problema. Pinheiro (2010), por exemplo, apresenta um Modelo Semântico Inferencialista (SIM) para expressão e raciocínio em Sistemas de Linguagem Natural. O modelo define os principais requisitos para expressar e manipular conhecimento semântico inferencialista de forma a capacitar os sistemas de linguagem natural para geração de premissas e conclusões de sentenças e textos. O modelo é usado como raciocinador semântico em um sistema de extração de informações sobre crimes (PINHEIRO, PEQUENO e FURTADO, 2010).

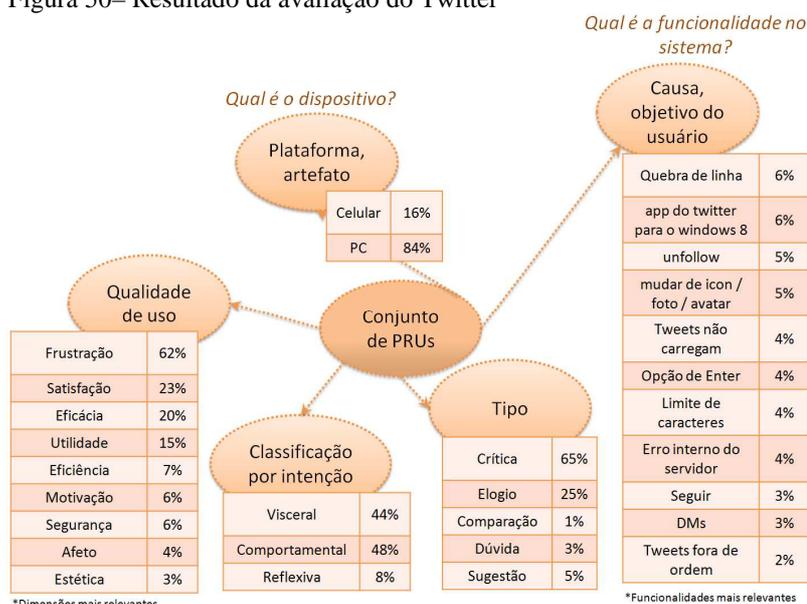
## 6.5 RESULTADO DE AVALIAÇÕES USANDO O MODELO MALTU

### 6.5.1 Resultado da avaliação do Twitter usando o modelo MALTU

Nesta seção será apresentado o resultado de avaliação do Twitter usando o MALTU. O objetivo foi avaliar os critérios de qualidade de uso UUX. A forma de extração foi automática e a classificação foi manual, resultante da investigação 4 (seção 4.4). Foram 350 PRUs classificadas por alunos de graduação e 150 por especialistas de IHC, totalizando 500 PRUs, corrigidas por dois especialistas de IHC.

A Figura 50 apresenta o resultado da avaliação do Twitter.

Figura 50– Resultado da avaliação do Twitter



Os gráficos e quadros apresentados a seguir, neste seção, apresentam o relacionamento das classificações obtidas, fornecendo uma visão geral do sistema avaliado. O Gráfico 9 ilustra as porcentagens obtidas em cada fator de usabilidade relacionado a PRUs do tipo crítica. O fator eficácia, por exemplo, obteve uma maior porcentagem (40%).

O Gráfico 10 ilustra as porcentagens obtidas em cada fator de UX relacionado a PRUs do tipo crítica. O fator frustração, por exemplo, obteve maior porcentagem (90%).

O Gráfico 11 ilustra as porcentagens obtidas em cada fator de UX relacionado a PRUs do tipo elogio. O fator encantamento, por exemplo, obteve maior porcentagem (46%).

O Quadro 39 apresenta as funcionalidades coletadas das PRUs do tipo crítica em cada fator de usabilidade. No fator eficiência, por exemplo, as funcionalidades (causas) mais citadas foram: “aplicativo muito pesado”, “demora para carregar tweets”, “não carrega mentions”, “resolução do avatar ruim”, “ruim para responder”, “travando”.

O Quadro 40 apresenta as porcentagens e funcionalidades coletadas das PRUs do tipo elogio em cada fator de usabilidade. A maior porcentagem não teve classificação de funcionalidade. A segunda maior porcentagem, fator satisfação, indica que os usuários estão satisfeitos com o Twitter pelos seguintes motivos: “interface bonita”, “desabafar”, “design”, “encontrar pessoas”, “ver novidades”.

O Quadro 41 apresenta as metas as funcionalidades coletadas das PRUs do tipo crítica em cada fator de UX. O fator frustração, por exemplo, apresenta uma maior quantidade de causas citadas nas PRUs.

O Quadro 42 apresenta as funcionalidades coletadas das PRUs por tipos quando o dispositivo é um celular. O resultado indica que o usuário tem dúvidas em “postar fotos”, mas acha a aplicação “bonita”, “carrega mais rápido” etc.

O Quadro 43 apresenta as principais funcionalidades que os usuários tiveram dúvidas e o Quadro 44 apresenta sugestões de funcionalidades para o sistema. De acordo com a sugestão de avaliação apresentada no Quadro 37, as PRUs do tipo dúvidas seriam formas de identificar problemas no sistema, enquanto as Prus do tipo sugestões seriam formas de possibilitar melhorias no sistema.

Gráfico 9 – Critério de qualidade de uso = usabilidade x tipo de PRU = crítica

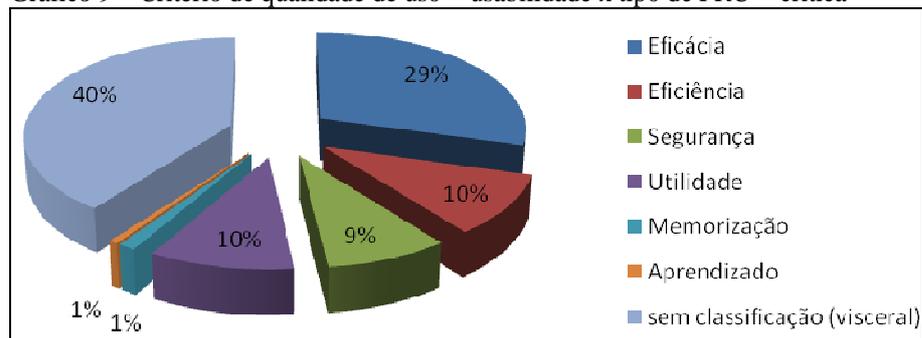


Gráfico 10 – Critério de qualidade de uso = UX x tipo de PRU = crítica

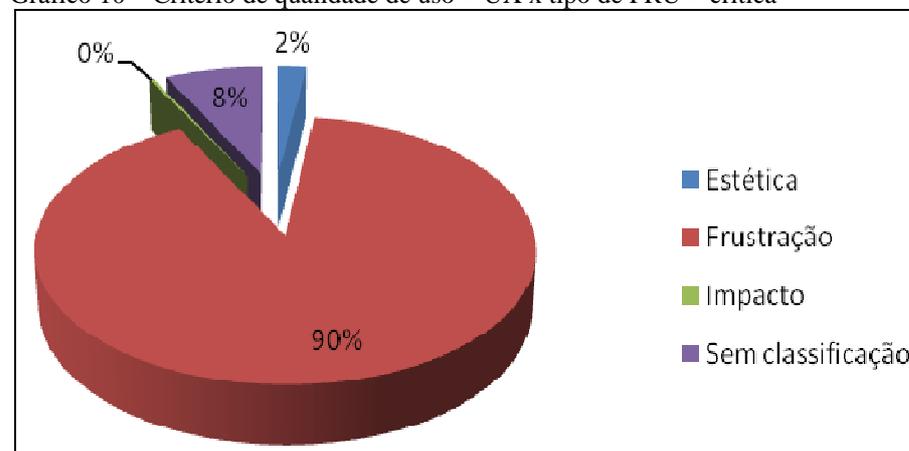
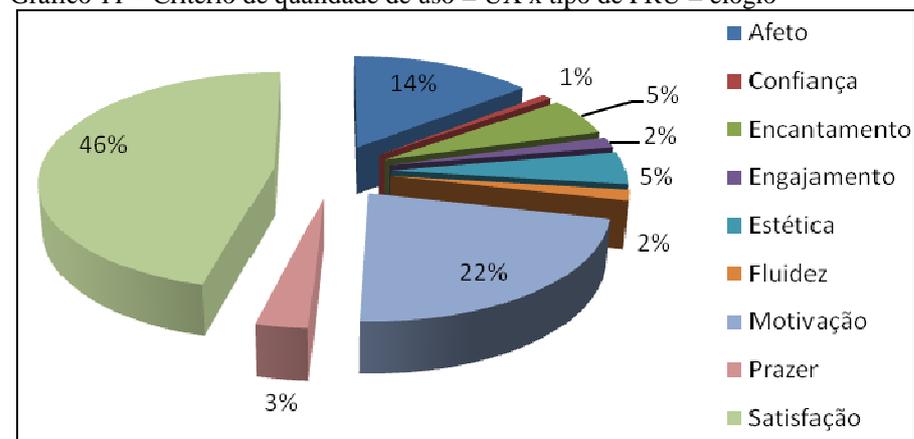


Gráfico 11 – Critério de qualidade de uso = UX x tipo de PRU = elogio



Quadro 39 – Critério de qualidade de uso = usabilidade x tipo de PRU = críticas x causa

<b>Critério de qualidade de uso = usabilidade x tipo de PRU = críticas x causa</b>	
<b>Metas de usabilidade</b>	<b>Causa (Funcionalidades)</b>
Eficácia	App – notificações que não existem, atualizar interface, O twitter me bloqueou, mudar capa/icon/avatar, DMs, editar <i>tweets</i> , erro interno do servidor, espaços, twitter fechando sozinho, twitter não atualiza a timeline, followers diminuindo, <i>tweets</i> fora de ordem
Eficiência	Aplicativo muito pesado, demora para carregar <i>tweets</i> , não carrega mentions, resolução do avatar ruim, ruim para responder, travando
Segurança	Erros quando atualiza a tela, bloqueio, Enter, errar <i>tweets</i> , espaços, excluir <i>tweets</i> , favoritar sem querer, follow sem querer, mudar status de privado, navegação, remover nome da lista, unfollow, responder, seguir, retweet
Utilidade	Barra de pesquisa, confirmação antes de seguir, editar <i>tweets</i> , emoticons, enter, Limite de caracteres, resolução do avatar, deveria melhorar a sugestão de contatos, retiraram a tradução automática
Memorização	Campo de texto do twitter, denunciar conta comprometida, design, interface, mudança de elementos da interface
Aprendizado	Elementos da interface

Quadro 40 – Critério de qualidade de uso = usabilidade x tipo de PRU = elogio x causa

<b>Critério de qualidade de uso = usabilidade x tipo de PRU = elogio x causa</b>		
<b>Porcentagem</b>	<b>Meta de usabilidade</b>	<b>Causa (Funcionalidades)</b>
1%	Aprendizado	mais fácil falar
2%	Eficiência	Fácil de escrever, carrega mais rápido
42%	Satisfação	interface bonita, desabafar, Design, encontrar pessoas, ver novidades
13%	Utilidade	bloquear, desabafar, dinâmico, Enter, digitar uma só vez, limite de caracteres, novo formato da timeline
43%	Sem classificação (visceral)	-

Quadro 41 – Critério de qualidade de uso = UX x tipo de PRU = crítica x causa

<b>Critério de qualidade de uso = UX x tipo de PRU = crítica x causa</b>	
<b>Meta de UX</b>	<b>Causa (Funcionalidades)</b>
Estética	Tudo na interface ficou grande demais, muito ruim escrever, Linhas diferentes, campo de texto do twitter
Frustração	Notificações que não existem, app androide, app windows 8, navegador opera, atualiza a tela, avatar, barra de pesquisa, bloqueio, capa, icon, carregar <i>tweets</i> , confirmação antes de seguir, denunciar conta comprometida, interface, digitação errada, DMs, Erro interno do servidor, espaços, excluir <i>tweets</i> , favoritar sem querer, ruim de fazer amigos, twitter fechando sozinho, não atualiza a tl, followers, Limite de caracteres, login, mudar a foto, capa, icon, não carrega a mentions, não consigo responder, não dá para ver o trend topics, não dá pra escrever muita coisa, não muda o status de privado, não segue e os <i>tweets</i> aparecem, navegação, pesado, propaganda na timeline, Quebra de linha, remover nome da lista, responder, seguir, só vejo alguns replies, Twitter lento, unfollow
Impacto	Mudanças de elementos na interface

Quadro 42 – Principais funcionalidades no celular

<b>Artefato = celular x tipos de PRUs x causa</b>		
<b>Porcentagem</b>	<b>Tipos de PRU</b>	<b>Causa (Funcionalidades)</b>
73%	Crítica	Aplicativo muito pesado, notificações que não existem, digitação errada, enter, favoritar sem querer, não dá para ver fotos, não carrega <i>tweets</i> , não dá para ver o trend topics, não tem como ver a timeline no iphone, Quebra de linha, ruim para responder, só vejo alguns replies, sumiu as letras, teclado ruim, travando, <i>tweets</i> compridos, <i>tweets</i> não enviados
25%	Elogio	Bonito, carrega mais rápido, bom, ver Twitter antigo sem precisa stalkear
1%	Dúvida	Postar fotos

Quadro 43 – Principais funcionalidades que os usuários tiveram dúvidas

<b>SS = Twitter; tipo de PRU = dúvida x funcionalidades</b>
Comentários direto na página do twitter, aviso sobre o recebimento de DMs, bloquear uma pessoa, compartilhar fotos do Instagram, denunciar conta comprometida, Favoritos, postar fotos, <i>tweets</i> Facebook

Quadro 44 – SS = Principais sugestões de funcionalidades

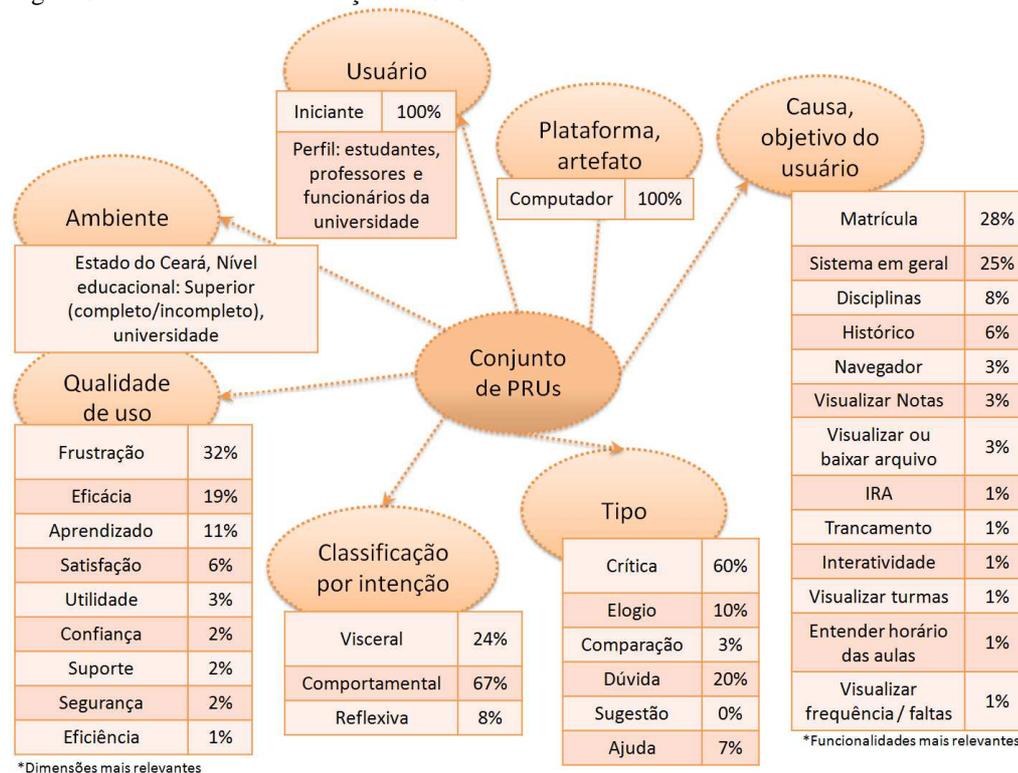
<b>SS = Twitter, tipo de PRU = sugestão x funcionalidades</b>
Bloquear palavras, botão curtir, botão de áudio, editar <i>tweets</i> , espaço para escrever devia ser maior, excluir contas que passam mais de um ano em entrar, layout, aumentar o limite de caracteres, lista de bloqueios igual ao Facebook, notificação antes de seguir, notificação de erros, quando bloquear um usuário, a conta ficar privada

### 6.5.2 Resultado da avaliação do SIGAA usando o modelo MALTU

Nesta seção, será apresentado o resultado de avaliação do SIGAA usando o MALTU. O objetivo foi avaliar a interação dos usuários usando os critérios de qualidade de uso UUX. A forma de extração foi automática e a classificação foi manual, resultante da investigação 4 (seção 4.4). Foram 500 PRUs classificadas por alunos de graduação e 150 por especialistas de IHC, totalizando 650 PRUs, corrigidas por dois especialistas de IHC.

A Figura 51 apresenta o resultado da avaliação do SIGAA.

Figura 51 – Resultado da avaliação do SIGAA



Os gráficos e quadros apresentados a seguir, nesta seção, apresentam o relacionamento entre as classificações obtidas, fornecendo uma visão geral do sistema avaliado. O Gráfico 12 ilustra as porcentagens obtidas em cada fator de usabilidade relacionado a PRUs do tipo crítica. O fator eficácia, por exemplo, obteve uma maior porcentagem (48%).

O Gráfico 13 ilustra as porcentagens obtidas em cada fator de UX relacionado a PRUs do tipo crítica. O fator frustração, por exemplo, obteve maior porcentagem (86%).

O Gráfico 14 ilustra as porcentagens obtidas em cada fator de UX relacionado a PRUs do tipo elogio. O fator confiança, por exemplo, obteve maior porcentagem (36%).

O Quadro 45 apresenta as funcionalidades coletadas das PRUs do tipo crítica em cada fator de usabilidade. No fator memorização, as críticas foram referentes a: “muita informação”, “como fazer a matrícula”, “visual”.

O Quadro 46 apresenta as porcentagens e funcionalidades coletadas das PRUs do tipo elogio em cada fator de usabilidade. A maior porcentagem, fator satisfação, indica que os usuários estão satisfeitos com o SIGAA pelos seguintes motivos: “comunicação”, “interação”, “beleza”, “novas funcionalidades”, “prático”, “sociável”.

O Quadro 47 apresenta as funcionalidades coletadas das PRUs do tipo crítica em cada fator de UX. O fator frustração, por exemplo, apresenta uma maior quantidade de causas

citadas nas PRUs. Os demais apresentam poucas funcionalidades, pois, pela análise realizada classificações de PRUs – UX, os usuários não apresentaram detalhes do sistema.

O Quadro 48 apresenta as principais funcionalidades que os usuários tiveram dúvidas e o Quadro 49 apresenta sugestões de funcionalidades para o sistema.

Gráfico 12 – Critério de qualidade de uso = usabilidade x tipo de PRU = crítica

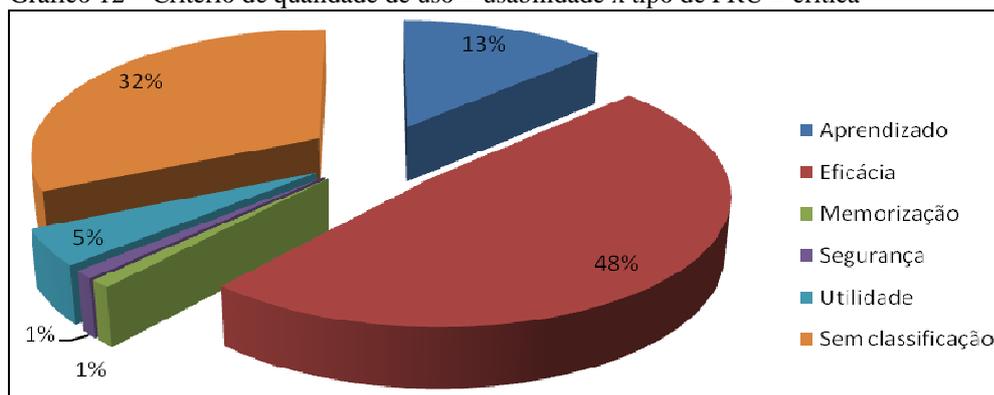


Gráfico 13 – Critério de qualidade de uso = UX x tipo de PRU = crítica

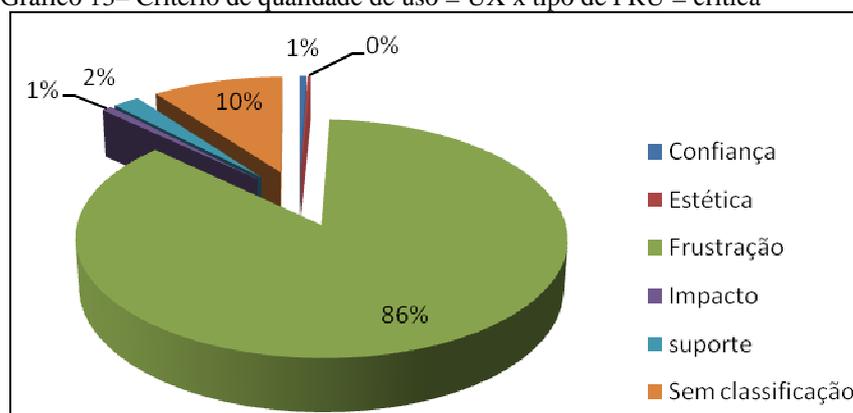
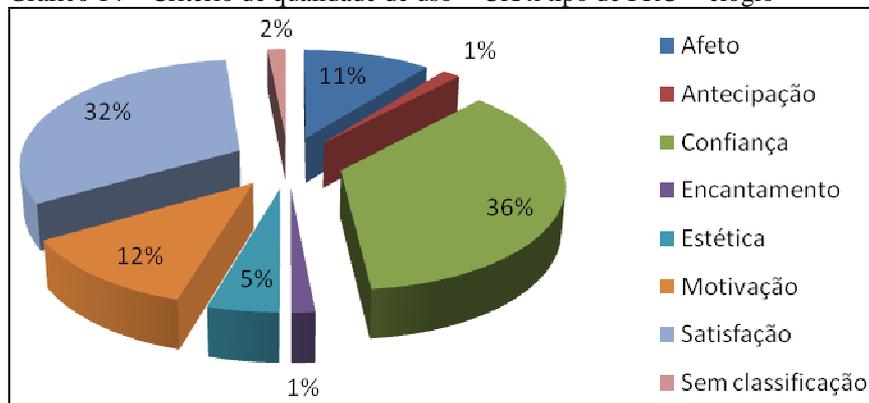


Gráfico 14 – Critério de qualidade de uso = UX x tipo de PRU = elogio



Quadro 45 – Critério de qualidade de uso = usabilidade x tipo de PRU = críticas x causa

<b>Critério de qualidade de uso = usabilidade x tipo de PRU = críticas x causa</b>	
<b>Metas de usabilidade</b>	<b>Causa (funcionalidades)</b>
Aprendizado	Visualizar, baixar ou inserir arquivo; visualizar ou erro nas disciplinas; editar informações; visualizar ou erro no histórico; erro no cálculo do IRA <sup>51</sup> ; realizar, visualizar ou erro na matrícula; visualizar notas, turmas, frequência ou faltas; realizar trancamento
Eficácia	Visualizar, baixar, abrir ou inserir arquivo; créditos a menos; visualizar ou erro no histórico; Realizar, visualizar ou erro na matrícula; visualizar notas ou horários; erro no cálculo do IRA; bloqueio no sistema, sistema em geral
Segurança	Realizar, visualizar ou erro na matrícula; visualizar ou erro nas disciplinas
Utilidade	Navegador; local da sala
Memorização	Muita informação; como fazer a matrícula; visual

Quadro 46 – Critério de qualidade de uso = usabilidade x tipo de PRU = elogio x causa

<b>Critério de qualidade de uso = usabilidade x tipo de PRU = elogio x causa</b>		
<b>Porcentagem</b>	<b>Meta de usabilidade</b>	<b>Causa (funcionalidades)</b>
2%	Aprendizado	Sistema em geral
5%	Eficácia	Realizar matrícula; Sistema em geral
64%	Satisfação	Comunicação; interação; Beleza; novas funcionalidades; prático; sociável
15%	Utilidade	Avisos; comunicação; interatividade; Foruns de discussão; sistema em geral
15%	Sem classificação (visceral)	-

Quadro 47 – Critério de qualidade de uso = UX x tipo de PRU = crítica x causa

<b>Critério de qualidade de uso = UX x tipo de PRU = crítica x causa</b>	
<b>Meta de UX</b>	<b>Causa (funcionalidades)</b>
Frustração	Menu indisponível; Visualizar, baixar ou inserir arquivo; Calendário; contabilização dos créditos; visualizar ou erro no histórico; Realizar matrícula; visualizar ou entender horários das disciplinas; acesso somente pelo navegador Firefox; erro no cálculo do IRA; visualizar turmas
Suporte	Realizar matrícula
Impacto	Sistema anterior
Confiança	Notas; Matrícula
Estética	Visual
Sem classificação	-

Quadro 48 – Principais funcionalidades que os usuários tiveram dúvidas

<b>Tipo de PRU = dúvida x funcionalidades</b>
<p>Editar informações; Visualizar, baixar ou inserir arquivo; Bloqueio no sistema; Visualizar ou erro nas disciplinas; como ocultar os números de matrícula do fórum; Visualizar ou erro no histórico; Visualizar ou entender o horário das aulas, notas, emenda das disciplinas; Erro de cálculo do IRA; Como fazer trancamento</p>

<sup>51</sup> IRA = Índice de Rendimento Acadêmico.

Quadro 49– Principais sugestões de funcionalidades

<b>Tipo de PRU = sugestão x funcionalidades</b>
opção “curtir”, fazer provas em casa; Mapa de localização da sala aliado as disciplinas; melhoria do sistema, explicação do código dos horários;

## 6.6 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou o modelo MALTU, bem como uma metodologia para a avaliação da interação em SS usando o modelo. Cada etapa da metodologia foi descrita e foram apresentados exemplos das categorias de classificação. O modelo apresentado possibilita uma nova forma de avaliação ainda não encontrada na literatura: avaliação de critérios de qualidade de uso (UUX) do sistema por meio dos textos espontâneos postados por usuários em SS. Esta avaliação possibilita conhecer a opinião do usuário sobre o sistema, considerando seu momento de uso, identificar problemas de interação, melhorar funcionalidades e ainda avaliar a apropriação de uma tecnologia. A partir dos estudos dos textos dos usuários também é relevante estudar o próprio conceito de IHC, a fim de entender como os usuários se expressam sobre o sistema.

Foi apresentada uma ferramenta de extração de PRUs — UUXPosts, que utiliza os padrões de extração fornecidos por esta tese. A ferramenta ainda deve ser evoluída, fornecendo um contexto de classificação manual e automático de postagens e interpretação por meio de gráficos e tabelas resultantes do cruzamento das categorias classificadas.

## 7 CONCLUSÃO

Este capítulo é dedicado às considerações finais. São apresentados os resultados alcançados, oportunidades de melhoria e trabalhos futuros. Nos resultados alcançados são apresentadas as contribuições desta tese, uma discussão sobre as questões da pesquisa e as publicações obtidas neste estudo. Em seguida, são apresentadas oportunidades de melhoria a serem realizados no modelo MALTU e na ferramenta UUX-POST. Por fim, são apresentados os trabalhos futuros oriundos desta tese de doutorado.

### 7.1 RESULTADOS ALCANÇADOS

As principais contribuições desta tese foram: (i) MALTU, um Modelo para Avaliação da interação em sistemas sociais a partir da Linguagem Textual do Usuário (Capítulo 6); (ii) uma metodologia para guiar o modelo de avaliação; (iii) padrões de extração de PRUs (Subseção 5.3.1); (iv) padrões de extração e classificação de categorias de PRUs (Subseções 5.3.2 e 5.3.3); (v) exemplos de PRUs classificadas por especialistas de IHC; e (vi) UUX-Post, uma ferramenta para extração de PRUs para apoio a futuras investigações (subseção 6.3).

Neste documento foi apresentada uma revisão sistemática (subseções 3.1, 3.2 e 3.3), o que possibilitou um maior rigor na revisão bibliográfica realizada. Essa revisão identificou trabalhos que realizaram análises a partir de textos de usuários para o estudo da UUX. Tais trabalhos não utilizam os textos do usuário para avaliar o próprio sistema durante seu uso e nem dispõem de um modelo para isso (subseção 3.4).

Nesta tese foi levantada a hipótese que as postagens dos usuários em SS fornecem dados relevantes para avaliação de aspectos da UUX (Seção 1.2). Os resultados deste trabalho confirmaram que essa hipótese foi aceita. As questões de pesquisa (QP) levantadas nesta tese são discutidas a seguir:

**QP1:** Quais são os principais métodos de avaliação da UUX em SS?

Em SS, com o resultado da revisão sistemática, os principais métodos encontrados em SS para avaliação da usabilidade foram: entrevista, questionário e teste com usuários. Nos estudos realizados (subseções 2.2.1 e 2.2.2) foram apresentados métodos de avaliação da usabilidade e UX, mas nenhum focou nas postagens dos usuários para coleta de opinião dos usuários para avaliação de UUX.

Foram encontrados trabalhos (subseção 3.4) de análise de textos de usuários para o estudo da UUX. Tais trabalhos consistem em analisar textos dos usuários ou a partir de relatos deles ou em sites de avaliação de produtos a fim de estudar a UUX. No entanto, tais trabalhos não fazem uma avaliação sobre o sistema a partir de PRUs. Nem apresentam um modelo de avaliação desses critérios (ou de outros) para tanto.

**QP2:** Os usuários “falam” sobre o sistema em uso?

A subseção 4.2 teve o objetivo de investigar se os usuários “falam” sobre o sistema durante seu uso. Foi verificado que sim. Durante sua interação no sistema, usuários postam mensagens referentes ao uso do sistema. Nessa investigação também foram definidos alguns padrões de extração de postagens a partir de algumas suposições. Tais padrões foram analisados mediante sua eficácia em extrair PRUs. Alguns padrões foram mais eficazes que outros. Essa investigação possibilitou fazer uma análise de quais padrões seriam mais eficazes para extração de PRUs em SS.

**QP3:** Como os usuários “falam”?

A subseção 4.3 teve o objetivo de investigar como os usuários “falam” sobre o sistema durante seu uso. Essa investigação foi dividida em dois subobjetivos: 1) identificar principais tipos de PRUs e 2) identificar como os usuários expressam seus sentimentos em relação ao sistema por meio de suas PRUs. No primeiro subobjetivo, foram analisados dois SS: Twitter e SIGAA e foram encontrados seis principais tipos de PRUs: elogios, críticas, dúvidas, sugestões, comparações e ajuda. No segundo subobjetivo, foram analisados três SS: Twitter, Facebook e TV, e foram identificadas PRUs com a presença ou ausência de sentimentos e a quantidade de detalhes sobre o uso do sistema. Com essa investigação, foi possível perceber formas de como o usuário fala do SS: por meio de críticas, dúvidas, sugestões, elogios, comparações, ajuda, com mais ou menos sentimentos, mais ou menos detalhes do SS. Todas essas PRUs são importantes para avaliar, de alguma forma, a UUX do sistema, seja seu contexto emocional (UX) ou mesmo o uso do sistema (usabilidade).

**QP4:** Como avaliar?

A subseção 4.4 teve o objetivo investigar como avaliar um SS a partir das PRUs de seus usuários. A metodologia utilizada foi a classificação das PRUs nos tipos definidos anteriormente e em metas de UUX e depois foi realizado um estudo da classificação e do entendimento dos participantes sobre o processo e o resultado da classificação.

Essa investigação foi aplicada, primeiramente como teste-piloto, com alunos de duas turmas da graduação que classificaram, cada um, um conjunto de 50 PRUs e, com isso,

puderam, por meio de um questionário fornecer suas impressões sobre o sistema avaliado. Em seguida, foi aplicada com 12 especialistas em IHC e, após a classificação das postagens, eles opinaram sobre o método e sobre as postagens analisadas.

A forma de avaliar a UUX de um SS a partir de suas postagens é classificando-as em categorias que permitam obter algum resultado sobre o sistema avaliado. Por exemplo, se o objetivo da avaliação for identificar problemas na interação e/ou na interface do sistema, classificar as PRUs em críticas e em funcionalidades seria uma opção.

Esta tese propôs o MALTU, um modelo de avaliação da interação a partir da linguagem textual do usuário. Nesse modelo, com a extração e análise das PRUs pode-se obter resultados de uma avaliação, de forma manual ou automática. O modelo permite ser estendido, possibilitando novas formas de extração e classificação por outros critérios de qualidades de uso.

A ferramenta apresentada nesta tese foi a UUX-Post. A UUX-Post permite a extração de PRUs a partir de um banco de postagens ou pelo endereço do SS. O objetivo da ferramenta é fornecer um apoio a essa nova forma de avaliação proposta neste trabalho.

O Quadro 50 apresenta as publicações produzidas no decorrer do curso de doutorado e aceitas para serem publicadas até o momento. O artigo 1 apresenta uma investigação a fim de apoiar a avaliação da usabilidade em SS; o artigo 2, um estudo sobre a avaliação da usabilidade em SS; o artigo 3, a proposta de tese em um workshop de teses e dissertações; o artigo 4, uma investigação a partir das postagens dos usuários no Twitter; o artigo 5, uma investigação sobre o sentimento expresso nas postagens relacionadas ao uso do sistema e suas principais características. O artigo 6 investiga a interação entre os usuários do Facebook diante de um problema sobre o sistema relatado por usuários. E, por fim, o artigo 7 apresenta uma investigação de como avaliar a UUX a partir das linguagem textual dos usuários em SS.

Quadro 50– Publicações relacionadas à tese

No.	Referência	Tipo
1	Uma investigação no apoio da avaliação da usabilidade em Sistemas Sociais usando Processamento da Linguagem Natural. In IX Brazilian Symposium in Information and Human Language Technology (STIL 2013) (MENDES, FURTADO e CASTRO, 2013)	Conferência
2	A Study about the usability evaluation of Social Systems from messages in Natural Language. In Congreso Lati-noamericano de la Interacción Humano-Computadora (CLIHC 2013) (MENDES et al., 2013)	Conferência
3	Framework de apoio na avaliação da usabilidade de <i>softwares</i> sociais a partir da Linguagem Natural. In Workshop de Teses e Dissertações - X Simpósio Brasileiro de Sistemas Colaborativos (WTD – SBSC 2013) (MENDES, FURTADO e CASTRO, 2013)	Conferência
4	Do users write about the system in use? An investigation from messages in Natural Language on Twitter. In 7th Euro American Association on Telematics and Information	Conferência

No.	Referência	Tipo
	Systems (EATIS 2014) (MENDES, FURTADO e CASTRO, 2014)	
5	How do users express their emotions regarding the social system in use? A classification of their postings by using the emotional analysis of Norman. In 16th International Conference on Human-Computer Interaction (HCII 2014) (MENDES et al., 2014)	Conferência
6	Hey, I have a problem in the system. Who can help me? An investigation of Facebook users interaction when facing privacy problems. In 17th International Conference on Human-Computer Interaction (HCII 2015). <sup>52</sup>	Conferência
7	Investigating Usability and User Experience from the user postings in Social Systems. In 17th International Conference on Human-Computer Interaction (HCII 2015) <sup>53</sup>	Conferência

## 7.2 OPORTUNIDADES DE MELHORIA

Esta tese abordou diversos aspectos de análises das postagens dos usuários em SS, apresentando um conjunto de contribuições para a área de avaliação. Ainda assim, algumas melhorias devem ser realizadas. Para uma melhor descrição, as melhorias foram separadas nas seguintes categorias: modelo MALTU e ferramenta UUX-Post.

### 7.2.1 Modelo MALTU

Mesmo com a aplicação do MALTU, o modelo ainda pode ser evoluído, bem como as suas atividades. Um detalhamento maior de cada atividade, para que seu uso seja mais intuitivo, e se destaque a importância de cada atividade para o projeto de experimentos e análise de PRUs, é um trabalho a ser desenvolvido. A repetição de sua aplicação permitirá um refinamento das atividades e, conseqüentemente, melhoria na sua execução. Além disso, a aplicação em diferentes contextos também permitirá sua evolução e adequação às possíveis situações específicas de uso que possam surgir.

O modelo permite ser estendido, possibilitando outras formas de extração e classificação de PRUs. O uso de regras baseadas na formação de sentenças pode ser um acréscimo interessante para o modelo. Observou-se, no estudo com o Twitter, que seus usuários se referiam ao sistema como se estivessem falando com o próprio sistema. Esta característica, assim, como outras apresentadas neste trabalho, podem ser estudadas para a elaboração de regras para o modelo.

---

<sup>52</sup> Aceito, a ser publicado em agosto de 2015.

<sup>53</sup> Aceito, a ser publicado em agosto de 2015.

O MALTU foi fundamentado usando os critérios de qualidades de uso UUX. No entanto, o modelo apresentado possibilita uma extensão a outras qualidades de uso, que geralmente são avaliadas em SS, como colaboração, sociabilidade, cultura, comunicabilidade etc. Em comunicabilidade, por exemplo, existe o Método de Avaliação da Comunicabilidade (MAC). Este método visa apreciar a qualidade da comunicação da metamsagem do designer para usuários (PRATES et al., 2000; DE SOUZA, 2005; PRATES; BARBOSA, 2007; DE SOUZA; LEITÃO, 2009 apud BARBOSA; SILVA, 2011). No MAC, a forma de identificar problemas de comunicação consiste em categorizar, com expressões de comunicabilidade, momentos em que o usuário encontra dificuldades de expressar sua intenção de comunicação na interface (BARBOSA; SILVA, 2011). Como exemplos de expressões de comunicabilidade, têm-se: “Cadê?”, “Epa!”, “E agora?”, “O que é isto?”, “Por que não funciona?” etc. A expressão “Por que não funciona?”, por exemplo, representa uma situação na qual o usuário esperava obter determinados resultados do sistema e não entende por que o sistema produziu os resultados diferentes do esperado (BARBOSA; SILVA, 2011). Um estudo interessante seria investigar o relacionamento entre as expressões de comunicabilidade e as PRUs.

Outro acréscimo ao modelo pode ser investigado com o estudo do trabalho de Silveira (2002). A autora, com o objetivo de apoiar os designers na construção de sistemas de ajuda, sugeriu o entendimento dos questionamentos dos usuários em relação ao sistema. A autora listou os tipos de dúvidas frequentes dos usuários em sistemas interativos, indicando formas de respondê-las. Por exemplo, as perguntas: “*Onde estou?*”, “*De onde vim?*” e “*O que eu fiz?*” requerem um acompanhamento do histórico de interação do usuário (SILVEIRA, 2002; BARBOSA; SILVA, 2010). Um estudo a ser realizado é analisar as PRUs do tipo dúvida, obtidas com o modelo MALTU, e associar às recomendações de Silveira (2001). Esse estudo possibilitaria uma tomada de decisão com os resultados do modelo.

Embora seu desenvolvimento tenha sido baseado em SS, a aplicação do MALTU em textos escritos por usuários sobre o sistema pode ser uma forma de verificar outras situações de uso do modelo. Outra forma para esta verificação é a comparação com outros modelos de avaliação.

### 7.2.2 Ferramenta UUX-Post

A ferramenta UUX-Post, desenvolvida neste trabalho, necessita de diversas melhorias e extensões: melhoria na automação da coleta de PRUs, automação da coleta de tipos de PRUs, automação da classificação das PRUs por metas de UUX, automação na identificação de funcionalidades, apresentação de gráficos e de estatísticas e relacionamento automatizado entre as categorias de classificação são alguns exemplos.

## 7.3 TRABALHOS FUTUROS

Esta tese abordou diversos aspectos de análises das postagens dos usuários em SS. Assim, diversos trabalhos futuros podem ser derivados desta tese.

Este trabalho forneceu um conjunto de padrões de extração e classificação de PRUs: por tipos de PRU, e metas de UUX. Porém não foram apresentados padrões para classificação automática por intenção, funcionalidade e artefato. Tais padrões podem ser definidos e melhorados por meio de outros experimentos com outros sistemas.

Diversos experimentos podem ser ainda realizados para estudar a linguagem textual do usuário ao se referir ao sistema, em suas PRUs. Experimentos com outras bases de dados, nas quais outras variáveis possam ser comparadas como: tipos de usuários, SS, contexto de uso, padrões de extração, entre outras.

Embora esta tese não tenha focado em estudos mais detalhados da sociolinguística (TARALLO, 1990), existem vários desafios a serem considerados em um SS com grande diversidade de usuários que implicam em variações de classe social, sexo, localidade etc. Para localidade, um estudo interessante seria considerar variantes regionais presentes nas PRUs, por exemplo, o termo “*Deu pau no sistema*” para se referir “*Problema no sistema*”. Estudos (SALGADO, 2011; SALGADO; LEITÃO; SOUZA, 2012) têm sido realizados relacionando IHC a cultura do usuário. O estudo de termos regionais, de forma textual, pode ser útil para o estudo de usuários multiculturais.

Este trabalho apresentou o conceito de PRUs, que foram classificadas nos critérios de qualidades de uso UUX. Um estudo interessante seria investigar formas de classificação de PRUs em outros critérios de qualidade de uso, como colaboração, sociabilidade, cultura, comunicabilidade etc.

O modelo apresentado permite a identificação de funções e a classificação de metas de UUX nas PRUs que podem ser associadas, a fim de identificar problemas nas funcionalidades do sistema. Experimentos com mineração web de uso ou *log* de interação poderiam ser feitos para checar se o que o usuário falou da interface faz sentido. Se o caminho que ele percorreu foi correto para poder dar essa opinião.

Outros estudos podem ser realizados para investigar a interação do usuário com o sistema. Pode-se fazer um estudo das PRUs a fim de identificar funcionalidades que não são claramente mencionadas, como: “Como eu faço para postar uma foto **aqui?**” ou “**Este campo** sempre dá erro!”. A resolução desse problema envolve conceitos como relacionamento proposicional, correferência ou expressões referenciais. O uso de PLN pode auxiliar em experimentos como esses.

Diversos estudos em PLN têm sido realizados em sentenças com ironias. Nesta tese, a ironia foi classificada como crítica. Com a aplicação de algoritmos de detecção de ironia, pode-se fazer um estudo para quantificar a ocorrência de PRUs com essa característica, como por exemplo: “Que lindo! O Twitter com erro novamente!” ou “O app do Twitter tem uma opção de denunciar conta comprometida, mas a usabilidade dele é tão incrível que eu não acho mais onde fica”.

Experimentos usando algoritmos de análise de sentimentos devem ser feitos a fim de investigar PRUs. Nesta tese foram observadas PRUs com diversas características: sentenças subjetivas, objetivas, com polaridades (positiva, neutra ou negativa) e com diferentes intensidades de sentimentos. Todos esses tipos foram considerados relevantes para avaliação de sistemas.

## REFERÊNCIAS

AGOSTO, Denise E.; ABBAS, June. High school seniors' social network and other ICT use preferences and concerns. **Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem (ASIS&T 2010)**, Pittsburgh, v. 47, n. 65, p. 1-10, October. 2010.

ALL ABOUT UX. **Information for user experience professionals**. Disponível em: <<http://www.allaboutux.org/>>. Acesso em: 12 jan. 2015.

ALLEN, James. **Natural Language Understanding**. 2nd. Redwood City, CA: The Benjamin/Cummings Publishing Company, Inc., 1994.

ARAMAKI, Eiji; MASKAWA, Sachiko; MORITA, Mizuki. Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter. **Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing**, Edinburgh, Scotland, UK, p. 1568–1576, July, 2011.

ARANHA, Christian; PASSOS, Emmanuel. A Tecnologia de Mineração de Textos. **RESI – Revista Elerônica de Sistemas de Informação**, Rio de Janeiro, n. 2, p. 1-8, 2006.

ASIAEE T., Amir; TEPPER, Mariano; BANERJEE, Arindam; SAPIRO, Guillermo. If you are happy and you know it... tweet. **Proceedings of the 21st ACM international conference on Information and knowledge management (CIKM 2012)**, Maui, HI, USA , p. 1602-1606, November, 2012.

BARBOSA, Simone. D. J.; DA SILVA, Bruno. S. **Interação Humano-Computador**. Rio de Janeiro: Elsevier, 2010.

BARGAS-AVILA, Javier A.; HORNBÆK, Kasper. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. **Proceedings of the 2011 annual conference on Human factors in computing systems (CHI 2011)**, Vancouver, p. 2689-2698, 2011.

BARION, E. C.N.; LAGO, D. Mineração de Textos. **Revista de Ciência Exatas e Tecnologias**, São Paulo, v. , n. 3, p. 123-140, 2008.

BARSADE, Sigal G. The Ripple Effect: Emotional Contagion and Its Influence on Group Behavior. **Administrative Science Quarterly**, v. 47, n. 4, p. 644-675, 2002.

BECKER, Karin; TUMITAN, Diego. Introdução à Mineração de Opiniões: Conceitos, Aplicações e Desafios. **Minicurso 2 do Simpósio Brasileiro de Banco de Dados (SBBDD 2013)**, Recife-PE, Brasil, 2013.

BENAMARA, F.; CESARANO, C.; PICARIELLO, A.; REFORGIATO, D. Sentiment analysis: adjectives and adverbs are better than adjectives alone. **Proceedings of international conference on weblogs and social media**, icwsm, boulder, 2007.

BERGMANN, Francine B.; SILVEIRA, Milene S. “Eu vi o que você fez... e eu sei quem você é!”: uma análise sobre privacidade no Facebook do ponto de vista dos usuários. **Proceedings of the Symposium on Human Factors in Computing Systems (IHC 2012)**, Cuiabá, MT, Brasil, novembro, 2012.

BEVAN, Nigel. Classifying and selecting ux and usability measures. In: **International Workshop on Meaningful Measures: Valid Useful User Experience Measurement**, 2008. p. 1318.

BEVAN, Nigel. What is the difference between the purpose of usability and user experience evaluation methods. **Proceedings of the Workshop UXEM’09 (Interact 09)**, Uppsala, Sweden, 2009.

BHAT, Sajid Y.; ABULAISH, Muhammad. A density-based approach for mining overlapping communities from social network interactions. **Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics (WIMS 2012)**, Craiova, Romania, 2012.

BIFET, Albert; FRANK, Eibe. Sentiment Knowledge Discovery in Twitter Streaming Data. **Proceedings of the 13th international conference on Discovery science (DS 2010)**, p. 1-15, 2010.

BISHOP, Matt.; GATES, Carrie.; YELLOWLEES, Peter; SILBERMAN, Gabriel. Facebook goes to the doctor. **Proceedings of the 2011 Workshop on Governance of Technology, Information, and Policies (GTIP 2011)**, Orlando, Florida, USA, p. 13-20, 2011.

BOSMA, Maarten; MEIJ, Edgar; WEERKAMP, Wouter. A framework for unsupervised spam detection in social networking sites. **Proceedings of the 34th European conference on Advances in Information Retrieval (ECIR 2012)**, p. 364-375, 2012.

BOYD, Danah M.; ELLISON, Nicole B. Social Network Sites: Definition, History, and Scholarship. **Journal of Computer-Mediated Communication**, p. 210-230, 2007.

BROOKE, Julian. **A semantic approach to automated text sentiment analysis**. PHD Thesis, Simon Fraser University, 2009.

BRUBAKER, Jed R.; KIVRAN-SWAIN, Funda; TABER, Lee; HAYES, Gillian R. Grief-Stricken in a Crowd: The language of bereavement and distress in social media. **Association for the advancement of artificial intelligence (AAAI 2012)**, 2012.

BUZZI, Maria C.; BUZZI, Marina; LEPORINI, Barbara. Web 2.0: Twitter and the blind. **Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction: Facing Complexity (CHIItaly)**, Alghero, Italy, p. 151-156, 2011.

CAMILO, Cássio O.; DA SILVA, João C. Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas. **Relatório Técnico**. Instituto de Informática, Universidade Federal de Goiás, 2009.

CARRILHO JUNIOR, João R. **Desenvolvimento de uma Metodologia para Mineração de Textos**. 96 f. Dissertação (Mestrado em Engenharia Elétrica) – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro-RJ, 2007.

CARRILLO DE ALBORNOZ, Jorge; PLAZA, Laura; GERVÁS, Pablo; DÍAZ, Alberto. **A Joint Model of Feature Mining and Sentiment Analysis for Product Review Rating**. P. Clough et al. (Eds.): ECIR 2011, LNCS 6611, p. 55-66, Springer-Verlag Berlin Heidelberg, 2011.

CARVALHO, Carlos R. M. Exploring the relationships between user experience engagement, user participation levels and individual knowledge building perception in a gamified collaborative system. 2013. 117f. **Dissertação (Mestrado em Informática Aplicada) – UNIFOR – Universidade de Fortaleza. Fortaleza – CE, Brasil, 2013.**

CARVALHO, Carlos R.; FURTADO, Elizabeth S. Wikimarks: an approach proposition for generating collaborative, structured content from social networking sharing on the web. **Proceedings of the Symposium on Human Factors in Computing Systems (IHC 2012)**, Cuiabá, MT, Brasil, p. 95-98, novembro, 2012.

CASTANHEIRA, Luciana. G. **Aplicação de técnicas de mineração de dados em problemas de classificação de padrões**. 95 f. 2008. Dissertação (Mestrado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica da UFMG, Belo Horizonte, Brasil, 2008.

CASTILLO, Carlos; MENDOZA, Marcelo; POBLETE, Barbara. Information credibility on twitter. **Proceedings of the 20th international conference on World wide web (WWW 2011)**, Hyderabad, Índia, p. 675-684, abril 2011.

CHEN, Lingji; ACHREKAR, Harshavardhan; LIU, Benyuan; LAZARUS, Ross. Vision: Towards Real Time Epidemic Vigilance through Online Social Networks. **Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond (MCS 2010)**, San Francisco, USA, 2010.

CHUNG, Namho; KOO, Chulmo; PARK, Seung-Bae. Why People Share Information in Social Network Sites? Integrating with Uses and Gratification and Social Identity Theories. **Intelligent Information and Database Systems**. Lecture Notes in Computer Science, v. 7197, p. 175-184, 2012.

COOLEY, R.; MOBASHER, B.; SRIVASTAVA, J. Web mining: information and pattern Discovery on the World Wide Web. **Proceedings of the 9th IEEE International Conference on Tools with Artificial Intelligence**, p. 558-567, 1997.

CORRÊA, Adriana C. G. **Recuperação de Documentos baseada em Informação Semântica no Ambiente AMMO**. 92 f. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Ciência da Computação, Departamento de Computação da Universidade Federal de São Carlos, São Carlos-SP, Brasil, 2003.

CORREA, Denzil; SUREKA, Ashish. Mining *tweets* for tag recommendation on social media. **Proceedings of the 3rd international workshop on Search and mining user-generated contents (SMUC 2011)**, Glasgow, Scotland, UK, p. 69-76, 2011.

CRAMPES, Michel; PLANTIÉ, Michel. Mining social networks and their visual semantics from social photos. **Proceedings of the International Conference on Web Intelligence, Mining and Semantics (WIMS 2011)**, 2011.

CRISTESCU, I. Emotions in human-computer interaction: the role of non-verbal behavior in interactive systems. **Revista In-formatica Economica**, v. 2, p. 110-116, 2008.

CULOTTA, Aron. Towards detecting influenza epidemics by analyzing Twitter messages. **Proceedings of the First Workshop on Social Media Analytics (SOMA 2010)**, p. 115-122, 2010.

CYBIS, Walter; BETIOL, Adriana H.; FAUST, Richard. **Ergonomia e usabilidade: conhecimentos, métodos e aplicações**. São Paulo: Novatec, 2007.

DE CHOUDHURY, Munmun; DIAKOPOULOS, Nicholas; NAAMAN, Mor. Unfolding the Event Landscape on Twitter: Classification and Exploration of User Categories. **Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW 2012)**, Seattle, Washington, USA, p. 241-244, 2012.

DE LERA, Eva; GARRETA-DOMINGO, Muriel. Ten Emotion Heuristics: Guidelines for Assessing the User's Affective Dimension Easily and Cost-Effectively. **Proceedings of the HCI07 Conference on People and Computers XXI**, p. 163-166, 2007.

DIAS DA SILVA, Bento C.; MONTILHA, Gisele; RINO, Lúcia H. M.; SPECIA, Lucia; NUNES, Maria das G. V. N.; OLIVEIRA JR, Osvaldo N. O.; MARTINS, Ronaldo T.; PARDO, Thiago A. S. Introdução ao Processamento das Línguas Naturais e Algumas Aplicações. In: **Série de relatórios do nilc. nilc-tr-07-10**, São Carlos-SP, 2007.

DIAS, Maria A. L.; MALHEIROS, Marcelo de G. Automatic Extraction of Keywords for the Portuguese Language. In: VIEIRA, Renata; QUARESMA, Paulo; NUNES, Maria da Graça Volpes; MAMEDE, Nuno J. Mamede; OLIVEIRA, Cláudia; DIAS, Maria Carmelita (Eds.). **Computational Processing of the Portuguese Language: 7th International Workshop, (PROPOR 2006)**. LNAI 3960, Berlin/Heidelberg: Springer Verlag. Itatiaia, Brasil, 2006. p. 204-207.

DITLEV, Joakim. Review Motivation Part 1 – The Real Reason People Write Negative Reviews. Trust Pilot, 12 mar. 2012. Disponível em: <<http://blog.trustpilot.com/review-motivation/>>. Acesso em: 12 maio. 2012.

EGELMAN, Serge; OATES, Andrew; KRISHNAMURTHI, Shriram. Oops, I did it again: mitigating repeated access control errors on facebook. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)**, Vancouver, BC, p. 2295-2304, 2011.

EIKVIL, L. **Information extraction from World Wide Web: A survey** Technical Report 945. Norwegian Computing Center, 1999.

EKMAN, Paul. An argument for basic emotions. **Cognition and Emotion**, p. 169-200, 1992.

ESULI, Andrea; SEBASTIANI, Fabrizio. Determining term subjectivity and term orientation for opinion mining. **Proceedings of EACL-06, 11th Conference of the European Chapter of the Association for Computational Linguistics**, Trento, IT, p. 193-200, 2006.

FAGUNDES DA SILVA, Cassiana. **Grupos gramaticais e sintáticos em categorização automática com Support Vector Machines**. 2004. 109 f. Dissertação (Mestrado em Computação Aplicada) – Universidade do Vale do Rio dos Sinos. Ciências Exatas e Tecnológicas Programa Interdisciplinar de Pós-Graduação em Computação Aplicada – PIPCA, São Leopoldo, 2004.

FAHL, Sascha; HARBACH, Marian; MUDERS, Thomas; SMITH, Matthew; SANDER, Uwe. Helping Johnny 2.0 to encrypt his Facebook conversations. **Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS 2012)**, Article No. 11, Washington, DC, USA, 2012.

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From Data Mining to Knowledge Discovery in Databases. **Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI 1996)**, Portland, Oregon, p. 37-54, 1996.

FETTER, Mirko; GROSS, Tom; SCHIRMER, Maximilian. CAESSA: Visual Authoring of ContextAware Experience Sampling Studies. **Conference on Human Factors in Computing Systems (CHI 2011)**, Work-in-Progress, Vancouver, BC, Canada, p. 2341-2346, May, 2011.

FOLMER, Elke; VAN GURP, Jilles; BOSCH, Jan. A framework for capturing the relationship between usability and software architecture. **Software Process: Improvement and Practice** 8, p. 67-87, 2003.

FORLIZZI, Jodi; BATTARBEE, Katja. Understanding experience in interactive systems. **Proceedings of the 2004 conference on Designing Interactive Systems (DIS 04): processes, practices, methods, and techniques (New York: ACM)**, Cambridge, Massachusetts, USA, p. 261-268, 2004.

FRANCO, José Wellington da Silva. **Aquisição de conhecimento de mundo para sistemas de processamento de linguagem natural**. 2013. 86f. Dissertação (Mestrado em Ciências da Computação) – Universidade Federal do Ceará. Mestrado e Doutorado em Ciências da Computação – Fortaleza, Ceará, 2013.

FURTADO, Elizabeth S.; FURTADO, Vasco; VASCONCELOS, Eurico. A Conceptual Framework for the Design and Evaluation of Affective Usability in Educational Geosimulation Systems. **Conference on Human-Computer Interaction (Interact 2007), Lecture Notes in Computer Science**, v. 4662, p. 497-510, 2007.

FURTADO, Elizabeth, S. (Org.). **Qualidade da interação de sistemas e novas abordagens para a avaliação**. Curitiba: CRV, Brasil, 2012.

FURTADO, MADEIRA e NÓBREGA. Uso de personas na avaliação das necessidades dos usuários. In **Qualidade da interação de sistemas e novas abordagens para a avaliação**. Curitiba: CRV, Brasil, 2012.

FURTADO, MILITAO e NÓBREGA. Avaliação de Softwares Sociais. In **Qualidade da interação de sistemas e novas abordagens para a avaliação**. Curitiba: CRV, Brasil, 2012.

FURTADO, Vasco; PINHEIRO, Vladia; FREIRE, Livio; FERREIRA, Caio. Knowledge-Intensive Word Disambiguation via Common-Sense and Wikipedia. **21st Brazilian Symposium on Artificial Intelligence**. SBIA 2012, Curitiba, PR. Lecture Notes in Artificial Intelligence (LNAI), 2012.

GAIZAUSKAS, Robert; WILKS, Yorick. Information extraction: beyond document retrieval. **Computational Linguistics and Chinese Language Processing**, v. 3, n. 2, p. 17-60, 1998.

GOMES, Roberto. M. **Desambiguação de Sentido de Palavras Dirigida por Técnicas de Agrupamento sob o enfoque da Mineração de Textos**. 118p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro-RJ, 2009.

GONÇALVES, Eduardo C. **Data Mining – Novos Recursos nos Sistemas de Banco de Dados**. Disponível em: <[www.devmedia.com.br](http://www.devmedia.com.br)>. Acesso em: 13 jan. 2015.

GOYAL, Amit; BONCHI, Francesco; LAKSHMANAN, Laks V.S. Learning influence probabilities in social networks. **Proceedings of the third ACM international conference on Web search and data mining (WSDM'10)**, New York City, USA, p. 241-250, 2010.

HASSENZAHN, Marc; DIEFENBACH, Sarah; GÖRITZ, Anja. Needs, affect, and interactive products. **Journal Interacting with Computers**. Volume 22 Issue 5, Pages 353-362, 2010.

HATFIELD, Elaine; CACIOPPO, John T.; RAPSON, Richard L. Emotional contagion. **Current Directions in Psychological Sciences**, Vol 2, p. 96-99, 1993.

HEDEGAARD, Steffen; SIMONSEN, Jakob G. Extracting Usability and User Experience Information from Online User Reviews. **ACM Conference on Human Factors in Computing Systems (CHI 2013)**, Paris, France, p. 2089-2098, 2013.

HUME, David. Emotions and Moods. In: ROBBINS, S.P., JUDGE, T.A. (Eds.). **Organizational Behavior**, p. 258-297. Disponível em: [http://www.pearsonhighered.com/assets/hip/us/hip\\_us\\_pearsonhighered/samplechapter/0132431564.pdf](http://www.pearsonhighered.com/assets/hip/us/hip_us_pearsonhighered/samplechapter/0132431564.pdf). Acesso em: 12 de janeiro de 2015.

ISBISTER, Katherine; HOOK, Kia; LAAKSOLAHTI, Jarmo; SHARPA, Michael. The sensual evaluation instrument: Developing a trans-cultural self-report measure of affect. **Journal of Human-Computer Studies archive**. Volume 65 Issue 4, pp. 315-328, Elsevier, 2007.

ISO 9241-11. **ISO 9241-11**. Geneve: International Organization For Standardization, 1998.

ISO DIS 9241 – 210:2008. **Ergonomics of human system interaction – Part 210: Human - centred design for interactive systems** (formerly known as 13407). International Standardization Organization (ISO). Switzerland, 2008.

JAMISON-POWELL, Sue; LINEHAN, Conor; DALEY, Laura; GARBETT, Andrew; LAWSON, Shaun. “I can’t get no sleep”: discussing #insomnia on Twitter. **Proceedings of**

the **SIGCHI Conference on Human Factors in Computing Systems (CHI 2012)**, Austin, Texas, USA, p. 1501-1510, May 2012.

JEHL, Laura; HIEBER, Felix; RIEZLER, Stefan. Twitter Translation using Translation-Based Cross-Lingual Retrieval. **Proceedings of the 7th Workshop on Statistical Machine Translation**, Montreal, Canada, p. 410-421, June 2012.

JIANG, Long; YU, Mo; ZHOU, Ming; LIU, Xiaohua; ZHAO, Tiejun. Target-dependent Twitter Sentiment Classification. **Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (HLT 2011)**, Portland, Oregon, v.1, p. 151-160, June 2011.

KAMAL, Noreen; FELLS, Sidney; MCGRENERE, Joanna; NANCE, Kara. Helping me helping you: designing to influence health behaviour through social connections. In: KOLZE, P.; MARSDEN, G.; LINDGAARD, G.; WESSON, J.; WINCKLER, M (Eds.). **Human-Computer Interaction (INTERACT 2013)**, Lecture Notes in Computer Science 8119, Springer Berlin/Heidelberg, p. 708-725, 2012.

KARNIK, Mayur; OAKLEY, Ian; VENKATANATHAN, Jayant; SPILIOPOULOS, Tasos; NISI, Valentina. Uses & Gratifications of a Facebook Media Sharing Group. **Proceedings of the 2013 conference on Computer supported cooperative work (CSCW 2013)**, San Antonio, Texas, USA, p. 821-826, 2013.

KARR, Douglas. Why Do People Write Online Reviews? **Social media today**, 18 Jun. 2012. Disponível em: <<http://www.socialmediatoday.com/content/infographic-why-do-people-write-online-reviews>>. Acesso em: 12 maio. 2012.

KATZIR, Liran; LIBERTY, Edo; SOMEKH, Oren. Estimating sizes of social networks via biased sampling. **Proceedings of the 20th international conference on World wide web (WWW 2011)**, Hyderabad, India, p. 597-606, April 2011.

KELLEY, Patrick G.; BREWER, Robin ; MAYER, Yael; CRANOR, Lorrie F.; SADEH, Norman. An Investigation into Facebook Friend Grouping. **Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction (INTERACT 2011)**, Lisbon, Portugal, p. 216-233, 2011.

KETOLA, Pekka; ROTO, Virpi. Exploring user experience measurement needs. **5th COST294-MAUSE Open Workshop on Valid Useful User Experience Measurement**, 2008.

KHAN, Vassilis J.; MARKOPOULOS, Panos. **EXPERIENCE SAMPLING: A workbook about the method and the tools that support it - A workbook on the Experience sampling method and tools**. A workbook on the Experience sampling method and tools. Disponível em: <<http://vjkh.com/old/content/publications/ExperienceSampling-A-workbook-about-the-method-and-the-tools-that-support-it.pdf>>. Acesso em: 12 Jan. 2015.

KIM, Jun H.; GUNN, Daniel V.; SCHUH, Eric; PHILLIPS, Bruce; PAGULAYAN, Randy J.; WIXON, Dennis. **Tracking Real-Time User Experience (TRUE): A comprehensive instrumentation solution for complex systems**. Conference on Human Factors in Computing Systems (CHI 2008), Florence, Italy, p. 443-452, 2008.

KIM, Soo-Min; HOVY, Eduard. Determining the sentiment of opinions. **Proceedings of the COLING conference**, Geneva, p. 1367–1373, 2004.

KINSELLA, Sheila; MURDOCK, Vanessa; O'HARE, Neil. “I'm eating a sandwich in Glasgow”: modeling locations with *tweets*. **Proceedings of the 3rd international workshop on Search and mining user-generated contents (SMUC 2011)**, Glasgow, Scotland, UK, p. 61-68, 2011.

KISILEVICH, Slava; MANSMANN, Florian. Analysis of privacy in online social networks of runet. **Proceedings of the 3rd international conference on Security of information and networks (SIN 2010)**, Tanganrog, p. 46-55 , september 2010.

KITCHENHAM, Barbara. **Procedures for performing systematic reviews**. Joint technical report. Department of computer sciences, Keele University (UK), Empirical Software Engineering National ITC Australia Ltd, Australia, Report No. 0400011T, July 2004.

KITCHENHAM, Barbara; CHARTERS, Stuart. **Guidelines for performing systematic literature review s in software engineering** (version 2.3). Technical report, Keele University and University of Durham, Durham, July 2007.

KORHONEN, Hannu; ARRASVUORI, Juha; VÄÄNÄNEN-VAINIO-MATTILA, Kaisa. Let users tell the story. **Proceedings of AMC CHI 2010 Extended Abstracts, The ACM Press**, Atlanta, p. 4051-4056, April 2010.

LAMPINEN, Airi; LEHTINEN, Vilma; LEHMUSKALLIO, Asko; TAMMINEN, Sakari. We're in it together: interpersonal management of disclosure in social network services. In **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011)**, Vancouver, BC, p. 3217-3226, May 2011.

LÁNYI, Cecília S.; NAGY, Eszter; SIK, Gergely. Accessibility testing of a healthy lifestyles social network. **Proceedings of the 13th international conference on Computers Helping People with Special Needs**, Volume Part I (ICCHP 2012), p. 409-416, 2012.

LAW, Effie L-C.; ROTO, Virpi.; HASSENZ AHL, Marc; VERMEEREN, Arnold P. O. S.; KORT, Joke. Understanding, Scoping and Defining User eXperience: A Survey Approach. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2009)**, Boston, MA, USA, p. 719-728, 2009.

LI, Haitao; LIU, Jiangchuan; XU, Ke; WEN, Song. Understanding video propagation in online social networks. **Proceedings of the 2012 IEEE 20th International Workshop on Quality of Service (IWQoS12)**, Article No. 21, Portland, Oregon, USA, 2012.

LIBRELOTTO, Solange R.; MOZZAQUATRO, Patrícia M.. Análise dos Algoritmos de Mineração J48 e Apriori Aplicados na Detecção de Indicadores da Qualidade de Vida e Saúde. **Revista Interdisciplinar de Ensino, Pesquisa e Extensão (RevInt)**, v.1, n.1, pp.26-37, 2013.

LIM, Kwan Hui; DATTA, Amitava. Finding twitter communities with common interests using following links of celebrities. **Proceedings of the 3rd international workshop on Modeling social media (MSM 2012)**, Milwaukee-WI, USA, p. 25-32, 2012.

LIPCZAC, Marek; SIGURBJORNSSON, Borkur; JAIMES, Alejandro. Understanding and leveraging tag-based relations in on-line social networks. **Proceedings of the 23rd ACM conference on Hypertext and social media (HT 2012)**, Milwaukee-WI, USA, p. 229-238, 2012.

LIU, Bing. **Sentiment analysis and opinion mining**. Morgan & Claypool Publishers, 2012.

LIU, Yabing; GUMMADI, Krishna P.; KRISHNAMURTHY, Balachander; MISLOVE, Alan. Analyzing facebook privacy settings: user expectations vs. reality. **Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference (IMC 2011)**, Berlin, Germany, p. 61-70, 2011.

LOPES, Maria C. S. **Mineração de Dados Textuais Utilizando Técnicas de Clustering para o Idioma Português**. 191 f. Tese (Doutorado em Engenharia Civil) – Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, Brasil, 2004.

MADEIRA, Kelma. Um Framework de Elaboração de Personas e sua Aplicação para a Elicitação de Requisitos e para Análise das Interações em Sistemas Sociais. 2010. 217f. **Dissertação (Mestrado em Informática Aplicada) – UNIFOR – Universidade de Fortaleza. Fortaleza – CE, Brasil, 2010.**

MADEIRA, Kelma; MILITÃO, Júlio; NOBREGA, Lenine; DA SILVA, Leandro; OLIVEIRA, Denilson; MATOS, Italo. Uma Avaliação do ORKUT utilizando Personas sob a ótica da Nova Usabilidade. In: **VIII Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais**, Porto Alegre, 2008.

MAIA, Camila L., B.; FURTADO, Elizabeth S. Uma Revisão Sistemática sobre Medição da Experiência do Usuário. **Proceedings of the Symposium on Human Factors in Computing Systems (IHC 2014)**, Foz do Iguaçu – PR, Brasil, 2014.

MAITLAND, Julie; CHALMERS, Matthew. Designing for peer involvement in weight management. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011)**, Vancouver, BC, p. 315-324, May, 2011.

MAO, Huina; SHUAI, Xin; KAPADIA, Apu. Loose *tweets*: an analysis of privacy leaks on twitter. **Proceedings of the 10th annual ACM workshop on Privacy in the electronic society (WPES 2011)**, Chicago, Chicago, IL, USA, p. 1-12, 2011.

MARINHO, Leandro B.; GIRARDI, Rosario. Mineração na Web. **Revista Eletrônica de Iniciação Científica**, São Luiz, 2003. Disponível em: <<http://paginas.fe.up.pt/~mgi03006/ARI/MineracaoNaWeb.pdf>>. Acesso em: 12 jan. 2015.

MARTINS, Claudia A. **Uma abordagem para pré-processamento de dados textuais em algoritmos de aprendizado**. 2013. Tese (Doutorado em Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação – ICMC – USP, São Carlos, 2003.

MASCETTI, Sergio; FRENI, Dario; BETTINI, Claudio; WANG, X. Sean; JAJODIA, Sushil. Privacy in geo-social networks: proximity notification with untrusted service providers and curious buddies. **Journal The VLDB Journal** – The International Journal on Very Large Data Bases archive, v. 20, Issue 4, p. 541-566, 2011.

MAZIERO, Erick G.; PARDO, Thiago A. S.; Di FELIPPO, Ariani; DIAS-DA-SILVA, Bento C. A base de dados lexical e a interface web do TeP 2.0: thesaurus eletrônico para o Português do Brasil. **Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web (WebMedia 2008)**, p. 390-392, 2008.

MAZZIA, Alessandra; LEFEVRE, Kristen; ADAR, Eytan. The PViz comprehension tool for social network privacy settings. **Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS 2012)**, Article No. 13, Washington, DC, USA, 2012.

MENDES, Marília S.; FURTADO, Elizabeth S.; CASTRO, Miguel F. Do users write about the system in use? An investigation from messages in Natural Language on Twitter. **7th Euro American Association on Telematics and Information Systems (EATIS 2014)**, Valparaiso, Chile, 2014.

MENDES, Marília S.; FURTADO, Elizabeth S.; CASTRO, Miguel F. Uma investigação no apoio da avaliação da usabilidade em Sistemas Sociais usando Processamento da Linguagem Natural. **IX Brazilian Symposium in Information and Human Language Technology (STIL 2013)**, Fortaleza, Brasil, 2013.

MENDES, Marília S.; FURTADO, Elizabeth S.; CASTRO, Miguel. F. Framework de apoio na avaliação da usabilidade de softwares sociais a partir da Linguagem Natural. **III Workshop de Teses e Dissertações em Sistemas Colaborativos (WTD-SBSC 2013)**, Manaus, Brasil, 2013.

MENDES, Marília S.; FURTADO, Elizabeth S.; FURTADO, Vasco; CASTRO, Miguel. F. How do users express their emotions regarding the social system in use? A classification of their postings by using the emotional analysis of Norman. **16th International Conference on Human-Computer Interaction (HCII 2014)**, Crete, Greece, 2014.

MENDES, Marília S.; FURTADO, Elizabeth S.; FURTADO, Vasco; CASTRO, Miguel. F.. Investigating Usability and User Experience from the user postings in Social Systems. **17th International Conference on Human-Computer Interaction (HCII 2015)**, Los Angeles, CA, USA, 2015.

MENDES, Marília S.; FURTADO, Elizabeth S.; MILITAO, Guido; CASTRO, Miguel. F.. Hey, I have a problem in the system, who helps me? An Investigation of the social interaction of Facebook users in front of usability problems. **17th International Conference on Human-Computer Interaction (HCII 2015)**, Los Angeles, CA, USA, 2015.

MENDES, Marília S.; FURTADO, Elizabeth; THEOPHILO, Fábio; CASTRO, Miguel F.. A Study about the usability evaluation of Social Systems from messages in Natural Language. In **Congreso Lati-noamericano de la Interacción Humano-Computadora (CLIHC 2013)**, Guanacaste, Costa-Rica, 2013.

MERRITT, Jessica. Why Do People Write Reviews? **Online Reputation Management**, 8 out. 2013. Disponível em: <<https://www.reputationmanagement.com/blog/why-do-people-write-reviews>>. Acesso em: 12 jan. 2015.

MESCHTSCHERJAKOV, Alexander; REITBERGER, Wolfgang; TSCHELIGI, Manfred. MAESTRO: Orchestrating User Behavior Driven and Context Triggered Experience Sampling. **Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research (MB'10)**, Eindhoven, The Netherlands, 2010.

MILLEN, David R. Rapid Ethnography: Time Deepening Strategies for HCI Field Research. **Proceedings of the ACM 2000 conference for Designing interactive systems: processes, practices, methods, and techniques (DIS'00)**. Brooklyn, NY, USA, p. 280-286, 2000.

MOONEY, Raymond J.; NAHM, Un Y..Text Mining with Information Extraction. **Multilingualism and Electronic Language Management: Proceedings of the 4th International MIDP Colloquium**. DAELEMANS, W., DU PLESSIS, T., SNYMAN, C. AND TECK, L. Bloemfontein, South Africa: Van Schaik Pub., p. 141-160 p., 2005.

MORAIS, Edison A. M.; AMBRÓSIO, Ana P. L. **Mineração de Textos**. Relatório Técnico, pp. 30, Instituto de Informática Universidade Federal de Goiás, 2007. Disponível em: <[http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF\\_005-07.pdf](http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_005-07.pdf)>. Acesso em: 12 jan. 2015.

NADAMOTO, Akiyo; ARAMAKI, Eiji; ABEKAWA, Takeshi; MURAKAMI, Yohei. Extracting the gist of social network services using Wikipedia. **Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services (iiWAS'10)**, Paris, France, p. 231-238, 2010.

NEWMAN, Mark W.; LAUTERBACH, Debra; MUNSON, Sean A.; RESNICK, Paul; MORRIS, Margaret E. It's not that i don't have problems, i'm just not putting them on facebook: challenges and opportunities in using online social networks for health. **Proceedings of the ACM 2011 conference on Computer supported cooperative work (CSCW'11)**, Hangzhou, China, p. 341-350, 2011.

NICOLACI-DA-COSTA, Ana M.; PIMENTEL, Mariano. Sistemas colaborativos para uma nova sociedade e um novo ser humano. In: PIMENTEL, Mariano; FUKS, Hugo (Org.). **Sistemas Colaborativos**. Rio de Janeiro: Campus-Elsevier, p. 3-15, 2011. v. 1.

NÓBREGA, Lenine. Um Framework de Elaboração de Persona Empresa para Suporte na Análise de Valor de Negócio na Aplicação em Sistemas de Redes Sociais. 129f. **Dissertação (Mestrado em Informática Aplicada) – UNIFOR – Universidade de Fortaleza**. Fortaleza – CE, Brasil, 2011.

NORMAN, Donald A. **Emotional Design: Why We Love (or Hate) Everyday Things**. New York: Basic Books, 2004.

NORMAN, Donald A.; MILLER, Jim; HENDERSON, Austin. What you see, some of what's in the future, and how we go about doing it: HI at Apple Computer. In: **Conference on Human Factors in Computing Systems (CHI 1995)**, Denver, Colorado, USA, 1995.

OBRIST, Marianna; ROTO, Virpi; VÄÄNÄNEN-VAINIO-MATTILA, Kaisa. User Experience Evaluation – Do You Know Which Method to Use? **Proceedings of the 2009 annual conference on Human factors in computing systems (CHI 2009)**, Special Interest Groups, Boston, Massachusetts, USA, 2009.

OLSSON, T.; SALO, M. Narratives of satisfying and unsatisfying experiences of current mobile augmented reality applications. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2012)**, p. 2779-2788, 2012.

ORENGO, Viviane M.; HUYCK, Chistian R. A stemming algorithm for the portuguese language. **8th International Symposium on String Processing and Information Retrieval (SPIRE)**, pp. 186-193, Los Alamitos, California, 2001.

OTHERO, Gabriel de A.; MENUZZI, Sérgio de M. **Linguística computacional – Teoria & prática**. São Paulo, Parábola Editorial, 2005.

PANG, Bo; LEE, Lillian; VAITHYANATHAN, Shivakumar. Thumbs up? sentiment classification using machine learning techniques. **Proceedings of EMNLP**, p. 79-86, 2002.

PEREIRA, Roberto; BARANAUSKAS, M. Cecília C.; SILVA, Sergio Roberto P. da. Softwares sociais: uma visão orientada a valores. **Proceedings of the IX Symposium on Human Factors in Computing Systems (IHC 2010)**, Belo Horizonte, Brasil, p. 149-158, outubro, 2010.

PEROTTI, Victor; HAIR, Neil. User experience in online social networks: A Qualitative analysis of key activities and associated features. **Proceedings of the 2011 44th Hawaii International Conference on System Sciences (HICSS 2011)**, p. 1-10, January, 2011.

PIMENTEL, Mariano; FUKS, Hugo. **Sistemas Colaborativos**. Rio de Janeiro: Campus-Elsevier, p. 264-293. v. 1, 2011.

PIMENTEL, Mariano; GEROSA, Marcos A.; FUKS, Hugo. Sistemas de Comunicação para colaboração. In: PIMENTEL, Mariano; FUKS, Hugo (Orgs.). **Sistemas Colaborativos**. Rio de Janeiro: Campus-Elsevier, 2011. p. 65-93. v. 1.

PINHEIRO, Vlória Célia Monteiro. SIM: Um Modelo Semântico Inferencialista para expressão e raciocínio em sistemas de linguagem Natural. 232f. **Tese (Doutorado) – Centro de Ciências, Universidade Federal do Ceará**, Fortaleza, 2010.

PINHEIRO, Vlória; PEQUENO, Tarcísio; FURTADO, Vasco; NOGUEIRA, Douglas. Semantic Inferentialist Analyser: Um Analisador Semântico de Sentenças em Linguagem Natural. **Proceedings of the 7th Brazilian Symposium in Information and Human Language Technology, STIL**, Brasil, 2009.

PRATES, Raquel O. Interação em Sistemas Colaborativos. In: PIMENTEL, Mariano; FUKS, Hugo Fuks (Orgs.). **Sistemas Colaborativos**. Rio de Janeiro: Campus-Elsevier, 2011. p. 264-293. v. 1.

PRATI, Ronaldo C. **Novas abordagens de aprendizado de máquina para a geração de regras, classes desbalanceadas e ordenação de casos**. 2006. 191f. Tese (Doutorado em

Ciências da Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação – IMC/USP, São Carlos-SP, Brasil, 2006.

PREECE, Jennifer; ROGERS, Yvonne; SHARP, Helen. **Design de interação**: além da interação homem computador. Trad. Viviane Possamai. Porto Alegre: Bookman, 2005.

PRIETO, Guillermo; LEAHY, Denise. Online social networks and older people. **ICCHP 2012 Proceedings of the 13th international conference on Computers Helping People with Special Needs**, Volume Part I, p. 666-672, 2012.

PRUITT, John; ADLIN, Tamara. The Persona Lifecycle: Keeping People. **Mind Throughout Product Design**. San Francisco: Morgan Kaufmann Publishers, 2006.

QIU, Lin; LIN, Han; RAMSAY, Jonathan; YANG, Fang. You are what you tweet: Personality expression and perception on Twitter. **Journal of Research in Personality**, p. 710-718, 2012.

RABELO, Juliano C.B.; PRUDÊNCIO, Ricardo C.B.; BARROS, Flávia A. Leveraging relationships in social networks for sentiment analysis. **Proceedings of the 18th Brazilian symposium on Multimedia and the web (WebMedia 2012)**, São Paulo-SP, Brasil, p. 181-188, 2012.

REITTER, David; LEBIERE, Christian. Did social networks shape language evolution?: a multi-agent cognitive simulation. **Proceedings of the 2010 Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2010)**, Uppsala, Sweden, p. 9-17, 2010.

REYNOLDS, Bernardo; VENKATANATHAN, Jayant; GONÇALVES, Jorge; KOSTAKOS, Vassilis. Sharing ephemeral information in online social networks: privacy perceptions and behaviours. **Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part III (INTERACT 2011)**, p. 204-215, 2011.

RHODES, Matt. Why do people write reviews? **Freshminds**, 12 mar. 2009. Disponível em: <<http://www.freshminds.net/2009/03/why-do-people-write-reviews/>>. Acesso em: 12 jan. 2015.

RIEDIGER, Michaela. Experience Sampling. **RatSWD Working Paper**, n. 62. Disponível em: <[http://library.mpib-berlin.mpg.de/ft/mr/MR\\_Experience\\_2009.pdf](http://library.mpib-berlin.mpg.de/ft/mr/MR_Experience_2009.pdf)>. Acesso em: 12 jan. 2015.

RILOFF, Ellen; JONES, Rosie. Learning dictionaries for information extraction by multi-level bootstrapping. **Proceedings of Association for the advancement of artificial intelligence (AAAI-99)**, Orlando, Florida, 1999.

ROCHA, Heloisa. V.; BARANAUSKAS, Maria. C. “**Design e Avaliação de Interfaces Humano-Computador**”, NIED / UNICAMP, Campinas-SP, Brasil, 2003.

ROSSI, Rafael G.; REZENDE, Solange O. FEATuRE - Ferramenta para a geração da representação Bag-of-related-words. **Technical Report**, Instituto de Ciências Matemáticas e Computação, São Carlos, n. 367, 2011. Disponível em: <[http://sites.labic.icmc.usp.br/ragero/feature/rt/RT\\_367.pdf](http://sites.labic.icmc.usp.br/ragero/feature/rt/RT_367.pdf)>. Acesso em: 12 jan. 2015.

ROTO, Virpi; OBRIST, Marianna; VÄÄNÄNEN-VAINIO-MATTILA, Kaisa. User Experience Evaluation Methods in Academic and Industrial Contexts. Position paper in **Interact 2009 conference, workshop on User Experience Evaluation Methods (UXEM'09)**, Uppsala, Sweden, 2009.

RUSSELL, Stuart; NORVIG, Peter. **Artificial Intelligence – A modern approach**. Prentice-Hall, 1995.

SALGADO, L.C.C. **Cultural Viewpoint Metaphors to explore and communicate cultural perspectives in cross-cultural HCI design**. 2011. 228 f. Tese (Doutorado em Informática) – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro-RJ, Brasil, 2011.

SALGADO, L.C.C.; LEITÃO, C.F.; DE SOUZA, C.S. **A Journey Through Cultures: Metaphors for Guiding the Design of Cross-Cultural Interactive Systems**. Heidelberg: Springer, 136 p, 2012.

SARDINHA, Tony B. Linguística de corpus: histórico e problemática. **D.E.L.T.A.**, v. 16, n. 2, p. 323-367, 2000.

SEFFAH, Ahmed; DONYAEE, Mohammad; KLINE, Rex B.; PADDA, Harkirat. Usability measurement and metrics: A consolidated model. **Software Quality Journal** 14, p. 159-178, 2006.

SHACKEL, Brian. Usability – context, framework, definition, design and evaluation. In. SHACKEL, Brian & RICHARDSON, Simon, eds. **Human factors for informatics usability**. pp. 21-38, Cambridge, Cambridge University Press, 1991.

SHARODA, A. Paul; LICHAN, Hong; CHI, Ed H. What is a question? crowdsourcing tweet categorization. **Workshop on Crowdsourcing and Human Computation at the Conference on Human Factors in Computing Systems (CHI 2011)**, Vancouver, BC, Canada, May, 2011.

SHI, Na; LEE, Matthew K.O.; CHEUNG, Christy M.K.; CHEN, Huaping. The Continuance of Online Social Networks: How to Keep People Using Facebook?. **Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS 2011)**, p. 2011.

SHNEIDERMAN, B., PLAISANT, C., COHEN, M., and JACOBS, S. **Designing the User Interface: Strategies for Effective Human-Computer Interaction**, 3rd ed. Addison-Wesley Publishing Company, 1998.

SIBONA, Christopher; WALCZAK, Steven. Unfriending on Facebook: Friend Request and Online/Offline Behavior Analysis. **Proceedings of the 2011 44th Hawaii International Conference on System Sciences (HICSS 2011)**, 2011.

SILBERCHATZ, Abraham; KORTH, Henry F.; SUDARSHAN, S. **Sistema de Banco de Dados**. 5. ed. Rio de Janeiro: Elsevier, 2006.

SILVA, Gamhewage C.; AIZAWA, Kiyoharu. Image-based dietary information mining for community creation in a social network. **Proceedings of second ACM SIGMM workshop on Social media (WSM 2010)**, p. 53-58, 2010.

SILVEIRA, M. **Metacomunicação Designer-Usuário na Interação Humano-Computador design e construção do sistema de ajuda**. 147 f. 2002. Tese (Doutorado em) – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro-RJ, Brasil, 2002.

SINGH, Ashutosh. K.; KUMAR, Ravi. A comparative study of page ranking algorithms for information retrieval. **International Journal of Electrical and Computer Engineering**, v. 4, p. 7, 2009.

SOARES, Matheus V. B.; PRATI, Ronaldo C.; MONARD, Maria C. PreText II: Descrição da reestruturação da ferramenta de pré-processamento de textos. **Technical Report, Instituto de Ciências Matemáticas e Computação**, São Carlos, n. 333, 2008. Disponível em: <[http://www.icmc.usp.br/CMS/Arquivos/arquivos\\_enviados/BIBLIOTECA\\_113\\_RT\\_333.pdf](http://www.icmc.usp.br/CMS/Arquivos/arquivos_enviados/BIBLIOTECA_113_RT_333.pdf)>. Acesso em: 12 jan. 2015.

STADDON, Jessica; HUFFAKER, David; BROWN, Larkin; SEDLEY, Aaron. Are privacy concerns a turn-off?: engagement and privacy in social networks. **Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS 2012)**, Article No. 10, Washington, DC, USA, 2012.

STEIN, Dieter; WRIGHT, Susan. **Subjectivity and subjectivisation – Linguistic perspectives**. Melbourne, Australia: Press Syndicate of the University of Cambridge, 1995.

STEIN, Tao; CHEN, Erdong; MANGLA, Karan. Facebook immune system. **Proceedings of the 4th Workshop on Social Network Systems (SNS 2011)**, Article No. 8, 2011.

STIBE, Agnis; KUKKONEN, Harri. Comparative analysis of recognition and competition as features of social influence using twitter. **Proceedings of the 7th international conference on Persuasive Technology: design for health and safety (PERSUASIVE 2012)**, p. 274-279, 2012.

SUN, Beiming; NG, Vincent T. Identifying influential users by their postings in social networks. **Proceedings of the 3rd international workshop on Modeling social media (MSM 2012)**, p. 1-8, 2012.

TABOADA, Maite; BROOKE, Julian; TOFILOSKI, Milan; VOLL, Kimberly; STEDE, Manfred. Lexicon-based methods for sentiment analysis. **Association for Computational Linguistics**, v. 37, n. 2, 2011.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. **Introdução ao Data Mining – Mineração de Dados**. Ciência Moderna Ltda, 900 p, Rio de Janeiro, 2009.

TARALLO, Fernando. **A pesquisa Sociolinguística**, São Paulo: Ática, 1990.

THELWALL, Mike; WILKINSON, David; UPPAL, Sukhvinder. Data mining emotion in social network communication: Gender differences in MySpace. **Journal of the American Society for Information Science and Technology**, v. 61 n. 1, p. 190-199, 2010.

TSOLMON, Bayar; KWON, A-Rong; LEE, Kyung-Soon. Extracting Social Events Based on Timeline and Sentiment Analysis in Twitter Corpus. **Natural Language Processing and Information Systems. Lecture Notes in Computer Science**, v. 7337, p. 265-270, 2012.

TURNEY, Peter D. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. **Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)**, Philadelphia, p. 417-424, 2002.

VIEIRA, Renata. Linguística computacional: fazendo uso do conhecimento da língua. **Entrelinhas**, São Leopoldo, Unisinos, ano 2, n. 4, 2002.

VIEIRA, Renata; LIMA, Vera. L. S. Linguística computacional: princípios e aplicações. In: **IX Escola de Informática da SBC-SI**. Passo Fundo, Maringá, São José. SBC Sul, 2001.

VOEFFRAY, Caroline. **Emotion-sensitive Human-Computer Interaction (HCI): State of the art – Seminar paper**. Seminar emotion Recognition. Disponível em: <<https://diuf.unifr.ch/main/diva/teaching/seminars/emotion-recognition>>, 2011. Acesso em: 12 de Janeiro de 2015.

WANG, Na; XU, Heng; GROSSKLAGS, Jens. Third-party apps on Facebook: privacy and the illusion of control. **Proceedings of the 5th ACM Symposium on Computer Human Interaction for Management of Information Technology (CHIMIT 2011)**, Boston, MA, USA, 2011.

WIEBE, Janyce M.; BRUCE, Rebecca F.; O'HARA, Thomas P. Development and use of a gold standard data set for subjectivity classifications. **Proc. 37th Annual Meeting of the Assoc. for Computational Linguistics (ACL-99)**, Association for Computational Linguistics, University of Maryland, p. 246-253, 1999.

WILSON, Theresa; WIEBE, Janyce; HOFFMANN, Paul. Recognizing Contextual Polarity: an exploration of features for phrase-level sentiment analysis. **Association for Computational Linguistics**, 2009.

XAVIER, Rogério A.C. **Uma abordagem híbrida para a avaliação da experiência emocional de usuários**. 161 f. 2013. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Ciência da Computação, Departamento de Computação da Universidade Federal de São Carlos, São Carlos-SP, Brasil, 2013.

XAVIER, Rogério A.C.; NÉRIS, Vânia. P.D.A.A Hybrid Evaluation Approach for the Emotional State of Information Systems Users. **14th International Conference on Enterprise Information Systems (ICEIS)**, Wroclaw, Poland, p. 45-53, 2012.

ZHANG, L.; ZHU, J.; YAO, T. An Evaluation of Statistical Spam Filtering Techniques. **ACM Transactions on Asian Language Information Processing** 3, p. 243-269, 2004.

## APÊNDICES

### APÊNDICE A – Questionário pós-análise das postagens

#### Questionário - Análise de postagens

Sistema Social analisado:

Nome:

Idade:

Curso:

Semestre:

**1) Você teve dificuldade em classificar as postagens se são ou não relacionadas ao uso?**

---

**2) Se sim, qual sua principal dificuldade?**

---

---

---

**3) Quanto tempo você levou para fazer a classificação?**

---

---

---

**4) Teve alguma postagem que lhe chamou atenção? Por quê?**

---

---

---

**5) O que você percebeu durante esta análise?**

---

---

---

---

**APÊNDICE B – Questionário de classificação de metas de UUX nas postagens**

Sistema Social analisado:

Nome:

Idade:

Curso:

Semestre:

**1) Qual o sentimento que você percebeu com maior frequência nas postagens?**

---

---

---

---

---

---

---

---

**2) Quais as principais reclamações (problemas encontrados no sistema) e os principais elogios (benefícios do sistema) percebido nas mensagens?**

---

---

---

---

---

---

---

---

**3) Outras observações percebidas:**

---

---

---

---

---

---

---

---

## APÊNDICE C – Quadro de metas de UUX

### Metas de UUX

Metas	Descrição	Referência
Eficácia	Ser eficaz no uso. O sistema precisa ser bom naquilo que se espera dele. Exatidão e completude com que os usuários atingem seus objetivos específicos.	Preece, Rogers, Sharp, 2005; ISO 9241-11, 2011
Eficiência	Ser eficiente no uso. Uma vez que os usuários já saibam operar o sistema, sua produtividade tem que ser melhorada. Recursos gastos em relação à exatidão e completude com que os usuários atingem seus objetivos.	Preece, Rogers, Sharp, 2005; ISO 9241-11, 2011
Segurança	Ser segura no uso. O sistema deve prevenir que o usuário cometa erros e, caso cometa erros, deve ser reversível.	Preece, Rogers, Sharp, 2005;
Utilidade	Ser de boa utilidade. O sistema deve fornecer funcionalidades úteis para que o usuário realize suas tarefas.	Preece, Rogers, Sharp, 2005;
Aprendizado	Ser fácil de aprender. As funções fundamentais do sistema devem ser de fácil assimilação.	Preece, Rogers, Sharp, 2005;
Memorização / Memorização	Ser fácil de lembrar como se usa. Os usuários devem poder lembrar ou, ao menos, rapidamente lembrados das funções fundamentais do sistema.	Preece, Rogers, Sharp, 2005;
Satisfação / Satisfatório	Atitudes positivas e ausência de desconforto em relação ao uso do produto.	ISO 9241-11 (2011)
Prazer, diversão	Medida em que os usuários estão satisfeitos com a sua realização percebida, por metas hedônicas de estimulação, identificação, evocação e respostas emocionais associados.	Bevan, 2008
Conforto	Medida em que os usuários estão satisfeitos com o conforto físico.	Bevan, 2008
Confiança	Medida em que os usuários estão satisfeitos que o produto vai se comportar conforme o esperado.	Bevan, 2008
Antecipação	O que os usuários esperam.	Ketola e Roto (2008)
Usabilidade	Medida em que um sistema, produto ou serviço pode ser usado por usuários específicos para se atingir objetivos específicos com eficácia, eficiência e satisfação em um determinado contexto de uso.	ISO 9241-11 (2011)
Hedônico	Satisfação das necessidades internas, tais como prazer, diversão, ou coisas preventivas disso, como frustração. Esta meta abrange totalmente enjoyment, fun e frustração, e seria considerado abrangido pela meta hedônica	Ketola e Roto (2008); Bargas-Avila e Hornbæk (2011)
Suporte (apoio)	Apoio do ser humano ou software de apoio disponíveis e como isso afeta a satisfação do usuário, por exemplo, possíveis devoluções de produtos, ou listas de desejos dos usuários.	Ketola e Roto (2008)
Impacto	se e como o novo dispositivo de mudar os padrões de utilização dos usuários.	Ketola e Roto (2008)

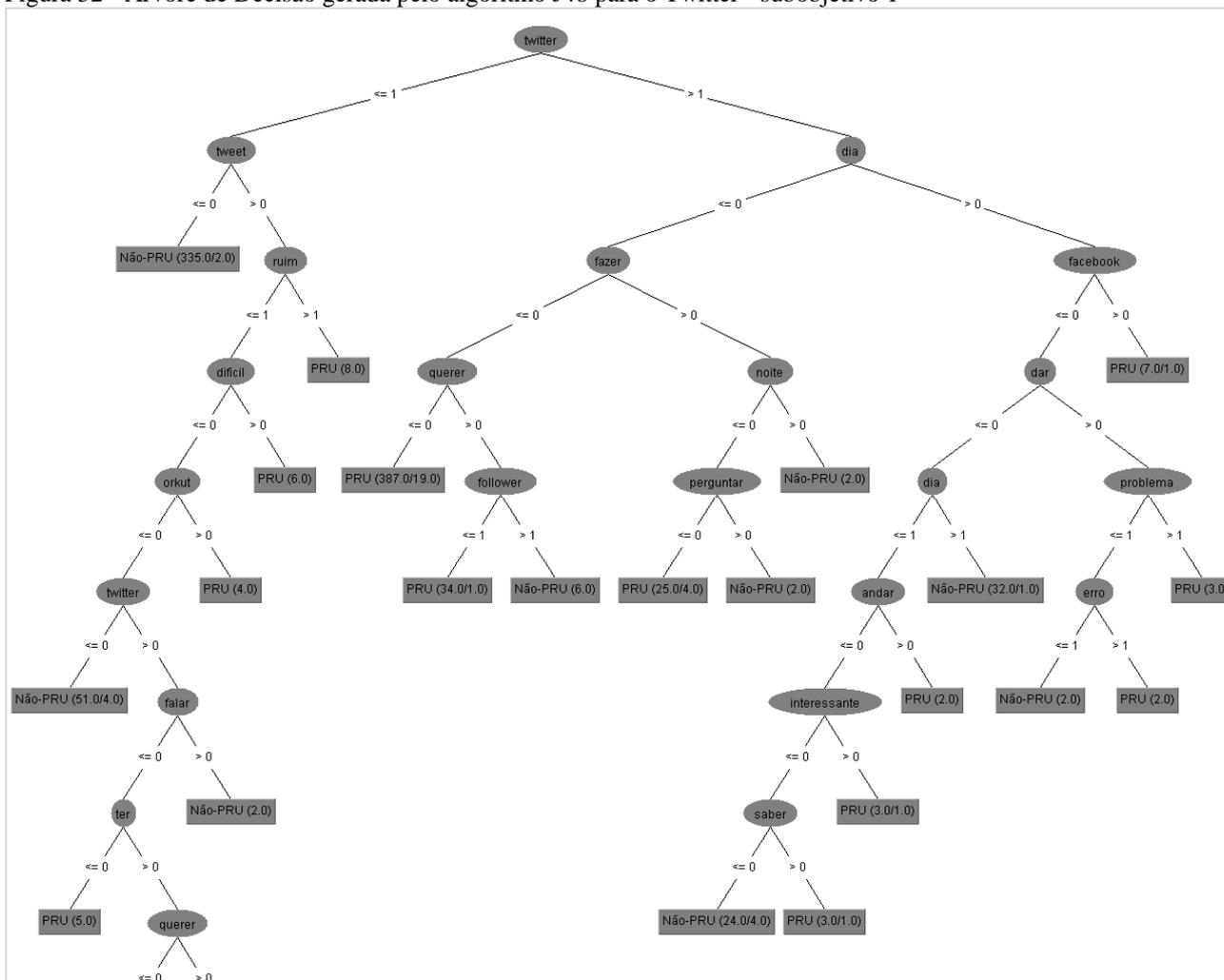
<b>Metas</b>	<b>Descrição</b>	<b>Referência</b>
Afeto e emoção	Afeto e emoção induzido pelo uso de um produto, incluindo outros aspectos como enjoyment, fun e frustração.	Bargas-Avila e Hornbæk (2011)
Estética, atração	Apreciação da beleza ou bom gosto. Normalmente associado com gráficos, som, interface.	Bargas-Avila e Hornbæk (2011)
Engajamento, fluidez	Quão engajado é o usuário no uso do produto esquecendo todo o resto? Também inclui desafios versus a habilidade necessária para alcançar o estado de fluxo.	Bargas-Avila e Hornbæk (2011)
Motivação	O que motiva o usuário no uso do produto (tarefa)	Bargas-Avila e Hornbæk (2011)
Encantamento	Ser “ambos apanhados e levados” na experiência, esquecendo todo o resto e causando uma desorientação associada a uma sensação prazerosa de plenitude e vivacidade que cobra atenção e concentração.	Bargas-Avila e Hornbæk (2011)
Frustração	Frustração: Frustração ou sofrimento induzido pela utilização do produto. Esta é também uma meta hedônica negativa.	Bargas-Avila e Hornbæk (2011)
Acessibilidade	Usabilidade de um produto, serviço, ambiente ou facilidade por pessoas com a mais ampla gama de capacidades	ISO 9241-171

## APÊNDICE D – Árvores de decisão geradas pelo algoritmo J48 no WEKA

### Subobjetivo 1: padrões de extração de PRUs

#### Twitter

Figura 52 - Arvore de Decisão gerada pelo algoritmo J48 para o Twitter - subobjetivo 1



## SIGAA

Figura 53 - Trecho 1 da árvore de decisão para o SIGAA - subobjetivo 1

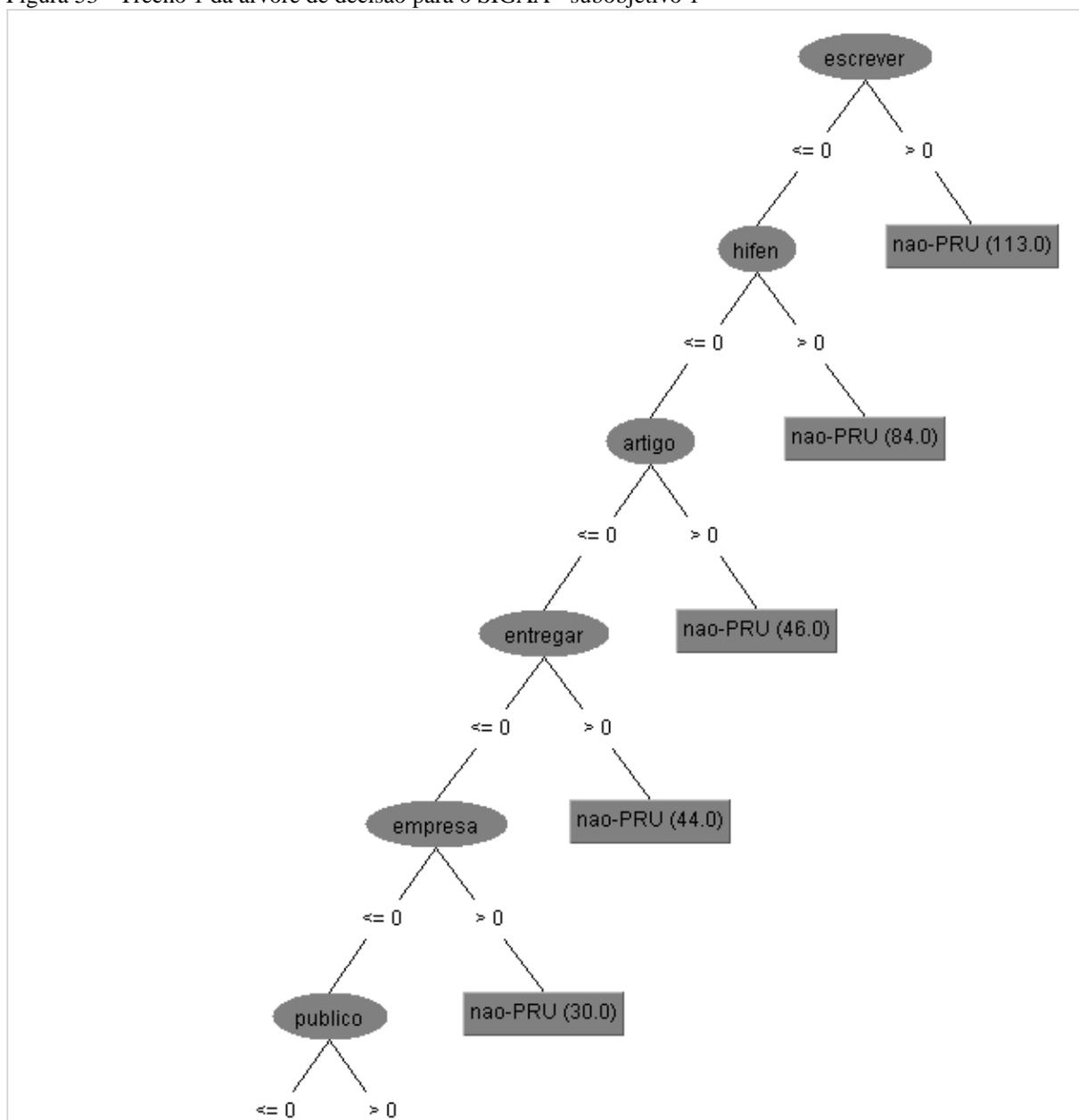


Figura 54 - Trecho 2 da árvore de decisão para o SIGAA - subobjetivo 1

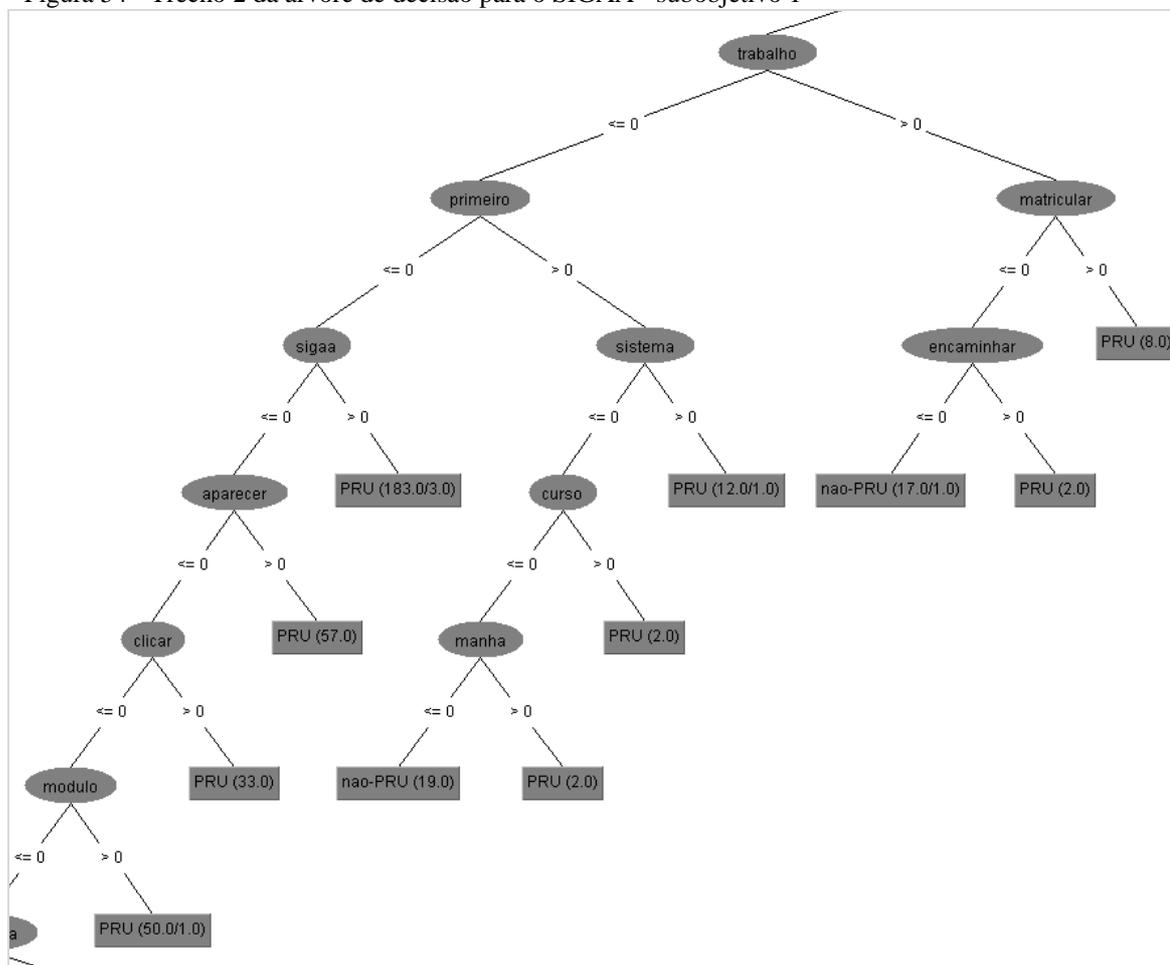
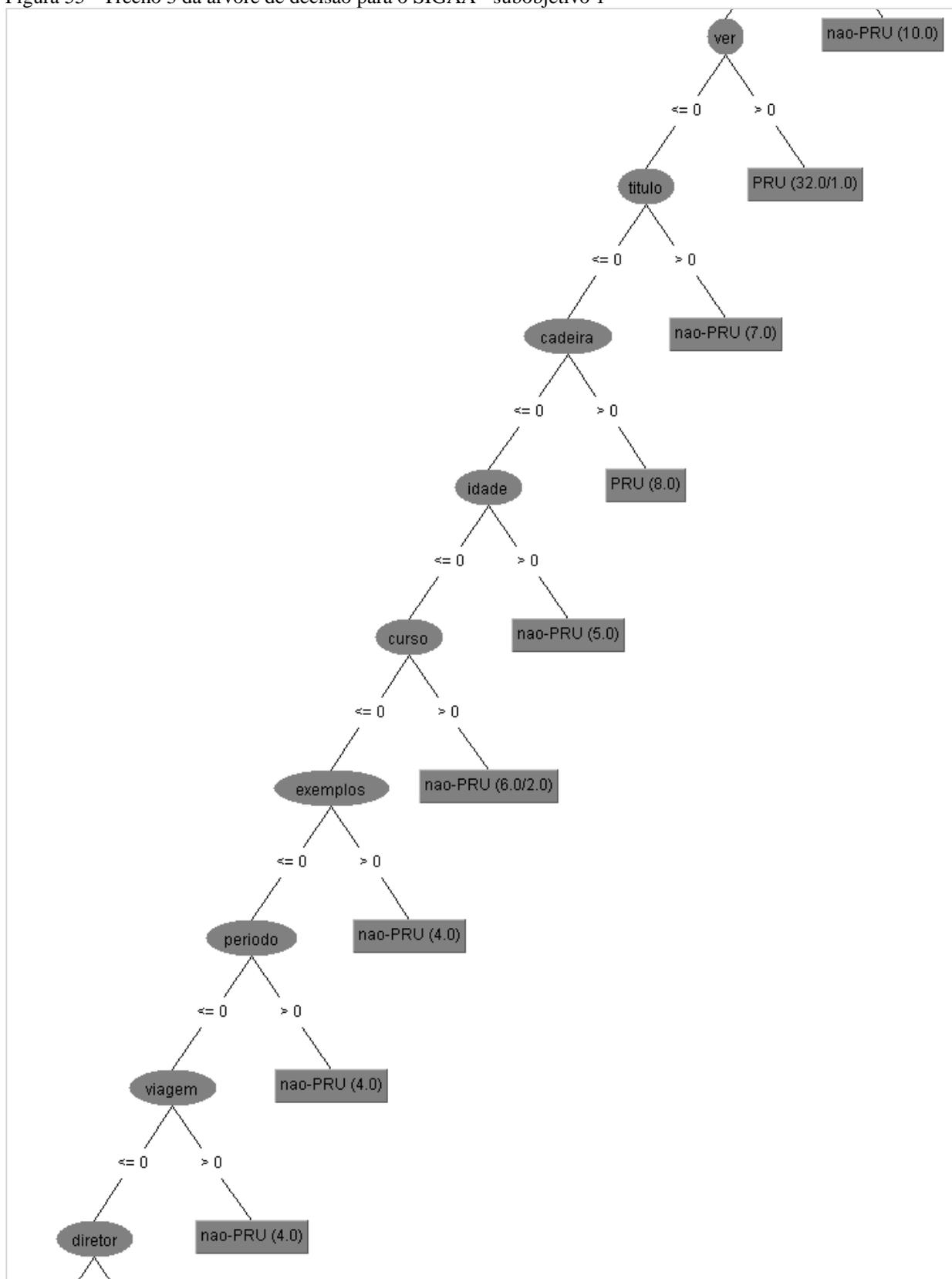


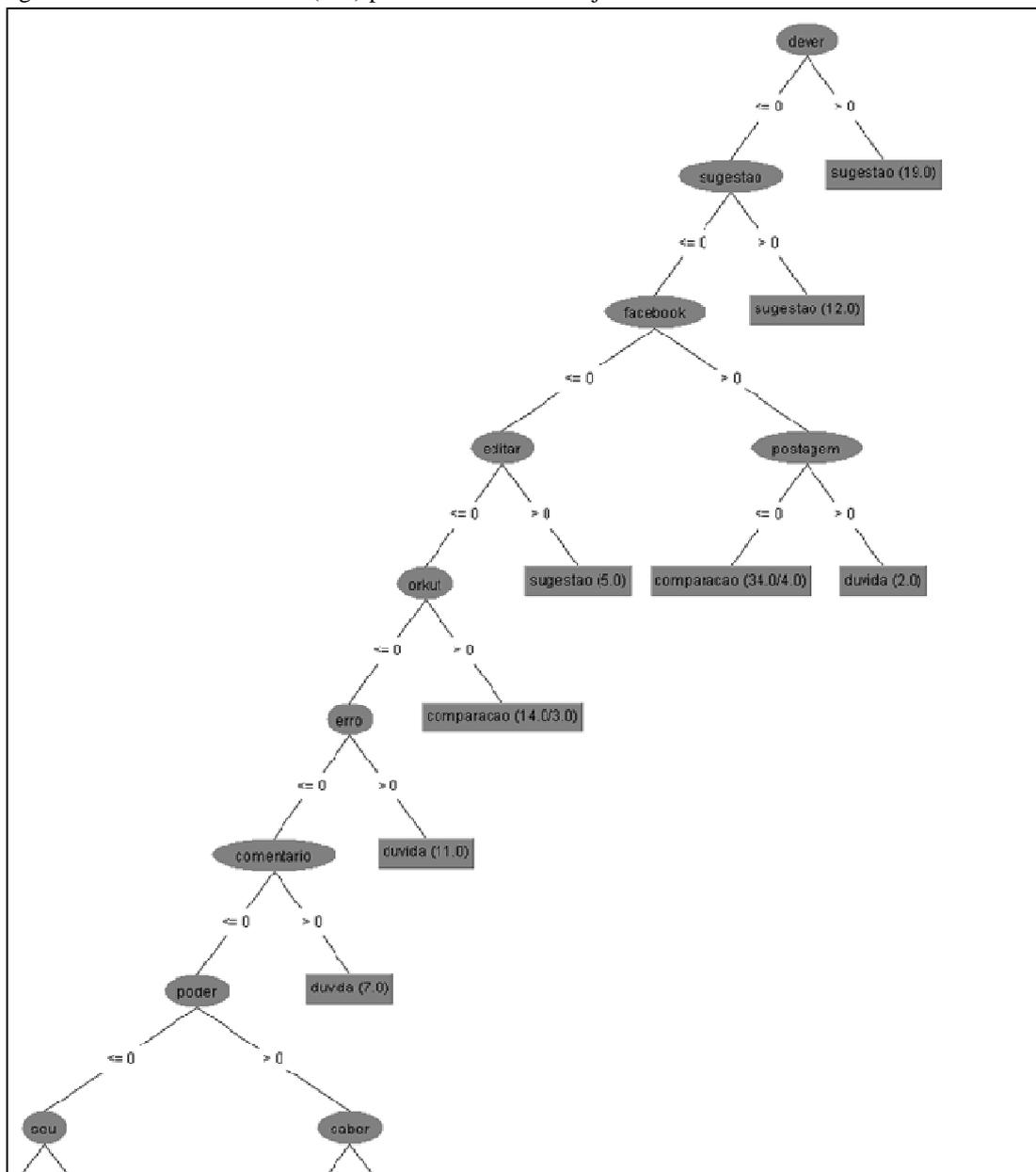
Figura 55 - Trecho 3 da árvore de decisão para o SIGAA - subobjetivo 1



## Subobjetivo 2: Extração de um vocabulário associado para cada tipo de PRU

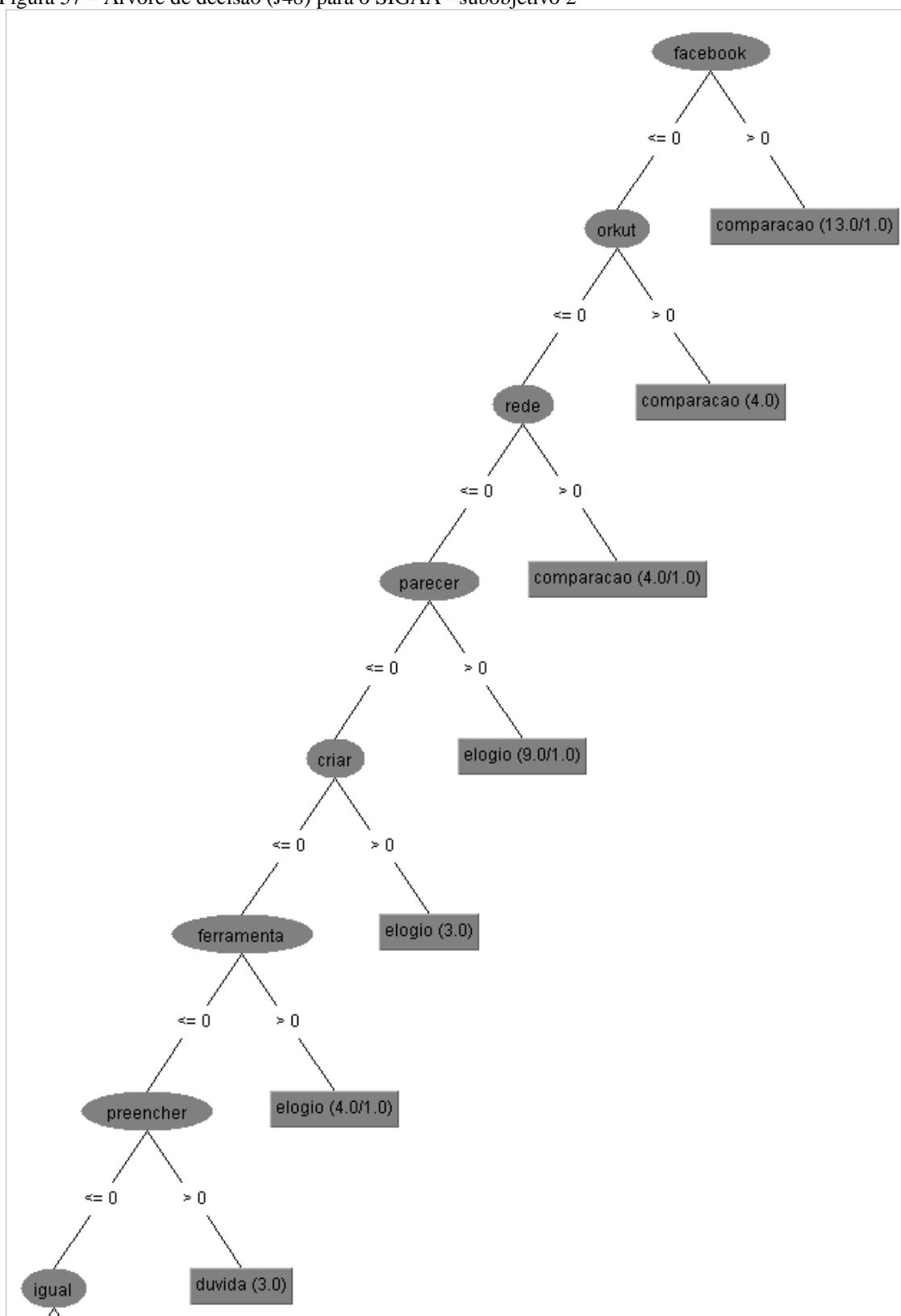
### Twitter

Figura 56 – Árvore de decisão (J48) para o Twitter - subobjetivo 2



## SIGAA

Figura 57 – Árvore de decisão (J48) para o SIGAA - subobjetivo 2



### Subobjetivo 3: extração de um vocabulário associado para cada dimensão UUX

Figura 58 – Trecho 1 da árvore de decisão (J48) para o subobjetivo 3 - Grupo 1

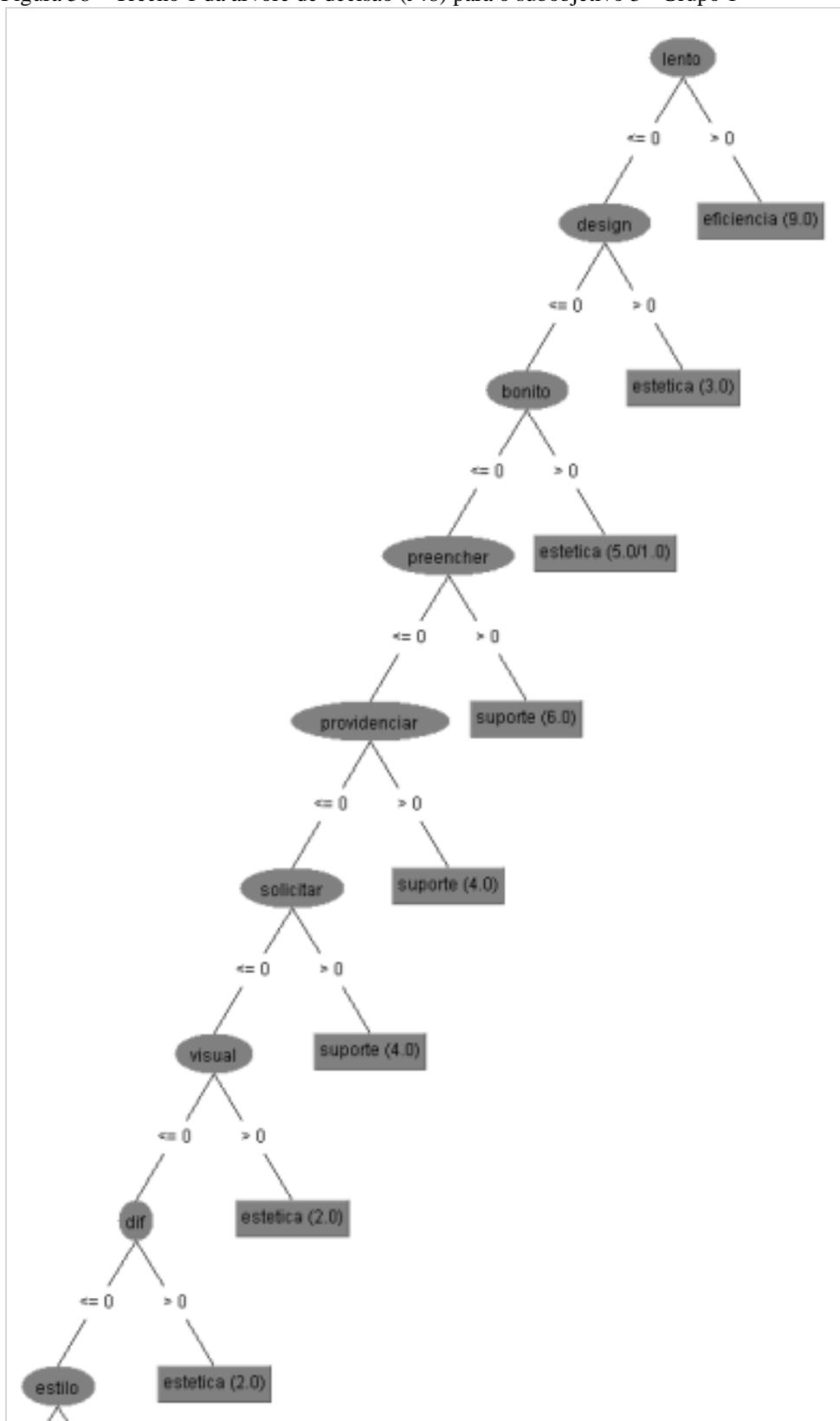


Figura 59 – Trecho 2 da árvore de decisão (J48) para o subobjetivo 3 - Grupo 1

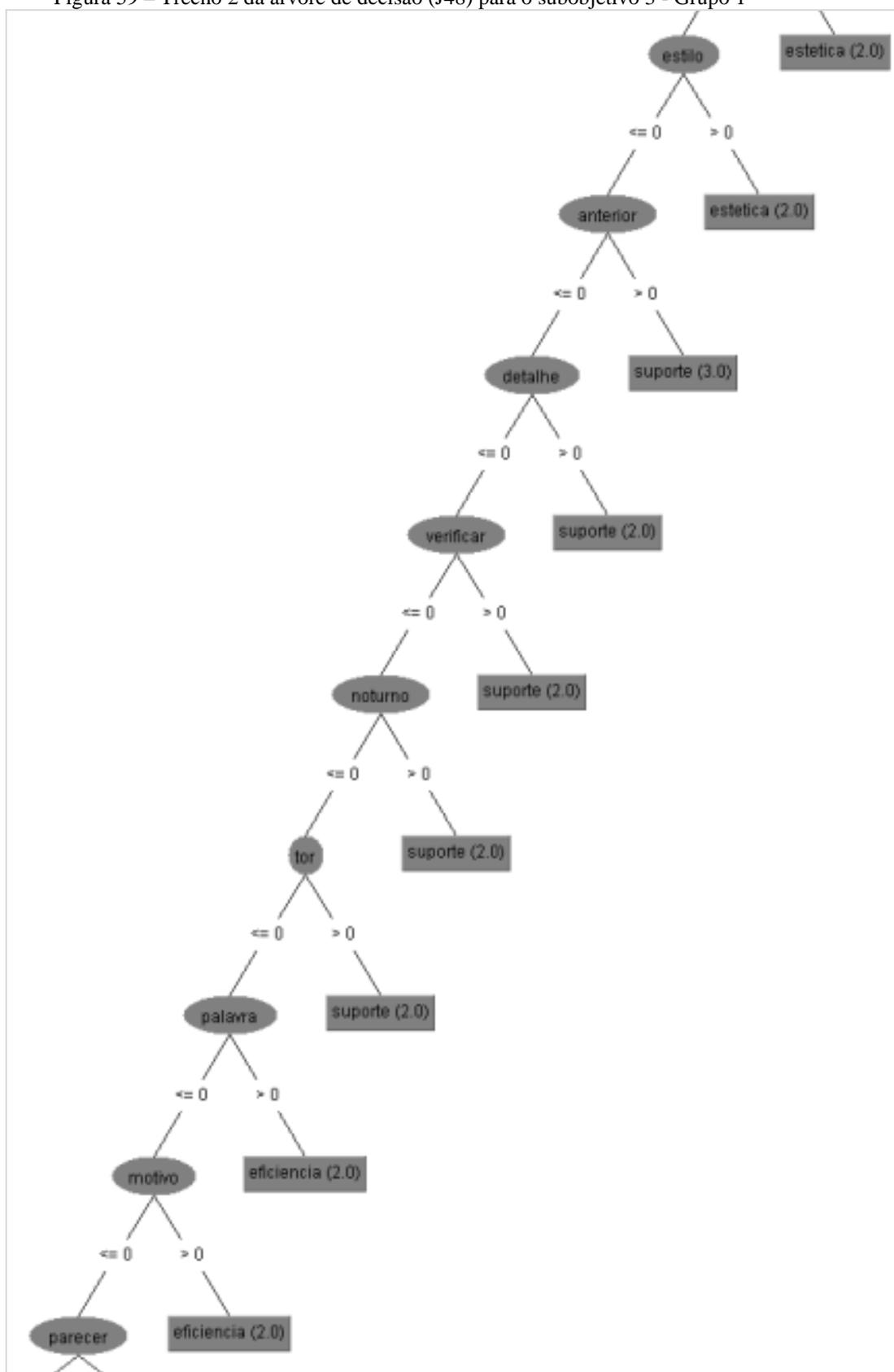


Figura 60 – Trecho da árvore de decisão (J48) para o subobjetivo 3 - Grupo 2

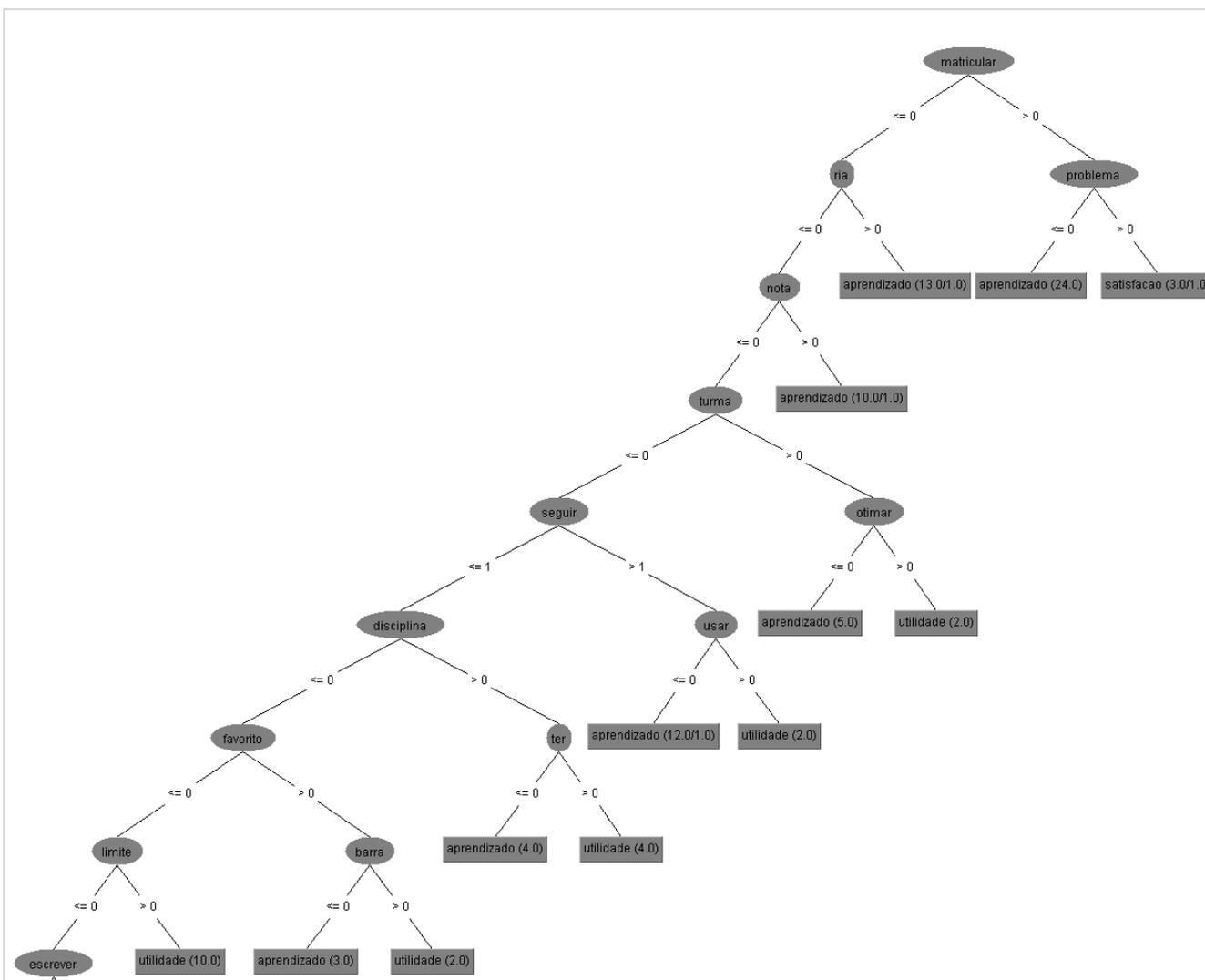


Figura 61 – Trecho da árvore de decisão (J48) para o subobjetivo 3 - Grupo 3

