

## Taxonomia de traços de personalidade de usuários das redes sociais

### Taxonomy of personality traits of social network users

DOI:10.34117/bjdv8n6-241

Recebimento dos originais: 21/04/2022

Aceitação para publicação: 31/05/2022

**Cecy Maria Martins Brito**

Mestra em Educação (UFVJM)

Instituição: Favenorte, Porteirinha- MG FAVEPORT

Endereço: Rua Reginaldo Ribeiro, 161, Sala 103 Centro, Montes Claros - MG

E-mail: cecymartins2004@hotmail.com

**Marcus Vinicius Carvalho Guelpli**

Dr. em Computação

Instituição: Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM)

Endereço: Rua Alexandrita 102, Diamantina - MG

E-mail: marcus.guelpli@ufvjm.edu.br

#### RESUMO

O presente trabalho de pesquisa trata-se de uma pesquisa de cunho bibliográfico e quantitativo, cujo *corpus* foi composto por textos escritos por usuários da rede social *Facebook* e que teve como objetivo geral avaliar a hipótese de a taxonomia de textos ser capaz de contribuir para a identificação de traços de personalidade em textos de usuários dessa rede social, por meio do modelo computacional Cassiopeia. Para isso, foram analisados textos publicados pelos usuários, por meio da seleção de grupos, fechados e secretos, de pessoas que se dizem usuárias do *Facebook*. Utilizando o Cassiopeia, tais textos foram classificados em um determinado traço de personalidade, de acordo com o Big Five – ou os cinco fatores da personalidade (extroversão / introversão; neuroticismo / estabilidade emocional; socialização / agradabilidade; escrupulosidade / conscienciosidade ou realização; e abertura para experiência). A partir dessa classificação, pôde-se identificar esses traços e, por meio deles, agrupar, de maneira automática, inúmeros textos de usuários, além de descobrir conhecimentos implícitos, criar um conjunto de palavras que classificam os grupos de usuários e identificar o traço de personalidade característico destes. Os textos foram selecionados utilizando a mineração de textos para criar uma taxonomia de traços de personalidade. Os resultados da pesquisa foram considerados positivos, porque o procedimento computacional acima descrito agrupou os 100 textos selecionados em 26 *clusters* (pastas) a partir da taxonomia de traços de personalidade criada.

**Palavras-chave:** redes sociais, taxonomia de perfil, mineração de textos, traços de personalidade.

#### ABSTRACT

This work is a bibliographic quantitative research, whose *corpus* is composed of written texts published by Facebook users. This research had as general objective to evaluate the hypothesis of a taxonomy of texts be capable of contributing to the identification of personality traits within the texts analyzed, through the computational model Cassiopeia.

In order to do it, several texts published were analyzed, through the selection of groups, both closed and secret, of people who identify themselves as Facebook users. By using Cassiopeia, those texts were classified into determined personality trait groups, according to the Big Five theory, also known as the five factors of personality (extroversion, agreeableness, openness, conscientiousness, and neuroticism). Based on that classification, it was possible to identify those personality traits, and, through them, automatically cluster countless texts from the *corpus*; to discover some implicit knowledge; to create a set of words that classify the users and identify their characteristic personality traits. The texts were selected by using text mining to create a personality traits taxonomy. The results of this research were considered positive, because the computational procedure described above clustered 100 selected texts into 26 clusters based on the personality traits taxonomy created.

**Keywords:** social networks, profile taxonomy, text mining, personality traits.

## 1 INTRODUÇÃO

A solução de problemas do dia a dia está atualmente pautada na constante busca das pessoas por economizar tempo. Nesse contexto, esta pesquisa foi desenvolvida com material coletado na rede social *Facebook*, utilizando o modelo computacional Cassiopeia, que é um agrupador de texto hierárquico<sup>1</sup>. Com base em fundamentações teóricas de autores da Psicologia, foi feita uma taxonomia de traços de personalidade de usuários do *Facebook*.

Para constituir os grupos deste estudo, realizou-se uma busca com os nomes específicos de cada traço de personalidade, além disso, foi solicitada a participação em pelo menos 15 grupos dessa rede social para cada traço de personalidade disponibilizado no modelo Big Five (ou cinco grandes fatores da personalidade). Esse modo de usar os traços de personalidade do Big Five nas redes sociais constitui-se em uma contribuição da presente pesquisa para áreas como a da Educação, sobretudo quanto ao ensino a distância – EaD.

A questão-problema que a presente pesquisa abordou foi: a taxonomia de textos, baseada no Cassiopeia, pode revelar traços da personalidade de usuários do *Facebook*? Portanto, traçaram-se ramificações de análise, as quais apontavam para os seguintes objetivos específicos: agrupar centenas de usuários, utilizando o Cassiopeia, a fim de contribuir para a compreensão de traços de personalidade dos usuários dessa rede social;

---

<sup>1</sup> É preciso destacar que há, atualmente, nas redes sociais, uma ferramenta de direcionamento de anúncios e propagandas, que pode ser utilizada pelo usuário apenas se ele pagar por esse serviço. Esse recurso direciona a publicidade para determinado público-alvo, conforme sua localização, interesses específicos, faixa etária e gênero. Essa informação foi extraída do próprio *Facebook*.

classificar traços de personalidade dos usuários do Facebook de maneira automática; detectar conhecimentos implícitos nos textos analisados; criar um conjunto de palavras que classifique os grupos de usuários.

A hipótese era que, com a elaboração da taxonomia de traços de personalidade, utilizando o modelo Big Five, seria possível ter uma visão categorizada dos usuários do Facebook. Para tanto, escolheu-se o Cassiopeia porque ele possibilita um agrupamento dos usuários em conjuntos, por meio de palavras, facilitando, assim, a busca de conhecimento na grande base textual dessa rede social.

## 2 FUNDAMENTAÇÃO TEÓRICA

O primeiro teórico a enfatizar o desenvolvimento da personalidade, especialmente a dos bebês, foi Freud, segundo Hall *et al.* (2000). Ele sustentava que a primeira infância desempenha um papel decisivo na formação da estrutura básica da personalidade das pessoas. De acordo com Hall *et al.* (2000), “a personalidade está relacionada a todos os comportamentos do indivíduo, tanto os manifestos quanto os que acontecem sob a pele”. (HALL, 2000, p. 258). Dessa forma, pode-se dizer que os traços de personalidade são perceptíveis nas pessoas, uma vez que elas demonstram diversos sentimentos.

Para Pervin e John (2004), a personalidade representa características que explicam padrões consistentes de sentimentos, pensamentos e comportamentos. O que é particularmente interessante é que esses pensamentos, sentimentos e comportamentos estão interligados para formar um indivíduo único. Assim, a personalidade está relacionada aos comportamentos de uma pessoa e é construída em seu desenvolvimento, acompanhando-a para o resto da vida.

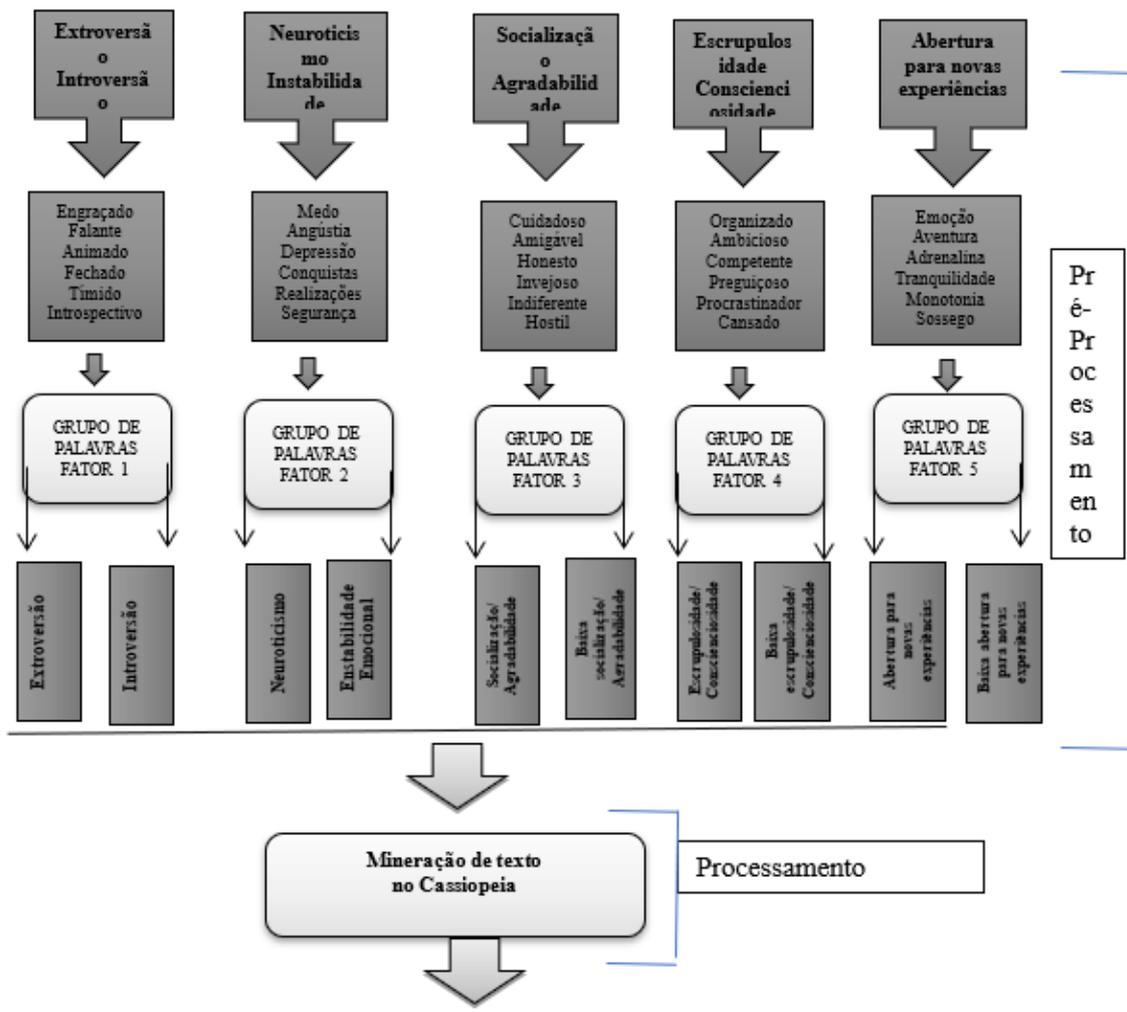
O modelo dos cinco grandes fatores da personalidade ou Big Five, apresenta cinco traços formadores da personalidade: extroversão / introversão; neuroticismo / estabilidade emocional; socialização / agradabilidade, escrupulosidade/ conscienciosidade ou realização; abertura para novas experiências.

## 3 MÉTODO PROPOSTO

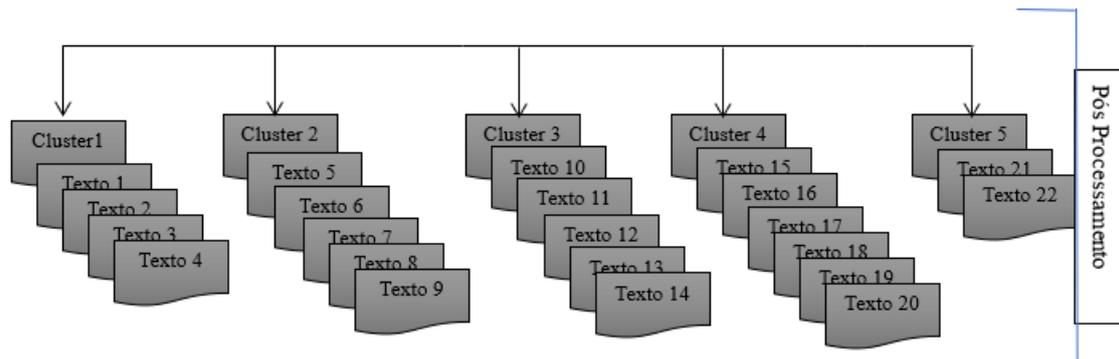
Para iniciar o estudo, no primeiro momento foi feita uma pesquisa bibliográfica na área da Psicologia para buscar características que apontassem os traços de personalidade, quando se chegou aos cinco fatores da personalidade ou Big Five. Em seguida, fez-se uma busca no *Facebook*, para procurar grupos que se enquadravam em cada um desses fatores.

Após o ingresso nos grupos específicos, por meio de buscas nos textos escritos pelos membros e ali postados, foi selecionado, para cada um dos cinco traços de personalidade, um conjunto de palavras que apareciam com maior frequência. Para escolher essas palavras, foi feito antes um estudo dos traços de personalidade de cada perfil do Big Five. Foram, então, extraídos dos grupos 10 textos para cada traço de personalidade, totalizando um *corpus* de 100 textos a serem analisados<sup>2</sup>.

Figura 1 – Modelo cognitivo computacional – Separação dos textos no modelo Cassiopeia



<sup>2</sup> É importante ressaltar que alguns grupos foram descartados por não corresponderem exatamente a esses critérios esperados, ou seja, os assuntos tratados entre os membros não estavam de acordo com o que expõem os teóricos estruturadores da pesquisa e/ou não havia frequência no uso de palavras características a cada traço de personalidade.



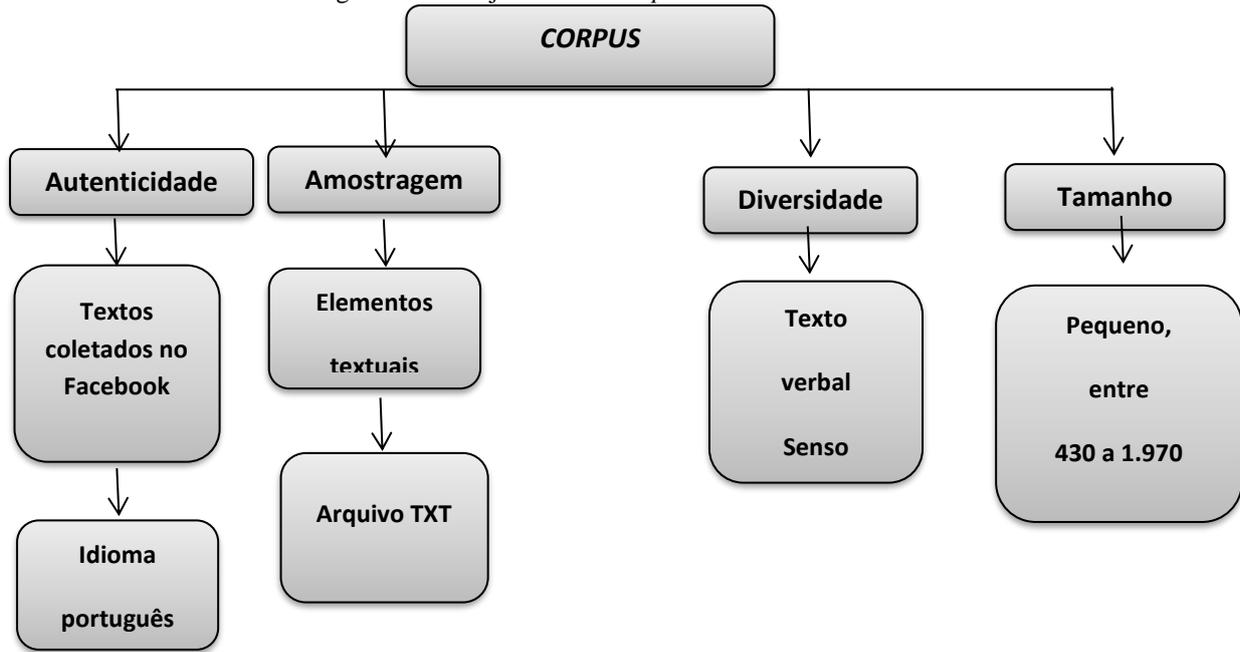
Fonte: Elaborado pela autora, 2020.

Por meio do Cassiopeia, realizou-se o processo de mineração dos textos selecionados, que foi iniciado com a entrada dos 100 textos no formato TXT. No início do procedimento, o modelo Cassiopeia identifica os textos de cada traço de personalidade por meio da base de dados, formada pelo conjunto das palavras escolhidas, realizando, assim, a taxonomia desses traços de personalidade.

### 3.1 CONSTRUÇÃO DO *CORPUS*

A partir desses traços de personalidade, foram criados conjuntos de palavras para cada um deles. Esses conjuntos foram criados a partir da análise de cada grupo, observando as palavras que mais se repetiam nos textos escritos pelos usuários e que apresentavam algum grau de afinidade com o traço de personalidade em questão, conforme a descrição feita pelos teóricos do Big Five. Formados os conjuntos de palavras, criou-se o *corpus* da pesquisa, extraindo-se 10 textos de cada grupo específico daquela rede social para cada traço de personalidade. Totalizaram-se 100 (cem) textos, que foram transformados para o formato TXT. Após essa etapa, iniciou-se a mineração de texto no Cassiopeia.

Figura 2 – Planejamento do *corpus* usado neste trabalho



Fonte: Elaborado pela autora, 2020.

Para o planejamento do *corpus*, observou-se: autenticidade, amostragem, diversidade e tamanho. Como se trata de textos retirados das redes sociais, não são longos e demonstram objetividade na mensagem que o usuário transmite aos leitores.

Para saber o tamanho total dos conjuntos de textos estatisticamente, em número de palavras e caracteres, foi usada a ferramenta FineCount Free. Foram utilizados 10 (dez) textos de cada traço de personalidade. A Tabela 1, a seguir, mostra a estatística do *corpus* utilizado neste trabalho.

Tabela 1 – Estatísticas dos 100 (cem) textos retirados da rede social Facebook, compostos pelos Cinco Fatores da Personalidade ou Big Five

Traços de personalidade	Arquivos	Caracteres	Caracteres e espaços	Apenas palavras	Palavras e números	% de números	Frases	Linhas	Páginas	Repetições de frases
Todos	1 (100 textos)	52.957	64.384	11.234	11.293	0,52 %	726	1.170,62	62,88	0,55%
Extroversão	1 (10 textos)	2.351	2.837	468	468	0,00 %	34	51,58	2,77	2,94%
Introversão	1 (10 textos)	8.841	10.714	1.853	1.855	0,11 %	104	194,80	10,46	0,00%
Neuroticismo	1 (10 textos)	8.784	10.774	1.970	1.981	0,56 %	119	195,89	10,52	0,00%
Estabilidade emocional	1 (10 textos)	5.373	6.552	1.147	1.149	0,17 %	70	119,13	6,40	1,43%
Socialização/	1 (10 textos)	4.657	5.637	973	981	0,82 %	64	102,49	5,50	0,00%

Agradabilidade										
Baixa socialização/ Agradabilidade	1 (10 textos)	4.438	5.391	925	926	0,11 %	43	98,02	5,26	0,00%
Escrupulosidade/ Conscienciosidade	1 (10 textos)	6.894	8.302	1.389	1.409	1,42 %	99	150,95	8,11	0,00%
Baixa escrupulosidade/ Baixa conscienciosidade	1 (10 textos)	2.557	3.140	562	562	0,00 %	44	57,09	3,07	0,00%
Abertura p/novas experiências	1 (10 textos)	7060	8.621	1517	1532	0,98 %	116	156,75	8,42	1,72%
Abertura p/novas experiências	1 (10 textos)	2002	2446	430	430	0,00 %	33	44,47	2,39	0,00%

Fonte: Elaborado pela autora

A primeira linha da Tabela 1 apresenta os 100 textos extraídos do Facebook, os quais tinham sido publicados no período de 20/08/2019 a 10/02/2021. Juntos, eles totalizaram 11.234 (onze mil, duzentas e trinta e quatro) palavras; 726 (setecentas e vinte e seis) frases; 1.170,62 linhas (mil cento e setenta linhas e sessenta e duas frações de linha). Pode-se perceber uma variação no tamanho do conjunto dos textos entre a quarta e a décima primeira linhas da Tabela 1. O maior conjunto de textos utilizado (neuroticismo) tem 1.970 (mil novecentas e setenta) palavras, 119 (cento e dezenove) frases e 195,89 linhas (cento e noventa e cinco linhas e oitenta e nove frações de linha). Já o menor conjunto de textos (baixa abertura para novas experiências), apresentou 430 (quatrocentas e trinta) palavras, 33 (trinta e três) frases e 44,47 linhas (quarenta e quatro linhas e quarenta e sete frações de linha).

### 3.2 MODELO COMPUTACIONAL CASSIOPEIA

Para que seja feita a mineração de textos. Morais e Ambrósio (2007) destacam que, após o surgimento da internet, os formatos de texto mais utilizados na mineração são HTML e XML. Outros fatores importantes a serem citados na mineração de textos são a

indexação e a normalização. Morais e Ambrósio (2007) ressaltam que esses processos “facilita[m] a identificação de similaridade de significado entre suas palavras, considerando [...] variações morfológicas e problemas de sinonímia”. (MORAIS; AMBRÓSIO, 2007, p.13).

Quanto ao Cassiopeia, de acordo com Guelpli (2012, p. 47), ele “foi proposto para ser um agrupador de texto hierárquico” e contém três grandes etapas: pré-processamento; processamento; pós-processamento.

Cada um dos 26 *clusters* fez uma separação diferente, por isso há uma variação de quantidade de textos por *cluster*, conforme demonstrado na Tabela 2.

Tabela 2 – Separação dos 26 *clusters*

Numeração do <i>cluster</i>	Quantidade de traços	Traços e respectivas quantidades de textos
<i>Cluster 1</i>	1	Neuroticismo – 2 textos
<i>Cluster 2</i>	1	Extroversão – 10 textos
<i>Cluster 3</i>	1	Neuroticismo – 2 textos
<i>Cluster 4</i>	2	Escrupulosidade – 1 texto / Abertura para experiências – 1 textos
<i>Cluster 5</i>	2	Escrupulosidade – 1 texto / Neuroticismo – 1 texto
<i>Cluster 6</i>	3	Baixa escrupulosidade – 4 textos / Baixa socialização – 3 textos / Escrupulosidade – 1 texto
<i>Cluster 7</i>	1	Baixa escrupulosidade – 2 textos
<i>Cluster 8</i>	3	Baixa socialização – 3 textos / Baixa escrupulosidade – 3 textos / Escrupulosidade – 1 texto
<i>Cluster 9</i>	2	Neuroticismo – 1 texto / Baixa socialização – 1 texto
<i>Cluster 10</i>	2	Estabilidade emocional – 4 textos / Abertura para experiências – 1 textos
<i>Cluster 11</i>	1	Estabilidade emocional – 2 textos
<i>Cluster 12</i>	1	Socialização – 4 textos
<i>Cluster 13</i>	2	Escrupulosidade – 2 textos / Introversão – 1 texto
<i>Cluster 14</i>	2	Socialização – 4 textos / Abertura para experiências – 1 texto
<i>Cluster 15</i>	2	Escrupulosidade – 1 texto / Estabilidade emocional – 1 texto
<i>Cluster 16</i>	1	Introversão – 2 textos
<i>Cluster 17</i>	2	Socialização – 2 textos / Abertura para experiências – 1 texto
<i>Cluster 18</i>	7	Baixa abertura para experiências – 10 textos / Baixa escrupulosidade – 1 texto / Baixa socialização – 3 textos / Escrupulosidade – 2 textos / Abertura para experiências – 1 texto / Introversão – 1 texto / Neuroticismo – 2 textos
<i>Cluster 19</i>	2	Introversão – 1 texto / Escrupulosidade – 1 texto
<i>Cluster 20</i>	2	Introversão – 1 texto / Neuroticismo – 1 textos
<i>Cluster 21</i>	2	Estabilidade emocional – 1 texto / Abertura para experiências – 1 texto
<i>Cluster 22</i>	2	Introversão – 1 textos / Abertura para experiências – 2 textos
<i>Cluster 23</i>	2	Abertura para experiências – 1 texto / Neuroticismo – 1 texto

Cluster 24	2	Introversão – 1 texto / Abertura para experiências – 1 texto
Cluster 25	1	Estabilidade emocional – 2 textos
Cluster 26	1	Introversão – 2 textos

Fonte: elaborado pela autora 2021.

#### 4 RESULTADOS

Neste tópico são apresentados os resultados da presente pesquisa, de acordo com o teste feito no modelo computacional Cassiopeia, que fez uma separação de 26 *clusters*, sendo que cada um deles fez uma separação de textos com seus grupos de palavras.

A seguir, são descritos cada um dos 26 *clusters* conforme os traços de personalidade do Big Five e a quantidade de textos que separou.

Tabela 3 – Resultados dos 26 *clusters*

Numeração do <i>cluster</i>	Quantidade de traços	Traços de personalidade e respectivas quantidades de texto	Quantidade de palavras para cada traço	Palavras encontradas nos textos
Cluster 1	1	Neuroticismo – 2 textos	23	sentimentos, corto, soffro, depressão, afetando, agradar, corta, culpa, decepção, acabando, afastei, afastou, afetado, aguento, atormentando, buraco, choro, conseguindo, consigo, constante, continuar, desistir, difícil.
Cluster 2	1	Extroversão – 10 textos	87	amigos, amizade, barzinho, carismático, corro, curto, divertido, esportes, extrovertido, fazer, gosto, música, legal, palhaço, passear, viver, vida, sair, rir, restaurante, bastante, cheio, pedalar, <i>bike</i> , encontrarmos, feliz, gente, precisamos, momento, apreciar, sábado, tomar uma, pagode, sertanejo, alegria, felicidade, sorriso, amores, distrair, conversar, comer, churrasco, amor, balada, bares, garantida, risadas, fim de semana, incrível, adoro, aproveita, pessoas, <i>vibe</i> , curtir, vivo, boas, bom, cervejinha, venham, vivendo, abraços, gostar, filmes, cinema, série, assistindo, sorrir, distrair, ser feliz, viagens, comida, companhias, curtir, praia, sol, vinho, amiga, alegria, alegre, amo, frescura, treino, sair, <i>bike</i> , jiu-jitsu, competir, divertida.
Cluster 3	1	Neuroticismo – 2 textos	19	acabar, acabando, acabo, afastei, afastou, afetado, afetando, suicídio, remédios, dor, hospital, choro, dormir, morrer, morte, depressão, difícil, perturbado, arrasada.
Cluster 4	2	Escrupulosidade – 1 texto	11	agenda, <i>business</i> , caro, agir, clientes, comece, começo, conta, corporativo, desafio, dinheiro.
		Abertura para experiências – 1 textos	9	mundo, adrenalina, arriscado, aventura, curiosa, desafiado, descobertas, desejo, emoção.
Cluster 5	2	Escrupulosidade – 1 texto	15	conhecimento, dinheiro, financeiro, gerar, investe, investem, investimento, investir, poupança, poupar, precisamos, próprio, públicos, rico, condições.
		Neuroticismo – 1 texto	8	matar, destruída, acabar, chorando, depressão, desistisse, fraca, lágrimas.

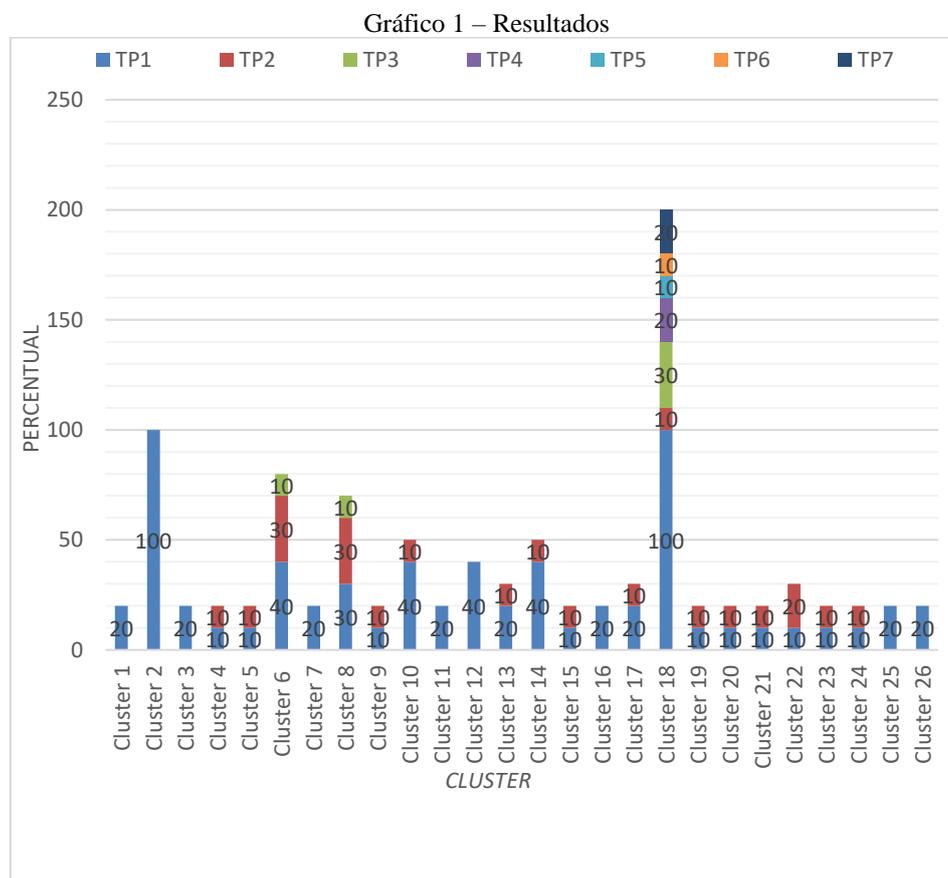
Cluster 6	3	Baixa escrupulosidad e – 4 textos	6	desenho, preguiça, sabotando, soneca, cansada, procrastinação.
		Baixa socialização – 3 textos	28	manipular, indiferente, interajo, vínculos, interesses, percepção, úteis, interação, empatia, calculista, interação social, crio, atrair, frágeis, sensação, submetem, submissas, dependentes, vontade, prazer, manipuladas, pornografia, usufruem, mentes, perversidade, influenciar, envolvimento, escolha.
		Escrupulosidade – 1 texto	17	adequar, evento, pagamentos, criação, mundo, burocrático, caros, clientes, dinheiro, baratos, empreendedores, financeiro, iniciativa, inovar, mercado financeiro, interessante.
Cluster 7	1	Baixa escrupulosidad e – 2 textos	7	preguiça, desanimada, sem coragem, cama, deitada, pijama, sofá.
Cluster 8	3	Baixa socialização – 3 textos	18	empatia, facilidade, identificar, indiferente, empatia, emoções, prazer, próximas, interagir, intuito, manipular, manipulo, vontade, falsidade, finjo, sangue, sensação, vontade.
		Baixa escrupulosidad e – 3 textos	6	moleza, preguiça, cama, serve pra nada, desânimo, não dou conta.
		Escrupulosidade – 1 texto	26	sucesso, metas, financeiro, objetivo, fundamental, financeira, concretizar, contas, sacrifícios, vantagem, estabelecer, realiza, prazos, foco, manter, observar, dinheiro, educação financeira, finanças, liberdade financeira, concretizar, consegue, contas, controle, eficiente, estratégia.
Cluster 9	2	Neuroticismo – 1 texto	6	depressão, desabafos, difícil, dor, dores, triste.
		Baixa socialização – 1 texto	–	
Cluster 10	2	Estabilidade emocional – 4 textos	34	acalmará, alma, autoconhecimento, autoestima, calma, confiança, constante, construção, coração, descobrir, descubra, emocional, encontrando, felicidade, verdade, agir, processo, constante, evolução, estabilidade emocional, aceitar, escolha, compreensão, alcançar, amando, aprendizado, buscar, descanso, alcançar, vida, equilíbrio, alma, espírito.
		Abertura para experiências – 1 textos	17	adrenalina, adoro, aventura, aventuras, aventureiro, barulho, correm, correndo, emoção, êxtase, explodir, fascinado, infinito, liberdade, vento, viver, forte emoção.
Cluster 11	1	Estabilidade emocional – 2 textos	6	aceitam, aprendizado, crescimento, acalma, calma, certeza.
Cluster 12	1	Socialização – 4 textos	29	ajuda, ajudando, ajudar, campanha, compartilhe, continua, continue, contribua, corrente, doe, gratidão, ajudar, solidariedade, prol, possível, continuarem, doarem, ajuda, ajudem, cura, divulguem, doem, favor, ajude, aquecendo, envolva, projeto, bem, causa.
Cluster 13	2	Escrupulosidade – 2 textos	13	pagar, objetivos, metas, atenção, contas, focado, comprar, compromisso, dinheiro, economizar, investir, bancárias, trabalho.
		Introversão – 1 texto	6	introversão, introvertido, interagir, dificuldade, estranhos, isolado.
Cluster 14	2	Socialização – 4 textos	22	aceitando, ajudar, ajuda, ajudem, cura, divulgue, doações, doem, doença, poder, quantia, doando, pedir, cura, amar, amor, conscientizar, projeto, doar, missões, obra, solidário.
		Abertura para experiências – 1 texto	5	altura, aventura, curtir, sensação, velocidade.

Cluster 15	2	Escurpulosidade – 1 texto	15	importante, momento, comemorar, meta, celebrar, vitória, importância, conquistas, treinamento, recompensa, reconhecer, desafios, comemoração, continuar, importante.
		Estabilidade emocional – 1 texto	10	alcance, conquistas, conquista, felicidade, feliz, objetivo, objetivos, realizações, reconhecendo, reconhecer.
Cluster 16	1	Introversão – 2 textos	16	diferentes, introvertidos, julgamentos, compreensão, pouco, diferente, sentimos, jogamos, julgamentos, apreço, dificuldade, inteligentes, introvertido, isolado, nervoso, poucos amigos.
Cluster 17	2	Socialização – 2 textos	17	acredita, acreditamos, acreditaram, acreditarem, Deus, doar, comunidade, projeto, espiritual, milagres, ajuda, família, carentes, ajuda, doação, ajudar, colaboração.
		Abertura para experiências – 1 texto	4	adrenalina, altura, aventura, experiências.
Cluster 18	7	Baixa abertura para experiências – 10 textos	32	sosego, meditação, saudável, comida, desenvolvimento, ficar em casa, café, monotonia, cozinha, estudar, filme, ler, tranquilo, série, aconchego, família, animaizinhos, rotina, comer, simples, privado, arte, cultura, companhia, vinho, massa, culinária, doguinho, jantar, calma, calma, tranquilidade.
		Baixa escurpulosidade – 1 texto	4	cama, indisposta, preguiça, procrastinar.
		Baixa socialização – 3 textos	23	manipular, relacionar, próximas, descarto, manipulador, manipula, articulados, convencer, culpa, culpados, descartar, desejos, disposição, inventar, inverdades, justificam, satisfazer, convencer, articuladas, empatia, manipulação, mentira, vítima.
		Escurpulosidade – 2 textos	35	acontecer, benefícios, bombar, conscientizar, conta, diferença, diretamente, ganhar, finalidade, investir, produto, processo, público-alvo, importante, preço, exatamente, dinheiro, reais, ações, conhecimento, treinamento, investe, tesouro, direto, rico, objetivo, absoluta, aprendizado, certeza, comprar, corretor, curso, empresas, financeira, ganhar.
		Abertura para experiências – 1 texto	10	emoção, loucura, vida, aproveitar, intenso, loucamente, adrenalina, ousadia, improvisado, agrada.
		Introversão – 1 texto	6	introversão, timidez, expressar, enturmar, atrapalha, críticas.
		Neuroticismo – 2 textos	22	<i>bullying</i> , cortei, decepções, depressão, problemas, rejeição, ruim, traumas, decepções, sofreram, dormir, dormindo, dores, sofrimento, acabar, crises, dói, dor, dores, dormindo, dormir, durmo.
Cluster 19	2	Introversão – 1 texto	1	sinto.
		Escurpulosidade – 1 texto	13	alto, apoio, clientes, confiança, dedicação, dinheiro, foco, profissionais, reconhecimento, esforço, trabalho, agradecer, confiança.
Cluster 20	2	Introversão – 1 texto	4	introvertida, tímida, atrapalhado, arrumar emprego.
		Neuroticismo – 1 texto	13	acabar, angústia, arrependo, chorar, coragem, culpa, crise, depressão, desabafo, desanimador, difícil, dói, sentido.
Cluster 21	2	Estabilidade emocional – 1 texto	5	reconhecer, objetivos, atitudes, entender, entendeu.
		Abertura para experiências –	6	adrenalina, aventura, desafio, campeonato, emoção, velocidade.

		1 texto		
Cluster 22	2	Introversão – 1 texto	1	ausência.
		Abertura para experiências – 2 textos	3	adrenalina, aventura, arriscado.
Cluster 23	2	Abertura para experiências – 1 texto	9	adrenalina, aventura, emoção, emoções, louco, loucura, velocidade, viciante, risco.
		Neuroticismo – 1 texto	12	sentindo, cortes, corta, depressão, abandonar, aguentamos, chorar, cortar, cortamos, corte, culpados, desabafar.
Cluster 24	2	Introversão – 1 texto	3	desconfortável, difícil, dificuldade.
		Abertura para experiências – 1 texto	6	viagens, loucos, vento, voar, adrenalina, inesquecíveis.
Cluster 25	1	Estabilidade emocional – 2 textos	5	encontrou, equilíbrio, espírito, estabilidade emocional, segura.
Cluster 26	1	Introversão – 2 textos	5	introversão, reservados, calados, calada, discreto.

Fonte: elaborado pela autora 2020.

Para uma melhor interpretação dos resultados da Tabela 3, eles foram organizados no Gráfico 1, que apresenta as dimensões estatísticas, representando cada *cluster* e seu percentual.



Fonte: elaborado pela autora, 2021.

*Cluster 1:*

Neuroticismo TEX 3.txt

Neuroticismo TEX 6.txt

Esse *cluster* realizou separação de um tipo de traço de personalidade. Separou 2 textos do traço neuroticismo, contendo 23 palavras do grupo correspondente a ele. O *cluster* fez, assim, uma separação de 20% dos textos desse tipo de traço.

*Cluster 2:*

Extroversão TEX 10.txt

Extroversão TEX 1.txt

Extroversão TEX 2.txt

Extroversão TEX 3.txt

Extroversão TEX 4.txt

Extroversão TEX 5.txt

Extroversão TEX 6.txt

Extroversão TEX 7.txt

Extroversão TEX 8.txt

Extroversão TEX 9.txt

O *cluster 2* fez a melhor separação do conjunto todo. Separou em 1 tipo, ou seja, em uma pasta, os 10 textos do traço de personalidade extroversão, contendo 87 vocábulos do grupo de palavras referentes a esse traço. Assim, esse *cluster* fez uma separação de 100% dos textos desse traço.

*Cluster 3:*

Neuroticismo TEX 2.txt

Neuroticismo TEX 1.txt

Já o *cluster 3* fez a separação somente de um tipo de traço de personalidade. Separou 2 textos do traço denominado neuroticismo, contendo 19 palavras do grupo correspondente a ele. O *cluster* fez uma separação de 20% dos textos desse traço.

*Cluster 4:*

Escrupulosidade TEX 8.txt

Abertura para experiências TEX 4.txt

Esse *cluster* separou 2 tipos de traços de personalidade, utilizando um texto do traço escrupulosidade, contendo 11 palavras do grupo correspondente, alcançando o percentual de 10% dos textos desse traço. Foi utilizado também um texto do traço de

personalidade abertura para novas experiências, contendo 9 palavras do grupo correspondente a ele e atingindo 10% dos textos desse traço.

*Cluster 5:*

Escrupulosidade TEX 9.txt

Neuroticismo TEX 8.txt

O *cluster 5* separou 2 tipos de traços de personalidade. Foi utilizado 1 texto do traço de personalidade escrupulosidade, com 15 palavras do grupo desse traço. Utilizou-se também 1 texto do traço neuroticismo, contendo 8 palavras do grupo correspondente. Ambos geraram o percentual de 10%.

*Cluster 6:*

Baixa escrupulosidade TEX 4.txt

Baixa escrupulosidade TEX 1.txt

Baixa escrupulosidade TEX 7.txt

Baixa escrupulosidade TEX 9.txt

Baixa socialização TEX 1.txt

Baixa socialização TEX 4.txt

Baixa socialização TEX 5.txt

Escrupulosidade TEX 1.txt

Esse *cluster* foi responsável pela separação de 8 textos. Foram explorados 3 tipos de traços de personalidade, sendo: 4 textos do traço baixa escrupulosidade, contendo 6 vocábulos do grupo de palavras relacionadas a esse traço e alcançando o percentual de 40% dos textos escolhidos; 3 textos do traço baixa socialização, com 28 palavras do grupo relativo a esse traço e atingindo o percentual de 30%; e, por fim, 1 texto do traço escrupulosidade, contendo 17 palavras do grupo específico desse traço e com percentual de 10% dos textos escolhidos.

*Cluster 7:*

Baixa escrupulosidade TEX 5.txt

Baixa escrupulosidade TEX 6.txt

O *cluster 7* fez a separação de 1 traço de personalidade. Selecionou 2 textos do traço baixa escrupulosidade, contendo 7 palavras do grupo correspondente a esse traço e gerando um percentual de 20% de textos para esse traço.

*Cluster 8:*

Baixa escrupulosidade TEX 3.txt

Baixa escrupulosidade TEX 8.txt

Baixa escrupulosidade TEX 2.txt

Baixa socialização TEX 7.txt

Baixa socialização TEX 8.txt

Baixa socialização TEX 9.txt

Escrupulosidade TEX 3.txt

O *cluster* 8 separou 7 textos de 3 tipos de traços de personalidade diferentes, sendo 3 textos para baixa escrupulosidade, contendo 6 palavras do grupo desse traço e gerando um percentual de 30%; 3 textos do traço para baixa socialização, contendo 18 palavras do grupo selecionado para esse traço, com percentual também de 30%; e 1 texto do traço de personalidade escrupulosidade, contendo 26 palavras desse traço, com percentual de 10% dos textos escolhidos para esse traço.

*Cluster* 9:

Neuroticismo TEX 10.txt

Baixa socialização TEX 10.txt

O *cluster* 9 fez a separação de 2 textos de 2 tipos de traços de personalidade, sendo 1 texto do traço neuroticismo, contendo 6 palavras e apresentando percentual de 10% dos textos escolhidos para ele; e 1 texto do traço baixa socialização, com 0 palavras do grupo de palavras desse traço de personalidade, assim gerando percentual de 10% dos textos escolhidos para esse traço.

*Cluster* 10:

Abertura para experiências TEX 5.txt

Estabilidade emocional TEX 6.txt

Estabilidade emocional TEX 8.txt

Estabilidade emocional TEX 9.txt

Estabilidade emocional TEX 10.txt

Esse *cluster* separou 5 textos de 2 traços de personalidade diferentes, sendo 1 texto do traço abertura para novas experiências, contendo 17 palavras do grupo desse traço, com percentual de 10% dos textos escolhidos; e 4 textos para estabilidade emocional, contendo 33 palavras relacionadas a esse traço, com percentual de 40% dos textos escolhidos para ele.

*Cluster* 11:

Estabilidade emocional TEX 2.txt

Estabilidade emocional TEX 7.txt

O *cluster* 11 fez a separação de 2 textos do traço de personalidade estabilidade emocional, contendo 6 palavras do grupo desse traço, com percentual de 20% dos textos escolhidos para ele.

*Cluster* 12:

Socialização TEX 10.txt

Socialização TEX 3.txt

Socialização TEX 4.txt

Socialização TEX 8.txt

O *cluster* 12 separou 4 textos do traço de personalidade socialização, contendo 29 palavras do grupo característico para esse traço e obtendo percentual de 40% dos textos escolhidos para ele.

*Cluster* 13:

Escrupulosidade TEX 7.txt

Escrupulosidade TEX 10.txt

Introversão TEX 5.txt

Esse *cluster* fez a separação de 3 textos, com 2 tipos diferentes de traços de personalidades. Foram 2 textos do traço escrupulosidade, contendo 13 palavras do grupo desse traço, com percentual de 20% dos textos escolhidos para ele. Também foi selecionado 1 texto do traço introversão, contendo 6 palavras do grupo ligadas a esse traço, com percentual de 10% dos textos desse traço.

*Cluster* 14:

Socialização TEX 5.txt

Socialização TEX 6.txt

Socialização TEX 7.txt

Socialização TEX 9.txt

Abertura para experiências TEX 1.txt

O *cluster* 14 separou 5 textos de 2 traços de personalidade, sendo 4 textos para socialização e contendo 22 palavras do grupo correspondente a esse traço, com percentual de separação dos textos escolhidos para ele de 40%. Foi também selecionado 1 texto do traço abertura para novas experiências, contendo 5 vocábulos do grupo de palavras escolhido para esse traço e com percentual de 10% dos textos escolhidos para ele.

*Cluster* 15:

Escrupulosidade TEX 5.txt

Estabilidade emocional TEX 1.txt

Esse *cluster* fez a separação de 2 textos de 2 tipos diferentes de traços de personalidades, sendo 1 texto para escrupulosidade, contendo 15 palavras do grupo correspondente a esse traço. Além disso, foi selecionado 1 texto do traço estabilidade emocional, contendo 10 palavras do grupo indicado para esse traço. Ambos os traços de personalidade tiveram um percentual de 10% dos textos escolhidos.

*Cluster 16:*

Introversão TEX 6.txt

Introversão TEX 8.txt

O *cluster* 16 obteve a separação de 1 tipo de traço de personalidade, sendo 2 textos do traço introversão, contendo 16 palavras do grupo correspondente a esse traço. Esse *cluster* teve um percentual de separação de 20% dos textos escolhidos para esse traço.

*Cluster 17:*

Socialização TEX 2.txt

Socialização TEX 1.txt

Abertura para experiências TEX 10.txt

O *cluster* 17 fez a separação de 3 textos, sendo 2 tipos de traços de personalidade, sendo 2 textos para socialização, com 17 vocábulos do grupo de palavras desse traço e tendo como percentual 20% dos textos dele; e 1 texto do traço abertura para novas experiências, com 4 vocábulos do grupo de palavras desse traço e tendo o percentual de 10% dos textos desse traço.

*Cluster 18:*

Baixa abertura para experiências TEX 1.txt

Baixa abertura para experiências TEX 10.txt

Baixa abertura para experiências TEX 2.txt

Baixa abertura para experiências TEX 3.txt

Baixa abertura para experiências TEX 4.txt

Baixa abertura para experiências TEX 5.txt

Baixa abertura para experiências TEX 6.txt

Baixa abertura para experiências TEX 7.txt

Baixa abertura para experiências TEX 8.txt

Baixa abertura para experiências TEX 9.txt

Baixa escrupulosidade TEX 10.txt

Abertura para experiência TEX 2.txt

Baixa socialização TEX 2.txt

Baixa socialização TEX 3.txt

Baixa socialização TEX 6.txt

Escrupulosidade TEX 2.txt

Escrupulosidade TEX 4.txt

Introversão TEX 9.txt

Neuroticismo TEX 4.txt

Neuroticismo TEX 7.txt

O *cluster* 18 fez a maior separação entre os demais *clusters*, já que separou 20 textos de 7 tipos diferentes de traços de personalidade do Big Five. Também fez uma boa separação, pois houve a seleção de 10 textos, ou seja, todos os textos do traço baixa abertura para novas experiências (100%), contendo 32 palavras do grupo desse traço. Separou também 1 texto do traço baixa escrupulosidade, contendo 4 palavras do grupo correspondente a esse traço, com percentual de 10% dos textos escolhidos; 1 texto do traço abertura para novas experiências, contendo 10 palavras do grupo dele e tendo como percentual 10% dos respectivos textos; 3 textos para baixa socialização, contendo 23 palavras do grupo característico desse traço, com percentual de 30% dos textos escolhidos para ele; 2 textos do traço escrupulosidade, contendo 35 palavras do grupo desse traço e com percentual de 20% dos textos escolhidos para esse traço; 1 texto de introversão, contendo 6 palavras do grupo relacionado a esse traço de personalidade, com percentual de 10% dos respectivos textos; e, por fim, 2 textos para neuroticismo, contendo 22 palavras do grupo desse traço, com percentual de 20% dos textos escolhidos para ele.

*Cluster* 19:

Introversão TEX 3.txt

Escrupulosidade TEX 6.txt

Esse *cluster* separou 2 textos de 2 traços de personalidade, sendo 1 texto para introversão, com 1 vocábulo do grupo de palavras desse traço; e 1 texto para escrupulosidade, com 13 vocábulos do grupo de palavras desse traço. Ambos tiveram um percentual de 10% de separação dos textos de cada traço de personalidade escolhido.

*Cluster* 20:

Introversão TEX 4.txt

Neuroticismo TEX 5.txt

Esse *cluster* separou 2 textos de 2 traços de personalidade, sendo 1 texto para introversão, com 4 palavras do grupo ligado a esse traço; 1 texto para neuroticismo, com

13 palavras do grupo correspondente a ele. Ambos tiveram um percentual de 10% de separação dos textos escolhidos para cada traço de personalidade.

*Cluster 21:*

Estabilidade emocional TEX 5.txt

Abertura para experiência TEX 3.txt

O *cluster* 21 também separou 2 textos de 2 traços de personalidade, sendo 1 texto para estabilidade emocional, com 5 palavras do grupo desse traço; 1 texto do traço abertura para novas experiências, com 6 vocábulos do grupo de palavras desse traço. Ambos tiveram um percentual de 10% de separação dos textos escolhidos para cada traço de personalidade.

*Cluster 22:*

Introversão TEX 2.txt

Abertura para experiências TEX 6.txt

Abertura para experiências TEX 8.txt

Esse *cluster* separou 3 textos de 2 traços de personalidade, sendo 1 texto para introversão, com 1 vocábulo do grupo de palavras desse traço e 2 textos do traço abertura para novas experiências com 3 vocábulos do grupo de palavras desse traço, com percentuais respectivamente de 10% e 20% de separação dos textos escolhidos para cada traço.

*Cluster 23:*

Abertura para experiências TEX 7.txt

Neuroticismo TEX 9.txt

O *cluster* 23 separou 2 textos de 2 traços de personalidade, sendo 1 texto do traço abertura para novas experiências, com 9 palavras do grupo correspondente a ele; e 1 texto para neuroticismo, com 12 vocábulos do grupo de palavras desse traço. Ambos tiveram um percentual de 10% de separação dos textos escolhidos para cada traço de personalidade.

*Cluster 24:*

Introversão TEX 1.txt

Abertura para experiências TEX 9.txt

Esse *cluster* separou 2 textos de 2 traços de personalidade, sendo 1 texto para introversão, com 3 vocábulos do grupo de palavras desse traço e 1 texto do traço abertura para novas experiências, com 6 vocábulos do grupo desse traço de personalidade. Ambos tiveram um percentual de 10% de separação dos textos escolhidos de cada traço.

*Cluster 25:*

Estabilidade emocional TEX 4.txt

Estabilidade emocional TEX 3.txt

Esse *cluster* separou 2 textos do traço de personalidade estabilidade emocional, contendo 5 vocábulos do grupo de palavras desse traço, com um percentual de 20% de separação dos textos escolhidos para ele.

*Cluster 26:*

Introversão TEX 7.txt

Introversão TEX 10.txt

O *cluster* 26 separou 2 textos do traço de personalidade introversão, contendo 5 vocábulos do grupo adequado a esse traço, com um percentual de 10% de separação dos textos escolhidos para ele.

As melhores separações feitas pelo Cassiopeia foram as dos *clusters* que têm o maior número de textos de um mesmo traço da personalidade. Assim, os *clusters* que fizeram a melhores separações foram o 2 e o 18, pois tiveram um percentual no seu resultado de 100%. A ferramenta FineCount Free, utilizada para fazer a estatística do tamanho dos textos, apresentou uma grande variação entre eles. Foram 100 (cem) textos, sendo 10 textos para cada traço de personalidade do Big Five, totalizando 11.234 (onze mil, duzentas e trinta e quatro) palavras.

Os traços de personalidade extroversão e baixa abertura para novas experiências são os traços de personalidades para os quais os 10 respectivos textos tiveram os menores números de palavras, sendo extroversão com 468 (quatrocentas e sessenta e oito) palavras e baixa abertura para novas experiências com 430 (quatrocentas e trinta).

## 5 DISCUSSÃO DOS RESULTADOS

Os 26 *clusters* fizeram a seguinte separação: 9 *clusters* classificaram os textos em 1 categoria; 14 *clusters* classificaram os textos em 2 categorias; 2 *clusters* classificaram os textos em 3 categorias; e 1 *cluster* classificou os textos em 7 categorias. As melhores separações foram para os seguintes traços de personalidade: extroversão, que está no *cluster* 2, cuja separação foi de 100% dos textos; e baixa abertura para novas experiências, que está no *cluster* 18 e também separou 100% dos textos do seu respectivo traço.

Além disso, esses dois *clusters* – que foram os que separaram as maiores quantidades de texto de um determinado traço de personalidade – obtiveram um resultado melhor também em relação ao número de palavras escolhidas anteriormente para o

respectivo traço de personalidade. Dessa forma, dentre todos os *clusters* gerados pelo Cassiopeia, esses dois foram os que obtiveram o resultado esperado, qual seja, de classificar 100% dos textos do mesmo traço de personalidade, na mesma categoria, ou seja, na mesma pasta.

## 6 CONCLUSÃO

Esta pesquisa testou a hipótese de que, com a elaboração de taxonomia de traços de personalidade utilizando o modelo Big Five, seja possível uma visão categorizada dos usuários do *Facebook*, utilizando o Cassiopeia para possibilitar um agrupamento desses usuários em categorias. Conclui-se, logo, as melhores separações feitas pelo Cassiopeia foram as dos *clusters* 2 e 18, os quais obtiveram maior número de textos correspondentes ao mesmo traço de personalidade do modelo Big Five. Consequentemente, revelou-se maior número de palavras que os descrevem.<sup>3</sup>

Para esse estudo, sugiro futuramente desenvolver trabalhos como: criação de um *software* com as cinco categorias dos traços de personalidade e com os conjuntos de palavras já existentes, para promover uma melhoria do funcionamento da ferramenta utilizada para fazer a mineração de texto no Cassiopeia e a taxonomia de traços de personalidade nas redes sociais automaticamente. Estudos mais aprofundados sobre esse tema trarão grandes contribuições para a Educação e demais áreas do conhecimento, uma vez que na atualidade as redes sociais são muito utilizadas por todas as idades, etnias e classes sociais.

---

<sup>3</sup> Para maior detalhe dos resultados, consultar dissertação completa.

## REFERÊNCIAS

AS 20 MAIORES redes sociais do mundo (2020 e 2021). *Maiores e Melhores* [site]. 2021. Disponível: <<https://www.maioresemelhores.com/maiores-redes-sociais-do-mundo/>>. Acesso em: 2 jan. 2021.

FREIRE, P. *Política e educação: ensaios*. 5. ed. São Paulo: Cortez, 2001. (Coleção Questões de Nossa Época, v. 23).

GUELPELI, M. V. C. *Cassiopeia: um modelo de agrupamento de textos baseado em sumarização*. Tese (Doutorado em Computação) – Universidade Federal Fluminense, Niterói, Brasil, 2012.

HALL, C. S.; LINDZEY, G.; CAMPBELL, J. B. Trad. Maria Adriana Veríssimo Veronese. *Teorias da personalidade*. 4. ed. Porto Alegre: Artmed, 2000.

MODELO BIG Five de personalidad y selección. *Aiteco Consultores Desarrollo y Gestión* [site]. 2021. Disponível em: <<https://www.aiteco.com/modelo-big-five-de-personalidad-y-seleccion/>>. Acesso em: 2 jan. 2021.

MORAIS, E. A. M.; AMBRÓSIO, A. P. L. Mineração de textos. *Technical Report – INF, 005/07* [Relatório Técnico]. Goiás: Instituto de Informática / Universidade Federal de Goiás, Dezembro, 2007. Disponível em: <[http://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF\\_005-07.pdf](http://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_005-07.pdf)>. Acesso em: 16 jun. 2020.

PERVIN, L. A.; JOHN, O. P. *Personalidade: teoria e pesquisa*. 8. ed. Trad. R. C. Costa. Porto Alegre: Artmed, 2004.