

MARCOS PAULO FERREIRA RODRIGUES

Orientador: Anderson Almeida Ferreira

**UM SISTEMA DE RECOMENDAÇÃO ENTRE  
ESTUDANTES E REPÚBLICAS BASEADO NA  
PERSONALIDADE**

Ouro Preto  
Dezembro de 2019

UNIVERSIDADE FEDERAL DE OURO PRETO  
INSTITUTO DE CIÊNCIAS EXATAS  
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

**UM SISTEMA DE RECOMENDAÇÃO ENTRE  
ESTUDANTES E REPÚBLICAS BASEADO NA  
PERSONALIDADE**

Monografia apresentada ao Curso de Bacharelado em Ciência da Computação da Universidade Federal de Ouro Preto como requisito parcial para a obtenção do grau de Bacharel em Ciência da Computação.

MARCOS PAULO FERREIRA RODRIGUES

Ouro Preto  
Dezembro de 2019

R618s      Rodrigues, Marcos Paulo Ferreira.  
Um sistema de recomendação entre estudantes e repúblicas baseado na  
personalidade [manuscrito] / Marcos Paulo Ferreira Rodrigues. - 2020.

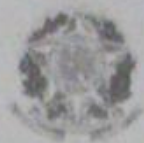
61f.: il.: color; grafs; tabs; Algoritmos.

Orientador: Prof. Dr. Anderson Almeida Ferreira.

Monografia (Graduação). Universidade Federal de Ouro Preto. Instituto de  
Ciências Exatas e Biológicas. Departamento de Computação.

1. WEB (Linguagem de programação). 2. Computação. I. Ferreira, Anderson  
Almeida. II. Universidade Federal de Ouro Preto. III. Título.

CDU: 004.738.5



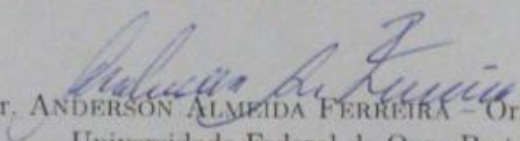
UNIVERSIDADE FEDERAL DE OURO PRETO

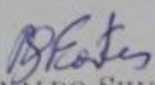
FOLHA DE APROVAÇÃO

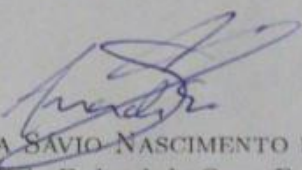
Um sistema de recomendação entre Estudantes e Repúblicas baseado na  
personalidade

MARCOS PAULO FERREIRA RODRIGUES

Monografia defendida e aprovada pela banca examinadora constituída por:

  
Dr. ANDERSON ALMEIDA FERREIRA - Orientador  
Universidade Federal de Ouro Preto

  
M.Sc REINALDO SILVA FORTES  
Universidade Federal de Ouro Preto

  
Dra. AMANDA SAVIO NASCIMENTO E SILVA  
Universidade Federal de Ouro Preto

Ouro Preto/4 Dezembro de 2019

# Resumo

A vida universitária pode oferecer diversos obstáculos durante sua trajetória. Para os estudantes que escolhem cursar uma universidade em uma cidade distante, um dos primeiros desafios é escolher a república para morar. Porém, é comum o caso de evasão de estudantes em repúblicas, muitas das vezes ocorrendo por incompatibilidade de personalidades. Assim, é interessante para o estudante conhecer a personalidade das pessoas que residem na república. Da mesma maneira, é interessante aos membros da fraternidade conhecerem o perfil do estudante que está interessado em conviver com eles. Desta maneira, este trabalho propõe um método de identificação de compatibilidade e incompatibilidade de perfis, por meio do desenvolvimento de um sistema web para manipulação de perfis de estudantes e repúblicas estudantis - representados por traços de personalidade - e indicar tais compatibilidades. Para avaliar as compatibilidades identificadas entre os tipos de usuários, estudantes e repúblicas, foram realizados experimentos com estudantes e repúblicas estudantis reais e, de acordo com os critérios utilizados, os resultados se mostraram promissores, chegando a ter, em alguns casos, acurácia de 100%.

# Abstract

University life can offer several obstacles during its trajectory. For students who choose to attend a university in a distant city, one of the first challenges is to choose the republic to live in. But it is trivial the case of student evasion in republics, often occurring because of incompatibility of personalities. Thus, it is interesting for the student to know the personality of the people residing in the republic. In the same way, it is interesting for members of the fraternity to know the profile of the student who is interested in living with them. Thus, this work proposes a recommendation method that addresses the problem of profile mismatch, in which a recommendation system that works with the student and republic profile - represented by personality traits - was developed and makes the appropriate recommendation. Metrics vary by the type of user requesting the recommendation. The experiments were submitted to republics and students resident and not residents of republics. For the experiments performed, in some cases, it is possible to realize an accuracy of up to 100%.

*Dedico este trabalho aos meus pais, irmãos, avós, primos e toda minha família que, com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa da minha vida. Decido também aos docentes do Departamento de Ciência da Computação da Universidade Federal de Ouro Preto, que foram de extrema importância na minha vida acadêmica, e em especial ao Professor Anderson Almeida Ferreira pela paciência na orientação e incentivo que tornaram possível a conclusão desta monografia.*

# Agradecimentos

À minha família pelo amor, incentivo e apoio incondicional. Ao meu orientador pelo empenho dedicado à elaboração deste trabalho. Aos professores do Departamento de Computação da Universidade Federal de Ouro Preto pela excelência no ensino. À República Maternidade pela acolhida, aprendizado e crescimento vividos nos últimos cinco anos. Aos meus amigos por sempre estarem presentes. A todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.



# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Contextualização	1
1.2	Justificativa	2
1.3	Objetivo Geral e Específicos	2
1.4	Método do Trabalho	3
1.5	Organização do Trabalho	3
<b>2</b>	<b>Revisão de Literatura</b>	<b>4</b>
2.1	Fundamentação Teórica	4
2.1.1	Entrada e saída	5
2.1.2	Técnicas de recomendação	5
2.1.3	Formação de perfis de usuários	10
2.2	Trabalhos Relacionados	17
2.2.1	Sistemas de Recomendação com Traços de Personalidade	17
2.2.2	Personalidade e Confiança Social em Grupos de Recomendação	18
2.2.3	Prever Traços de Personalidade com Imagens do Instagram	19
<b>3</b>	<b>Desenvolvimento</b>	<b>21</b>
3.1	Concepção do sistema	22
3.1.1	Objetivo	22
3.1.2	Missão e Descrição do Produto	22
3.1.3	Escopo	22
3.1.4	Diagrama de Contexto	23
3.1.5	Hierarquia dos Atores	23
3.1.6	Diagrama Relacional	25
3.1.7	Storyboard	25
3.1.8	Backlog	25
3.1.9	Estória de Usuário	29
3.1.10	Cenário de Teste	32
3.2	Modelagem do Banco de Dados	35

3.2.1	Diagrama Entidade Relacionamento	35
3.2.2	Dicionário de Dados	35
3.3	Sistema	38
3.3.1	Tecnologias	39
3.3.2	Ferramenta	39
3.3.3	Personalidade	39
3.3.4	Perfil do Usuário	40
3.3.5	Perfil da República	40
3.3.6	Recomendação	42
<b>4</b>	<b>Experimentos Computacionais</b>	<b>44</b>
4.1	Descrição	44
4.1.1	Experimentos	44
4.1.2	Critérios de avaliação	45
4.2	Estudos de Caso	46
4.2.1	Estudantes sem vínculo com República	46
4.2.2	Estudantes moradores de República	47
4.2.3	Recomendação aplicada a Repúblicas	47
4.3	Análise dos Resultados	49
<b>5</b>	<b>Conclusão</b>	<b>50</b>
	<b>Referências Bibliográficas</b>	<b>61</b>

# Lista de Figuras

2.1 Filtragem Híbrida. Fonte: (Cazella et al., 2010)	9
2.2 Dimensões básicas do usuário segundo modelo GUMO FONTE: (Cazella et al., 2010)	11
2.3 iGoogle Portal utilizando a coleta de dados de forma explícita.	12
2.4 Site AliExpress utilizando a coleta de dados implícita para recomendar produtos	12
3.1 Diagrama de Contexto do Sistema	23
3.2 Hierarquia dos Atores	24
3.3 Diagrama Relacional do Sistema	25
3.4 Index	26
3.5 Recomendação	26
3.6 Login/Cadastro	26
3.7 Perfil	27
3.8 Inventário	27
3.9 Detalhes	27
3.10 Diagrama Entidade Relacionamento.	36
4.1 Relação de Acerto e Erro da Recomendação aplicada aos Moradores	48
4.2 Relação de Acerto e Erro da Recomendação aplicada às Repúblicas	48

# Lista de Tabelas

2.1 Avaliação de filmes . . . . .	8
3.1 Descrição do Diagrama de Contexto . . . . .	24
3.2 Descrição dos Atores . . . . .	24
3.3 Descrição do Diagrama Relacional . . . . .	25
3.4 Descrição das telas do sistema . . . . .	28
4.1 Relação de repúblicas e quantidade de moradores . . . . .	47

# Lista de Algoritmos

1	Personalidade da república	42
2	Calcula Moda Da Faceta	42
3	Função Max (vetor moda)	42

# Capítulo 1

## Introdução

Em um mundo onde a informação disponível na rede virtual é incalculável e cresce em alta velocidade, graças ao avanço da tecnologia e comunicação, todo usuário que acessa a Web por meio do seu computador ou possui algum dispositivo conectado à Internet é contribuinte e consumidor de informação. É, então, interessante saber que tipo de informação os usuários consomem, possibilitando a criação de perfis de acordo com seus gostos musicais, visões políticas e interesse geral.

Um possível uso desses perfis seria na recomendação de pessoas a grupos (Nuno Miranda, 2015). Por exemplo, um calouro de uma universidade, que está à procura de um lugar para morar, procurará por uma casa onde os moradores tenham um perfil parecido com o seu. Da mesma maneira, há empresas que, nos seus processos seletivos, além de buscarem somente perfis técnicos, procuram também por profissionais que apoiem políticas de inclusão social.

Este trabalho consiste em buscar uma forma eficaz de recomendar grupos à pessoas. A sequência do capítulo está organizado como segue. A Seção 1.1 contextualiza o trabalho. A Seção 1.2 apresenta a justificativa deste trabalho. A Seção 1.3 descreve seus objetivos geral e específicos. A Seção 1.4 aborda a metodologia utilizada para o desenvolvimento deste trabalho. A Seção 1.5 apresenta a continuação do trabalho.

### 1.1 Contextualização

Em diversos locais do mundo é comum ver casos de moradia compartilhada, comumente chamada de república, principalmente quando se trata de alunos de uma universidade, que muitas vezes são oriundos de outras cidades e moram temporariamente em uma casa com outros alunos. Quando um calouro decide ingressar em uma universidade, longe do local onde mora, geralmente sua primeira opção é procurar uma república para morar. Além da facilidade de gerir as contas, o peso da responsabilidade de se manter uma casa é dividido entre todos. É normal o caso do estudante que se muda para uma cidade não conhecida por ele antes de ingressar na universidade. Assim, escolher uma nova casa não é uma tarefa trivial,

considerando que na maioria das vezes o calouro não conhece os indivíduos que ali residem. Desta maneira, a situação onde há incompatibilidade da personalidade dos moradores da república com a personalidade do calouro pode gerar evasão. Nesse sentido, é interessante que o calouro conheça o perfil da república, como também é da mesma importância que a república conheça o calouro interessado na vaga

## 1.2 Justificativa

O problema da evasão por diferença de perfil, que acontece nas repúblicas estudantis, pode ser reduzido a outros problemas do cotidiano, como: busca de uma empresa por profissionais que defendam os seus ideais sociais; algoritmos para sugestão de novas amizades; e outros.

É interessante notar que esses problemas são resolvidos da mesma maneira: recomendando corretamente o perfil do usuário que se encaixa com a busca desejada. A recomendação adequada pode fazer a diferença entre conquistar um usuário ou perdê-lo. Uma vez vivendo em um mundo com acesso fácil à informação, onde, quando uma informação não serve, ela é descartada e o usuário não sendo “cativado” pela recomendação, dificilmente irá voltar a usá-la no futuro.

Com a recomendação apropriada de pessoas e grupos, o problema de evasão nas repúblicas estudantis pode ser amenizado, bem como a busca de empresas por profissionais específicos.

## 1.3 Objetivo Geral e Específicos

Este trabalho de monografia possui como objetivo principal a concepção e o desenvolvimento de um sistema de recomendação de pessoas a grupos aplicado no contexto de estudantes e repúblicas. A recomendação será baseada na personalidade dos estudantes e das repúblicas. Desta forma, se espera recomendar repúblicas à estudantes de acordo com a similaridade entre suas personalidades. Para isso, neste projeto, será definido a Modelagem do Banco de Dados, a Concepção do Sistema de Recomendação e o desenvolvimento do sistema, detalhando o funcionamento da ferramenta.

De modo geral, os objetivos deste trabalho são:

- Investigar formas de detectar e representar personalidades;
- Desenvolver uma estratégia de indicar a personalidade de um grupo por meio de suas personalidades individuais;
- Desenvolver e avaliar uma estratégia de recomendação de pessoas a grupos;
- Implementar um sistema que identifique a personalidade de pessoas e grupos e possibilite recomendações.

## 1.4 Método do Trabalho

Para a execução deste projeto, será necessário levantar o Perfil do Usuário, onde a identificação do usuário é feita por meio de um e-mail e senha fornecidos. O próximo passo é, então, sugerir ao usuário que responda o questionário do Teste de Personalidade, detalhado na Seção 2.1.3.5, para computar os cinco fatores da sua personalidade, de acordo com o Modelo Big Five, descrito na Seção 2.1.3.4. O trabalho propõe a criação da Identidade do Usuário explícita do perfil, sendo adquirida através do Teste de Personalidade.

Posto que o Perfil do Usuário está definido, a próxima etapa é definir o Perfil do Grupo. O grupo poderá ser composto por usuários de diversas personalidades. Assim, a elaboração do Perfil do Grupo está descrita na subseção 3.3.5, sendo constantemente mutável.

Definido o Perfil do Usuário e o Perfil dos Grupos, vê-se a necessidade de recomendar um usuário à um grupo de usuários e vice-versa. Desta maneira, a recomendação feita a um usuário  $U_i$  será uma lista de grupos de usuários  $G$ , ordenado de acordo com a porcentagem de similaridade entre a personalidade de  $U_i$  e a personalidade do grupo  $G_j$ .

Para a validação das estratégias utilizadas, serão elaborados dois casos de teste. O primeiro é composto por dados sintéticos, com o objetivo de explorar casos extremos que podem vir a acontecer em situações de conflito. Com dados sintéticos, é possível ter controle da distribuição dos dados, permitindo explorar a composição de cada grupo e a distribuição de personalidade dentro do grupo, além de permitir reproduzir o comportamento de grandes grupos, que são difíceis de serem organizados com dados reais. O segundo caso de teste é composto por dados reais, com o objetivo de verificar os resultados obtidos no primeiro caso.

## 1.5 Organização do Trabalho

O restante deste trabalho está organizado da seguinte maneira: O Capítulo 2 apresenta a revisão da literatura e os trabalhos relacionados, necessária para a compreensão desta monografia, abordando as principais ferramentas de desenvolvimento utilizadas no projeto. O Capítulo 3 aborda o desenvolvimento do projeto. O Capítulo 4 apresenta os resultados obtidos no desenvolvimento deste projeto. O Capítulo 5 aborda a conclusão desta monografia.



## Capítulo 2

# Revisão de Literatura

Neste capítulo, é apresentada a revisão de literatura desta monografia. A Seção 2.1 aborda a fundamentação teórica necessária para entendimento deste trabalho. A Seção 2.2 mostra os trabalhos relacionados ao tema deste projeto.

### 2.1 Fundamentação Teórica

Sistemas de recomendação são filtros de informação para apresentar itens ou objetos - como páginas Web, filmes, músicas, livros, medicamentos, lojas, artigos - que provavelmente são do interesse do usuário (Motta et al., 2011). Um Sistema de Recomendação pode ser um sistema colaborativo, porque a recomendação é feita a partir da organização, manipulação, sumarização e agrupamento das avaliações individuais. Os principais componentes de um Sistema de Recomendação são cliente (usuário) e produto (item). Um produto é um recurso que pode ser de diferentes naturezas, por exemplo: um conteúdo, um arquivo, uma informação, uma pessoa, um objeto. A recomendação é uma função de mapeamento de interesses do cliente para obtenção de um ou mais produtos (Motta et al., 2011).

Uma recomendação  $R$  para a escolha de  $p$  segue a seguinte regra:

$$R(c, P) = \max F(c, p_i) \quad (2.1)$$

Sendo:

- $F$  é a função que determina a relevância de  $p_i$  em relação a  $c$
- $c$  representa o cliente que usa o Sistema de Recomendação
- $P$  é um conjunto de produtos disponíveis à avaliação
- $p_i \in P$

Sistemas de Recomendação são caracterizados de acordo com três eixos (Motta et al., 2011): (a) Tipos de entrada e saída; (b) Itens do projeto e; (c) Método de recomendação.

### 2.1.1 Entrada e saída

Deve-se considerar se a recomendação é para um cliente ou uma comunidade. As informações podem ser providas pelo usuário ou de dados gerados pela comunidade. Uma opção válida é perguntar ao cliente suas preferências, entretanto, intervenções podem retardar a obtenção dos dados. Por isso, algumas informações sobre o cliente são obtidas de forma indireta (Motta et al., 2011).

Motta et al. (2011) descrevem que a navegação explícita pelas páginas da Internet, registrada no histórico de páginas visitadas, é uma indicação do interesse do usuário. A extração de palavras-chaves de consultas realizadas utilizando o histórico de acesso também pode ser uma indicação de interesse do usuário. O histórico de compras realizadas ou informações baixadas são excelentes indícios de avaliação positiva do cliente sobre um produto. Avaliações deixadas por clientes podem ser usadas como insumo para avaliar o produto e também entender o usuário. Essas informações podem dar pistas a um Sistema de Recomendação sobre os gostos dos clientes.

As informações podem ficar registradas no perfil do cliente, caso este tenha se identificado, ou podem servir para a manutenção de um perfil genérico da comunidade, sem as especificações do usuário.

### 2.1.2 Técnicas de recomendação

De acordo com Motta et al. (2011), existem quatro grandes tipos de filtragem, sendo elas: Colaborativa, Baseada em Conteúdo, Híbrida e Baseada por Descoberta de Conhecimento.

#### 2.1.2.1 Filtragem Colaborativa

Filtragem colaborativa é um método de geração de recomendação que tenta prever o grau de interesse de um cliente em determinados produtos, a partir de correlações entre as avaliações feitas por este cliente e as avaliações fornecidas por outros clientes (Motta et al., 2011).

Motta et al. (2011) diz que essa técnica consiste em correlacionar as avaliações fornecidas por um cliente com as avaliações fornecidas pelos demais usuários, a fim de prever o grau de interesse de um cliente em determinado produto.

O método consiste em três passos:

1. Calcular a similaridade/distância entre os usuários, usando a avaliação fornecida pelos outros usuários como parâmetro.

$$D_{u,w} = 1 - C_{u,w} \quad (2.2)$$

sendo  $D_{u,w}$  a distância entre os usuários  $u$  e  $w$  e  $C_{u,w}$  é a correlação das avaliações desses usuários.

A correlação indica proximidade entre avaliações calculadas. Considere uma tabela, onde cada coluna representa um usuário e cada linha um produto. A célula contém a avaliação do usuário  $u$  sobre o produto  $p$ .  $D$  mede a distância das avaliações de  $u_1$  e  $u_2$  a partir do conjunto de avaliações que deram sobre os mesmos produtos.

2. Selecionar os vizinhos mais próximos.
3. Fazer a previsão do usuário sobre um produto não consumido. A previsão mostra quão válido o produto será para o cliente. A previsão  $P$  se dá pela seguinte fórmula:

$$P_{u,i} = \bar{r}_u + \frac{\sum_{j=1}^n w_{u,j}(r_{j,i} - \bar{r}_j)}{\sum_{j=1}^n |w_{u,j}|} \quad (2.3)$$

Sendo:

- a)  $P_{u,i}$  previsão de avaliação do usuário  $u$  sobre o item  $i$ .
- b)  $n$  número de vizinhos
- c)  $\bar{r}_u$  avaliação média do usuário  $u$  sobre todos os produtos que já avaliou.
- d)  $\bar{r}_j$  avaliação média do produto  $j$ .
- e)  $w_{u,j}$  similaridade entre o usuário  $u$  e  $j$ .
- f)  $r_{j,i}$  avaliação do usuário  $j$  sobre o produto  $i$
- g)  $r_{j,i} - \bar{r}_j$  diferença entre avaliação do usuário  $j$  sobre o produto  $i$  e a avaliação média do usuário  $j$

A fórmula retrata a expectativa sobre a avaliação de um novo produto com base nos  $n$  clientes vizinhos que avaliaram o produto. Ignorando a filtragem colaborativa, essa expectativa é a média das avaliações dadas pelo usuário.

Motta et al. (2011) especifica que a filtragem colaborativa é uma técnica eficiente devido ao seu poder discriminatório, embora seja computacionalmente cara, pois a matriz de similaridade deve ser calculada e armazenada. Além da complexidade, com poucos usuários e poucas avaliações o método tende a gerar resultados ruins. Também, pelo fato de serem necessárias avaliações prévias, os lançamentos não entram nas recomendações. Entretanto, este método possibilita apresentar recomendações inesperadas aos usuários, além da possibilidade de formação de grupos de usuários agrupados por seus gostos e interesses similares (Reategui e Cazella, 2005).

### 2.1.2.2 Filtragem baseada em Conteúdo

Na filtragem baseada em conteúdo, a sugestão é feita de acordo com o que o usuário demonstrou interesse no passado. Assim, são personalizadas para cada usuário. Essa técnica consiste em aplicar algoritmos de aprendizagem de máquina para induzir um perfil de preferências de um usuário a partir de exemplos (Motta et al., 2011). Por exemplo, filmes são categorizados como ação, comédia e terror.

Esta técnica consiste em quatro passos:

1. Classificar os itens de acordo com categorias pré-estabelecidas.
2. Para cada categoria, calcular a avaliação média de cada avaliador.
3. Ordenar os itens avaliados.
4. Calcular a previsão de avaliação. A previsão de utilidade do produto  $i$  para o usuário  $u$  é a avaliação média de  $i$  ajustada pela avaliação média normalizada de  $u$  para produtos da categoria de  $i$ .

$$P_{u,i} = \frac{\bar{r}_i * \bar{r}_{cat,u}}{\bar{r}_{cat}} \quad (2.4)$$

Sendo:

- a)  $P_{u,i}$  previsão de avaliação do usuário  $u$  sobre o item  $i$
- b)  $\bar{r}_{cat}$  avaliação média da categoria  $cat$  do item  $i$
- c)  $\bar{r}_i$  avaliação média do item  $i$
- d)  $\bar{r}_{cat,u}$  avaliação média do usuário  $u$  sobre a categoria  $cat$

Uma das vantagens desse método é que não necessita de dados de outros usuários, permitindo recomendações a usuários com gostos exclusivos e itens novos ou não populares. Além de ser uma tecnologia madura, pode fornecer explicações sobre os itens recomendados listando as características do conteúdo que conduziram à recomendação.

Por outro lado, as recomendações são estáticas, desprezando a opinião dos usuários, além da necessidade do conteúdo possuir características inteligíveis. Muitas vezes o conteúdo é analisado automaticamente para se identificar categorias. Também possui baixa eficiência caso o conteúdo seja pouco informativo.

### 2.1.2.3 Filtragem baseada em Descoberta de Conhecimento

Os sistemas de recomendação baseados em descoberta de conhecimento fundamentam as recomendações a partir da correspondência entre as necessidades do usuário e o conjunto de

itens disponíveis (de Souza, 2013). O sistema utiliza o conhecimento que tem dos usuários e dos itens disponíveis para fazer inferências sobre as preferências do cliente. A recomendação é baseada na frequência com que as ações acontecem conjuntamente (Motta et al., 2011).

Considere um site de vendas e que cada interação do cliente seja uma transação, onde uma transação seja um conjunto de itens. Considere  $X$  e  $Y$  transações únicas, que podem ser uma avaliação, uma compra ou até a navegação. A regra de associação  $X \rightarrow Y$  significa que após o evento  $X$  acontecer é esperado que o evento  $Y$  aconteça.

A regra  $X \rightarrow Y$  possui dois critérios na geração da associação: suporte e confiança. Suporte representa a frequência com que os eventos  $X$  e  $Y$  aconteceram simultaneamente e a confiança representa a probabilidade de acontecer  $Y$  uma vez que aconteceu  $X$  (Motta et al., 2011).

Considere a Tabela 2.1 uma base de dados sobre avaliação dos filmes. Cada linha representa um usuário e cada coluna representa o filme. A marcação  $\checkmark$  representa “gostei” e a marcação  $\times$  representa “não gostei”.

Tabela 2.1: Avaliação de filmes

ID	Godfather	O Rei Leão	Interstellar	Harry Potter	American Pie	It Follows
1	$\checkmark$	$\checkmark$		$\checkmark$		
2	$\checkmark$	$\times$	$\checkmark$	$\times$	$\checkmark$	$\times$
3	$\times$	$\checkmark$		$\checkmark$		
4	$\checkmark$		$\checkmark$	$\checkmark$	$\times$	$\checkmark$
5	$\times$			$\times$		
6		$\checkmark$	$\times$	$\checkmark$	$\checkmark$	
7	$\checkmark$	$\times$		$\times$	$\times$	$\checkmark$
8			$\times$	$\times$	$\checkmark$	
9	$\checkmark$	$\times$		$\checkmark$	$\times$	$\checkmark$
10	$\times$	$\checkmark$		$\checkmark$	$\times$	

Analisando a tabela, por intuição, cria-se a expectativa errada de que usuários que gostam de Rei Leão também gostam de Harry Potter, e vice versa. Usando a filtragem baseada em descoberta de conhecimento, entende-se que tais expectativas deveriam ser diferentes. Considere que se deseja um suporte superior a 25% e uma confiança superior a 80%. Isso significa que se deseja regras que reflitam padrões que aparecem em pelo menos 25% das transações e que se tenha uma expectativa de pelo menos 80% de sucesso.

Adotando a sigla  $RL$  para denominar o filme O Rei Leão e a sigla  $HP$  para denominar o filme Harry Potter. Assim, o suporte e a confiança são calculados da seguinte maneira:

$$F_{sup}(RL, HP) = \frac{4}{10} = 40\% \quad (2.5)$$

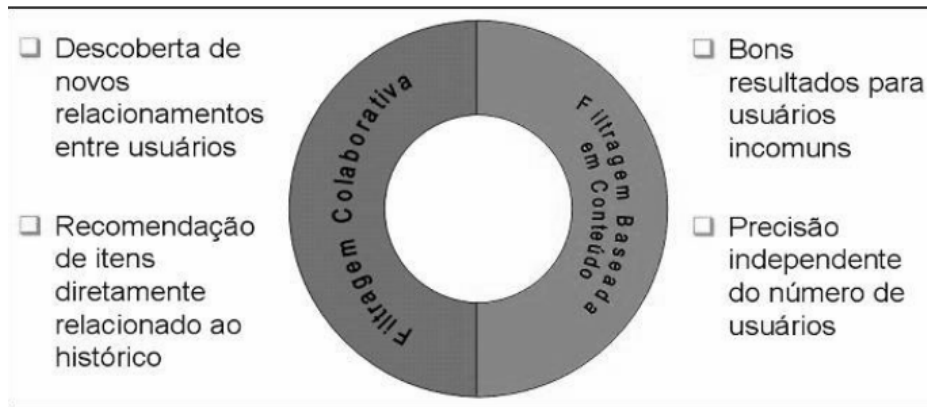


Figura 2.1: Filtragem Híbrida. Fonte: (Cazella et al., 2010)

$$F_{conf}(RL- > HP) = \frac{40\%}{40\%} = 100\% \quad (2.6)$$

$$F_{conf}(HP- > RL) = \frac{40\%}{60\%} = 66\% \quad (2.7)$$

Após a aplicação da técnica da descoberta de conhecimento, nota-se que os clientes que gostaram de Rei Leão também gostem do filme Harry Potter, mas não vice-versa.

Como a recomendação é gerada pela correlação entre itens, pode-se calcular essa correlação a priori. Entretanto, o processamento computacional é caro e cresce exponencialmente com o aumento de itens (Motta et al., 2011). de Souza (2013) argumenta que a principal vantagem dessa técnica é aumentar a precisão da recomendação e evitar limitações intrínsecas das abordagens colaborativas e baseadas em conteúdo. Entretanto, requer a necessidade de se adquirir conhecimento - que é a dificuldade de muitas aplicações que utilizam inteligência artificial.

#### 2.1.2.4 Filtragem Híbrida

Cazella et al. (2010) diz que a filtragem híbrida procura combinar os pontos fortes da filtragem baseada em conteúdo e a filtragem colaborativa. A filtragem híbrida utiliza as vantagens proporcionais dos métodos baseados em conteúdo e filtragem colaborativa, unificando ambas as técnicas e descartando suas desvantagens, como ilustra a Figura 2.1.

Burke (2007) define filtragem híbrida como qualquer sistema de recomendação que combine múltiplas técnicas de recomendação juntas para produzir uma saída. Não há razão para que diversas técnicas do mesmo tipo não possam ser caracterizadas como híbrida. Dois diferentes recomendadores baseados em conteúdo podem trabalhar juntos e diversas aplicações utilizam este tipo de filtragem híbrida (Burke, 2007).

### 2.1.3 Formação de perfis de usuários

Para que seja possível recomendar produtos, serviços ou pessoas a um usuário é necessário ter conhecimento sobre quem é este usuário. Antes mesmo de pensar em capturar e armazenar suas informações pessoais e comportamentais é necessário identificar qual o tipo de informação será relevante para a geração da recomendação, visando uma eficiente personalização dos produtos, serviços e pessoas. Para a correta geração da recomendação a definição do perfil do usuário e coleta de informação é imprescindível (Cazella et al., 2010).

#### 2.1.3.1 Perfil de Usuário

Cazella et al. (2010) relata que, para a criação de uma Identidade Visual, é necessário que o usuário tenha definida a sua Identidade Social, também chamada Reputação, e Identidade Interna. A primeira é definida pelos outros usuários e a segunda definida pelo próprio usuário, ou através de técnicas de *Machine Learning*. Ambas as identidades são armazenadas no Perfil do Usuário.

Os perfis de usuário, que podem ser identificados como a Identidade Interna, refletem o interesse do usuário em momentos particulares. Cada componente do Perfil do Usuário é uma característica particular. Cazella et al. (2010) diz que o Perfil do Usuário pode ser visto como uma base de dados onde a informação, interesse e preferência sobre o usuário é armazenada e dinamicamente mantida.

Enquanto na vida real normalmente as pessoas confiam em recomendações de outras pessoas, oriundas de cartas de recomendação ou comentários de filmes, na vida digital as recomendações começam a ser usadas como informações confiáveis de opiniões (reputação) de outras pessoas e produtos utilizados por elas (Nunes et al., 2008).

Existem diversos tipos de perfis de usuário de diferentes complexidades na *Web*. Eles são encontrados principalmente em *e-commerce*, *e-learning* e *e-community*. Cazella et al. (2010) cita uma Modelagem Genérica de Usuário usada como *shell* para criação de categorias de informações sobre o usuário objetivando personalizar as aplicações web.

Nas definições de modelos de usuário, Cazella et al. (2010) menciona uma Ontologia<sup>1</sup> de um Modelo de Usuário Geral, chamada *GUMO*, e descreve como um modelo ubíquo de Modelo de Usuário, que inclui várias informações básicas do usuário, como informações de contato, informações demográficas, habilidades fisiológicas e psicológicas, estado emocional, estado mental e nutrição. A Figura 2.2 mostra as dimensões básicas propostas nessa Ontologia.

#### 2.1.3.2 Coleta de informações

Para que se possa recomendar itens a um usuário, é necessário ter conhecimento sobre quem é esse usuário. É necessário capturar e armazenar seus dados pessoais e comportamentais

---

<sup>1</sup>Especificação de um conceito

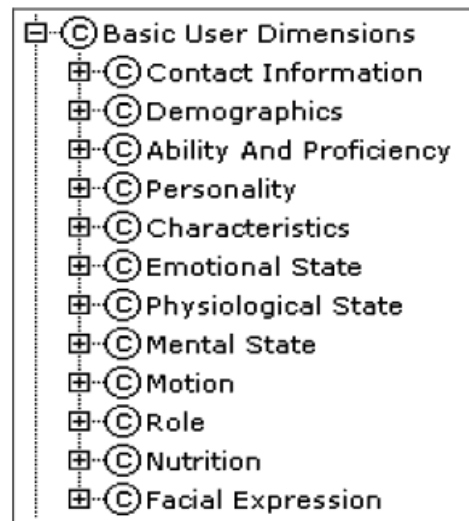


Figura 2.2: Dimensões básicas do usuário segundo modelo GUMO FONTE: (Cazella et al., 2010)

relativos. Portanto, é essencial que se possa identificar o usuário no momento em que ele acessa o sistema onde foram implantadas as rotinas de recomendação (Reategui e Cazella, 2005).

As duas principais formas de identificação do usuário são:

- Identificação no servidor: disponibiliza ao usuário um formulário para preenchimento de alguns dados pessoais. Solicita obrigatoriamente um login e senha. Essas informações ficam armazenadas em um banco de dados no servidor e são utilizadas para realizar a identificação do usuário conectado;
- Identificação no cliente: utiliza cookies, que são arquivos de texto simples utilizado na comunicação do site com o navegador e contém dados de navegação. Este método assume que cada computador é utilizado por uma única pessoa, sendo possível identificar o usuário a partir da máquina que está utilizando o sistema.

Após a identificação do usuário, a manutenção do perfil se dá através da coleta de dados, que podem ser adquiridas de forma explícita ou implícita (Reategui e Cazella, 2005).

Na coleta explícita, conhecida como personalização, o cliente indica de forma espontânea os seus interesses. A Figura 2.3 exemplifica o uso dessa abordagem.

Na modalidade implícita, as informações sobre as necessidades e preferências do usuário são adquiridas de acordo com as ações do cliente (Cazella et al., 2010). A Figura 2.4 mostra o uso do interesse implícito bem como a utilização da filtragem baseada em conteúdo.



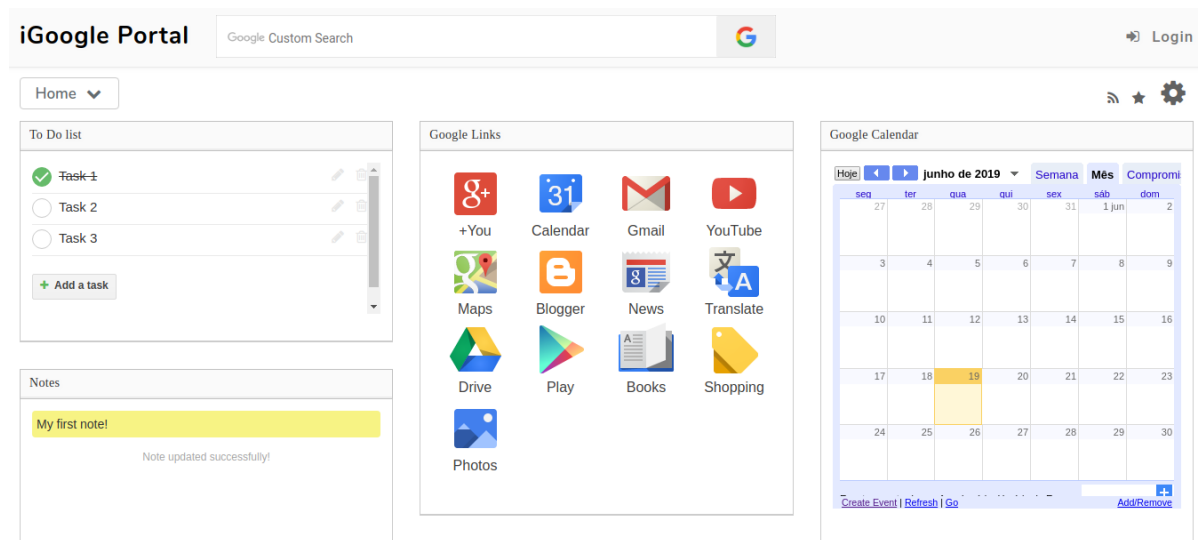


Figura 2.3: iGoogle Portal utilizando a coleta de dados de forma explícita.

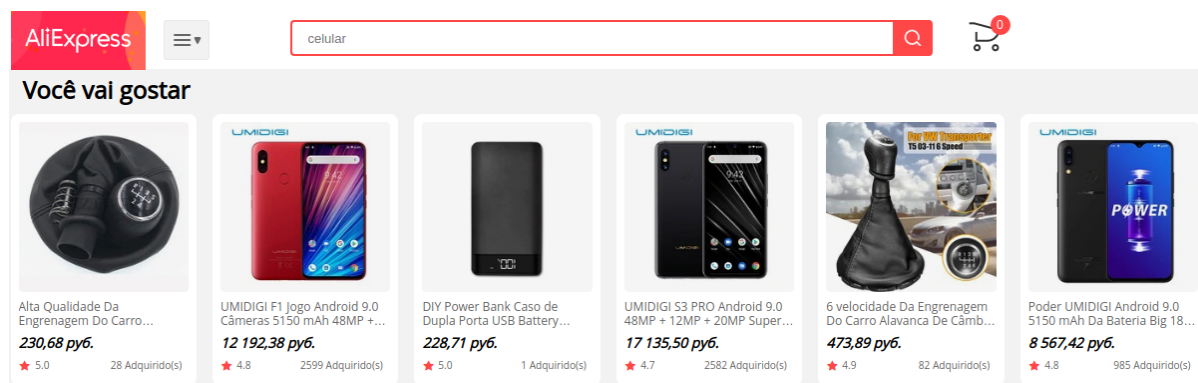


Figura 2.4: Site AliExpress utilizando a coleta de dados implícita para recomendar produtos

### 2.1.3.3 Traço de Personalidade

A teoria da atração interpessoal pode ditar a maneira como personalidades humanas se interagem. Na literatura psicológica, há fortes indicações de que as pessoas preferem interagir com outros que tenham personalidade similares (Cazella et al., 2010). Personalidades são relativamente estáveis e previsíveis, entretanto, não são rígidas e imutáveis, e podem ser definidas por diferentes abordagens. Traços podem ser definidos como padrões habituais de comportamento, pensamento e sentimento. A abordagem do uso de traços para definir a personalidade pode diferir psicologicamente pessoas usando conceituação e mensuráveis traços, chamados Traços de Personalidade (Nunes et al., 2008).

Traços de Personalidade foram descritos pela primeira vez com a criação de 17.953 traços (traços comuns e traços individuais) diferentes para descrever a personalidade de um indivíduo. Considerando que diversas diferenças entre indivíduos não afeta suas interações diárias e a fim

de definir um limite de traços de uma forma exponencial, pesquisadores reduziram 99% dos traços pois concluíram que apenas cinco traços eram replicáveis. Desta maneira foi criado o Modelo Big Five, que apesar de representar um alto nível de estrutura de personalidade, não aborda todas as dimensões significativas da personalidade. Assim, psicólogos usam facetas para enriquecer os fatores do Big Five com características mais refinadas (Nunes et al., 2008).

Nunes et al. (2008) propõem armazenar traços psicológicos em Perfil de Usuário, na intenção de oferecer recomendações mais interessantes para os usuários. Como consequência do uso de características psicológicas no Perfil do Usuário, o conhecimento do usuário será capaz de integrar dinamicamente a serviços de conhecimento contextualizados. O raciocínio e a tomada de decisão de seres humanos são fortemente afetadas por aspectos psicológicos. Desta maneira, para manter o mesmo nível de personalização prestado pelos humanos, os computadores devem considerar a “razão” inserida nos aspectos psicológicos do Perfil do Usuário.

A fim de extrair traços humanos (como os fatores do Big Five), são utilizados questionários eletrônicos, chamados Testes de Personalidade. Nunes et al. (2008) propõem a utilização do teste de personalidade NEO-IPIP<sup>2</sup>, que permite avaliar 5 fatores do Big Five, incluindo 6 facetas em cada dimensão, usando descrições refinadas dos traços de personalidade das pessoas e uma consequente maior representação desses traços.

Nunes et al. (2008) argumentam que Teste de Personalidade é um instrumento computacional capaz de medir a diferença individual das pessoas. Essas diferenças são os Traços de Personalidade e revelam pistas sobre a Identidade Interna e a Reputação do usuário. Como a Identidade do Usuário é um canal onde a personalidade aparece, seus traços darão pistas sobre seus futuros comportamentos e necessidades em uma comunidade.

Nunes et al. (2008) definem a Reputação como uma extensão da Identidade Interna. Isso significa que é usado o mesmo tipo de informação presente no Perfil do Usuário (Identidade Interna). Entretanto, a informação é preenchida por um amigo e não por qualquer usuário. Assim, a Identidade é definida como os Traços de Personalidade armazenados no Perfil do Usuário e por seus amigos armazenados na Reputação do Usuário. A Identidade do usuário é muito útil na interação social dentro de uma comunidade, visto que pode fornecer a previsão de comportamento e das necessidades do usuário, enquanto a Reputação também permite a criação de uma relação de confiabilidade entre membros de uma comunidade.

#### 2.1.3.4 O Modelo Big Five

O Modelo Big Five descreve as características humanas de acordo com cinco principais traços: neuroticismo, extroversão, socialização, realização e abertura para o novo. A combinação desses cinco fatores reflete os principais aspectos da personalidade humana. Entretanto, essas

---

<sup>2</sup>O Inventário NEO-IPIP foi criado com base nos inventários de personalidade NEO-PI-R (NEO Personality Inventory), um dos inventários comerciais mais robustos, usados e bem validados do mundo. Enquanto o NEO-PI-R utiliza 240 itens em seu inventário, a proposta do NEO-IPIP traz um inventário de 300 perguntas e é mais robusto comparado ao NEO-PI-R (Nunes et al., 2008).

cinco dimensões são o máximo de abstração da personalidade e cada dimensão é composto de um grande número de características mais específicas da personalidade (Sandy M. Porto, 2011).

O modelo Big Five é um dos modelos mais utilizados até então. Os cinco fatores são considerado a base do espaço da personalidade. De acordo com Sandy M. Porto (2011), as dimensões podem ser descritas da seguinte maneira:

- **Abertura para o novo:** apreciação de arte, emoção, aventura, ideias incomuns, curiosidade e variedade de experiência.
- **Realização:** tendência a mostrar autodisciplina, agir literalmente e realizar objetivos. Rígido em vez de um comportamento espontâneo.
- **Extroversão:** energia, emoções positivas, urgência e uma tendência a estimular o ambiente na companhia de outras pessoas.
- **Socialização:** tendência para ser mais compassivo e cooperativo do que desconfiado e antagonista em relação aos outros.
- **Neuroticismo:** tendência a ter emoções negativas facilmente, como raiva, ansiedade, depressão e vulnerabilidade.

Sandy M. Porto (2011) cita que o inventário de personalidade que usa cada um desses cinco fatores pode ser ainda subdividido em outros fatores, as quais chamaram de facetas.

### 2.1.3.5 Teste de Personalidade

Teste de Personalidade pode ser visto como uma ferramenta computacional capaz de medir as diferentes individuais das pessoas (Nunes et al., 2008). Dentre os diversos testes de personalidade existentes na literatura, o Big Five Inventory (BFI)<sup>3</sup> é um dos inventários mais utilizados (Oliver E John, 1998; John, 1991; Oliver E John, 2008). O BFI é um inventário de autorrelato projetado para medir os cinco fatores do Modelo Big Five. É bastante curto para um inventário de personalidade, composto por 44 questões, e consiste em frases curtas com vocabulário bastante acessível. O questionário não está disponível para domínio público, isto é, pode ser utilizado somente para fins acadêmicos.

O questionário é ideal para ser aplicado em pessoas na faixa etária entre 20 e 60 anos. Caso se deseje aplicar o teste em pessoas com idade inferior a 20 anos, há uma versão com a linguagem adaptada para crianças e adolescentes. Não há versão do teste para pessoas com idade superior a 60 anos (Oliver E John, 1998; John, 1991; Oliver E John, 2008).

<sup>3</sup>Big Five Inventory é um questionário composto por 44 perguntas objetivas. É tido pela academia como o mais viável a ser implementado computacionalmente. Informações sobre direito de uso podem ser adquiridas no site da metodologia, hospedado na Universidade de Berkeley e disponível através do link <https://www.ocf.berkeley.edu/~johnlab/bfi.htm>.

Além do BFI original, composto por 44 itens, e da versão para crianças e adolescentes, há uma versão mais curta do BFI original (Oliver E John, 1998; John, 1991; Oliver E John, 2008). Entretanto, como o BFI é um dos questionários mais curtos para aplicação de Testes de Personalidade e a média de conclusão do teste é de menos de 10 minutos, é altamente recomendável o uso do BFI original (Oliver E John, 1998; John, 1991; Oliver E John, 2008).

O Big Five Inventory e suas respectivas regras de pontuação estão anexadas ao final desta monografia, no Capítulo 5.

### 2.1.3.6 Recomendação para Grupos de Pessoas

Sistemas de recomendação tradicionalmente recomenda produtos a um usuário individual. Quando se busca sair do âmbito individual e migrar para a recomendação de grupos, alguns problemas são encontrados, por exemplo, adquirir as preferências do grupo ou ajudar o grupo em tomadas de decisão. Quijano-Sánchez et al. (2010) citam três principais abordagens para gerar uma agregação de preferências baseado nas preferências individuais do usuário: (a) fundindo as recomendações feitas para os indivíduos; (b) agregação de classificações para indivíduos e (c) construir um modelo de preferência de grupo, sendo (b) e (c) as abordagens mais comumente usadas.

Quijano-Sánchez et al. (2010) determinam que satisfação geral de um grupo não é sempre a agregação da satisfação de seus membros, visto que diferentes pessoas possuem expectativas e comportamentos distintos em situação de conflito, e propõe caracterizar os usuários usando o Instrumento de Modelo de Conflito Thomas-Kilmann (TKI), que descreve o comportamento pessoal em situações de conflito em duas dimensões: assertiva e cooperativa, sendo usadas para definir cinco modelos de personalidade lidando com conflitos: competindo, colaborando, evitando, acomodando e comprometido.

Quijano-Sánchez et al. (2010) defendem a utilização do teste TKI para determinar as personalidades dos usuários, onde cada  $u_i$  calculado representa o quão egoísta ou cooperativista o usuário é. Essa representação é chamada  $CMW_i$  e significa o comportamento predominante do usuário de acordo com a avaliação do TKI. A função CMW é um valor em  $[0,1]$ , sendo que 0 representa uma pessoa muito cooperativa e 1 representa uma pessoa muito egoísta.

Quijano-Sánchez et al. (2010) criam a recomendação de grupos baseada na mistura de recomendações individuais. A recomendação individual é um sistema colaborativo baseado na avaliação de outros usuários. Para se obter os usuários similares, se usa a avaliação dos usuários sobre um produto, sendo que os usuários com gostos similares avaliaram o mesmo produto de uma maneira similar. Então, a avaliação de usuários similares são usadas para inferir avaliação desconhecidas e para criar novas recomendações.

A abordagem utiliza o tipo de personalidade como peso na influência das avaliações durante o processo de recomendação. Para computar as recomendações baseadas em personalidade, é usado uma função agregação de satisfação média:

$$pbr(u_i, m) = \frac{1}{|G|} \sum_{j \in G \& j \neq i} (ir_{i,m} + pd(u_i, u_j)) \quad (2.8)$$

Sendo:

- $pbr(u_i, m)$  é a recomendação baseada em personalidade do usuário  $u_i$  em relação ao produto  $m$ .
- $|G|$  representa o número de componentes do grupo.
- $ir_{i,m}$  é a avaliação do usuário  $u_i$  em sobre o produto  $m$ .
- $pd(u_i, u_j) = (CMW_j - CMW_i) * \alpha$ . Representa a diferença de personalidade.
- $CMW_j$  é o peso no modo do conflito do usuário.
- $\alpha$  é experimentalmente selecionado e é usado para modificar o impacto das diferenças das personalidades na classificação modificada.

O que se observa é a diferença entre pares de personalidade dos indivíduos de um grupo e se baseia em uma modificação no método de agregação de satisfação média. Tal estratégia mostra que os personagens assertivos terão mais influência na satisfação média quando comparado com os cooperativos. Este fator é calculado como a distancia dos valores de CMW das diferentes personalidades.

Quijano-Sánchez et al. (2010) propõem o uso de confiança social na recomendação para grupos, denominada *trust*, que analisa a relação de amizade entre dois usuários. Considerando o uso de redes sociais, a confiança entre dois usuários pode ser definida como uma análise de diversos fatores que descrevem sua relação. Quijano-Sánchez et al. (2010) propõem o uso dos seguintes fatores de confiança:

- $f_1$ : Distância na rede social.
- $f_2$ : Número de amigos em comum
- $f_3$ : Intensidade da relação: quantas vezes eles escrevem um ao outro.
- $f_4$ : Intimidade da relação: palavras-chave que representam o nível da intimidade.
- $f_5$ : Duração: quanto tempo eles se conhecem.
- $f_6$ : Serviços recíprocos: número de vídeos/músicas postados e jogos/aplicações compartilhados.
- $f_7$ : Variável estrutural: interesses em comum descritos no perfil do usuário, como filmes, livros, etc.

- $f_8$ : Distância social: quantas das seguintes propriedades são compartilhadas: informações políticas, educacionais, religiosas e demográficas.
- $f_9$ : Status: valor dependendo do tipo do status: casal, família, melhores amigos.
- $f_{10}$ : Imagem: porcentagem de imagem que eles aparecem juntos.

O valor final de confiança  $trust(u_i, u_j)$  é uma média ponderada dos fatores descritos anteriormente.

$$trust(i, j) = t_{ij} = \sum_{k=1}^{10} (\alpha_k * f_k(u_i, u_j)) \quad (2.9)$$

sendo que  $\alpha$  representa um vetor de pesos. Estes pesos podem ser obtidos utilizando um algoritmo genético. O objetivo é maximizar a precisão da recomendação do grupo.

Quijano-Sánchez et al. (2010) adicionam o fator de confiança na função agregação de satisfação média. Com a modificação, a nova função é definida por:

$$sgr(u_i, m) = \frac{1}{|G|} \sum_{j \in G, j \neq i} (ir_{i,m} + pd(u_i, u_j) + trust(u_i, u_j)) \quad (2.10)$$

## 2.2 Trabalhos Relacionados

Este capítulo descreve os trabalhos relacionados ao tema desta monografia e se organiza da seguinte maneira: na Subseção 2.2.1, apresentam-se Sistemas de Recomendação utilizando traços de personalidade, seguido do trabalho Personalidade e Confiança Social em Grupos de Recomendação na Subseção 2.2.2. Finalmente, a Subseção 2.2.3 apresenta a extração de Traços de Personalidade utilizando imagens no Instagram.

### 2.2.1 Sistemas de Recomendação com Traços de Personalidade

Em sua pesquisa, Nunes et al. (2008) buscam aprimorar os Sistemas de Recomendação utilizando Traços de Personalidade no Perfil do Usuário. Na recomendação, o cliente é um usuário e o produto é um nome, que deve ser considerado na tomada de decisão do usuário. Enquanto seu foco é na Reputação Psicológica para recomendar uma pessoa ao usuário, este projeto buscar utilizar Traços de Personalidade na Identidade Interna, a fim de recomendar ao usuário, com base nas características psicológicas, os grupos mais semelhantes ao seu perfil.

Nunes et al. (2008) buscam provar em seus experimentos que Sistemas de Recomendação podem ser mais eficientes se forem utilizados Traços de Personalidade, ao invés do uso convencional. É contemplado o Sistema de Recomendação, mostrando sua capacidade de recomendar pessoas, considerado como produto a ser entregue de acordo com uma visão de

produto<sup>4</sup>. O produto gerado é o nome de uma pessoa. Este nome pode ser usado como suporte de conhecimento durante o processo de tomada de decisão para uma pessoa e sua comunidade.

O experimento de Nunes et al. (2008) foi apresentado para as Eleições para Presidente na França, ocorrida em Abril de 2007. O Sistema de Recomendação foi usado para gerar uma recomendação privada considerando a melhor escolha de um dentre dois candidatos as presidente para um usuário votar. O foco do experimento é na Reputação Psicológica do Usuário (Perfil Psicológico do Usuário de acordo com a visão de seus amigos) com base no *feedback* das pessoas sobre os candidatos em um caso específico da presidência francesa. Para a criação do Perfil/Reputação Psicológico do Usuário foi usado três Inventários NEO-IPIP, um para o "presidente ideal" e um para cada candidato à presidência, com aproximadamente 100 pessoas.

Com as respostas, foi possível modelar aspectos psicológicos de dois candidatos à presidência na França, além de um Presidente Ideal imaginário. A recomendação feita se baseou na reputação dos candidatos à presidência e em um personagem imaginário, que seria o presidente ideal do participante.

Nunes et al. (2008) realizaram dois tipos de recomendações: ambas baseadas em 5 fatores do Big Five, sendo uma também baseada em 30 facetas. Quando utilizado as 30 facetas em conjunto com os 5 fatores do Big Five, a recomendação é mais refinada, tendo como resultado uma precisão de 100%. Isto quer dizer que 100% dos casos recomendados pelo Sistema foram compatíveis com o candidato escolhido pelo usuário durante as eleições. No experimento utilizando somente 5 fatores do Big Five, o resultado foi de 80%. Apesar da dificuldade de implementação de um questionário detalhado (com 30 facetas) o resultado final de uma recomendação é 25% superior se comparado somente com a utilização do Big Five.

## 2.2.2 Personalidade e Confiança Social em Grupos de Recomendação

Em seu trabalho, Quijano-Sánchez et al. (2010) descrevem novas ideias para melhorar Sistemas de Recomendação destinadas à grupos de pessoas. Sua abordagem maximiza a satisfação dos grupos usando a personalidade das pessoas e relações sociais entre pessoas do grupo. São usados dois casos de teste baseado no domínio de recomendação de filmes com grupos heterogêneos. É apresentado a comparação de três sistemas de recomendação: o primeiro utiliza a recomendação básica, considerando somente as preferências do usuário, o segundo usando a personalidade no perfil do usuário e o terceiro recomendando com base na personalidade do usuário e confiança social.

---

<sup>4</sup>O Sistema de Recomendação gera como produto uma pessoa, e este produto é considerado um pacote fechado representado por seu nome. A visão do produto significa que um usuário recebe uma recomendação de outro usuário, como um nome a ser levado em consideração no seu processo de tomada de decisão. O SR fornece uma resposta passiva, diferente de uma exibição de serviço em que a resposta gera uma interação dinâmica.

A abordagem cria uma recomendação de grupo baseada na mistura de recomendações individuais. Cada recomendação individual é um sistema colaborativo baseado na avaliação de outros usuários. A proposta é utilizar a personalidade como peso no processo de recomendação. Quijano-Sánchez et al. (2010) fazem uso do Teste de Personalidade TKI, composto por 30 perguntas, para a criação do Perfil do Usuário.

Os experimentos em (Quijano-Sánchez et al., 2010) utilizam um conjunto de 100 pessoas fictícias. A cada pessoa é atribuído aleatoriamente um valor CMW que reflete a sua personalidade - trabalhando com dados reais, este é o valor computado usando os resultados do teste de personalidade. É definido cinco tipos diferentes de personalidade: muito egoísta, egoísta, tolerante, cooperativo e muito cooperativo. É utilizado 20 usuários para cada tipo de personalidade.

(Quijano-Sánchez et al. (2010) selecionam 50 filmes heterogêneos. Cada filme recebe uma avaliação de um usuário, em uma escala de 0.0 até 5.0 e baseada no perfil do usuário. É obtido quais filmes seriam classificados com 4.5 ou mais e definido o conjunto de filmes favoritos pelo usuário. Organiza a listagem do cinema em ordem de preferência de acordo com um teste de similaridade baseada em conteúdo. Então é reproduzido uma situação de vida real, onde os membros do grupo argumentariam suas preferências levando em consideração a personalidade e amizade e assim obtém os verdadeiros filmes favoritos para o grupo.

(Quijano-Sánchez et al. (2010) mostram que a utilização de confiança social em sistemas de recomendação influencia na tomada de decisão de um grupo. Enquanto Quijano-Sánchez et al. (2010) propõem aprimorar sistemas de recomendação para grupos de usuários com auxílio da confiança social dos usuários, este projeto busca recomendar usuários à grupos de acordo com a personalidade.

### 2.2.3 Prever Traços de Personalidade com Imagens do Instagram

(Ferwerda et al. (2015) propõem a extração de personalidade focalizando especificamente a relação entre as características da imagem de uma coleção do Instagram<sup>5</sup> e os traços de personalidade do usuário. A pesquisa tem como objetivo estimar as preferências do usuário sem usar extensivos questionários ou observações. Para categorizar as personalidades, o modelo Big Five é aplicado. O experimento foi aplicado a 113 usuários que residem nos Estados Unidos, coletando 22,398 imagens.

(Ferwerda et al. (2015) extraem diversas características das imagens dos usuários - brilho, saturação, prazer/excitação/domínio, matiz e recursos baseado em conteúdo. As características da imagem são utilizadas para criar uma correlação entre o usuário e sua personalidade. Os resultados sugerem a existência de uma relação entre os traços de personalidade e o modo como usuários publicam as fotos. Assim, com a análise de imagens do Instagram, traços de personalidade podem ser inferidos.

---

<sup>5</sup><https://www.instagram.com/>



Os resultados em (Ferwerda et al., 2015) são satisfatórios para três tipos de personalidade: Abertura a experiência, Conscienciosidade e Amabilidade. Para os tipos de personalidade Extroversão e Neuroticismo, a pesquisa demonstra menor confiança na correlação entre as imagens analisadas e a personalidade do usuário.

Ferwerda et al. (2015) buscam obter imagens publicadas no Instagram de um usuário, correlacionando as características da imagem extraída com a personalidade do usuário. Entretanto, o método utilizado se limita a prever traços de personalidade, ignorando a recomendação baseada na personalidade prevista.

## Capítulo 3

# Desenvolvimento

O objetivo deste trabalho é desenvolver um sistema de recomendação capaz de recomendar, de maneira adequada, repúblicas a estudantes, baseado na similaridade entre suas personalidades. A ferramenta é acessada via *Web*, sendo assim, será usado a Identificação no Servidor. Para modelar as personalidades dos usuários, representadas nesta monografia exclusivamente pela Identidade Interna, foi utilizado como base o Modelo de Personalidade Big Five, descrito na Subseção 2.1.3.4, sendo aplicado através do Big Five Inventory (BFI), apresentado na Subseção 2.1.3.5. Uma vez definida a Identidade do Usuário, é possível associá-la a uma república, a qual é representada por meio de uma personalidade, nomeada Perfil do República. A Identidade Interna da República é formada de acordo com a Identidade Interna dos estudantes que o compõem e pode ser constantemente mutável.

Com a posse da Identidade Interna do estudante e do Perfil das Repúblicas, a recomendação se baseia na distância dos atributos de cada personalidade. Como o Modelo Big Five reduz a personalidade à cinco fatores, a distância de dois fatores  $f_1$  e  $f_2$  significa a diferença entre  $f_1$  e  $f_2$ . Em outras palavras:

$$D_{u_i, u_j} = \sum_{k=1}^{k=5} |f_{k_i} - f_{k_j}| \quad (3.1)$$

Sendo:

- $D_{u_i, u_j}$  representa a distância da personalidade do usuário  $i$  em relação ao usuário  $j$ ;
- $k \in \{\text{Abertura para o novo, Realização, Extroversão, Socialização e Neuroticismo}\}$
- $f_k$  representa o valor obtido através da aplicação do BFI para o fator  $k$ .

A divisão deste capítulo está organizada da maneira como se segue. A Seção 3.1 apresenta a Concepção do Sistema. A Seção 3.2 apresenta a modelagem do banco de dados, detalhando as ferramentas a serem utilizadas e as entidades do esquema, seguido da apresentação da ferramenta descrita na Seção 3.3.

## 3.1 Concepção do sistema

Nesta seção, é abordada a concepção do sistema (Sommerville, 2011). A seção está organizada como segue. Na Subseção 3.1.1, é descrito o objetivo desta concepção, seguido da missão e descrição do produto, descrito na Subseção 3.1.2. A Subseção 3.1.3 aborda o escopo positivo e negativo do sistema. O Diagrama de Contexto, Hierarquia de Atores e Diagrama Relacional são apresentados nas Subseções 3.1.4, 3.1.5 e 3.1.6, respectivamente. Os Storyboards são mostrados na Subseção 3.1.7, seguidos dos Backlogs na Subseção 3.1.8 e as Estórias de Usuário sendo descritas na Subseção 3.1.9. Finalmente, a Subseção 3.1.10 relata os cenários de teste.

### 3.1.1 Objetivo

O objetivo desta concepção é especificar um sistema *Web* capaz de recomendar ao estudante as repúblicas que mais possuem compatibilidade com sua personalidade. A ideia é construir um software responsável por armazenar a personalidade dos calouros e das repúblicas, realizando a recomendação baseada na distância das personalidades.

### 3.1.2 Missão e Descrição do Produto

#### 3.1.2.1 Nome e Componentes

A proposta para o nome do produto é que se chame FindMyRepublic (FMR). O sistema é dividido em dois componentes, a depender do ator que está acessando o software (Estudante e República). A interface a ser exibida dependerá do tipo do usuário logado.

#### 3.1.2.2 Missão

Auxiliar na redução da evasão de estudantes nas repúblicas. Contribuir para que o estudante encontre uma república com o perfil que deseje. Ajudar repúblicas a encontrarem estudantes que complementem o perfil da república.

#### 3.1.2.3 Descrição

O produto é um sistema *Web* que oferece serviços de relatórios de personalidade de estudantes e repúblicas obtidos a partir do formulário Big Five Inventory. Também pode oferecer serviços de exibição de informação dos usuários, uma vez que eles permitam a exibição.

### 3.1.3 Escopo

#### 3.1.3.1 Escopo positivo

O sistema permite ao estudante consultar e visualizar dados que representam a personalidade das repúblicas. Essa representação é de extrema importância, pois diz respeito a

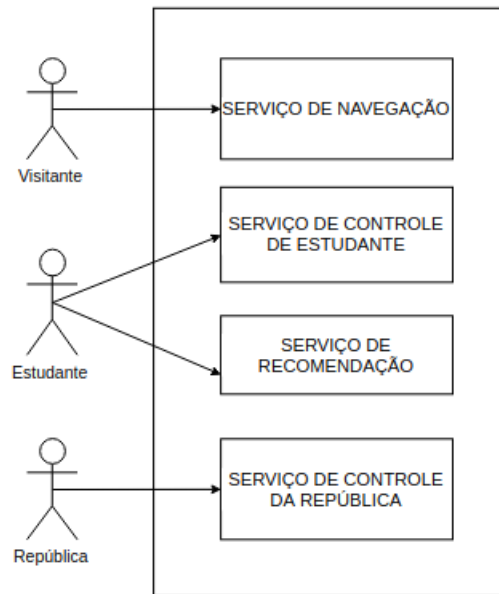


Figura 3.1: Diagrama de Contexto do Sistema

agregação das personalidades dos que ali residem, podendo auxiliar na tomada de decisão do estudante sobre qual república escolher para morar.

### 3.1.3.2 Escopo negativo

O sistema fornece a recomendação de repúblicas para estudantes baseado na personalidade. Entretanto, a limitação se diz respeito à recomendação do estudante para a república. Esta funcionalidade não será desenvolvida.

### 3.1.4 Diagrama de Contexto

A Figura 3.1 mostra o diagrama de contexto da aplicação bem como seus atores e serviços. A Tabela 3.1 mostra a descrição dos serviços.

### 3.1.5 Hierarquia dos Atores

A Figura 3.2 mostra a hierarquia dos atores da aplicação, mostrando que o Estudante e a República podem realizar tanto as suas tarefas como as tarefas do Visitante. A Tabela 3.2 mostra a descrição dos atores.

Tabela 3.1: Descrição do Diagrama de Contexto

<b>Serviço</b>	<b>Descrição</b>
Navegação	Permite ao usuário navegar pelas abas da aplicação e realizar buscas por repúblicas a partir de palavras-chave.
Controle do Estudante	Permite ao estudante gerenciar seus dados demográficos
Recomendação	Permite ao estudante solicitar uma recomendação ao sistema
Personalidade	Permite ao estudante realizar o Teste de Personalidade aplicado através do BFI para calcular os fatores da personalidade. Permite à república a calcular a personalidade com base nos estudantes vinculados a mesma.
Controle da República	Permite à república gerenciar seus dados demográficos.

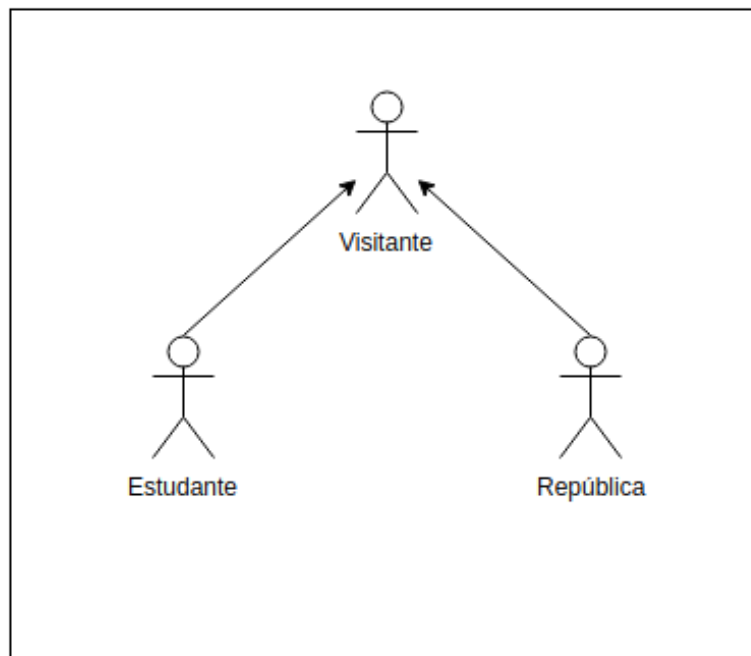


Figura 3.2: Hierarquia dos Atores

Tabela 3.2: Descrição dos Atores

<b>Ator</b>	<b>Descrição</b>
Visitante	Qualquer pessoa física interessada em utilizar o sistema.
Estudante	Pessoa específica interessada em conhecer a personalidade de possíveis repúblicas com o objetivo de moradia.
República	Grupo de estudantes específicos interessados em compartilhar a personalidade para que Estudantes possam encontrá-la.

Tabela 3.3: Descrição do Diagrama Relacional

Relação	Descrição
Estudante	É um usuário e pode se relacionar com República
República	É um usuário e pode ser composta por diversos Estudantes
Usuário	Pode ser um Estudante ou uma República e possui um Perfil
Questionario	Compõe o Inventário
Inventário	É vinculado a um Perfil e composto por uma lista de Questionário
Personalidade	Pertence a um Perfil
Perfil	É de um Usuário, possui um Inventário e uma Personalidade

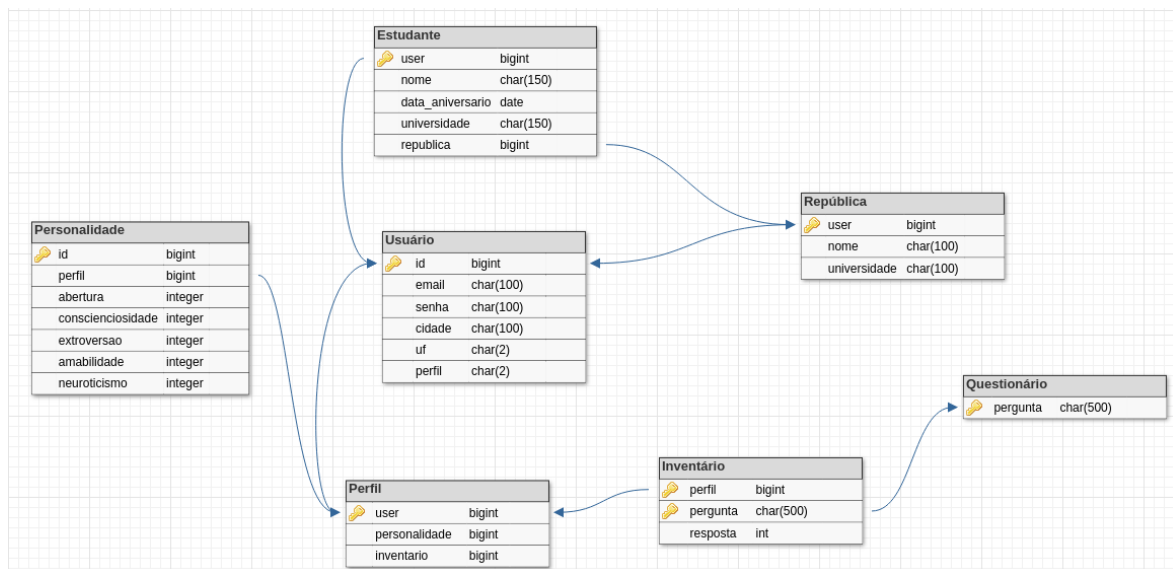


Figura 3.3: Diagrama Relacional do Sistema

### 3.1.6 Diagrama Relacional

A Figura 3.3 mostra o Diagrama Relacional do Sistema e a Tabela 3.3 apresenta sua descrição.

### 3.1.7 Storyboard

Os storyboards foram obtidos a partir de cópias das telas da interface do sistema. Todas as páginas contêm o menu.

A Tabela 3.4 apresenta a descrição de cada figura.

### 3.1.8 Backlog

**Serviço de Navegação** permite:



Figura 3.4: Index

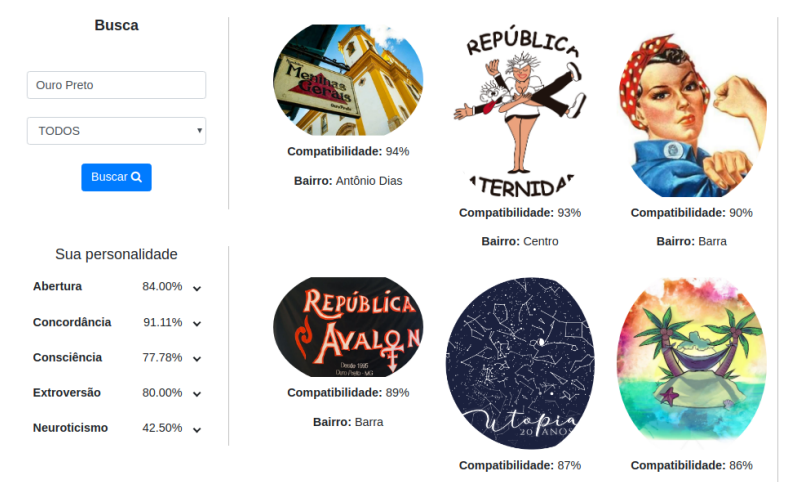


Figura 3.5: Recomendação

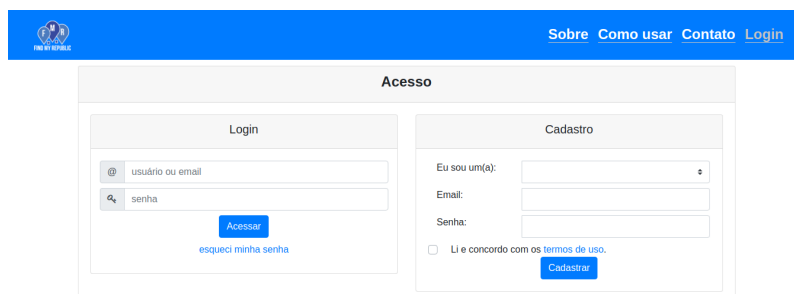


Figura 3.6: Login/Cadastro



Figura 3.7: Perfil



Figura 3.8: Inventário

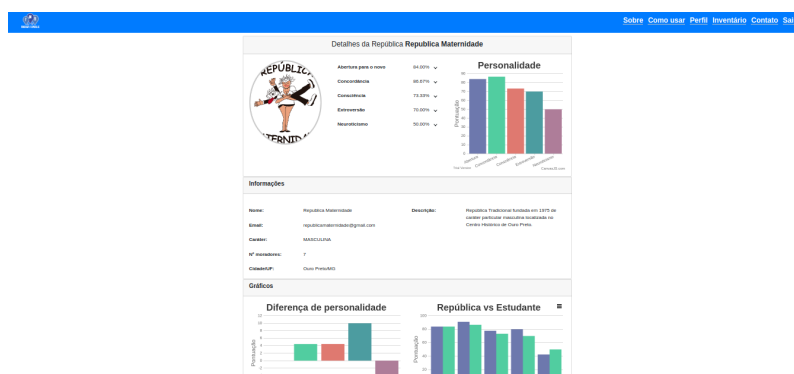


Figura 3.9: Detalhes



Tabela 3.4: Descrição das telas do sistema

<b>Figura</b>	<b>Nome</b>	<b>Conteúdo</b>	<b>Acesso</b>	<b>Direciona para</b>
Figura 3.4	Index	Página inicial do sistema, contém um campo de busca que direciona para Recomendação	Tela Inicial	Recomendação
Figura 3.5	Recomendação	Página de recomendação baseada na personalidade. Contém a lista de repúblicas ordenada de acordo com a distância das personalidades	Index	Detalhes.
Figura 3.6	Login	Página de criação de usuário e acesso ao sistema, contendo seus respectivos formulários.	Menu	Index
Figura 3.7	Perfil	Página de Perfil do usuário que está acessando o sistema, contendo seu formulário.	Menu	-.
Figura 3.8	Inventário	Página que contém o Teste de Personalidade a ser respondido pelo Estudante	Menu	-
Figura 3.9	Detalhes	Página de Detalhes da República, contendo algumas informações e gráficos comparativos de personalidade	Recomendação	-

1. Efetuar consulta
2. Acessar república
3. Efetuar cadastro

**Serviço de Controle do Estudante** permite:

1. Efetuar login
2. Alterar dados
3. Visualizar dados
4. Excluir conta

5. Atualizar personalidade
6. Vincular a uma República

**Serviço de Recomendação** permite:

1. Executar recomendação

**Serviço de Controle da República** permite:

1. Efetuar login
2. Alterar dados
3. Visualizar dados
4. Excluir conta
5. Atualizar personalidade
6. Vincular estudante

### 3.1.9 Estória de Usuário

**Serviço de Navegação**

1. Efetuar consulta
  - **Como um** visitante
  - **Com o objetivo de** obter uma lista de repúblicas
  - **Gostaria de** efetuar uma consulta
2. Acessar república
  - **Como um** visitante
  - **Com o objetivo de** obter dados de uma república
  - **Gostaria de** acessar os dados de uma república
3. Efetuar cadastro
  - **Como um** visitante
  - **Com o objetivo de** inserir dados no sistema
  - **Gostaria de** efetuar um cadastro

### Serviço de Controle do Estudante

#### 1. Efetuar login

- **Como um** estudante
- **Com o objetivo de** acessar a página inicial do sistema
- **Gostaria de** efetuar o login

#### 2. Atualizar dados

- **Como um** estudante
- **Com o objetivo de** alterar meus dados cadastrados
- **Gostaria de** atualizar meus dados

#### 3. Visualizar dados

- **Como um** estudante
- **Com o objetivo de** obter meus dados cadastrados
- **Gostaria de** visualizar meus dados.

#### 4. Excluir conta

- **Como um** estudante
- **Com o objetivo de** remover meus dados do sistema
- **Gostaria de** excluir minha conta

#### 5. Atualizar personalidade

- **Como um** estudante
- **Com o objetivo de** alterar os dados da personalidade
- **Gostaria de** preencher o BFI

#### 6. Vincular a uma República

- **Como um** estudante
- **Com o objetivo de** contribuir com a personalidade da república
- **Gostaria de** vincular meu perfil ao perfil de uma república

### Serviço de Recomendação

1. Executar recomendação
  - **Como um** estudante
  - **Com o objetivo de** obter uma lista de repúblicas com fatores de personalidade semelhante ao meu
  - **Gostaria de** solicitar uma recomendação

### Serviço de Controle da República

1. Efetuar login
  - **Como uma** república
  - **Com o objetivo de** acessar a página inicial do sistema
  - **Gostaria de** efetuar o login
2. Alterar dados
  - **Como uma** república
  - **Com o objetivo de** alterar meus dados cadastrados
  - **Gostaria de** atualizar meus dados
3. Visualizar dados
  - **Como uma** república
  - **Com o objetivo de** obter meus dados cadastrados
  - **Gostaria de** visualizar meus dados.
4. Excluir conta
  - **Como uma** república
  - **Com o objetivo de** remover meus dados do sistema
  - **Gostaria de** excluir minha conta
5. Atualizar personalidade
  - **Como uma** república
  - **Com o objetivo de** alterar os dados da personalidade
  - **Gostaria de** calcular minha personalidade
6. Vincular estudante

- **Como uma** república
- **Com o objetivo de** adicionar o estudante a lista de estudantes
- **Gostaria de** adicionar o estudante a lista de estudantes

### 3.1.10 Cenário de Teste

#### Serviço de Navegação

##### 1. Efetuar consulta

- DADO que sou um visitante E desejo realizar uma consulta a partir de uma cidade QUANDO eu preencher o campo de pesquisa E clicar no botão pesquisar ENTÃO os dados das repúblicas começarão a ser exibidos
- DADO que eu escreva a minha consulta E a busca não retorne resultado ENTÃO o sistema exibe a mensagem "SEM RESULTADO PARA A BUSCA"

##### 2. Acessar república

- DADO que sou um visitante e escreva a minha consulta E a busca retorne uma lista de repúblicas QUANDO eu clicar na república ENTÃO o sistema irá mostrar a página referente à república

##### 3. Efetuar cadastro

- DADO que sou um visitante E a tela atual seja a tela de cadastro E eu preencha os campos obrigatórios de cadastro E aperte o botão cadastrar ENTÃO o sistema irá inserir os dados no banco de dados E mostrar a página inicial do software.

#### Serviço de Controle do Estudante

##### 1. Efetuar login

- DADO que sou um estudante E que esteja desconectado do sistema E esteja na tela de login E preencha os campos com dados autenticados ENTÃO o sistema irá mostrar a página inicial do software
- DADO que sou um estudante E que esteja desconectado do sistema E esteja na tela de login E preencha os campos com dados inválidos ENTÃO o sistema irá exibir a mensagem "DADOS INVÁLIDOS"
- DADO que sou um estudante E que esteja logado no sistema E aperte o botão Logout ENTÃO o sistema irá mostrar a página de login.

##### 2. Alterar dados

- DADO que sou um estudante E que esteja logado no sistema E esteja na página do alterar meu perfil E altere os dados do formulário E aperte o botão Salvar ENTÃO o sistema irá mostrar a página de visualizar meu perfil com os dados atualizados
3. Visualizar dados
    - DADO que sou um estudante E que esteja logado E que eu aperte o botão Meu Perfil ENTÃO o sistema irá mostrar a página referente ao meu perfil
  4. Excluir conta
    - DADO que sou um estudante E que esteja logado E que eu não tenha vinculo com República E aperte no botão Excluir Conta ENTÃO o sistema irá remover todos os meus dados armazenados do banco de dados
    - DADO que sou um estudante E que esteja logado E que eu tenha vínculo com República E aperte o botão Excluir Conta ENTÃO o sistema irá exibir a mensagem "VOCÊ NÃO PODE EXCLUIR UMA CONTA VINCULADA A UMA REPÚBLICA"
  5. Atualizar personalidade
    - DADO que sou um estudante E que esteja na página do BFI E preencha o questionário E aperte no botão Salvar ENTÃO o sistema irá calcular os fatores da personalidade com base no formulário preenchido E irá armazenar os valores no banco de dados E mostrar a página do estudante com os dados
  6. Vincular a uma República
    - DADO que sou um estudante E que não tenha república vinculada E que esteja na página de detalhe de uma república E aperte no botão "Sou membro da república" ENTÃO o sistema irá vincular meu perfil ao da república E irá armazená-lo no banco de dados E irá recarregar a página de detalhe da república
    - DADO que sou um estudante E que tenha república vinculada E que esteja na página de detalhe de uma república E aperte no botão "Sou membro da república" ENTÃO o sistema irá exibir a mensagem "VOCÊ JÁ FAZ PARTE DE OUTRA REPÚBLICA"

### Serviço de Recomendação

1. Executar recomendação
  - DADO que sou um estudante E esteja na página de Recomendação E aperte o botão Pesquisar ENTÃO o sistema irá exibir uma lista de repúblicas ordenada de acordo com a distância entre a minha personalidade e a da república

## Serviço de Controle da República

### 1. Efetuar login

- DADO que sou uma república E que esteja desconectado do sistema E esteja na tela de login E preencha os campos com dados autenticados ENTÃO o sistema irá mostrar a página inicial do software
- DADO que sou uma república E que esteja desconectado do sistema E esteja na tela de login E preencha os campos com dados inválidos ENTÃO o sistema irá exibir a mensagem "DADOS INVÁLIDOS"
- DADO que sou uma república E que esteja logado no sistema E aperte o botão Logout ENTÃO o sistema irá mostrar a página de login.

### 2. Alterar dados

- DADO que sou uma república E que esteja logado no sistema E esteja na página de alterar meu perfil E altere os dados do formulário E aperte o botão Salvar ENTÃO o sistema irá mostrar a página de visualizar meu perfil com os dados atualizados

### 3. Visualizar dados

- DADO que sou uma república E que esteja logado E que eu aperte o botão Meu Perfil ENTÃO o sistema irá mostrar a página referente ao meu perfil

### 4. Excluir conta

- DADO que sou uma república E que esteja logado E aperte no botão Excluir Conta ENTÃO o sistema irá remover todos os meus dados armazenados do banco de dados

### 5. Atualizar personalidade

- DADO que sou uma república E que esteja na página do perfil da república E que aperte o botão Calcular Fatores ENTÃO o sistema irá calcular os fatores da personalidade da república baseado nos fatores da personalidade de cada estudante vinculado E armazenar os valores E exibir a página da república

### 6. Vincular estudante

- DADO que sou uma república E que esteja na página de um estudante E o estudante não tenha república vinculada E eu aperte no botão "Adicionar estudante" ENTÃO o sistema irá adicionar o estudante à lista de estudantes E irá armazená-lo no banco de dados E irá recarregar a página de detalhe do estudante
- DADO que sou uma república E que esteja na página de um estudante E o estudante tenha república vinculada ENTÃO o sistema irá ocultar o botão "Adicionar estudante"

## 3.2 Modelagem do Banco de Dados

Nesta seção, é descrita brevemente o esquema de Entidade e Relacionamento (ER) do banco de dados (Hector Garcia-Molina, 2001), bem como seus atributos. O esquema ER conta com 7 entidades, 28 atributos e 7 relacionamentos. O sistema de gerência de banco de dados (SGBD) utilizado neste trabalho é o MySQL, pelo fato de ser um SGBD *open-source*, com o auxílio do SGBD MySQL Workbench. O diagrama foi feito utilizando a ferramenta brModelo<sup>1</sup>. A Subseção 3.2.2 mostra o Dicionário de Dados, dado o Diagrama Entidade Relacionamento, mostrado na Subseção 3.2.1

### 3.2.1 Diagrama Entidade Relacionamento

A empresa FindMyRepublic é uma empresa digital que pode ser acessada via página web<sup>2</sup>. O usuário pode ser um Estudante ou uma República. Uma república pode ser composta por diversos estudantes, enquanto um estudante pode compor somente uma república.

O usuário, república ou estudante, possui um perfil. O perfil pode representar a personalidade de um usuário. Caso o usuário seja um estudante, o perfil deve responder a um inventário (que fica vinculado ao perfil) contendo questionários (perguntas) e as respostas. Caso o usuário seja uma república, a personalidade do seu perfil é calculada com base na personalidade dos seus moradores.

A Figura 3.10 mostra o Diagrama Entidade Relacionamento construído para a criação do sistema proposto. O esquema ER conta com 7 entidades, 28 atributos e 7 relacionamentos, contemplando a organização necessária para o desenvolvimento do software.

### 3.2.2 Dicionário de Dados

#### Usuário

Tipo de entidade que representa um usuário do sistema, representado por seus atributos demográficos e personalidade. Se relaciona com Pessoa, República e Perfil.

- **id (primary key):** atributo numérico, monovalorado e armazenado que identifica um usuário.
- **email:** atributo textual, monovalorado e armazenado que identifica um usuário.
- **senha:** atributo textual, monovalorado e armazenado, usado para autenticação.
- **cidade:** atributo simples, monovalorado e armazenado.
- **uf:** atributo simples, monovalorado e armazenado.

<sup>1</sup><http://www.sis4.com/brModelo/>

<sup>2</sup><https://fmr-web.appspot.com/>



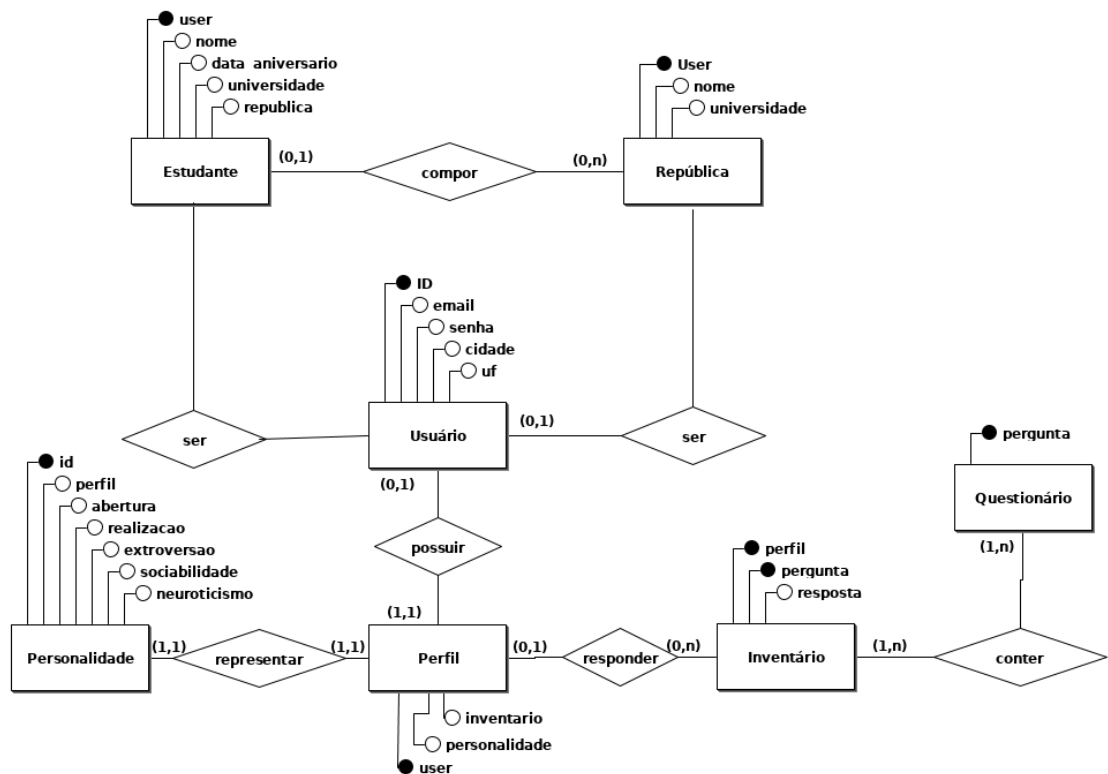


Figura 3.10: Diagrama Entidade Relacionamento.

## Estudante

Tipo de entidade que representa um estudante, calouro ou veterano, representado por seus atributos. Se relaciona com Usuário e República.

- **user (primary key):** atributo simples, monovalorado e armazenado que representa o usuário da pessoa. Faz referencia ao atributo *id* da entidade Usuário.
- **nome:** atributo textual, monovalorado e armazenado que nomeia um usuário.
- **data\_aniversario:** atributo de data, monovalorado e armazenado.
- **universidade:** atributo simples, monovalorado e armazenado.
- **republica:** atributo simples, monovalorado e armazenado, que representa a República que a pessoa reside.

## Republica

Tipo de entidade que representa uma república, representada por seus atributos. Se relaciona com Usuário e Pessoa

- **user (primary key)**: atributo simples, monovalorado e armazenado que representa o usuário da república. Faz referencia ao atributo *id* da entidade Usuário.
- **nome**: atributo textual, monovalorado e armazenado que nomeia a república.
- **universidade**: atributo simples, monovalorado e armazenado.
- **morador**: atributo simples, multivalorado e armazenado, que representa os moradores da república.

## Perfil

Tipo de entidade que representa uma personalidade, representada por seus atributos. Se relaciona com Usuário, Personalidade e Inventário.

- **user (primary key)**: atributo simples, monovalorado e armazenado que representa o usuário do perfil. Faz referencia ao atributo *id* da entidade Usuário.
- **personalidade**: atributo simples, monovalorado e armazenado, que representa a personalidade do perfil.
- **inventario**: atributo simples, monovalorado e armazenado, que representa as resposta do usuário ao inventário.

## Personalidade

Tipo de entidade que representa a pontuação de um usuário de acordo com o BFI. Se relaciona com o Perfil.

- **id (primary key)**: atributo simples, monovalorado e armazenado que identifica a personalidade.
- **perfil**: atributo simples, monovalorado e armazenado que representa o perfil da personalidade. Faz referencia ao atributo *user* da entidade Perfil.
- **abertura**: atributo simples, monovalorado e armazenado que representa a pontuação obtida através do BFI para o fator "Abertura à experiência".
- **realização**: atributo simples, monovalorado e armazenado que representa a pontuação obtida através do BFI para o fator "Realização".
- **extroversão**: atributo simples, monovalorado e armazenado que representa a pontuação obtida através do BFI para o fator "Extroversão".

- **sociabilidade**: atributo simples, monovalorado e armazenado que representa a pontuação obtida através do BFI para o fator "Sociabilidade".
- **neuroticismo**: atributo simples, monovalorado e armazenado que representa a pontuação obtida através do BFI para o fator "Neuroticismo".

## Questionário

Tipo de entidade que representa o questionário de perguntas do BFI. Se relaciona com Inventário

- **pergunta (primary key)**: atributo simples, monovalorado e armazenado, que identifica a pergunta.

## Inventário

Tipo de entidade que representa as respostas de um Perfil às perguntas do Questionário. Tem relação com a entidade Perfil e a entidade Questionário.

- **user (primary key)**: atributo simples, monovalorado e armazenado que diz respeito ao usuário que respondeu ao questionário. Faz referência ao atributo *user* da entidade Perfil.
- **pergunta (primary key)**: atributo simples, monovalorado e armazenado que representa a pergunta respondida. Faz referência ao atributo *pergunta* da entidade Questionário.
- **resposta**: atributo numérico, monovalorado e armazenado que representa a resposta do usuário à pergunta. (1, 2, 3, 4 ou 5)

## 3.3 Sistema

Nesta seção, são descritas as funcionalidades e o comportamento detalhado do sistema sobre cada etapa necessária para a realização da recomendação. Inicialmente, são apontadas as tecnologias utilizadas para o desenvolvimento do sistema, contido na Subseção [3.3.1](#). Na Subseção [3.3.2](#) é apresentada a ferramenta e suas funcionalidades, seguido dos detalhes do Perfil do Usuário abordado na Subseção [3.3.4](#). Então, é discutido o Perfil da República na Subseção [3.3.5](#) e, por fim, é debatido a Recomendação na Subseção [3.3.6](#).

### 3.3.1 Tecnologias

Diversas são as tecnologia que envolvem a completude e complexidade de um sistema. Segmentando o sistema como três principais módulos: (a) front-end+ (b) back-end e; (c) servidor, as principais tecnologias usadas são:

- (a) Angular 8.1.3v; Node 11.6v; Bootstrap 4.0v<sup>3</sup>;
- (b) Java 8; SpringBoot 2.1.7v; MySQL 5.7v<sup>4</sup>;
- (c) Google Cloud App Engine; Google Cloud Storage; Google Cloud SQL<sup>5</sup>;

### 3.3.2 Ferramenta

A ferramenta desenvolvida é um sistema de recomendação web que fornece, principalmente, o serviço de recomendar repúblicas a estudantes de acordo com a personalidade. Os dois principais componentes do FMR são Estudantes (cliente) e Repúblicas (produto), e são responsáveis por controlar as funcionalidades de cada tipo de usuário. A identificação do usuário é feita pelo lado do servidor.

As funcionalidades podem variar de acordo com tipo de usuário, porém algumas dessas são compartilhadas. Ambos os usuários podem fornecer espontaneamente algumas informações para o sistema e cadastrá-las, assim como podem realizar a busca por Repúblicas de acordo com a personalidade. A definição da Personalidade, detalhada na Subseção 3.3.3, também é aplicada para ambos os usuários. Por outro lado, somente o estudante pode definir um Perfil do Usuário, descrito na Subseção 3.3.4, enquanto somente a república pode definir o Perfil da República, descrito na Subseção 3.3.5. Ambos os usuários podem aproveitar o sistema de recomendação, entretanto, o produto sempre é uma república.

### 3.3.3 Personalidade

A personalidade é calculada de acordo com as métricas e facetas do modelo Big Five, conforme relatado na Subseção 2.1.3.4. O resultado da personalidade é composto pelos cinco traços levantadas pelo modelo, representados por valores numéricos. Considerando as regras de pontuação do modelo, cada característica da personalidade pode atingir um valor mínimo e máximo, sendo os intervalos de cada faceta definido como segue: abertura para o novo: [10,50]; concordância: [9,45]; consciência: [9,45]; extroversão: [8,40] e; neuroticismo: [8,40].

A normalização proposta pelo sistema de recomendação considera cada faceta da personalidade. A personalidade normalizada do usuário  $u$  sobre a faceta  $i$  é resultado da divisão

---

<sup>3</sup>Disponível em [https://github.com/marcosrodriigues/ufop/tree/master/conteudo/Monografia I/Sistema/FMR](https://github.com/marcosrodriigues/ufop/tree/master/conteudo/Monografia%20I/Sistema/FMR)

<sup>4</sup>Disponível em: [https://github.com/marcosrodriigues/fmr\\_api](https://github.com/marcosrodriigues/fmr_api)

<sup>5</sup>Aplicação disponível em: <https://fmr-web.appspot.com>

do valor obtido da pelo usuário  $u$  sobre a faceta  $i$  pelo valor máximo que é possível de ser obtido para a faceta. Após realizar a divisão, é feita a multiplicação por 100 para se obter a personalidade em porcentagem. Em outras palavras:

$$PN_{u_i} = \frac{P_{u_i}}{\max(P_i)} * 100\% \quad (3.2)$$

Onde:

- $PN_{u_i}$ : personalidade normalizada do usuário  $u$  sobre a faceta  $i$
- $P_{u_i}$ : personalidade obtida no teste pelo usuário  $u$  sobre a faceta  $i$
- $\max(P_i)$ : máximo valor que pode ser obtido da faceta  $i$
- $i \in [1, 2, 3, 4, 5]$ : 1 = abertura; 2 = concordância; 3 = consciência; 4 = extroversão; 5 = neuroticismo.

### 3.3.4 Perfil do Usuário

A funcionalidade de Perfil do Usuário está disponível para os usuários identificados como Estudantes. O módulo contém dois quesitos importantes para a elaboração da personalidade: o inventário do teste de personalidade aplicado na ferramenta e a regra de pontuação do respectivo instrumento. O teste de personalidade aplicado é o Big Five Inventory, conforme descrito na subseção 2.1.3.4. As regras de pontuação está contida no Capítulo 5, na seção ?? desta monografia.

O teste de personalidade é aplicado internamente no sistema e pode ser realizado a qualquer momento. Ao completar o teste, a personalidade do estudante é calculada seguindo as regras de pontuação apresentadas anteriormente. A pontuação obtida é mantida e jamais alterada, com exceção do caso do estudante refazer o teste de personalidade. Entretanto, a pontuação é normalizada antes de ser exibida ao usuário.

### 3.3.5 Perfil da República

A funcionalidade de Perfil da República está disponível somente para os usuários identificados como República. Diferente do Perfil do Usuário, este tipo de personalidade é calculado de acordo com a personalidades dos estudantes vinculados à república.

A agregação proposta para a elaboração da personalidade da república pode ser descrita da seguinte maneira:

1. Obtêm-se as Personalidades  $ps$  dos estudantes vinculados à república
2. Para cada estudante  $u$ , categoriza cada Faceta  $i$  da Personalidade  $p$  seguindo a regra:

- a) Calcula a Personalidade Normalizada  $PN$  da faceta  $i$
- b) Obtém a Categoria  $C$  da norma personalidade do usuário  $u$  para a faceta  $i$ , sendo:

$$C_{u_i} = \left\lceil \frac{PN_{u_i}}{20} \right\rceil \quad (3.3)$$

3. Calcula as modas das Categorias  $c$  de todos os estudantes  $u$  com base em cada faceta  $i$ .
4. Obtém a maior moda para cada faceta  $i$ .
5. Calcula a média das personalidades dos estudantes para a faceta correspondente à categoria.
6. A Personalidade  $P$  para a faceta  $i$  é resultado da média da faceta  $i$  de todos os estudantes  $u$  que foram obtidos através da maior moda das categorias para a faceta  $i$ .

A Personalidade da República também pode ser descrita através da fórmula:

$$PR_i = \frac{\sum_{u=1}^N \alpha_u * P_{u_i}}{|max(modas(C_i))|} \quad (3.4)$$

com:

$$\alpha_k = \begin{cases} 1, & \text{se } P_{u_i} \in max(modas(C_i)) \\ 0, & \text{caso contrário} \end{cases} \quad (3.5)$$

Sendo:

- $i \in [1, 2, 3, 4, 5]$ : 1 = abertura; 2 = concordância; 3 = consciência; 4 = extroversão; 5 = neuroticismo;
- $PR_i$ : Personalidade da República sobre a Faceta  $i$
- $N$ : número de estudantes vinculados à República
- $P_{u_i}$ : Personalidade do estudante  $u$  sobre a faceta  $i$ .
- $C_i$ : Vetor de categoria com as personalidades dos estudantes categorizadas correspondente a faceta  $i$
- $modas(C_i)$ : lista de personalidades dos estudantes que compõem as modas da categoria  $i$ .
- $max(modas(C_i))$ : lista de estudantes que foram obtidos a partir da maior moda da categoria  $C$  para a faceta  $i$ .

O pseudo-código para calcular a personalidade da república está descrito abaixo:

---

**Algoritmo 1:** Personalidade da república
 

---

**Entrada:**  $M[N][5]$  // matriz de morador e as cinco facetas da personalidade

**Saída:**  $PR[5]$  // personalidade da república

```

1  $PR[5] = [0,0,0,0,0];$ 
2
3 for  $i \leftarrow 1$  to 5 do
4   |  $vetor\_moda = calculaModaDaFaceta(M, i);$ 
5   |  $estudantes\_maximo\_moda = \max(vetor\_moda);$ 
6   |  $PR[i] = mediaAritmetica(estudantes\_maximo\_moda);$ 
7 end
```

---



---

**Algoritmo 2:** Calcula Moda Da Faceta
 

---

**Entrada:**  $M[N][5]; i$  // faceta pra calcular a moda

**Saída:**  $V[N][5]$  // vetor da moda com a lista de personalidades

```

1  $V[5] = [0,0,0,0,0];$ 
2 for  $k \leftarrow 1$  to  $N$  do
3   |  $position = arredondarParaCima(normaliza(M[k][i]) / 20);$ 
4   |  $V[position].add(M[k][i]);$ 
5 end
```

---



---

**Algoritmo 3:** Função Max (vetor\_moda)
 

---

**Entrada:**  $V[N][5];$  vetor com a moda//

**Saída:**  $E[N]$  // vetor de estudantes que compõem a máxima moda do vetor  $V$

```

1  $E[N] = [];$ 
2  $j = 0;$  //j guarda a posição do vetor com a maior moda
3 for  $k \leftarrow 1$  to  $N$  do
4   | if  $V[j].length \leq V[k].length$  then
5   |   |  $j = k;$ 
6   | else
7   | end
8 for  $k \leftarrow 1$  to  $N$  do
9   |  $E[k] = V[j][k]$ 
10 end
```

---

### 3.3.6 Recomendação

A funcionalidade de Recomendação pode se aplicar para ambos os tipos de usuários. Um estudante pode pesquisar por repúblicas com o intuito de encontrar um ambiente com pessoas que compartilham da sua personalidade. Já uma república pode buscar por uma recomendação

de novas repúblicas para propor algum vínculo.

Assim, a recomendação é calculada de acordo com a distância entre duas personalidades:

$$R_{u,v} = |D_{u,v}| \quad (3.6)$$

sendo:

- $R_{u,v}$ : recomendação do usuário  $u$  para o usuário  $v$
- $D_{u,v}$ : distância definida pela fórmula:

$$D_{u,v} = 1 - |Eq_{u,v}| \quad (3.7)$$

sendo:

$$Eq_{u,v} = \begin{cases} \frac{\sum_{i=1}^5 P_{v,i}}{\sum_{i=1}^5 P_{u,i}}, & \sum_{i=1}^5 P_{u,i} \geq \sum_{i=1}^5 P_{v,i} \\ \frac{\sum_{i=1}^5 P_{u,i}}{\sum_{i=1}^5 P_{v,i}}, & \text{caso contrário} \end{cases} \quad (3.8)$$

Em outras palavras, a recomendação é a diferença da soma dos cinco fatores de personalidades apresentados pelo modelo Big Five comparados entre dois usuários.



## Capítulo 4

# Experimentos Computacionais

Este capítulo descreve e analisa os experimentos computacionais realizados nesta monografia. A Seção [4.1](#) contempla a descrição dos experimentos e as métricas de avaliação de resultados. A Seção [4.2](#) descreve e analisa o estudo de caso realizados sobre dados obtidos. A Seção [4.3](#) analisa os resultados obtidos por meio do estudo de caso.

### 4.1 Descrição

Esta seção apresenta a descrição dos experimentos, descritos na Subseção [4.1.1](#), bem como as métricas de avaliação dos resultados, na Subseção [4.1.2](#).

#### 4.1.1 Experimentos

Os experimentos foram realizados com dados reais, isto é, após a extração de dados realizada pelo sistema de recomendação proposto.

Os passos detalhados dos experimentos estão enumerados a seguir:

1. Obtenção das personalidades dos estudantes;
2. Atribuir os vínculos entre os estudantes e as repúblicas;
3. Calcular a personalidade das repúblicas;
4. Aplicar separação selecionada entre estudantes e repúblicas;
5. Aplicar recomendação;
6. Analisar a recomendação com base nos critérios especificados.

Os experimentos podem prever algumas situações, como:

- Proximidade dos estudantes às repúblicas;

- Importância relativa dos estudantes às repúblicas;
- Eficácia da recomendação nos critérios analisados.

#### 4.1.2 Critérios de avaliação

Os critérios utilizados nos experimentos dizem respeito aos dados coletados sobre as repúblicas e a seus moradores sobre um curto período de tempo, isto é, o cenário onde o estudante solicita a recomendação da república, toma a decisão de ir morar nesta república e após um período de tempo fornece um *feedback* à recomendação, não foi experimentado. Esta é uma forma interessante de medir a eficácia do sistema de recomendação, entretanto o período de experimentos seria de um a dois semestres, o que tornou inviável a esta análise.

Foi, então, utilizado nos experimentos, critérios que variam de acordo com o tipo de usuário que está solicitando a recomendação. Os critérios usados para a recomendação ao Estudante está contida na Subseção [4.1.2.1](#) e os critérios para a recomendação aplicada à República estão descritos na Subseção [4.1.2.2](#).

##### 4.1.2.1 Critério de avaliação do Estudante

Primeiramente, se obtém o perfil completo do Estudante, isto é, seu Teste de Personalidade preenchido, e após a Personalidade é atribuída e vinculada a uma república. O vínculo com a república é opcional, visto que o estudante está associado a uma república, ou seja, ser morador de alguma república, ou não estar vinculado a república alguma, isto é, está interessado em ser recomendado para uma república.

O experimento seguiu o seguinte procedimento:

1. Retira o vínculo do Estudante à República, caso ele esteja vinculado.
2. Aplica a recomendação ao Estudante.

A análise então é feita de acordo com a existência do vínculo do estudante com a república, sendo: (a) o estudante está vinculado à república. Neste caso, espera-se que a república do estudante pertença ao conjunto de repúblicas com compatibilidade de personalidade igual ou superior a 80% ou que a república do estudante pertença ao conjunto das 3 primeiras repúblicas recomendadas, ou; (b) o estudante não está vinculado a república. Neste caso, considerando que se conheça a existência de um vínculo de amizade entre o estudante e as repúblicas, espera-se que as repúblicas recomendadas possuam esse vínculo de amizade.

##### 4.1.2.2 Critério de avaliação da República

As análises utilizadas para a recomendação aplicada à República são similares às utilizadas para o caso do estudante não estar vinculado à República. Desta maneira, considerando que se

conheça a existência ou não de uma amizade entre duas repúblicas, espera-se que as repúblicas amigas estejam presentes no conjunto das primeiras repúblicas recomendadas.

## 4.2 Estudos de Caso

Esta seção apresenta o estudo de caso realizados sobre a recomendação proposta nesta monografia e os resultados obtidos aplicando a ferramenta em três cenários: (a) quando o estudante não tem vínculo com repúblicas; (b) quando o estudante é morador de uma república e; (c) quando a república solicita recomendação.

Foram coletados 120 usuários, sendo 23 repúblicas e 97 estudantes. Apesar do número razoável de usuários coletados, aproximadamente 50 usuários foram ignorados na análise dos dados, pois diz respeito à usuários que não concluíram a pesquisa. 97.96% dos usuários cadastrados no sistema são da cidade de Ouro Preto e Mariana, em Minas Gerais, entretanto não representam todas as repúblicas localizadas nas duas cidades.

A análise foi feita sobre 62 personalidades, sendo 11 repúblicas e 51 estudantes, onde 11 estudantes não estão vinculados à nenhuma república e 40 estudantes são moradores de alguma república.

A subseção [4.2.1](#) apresenta os resultados obtidos a partir dos estudantes que não possuem vínculo com República. A subseção [4.2.2](#) apresenta os resultados sobre os estudantes que são moradores de repúblicas. Por fim, a subseção [4.2.3](#) apresenta os resultados das recomendações aplicadas às repúblicas

### 4.2.1 Estudantes sem vínculo com República

A recomendação foi aplicada a 11 estudantes que não possuem vínculo com repúblicas. Para a análise desses dados, foi considerado a existência de algum grau de amizade entre o estudante que recebe a recomendação e a república que está sendo recomendada. A existência de amizade foi coletada de forma direta, isto é, através de *feedback* verbal. Assim, espera-se que as repúblicas consideradas amigas estejam contidas na recomendação. O resultado atingiu uma acurácia de 76.36% na sugestão de repúblicas.

A Tabela ?? relata o estudante aplicado à recomendação, o número de repúblicas recomendadas que são consideradas amigas, o número de repúblicas recomendadas que não são consideradas amigas e a precisão nas 5 primeiras recomendações, ou seja, a porcentagem de acerto para cada caso nas 5 primeiras recomendações. Neste cenário, foi considerado as cinco primeiras repúblicas recomendadas.

É importante relatar que a precisão de 76.36% sobre as recomendações significa a porcentagem geral de repúblicas amigas recomendada, ou seja, não quer dizer que 23.64% das recomendações foram ruins, apenas que as repúblicas recomendadas não são amigas ou não são conhecidas pelos estudantes que foram recomendados.

### 4.2.2 Estudantes moradores de República

A recomendação foi aplicada a 51 estudantes que possuem vínculo com repúblicas. Para a análise dos dados, foi considerado que a república na qual o estudante está vinculado deve estar entre as 3 primeiras repúblicas recomendadas por meio do sistema ou possuir uma compatibilidade igual ou superior a 80%.

A Tabela 4.1 mostra a relação de repúblicas e a quantidade de moradores vinculados à repúblicas. A Figura 4.1 ilustra a porcentagem de acerto e de erro da recomendação quando aplicada aos moradores das repúblicas. O resultado de cada república é a média do resultado dos seus moradores.

Tabela 4.1: Relação de repúblicas e quantidade de moradores

República	Nº de moradores
República 1	9
República 2	2
República 3	3
República 4	2
República 5	2
República 6	4
República 7	7
República 8	2
República 9	3
República 10	4
República 11	3
<b>Total</b>	<b>41</b>

### 4.2.3 Recomendação aplicada a Repúblicas

A recomendação foi aplicada às 11 repúblicas que completaram seu cadastro, isto é, concluíram o cadastro dos moradores e sua personalidade, bem como a personalidade da República. Para a análise dos dados, foi considerado os mesmos critérios dos estudantes que não possuem vínculo com repúblicas, descritos na Seção 4.2.1.

Ao se recomendar uma república a outra, é considerado o vínculo de amizade entre essas duas repúblicas e espera-se que repúblicas consideradas amigas pertençam ao conjunto das cinco primeiras repúblicas recomendadas.

A Figura 4.2 mostra a porcentagem de acerto e erro da recomendação quando aplicada às repúblicas.

Assim como nas recomendações aplicadas a moradores sem república, a eficácia da recomendação diz respeito à quantidade de repúblicas que são consideradas amigas e que foram recomendadas pela ferramenta. Isso quer dizer que, para cada república que foi aplicada a re-

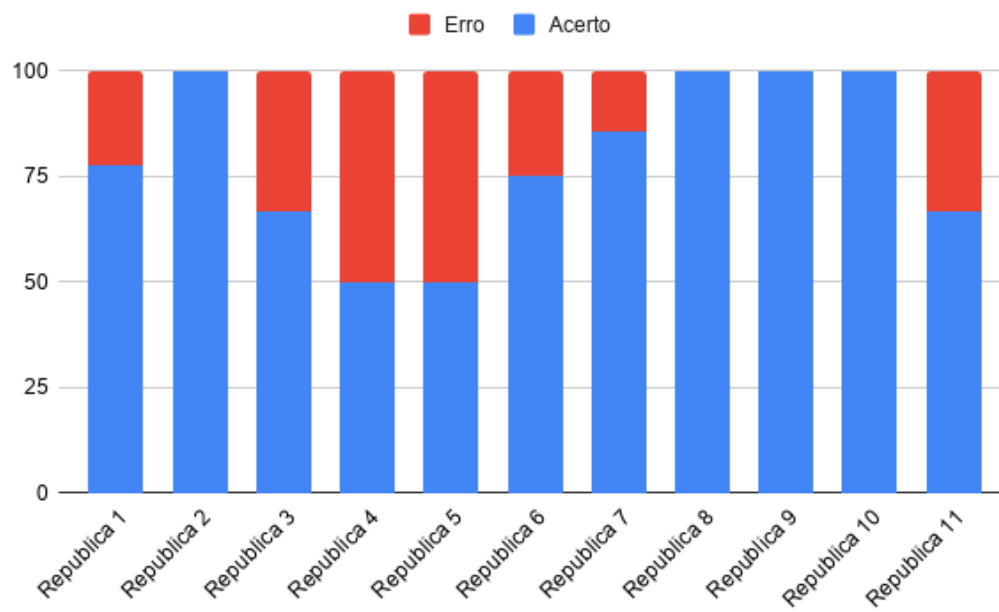


Figura 4.1: Relação de Acerto e Erro da Recomendação aplicada aos Moradores

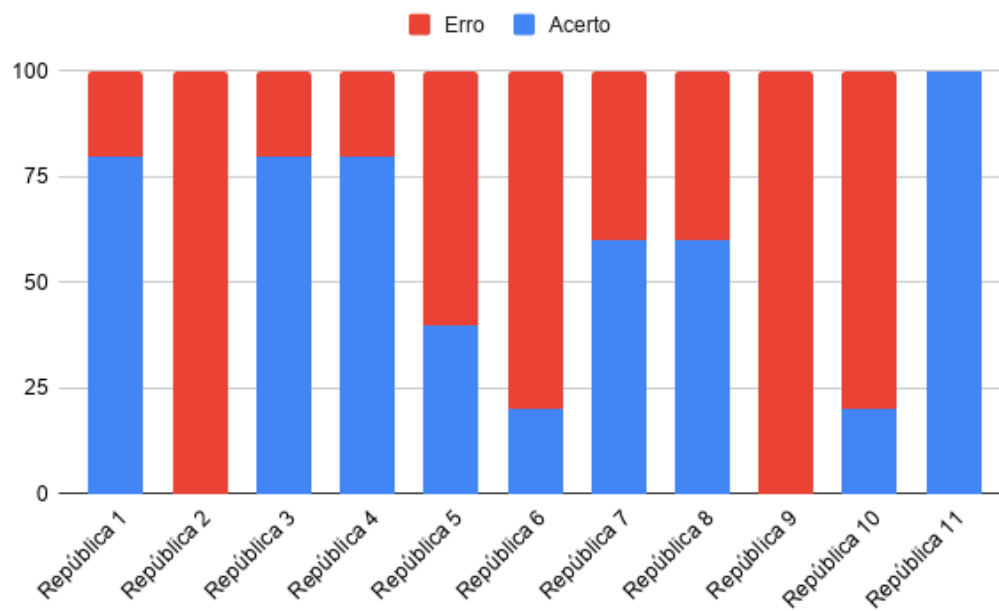


Figura 4.2: Relação de Acerto e Erro da Recomendação aplicada às Repúblicas

comendação, o acerto diz respeito às repúblicas consideradas amigas desta e o erro diz respeito às repúblicas que não são consideradas amigas.

### 4.3 Análise dos Resultados

Nesta subseção, são abordados os resultados obtidos a partir da análise dos dados reais aplicada aos três cenários apresentados. É importante separar o cenário de cada experimento realizados, de acordo com os critérios definidos para cada um.

Quando aplicada a recomendação aos estudantes que não moram em república, o objetivo é obter as repúblicas que são amigas como primeiras repúblicas sugeridas, analisando a eficácia do método por meio deste vínculo. Neste contexto, a ferramenta obteve um acerto de 76.36%.

Ao aplicar a recomendação aos estudantes que já estão vinculados a uma república, o objetivo é obter como sugestão a república onde o estudante está vinculado, na intenção de validar a recomendação caso o estudante estivesse procurando uma república. É possível perceber que o número de moradores influencia significativamente no valor obtido de cada república. Analisando os dados da Tabela 4.1 e mostrados na Figura 4.1, as repúblicas 3, 4, 5, 6 e 11, que possuem respectivamente 3, 2, 2 e 4 moradores, ao desvincular um morador dessas repúblicas, a personalidade da república pode sofrer grandes alterações, influenciando diretamente na recomendação. Os casos das repúblicas 2, 8, 9 e 10, que possuem respectivamente 2, 2, 3 e 4 moradores, o resultado mostra uma compatibilidade de personalidade entre os moradores dessas repúblicas, concluindo que este é o cenário ideal de uma república, onde todos os moradores tem compatibilidade. Nos cenários onde a república possui 7 ou mais moradores, como o caso das repúblicas 1 e 7, o método obteve uma acurácia igual ou superior a 80%, concluindo ser uma recomendação eficiente, porém com possibilidade de melhoria.

Aplicando a recomendação às repúblicas, espera-se obter como sugestão repúblicas que possuam um vínculo de amizade com a república que está utilizando o sistema. Neste contexto, os dados apresentados na Figura 4.2 mostram a acurácia do método no que diz respeito aos critérios explicados na Seção 4.1.2.2. Dentre os resultados obtidos pela recomendação de repúblicas a repúblicas, observa-se que as Republicas 2 e 9 obtiveram 0% de acerto. O *feedback* relatado por estas repúblicas diz que a ineficiência do método se deu pela localidade das repúblicas. Enquanto uma das repúblicas se encontra em um bairro afastado das principais repúblicas que houve coleta de dados, a outra república se localiza em outra cidade. Assim, é natural que não haja vínculos de amizade, conduzindo o sistema a aplicar a recomendação sem possibilidade de validar a eficácia para as duas repúblicas. As repúblicas 4 e 5 não relataram *feedback* sobre a recomendação.

## Capítulo 5

# Conclusão

Após o início das pesquisas para o desenvolvimento deste projeto, por meio de estudos sobre a personalidade do indivíduo e a identificação de compatibilidade entre estudantes e repúblicas estudantis, foi possível compreender alguns fatores interessantes a serem abordados. Tal identificação de compatibilidade pode contribuir para o problema de evasão de moradores nas repúblicas estudantis, bastando relacionar o estudante certo para a república certa. Para obter um relacionamento com mais acurácia, pode-se aplicar Testes de Personalidade para obter traços de personalidade dos estudantes, a fim de levantar um Perfil de Usuário e conectá-lo aos grupos de usuários com traços de personalidade semelhantes. Testes de Personalidade podem ser complexos de serem implementados computacionalmente, entretanto, possuem uma exatidão significativa no que diz respeito aos resultados calculados, sendo bastante utilizados para comprovação de resultados computacionalmente obtidos.

Neste trabalho, foram apresentadas as etapas que compõem o desenvolvimento de um sistema baseado em traços de personalidade cujo foco é a recomendação/relacionamento de compatibilidade de uma república a um estudante, de acordo com a similaridade entre seus perfis. Para a execução deste trabalho, foi realizada uma pesquisa na literatura a fim de obter a fundamentação teórica e a revisão de trabalhos relacionados ao mesmo tema. Em seguida, elaborada a documentação do sistema, definido o esquema do banco de dados relacional, bem como seus relacionamentos e atributos, o escopo do sistema e o detalhamento do software, com a proposta da recomendação e identificação de compatibilidade implementada. Finalmente, foram realizados e analisados experimentos para a conclusão desta monografia.

Como principais resultados da avaliação experimental, pode-se concluir que a ferramenta de identificação de compatibilidade e recomendação proposta é aplicável e pode auxiliar na solução do problema de evasão de repúblicas. Entretanto, para recomendações mais eficientes, é recomendado que as repúblicas estudantis possuam pelo menos 7 moradores, uma vez que a personalidade da república pode ficar muito versátil.

Como trabalhos futuros, pretende-se aplicar o sistema de recomendação a calouros que estejam procurando por uma república estudantil e possibilitar às repúblicas a procura por

---

estudantes, além de disponibilizar aos estudantes meios para avaliar explicitamente as indicações recebidas. Também espera-se validar a personalidade calculada para as repúblicas, através de coleta direta, na intenção de verificar a eficiência da fórmula aplicada às repúblicas. As principais atividades a serem desenvolvidas estão relacionadas com melhorias do software em si. As tarefas dizem respeito a implementações no sistema sugeridas por usuários.



# Anexo I - Formulário Big Five Inventory

A seguir é mostrado um questionário do Big Five Inventory (Oliver E John, 1998; John, 1991; Oliver E John, 2008).

Abaixo encontram-se um número de características que podem ou não descrever-te. Por exemplo, achas que és uma pessoa que gosta de estar com os outros? Assinala com uma X o número que corresponde ao grau em que acreditas que a frase te caracteriza. Não existem respostas certas nem erradas, responde com sinceridade, de acordo com o teu modo de ser habitual.

Discordo totalmente	Discordo moderadamente	Nem concordo nem discordo	Concordo moderadamente	Concordo totalmente
1	2	3	4	5

**Vejo-me como alguém que...**

1. É falador(a).	1	2	3	4	5
2. Tende a encontrar os defeitos dos outros.	1	2	3	4	5
3. Faz um trabalho exaustivo.	1	2	3	4	5
4. É deprimido(a), triste.	1	2	3	4	5
5. É original, tem sempre novas ideias.	1	2	3	4	5
6. É reservado(a).	1	2	3	4	5
7. É prestável e não inveja os outros.	1	2	3	4	5
8. Por vezes pode ser um pouco descuidado(a).	1	2	3	4	5
9. É relaxado(a), lida bem com o stress.	1	2	3	4	5
10. Tem curiosidade em relação a várias coisas.	1	2	3	4	5
11. Tem muita energia.	1	2	3	4	5
12. Inicia muitas disputas com os outros	1	2	3	4	5
13. É um trabalhador(a) de confiança.	1	2	3	4	5
14. Pode ficar tenso.	1	2	3	4	5
15. É engenhoso(a), um(a) pensador(a) profundo.	1	2	3	4	5
16. Gera muito entusiasmo.	1	2	3	4	5
17. Perdoa com facilidade.	1	2	3	4	5
18. Tende a ser desorganizado(a).	1	2	3	4	5
19. Se preocupa muito.	1	2	3	4	5
20. Tem uma imaginação activa.	1	2	3	4	5
21. Tende a ser sossegado(a).	1	2	3	4	5
22. Geralmente é de confiança.	1	2	3	4	5
23. Tende a ser preguiçoso(a).	1	2	3	4	5

Discordo totalmente	Discordo moderadamente	Nem concordo nem discordo	Concordo moderadamente	Concordo totalmente
1	2	3	4	5

24. É emocionalmente estável, não se aborrece com facilidade.	1	2	3	4	5
25. É inventivo(a).	1	2	3	4	5
26. Tem uma personalidade assertiva.	1	2	3	4	5
27. Pode ser frio(a) e indiferente.	1	2	3	4	5
28. É perseverante até a tarefa estar concluída.	1	2	3	4	5
29. Pode ter um humor instável.	1	2	3	4	5
30. Valoriza experiências artísticas, estéticas.	1	2	3	4	5
31. Por vezes é tímido(a), inibido(a).	1	2	3	4	5
32. É atencioso(a) e bondoso(a) com quase toda a gente.	1	2	3	4	5
33. Faz as coisas de modo eficaz.	1	2	3	4	5
34. Permanece calmo(a) em situações tensas.	1	2	3	4	5
35. Prefere o trabalho rotineiro.	1	2	3	4	5
36. É sociável, amigável.	1	2	3	4	5
37. Por vezes é rude para com os outros.	1	2	3	4	5
38. Faz planos e cumpre-os.	1	2	3	4	5
39. Fica nervoso(a) facilmente.	1	2	3	4	5
40. Gosta de reflectir, brincar com as ideias.	1	2	3	4	5
41. Tem poucos interesses artísticos.	1	2	3	4	5
42. Gosta de cooperar com os outros.	1	2	3	4	5
43. Se distrai com facilidade.	1	2	3	4	5
44. É sofisticado(a) na arte, música ou literatura.	1	2	3	4	5

**OBRIGADO PELA TUA COLABORAÇÃO**

## Big 5 Inventory – Scoring Key

**Instructions for scoring:** Key items that are bold, italicized, and underlined (e.g., **E**) should be reverse scored. To reverse score an item, change 1 to 5, 2 to 4, 4 to 2, and 5 to 1. After reverse scoring, find the total score for E (Extroversion), A (Agreeableness), C (Conscientiousness), N (Neuroticism), and O (Openness). Important note: Please be sure to submit original responses on each item to the Psi Beta spreadsheet. So please do NOT report reverse-scored data, just original responses made by your participants.

__1	Is talkative	E	__23	Tends to be lazy	<u><b>C</b></u>
__2	Tends to find fault with others	<u><b>A</b></u>	__24	Is emotionally stable, not easily upset	<u><b>N</b></u>
__3	Does a thorough job	C	__25	Is inventive	O
__4	Is depressed, blue	N	__26	Has an assertive personality	E
__5	Is original, comes up with new ideas	O	__27	Can be cold and aloof	<u><b>A</b></u>
__6	Is reserved	<u><b>E</b></u>	__28	Perseveres until the task is finished	C
__7	Is helpful and unselfish with others	A	__29	Can be moody	N
__8	Can be somewhat careless	<u><b>C</b></u>	__30	Values artistic, aesthetic experiences	O
__9	Is relaxed, handles stress well	<u><b>N</b></u>	__31	Is sometimes shy, inhibited	<u><b>E</b></u>
__10	Is curious about many different things	O	__32	Is considerate and kind to almost everyone	A
__11	Is full of energy	E	__33	Does things efficiently	C
__12	Starts quarrels with others	<u><b>A</b></u>	__34	Remains calm in tense situations	<u><b>N</b></u>
__13	Is a reliable worker	C	__35	Prefers work that is routine	<u><b>O</b></u>
__14	Can be tense	N	__36	Is outgoing, sociable	E
__15	Is ingenious, a deep thinker	O	__37	Is sometimes rude to others	<u><b>A</b></u>
__16	Generates a lot of enthusiasm	E	__38	Makes plans and follows through with them	C
__17	Has a forgiving nature	A	__39	Gets nervous easily	N
__18	Tends to be disorganized	<u><b>C</b></u>	__40	Likes to reflect, play with ideas	O
__19	Worries a lot	N	__41	Has few artistic interests	<u><b>O</b></u>
__20	Has an active imagination	O	__42	Likes to cooperate with others	A
__21	Tends to be quiet	<u><b>E</b></u>	__43	Is easily distracted	<u><b>C</b></u>
__22	Is generally trusting	A	__44	Is sophisticated in art, music, or literature	O

## Big 5 Inventory – Scoring Key

---

### References

- McCrae, R. R. (2004). Human nature and culture: A trait perspective. *Journal of Research in Personality, 38*, 3-14.
- McCrae, R. R., & Costa, P. T. Jr. (1999). A five-factor theory of personality. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research*. New York: Guilford Press.

## Anexo II - Parecer comitê de ética

**PARECER CONSUBSTANCIADO DO CEP**

**DADOS DO PROJETO DE PESQUISA**

**Título da Pesquisa:** Desenvolvimento de um sistema de recomendação entre pessoas e grupos baseado na personalidade dos usuários

**Pesquisador:** Anderson Almeida Ferreira

**Área Temática:**

**Versão:** 2

**CAAE:** 20556019.5.0000.5150

**Instituição Proponente:** Universidade Federal de Ouro Preto

**Patrocinador Principal:** Financiamento Próprio

**DADOS DO PARECER**

**Número do Parecer:** 3.657.800

**Apresentação do Projeto:**

A vida universitária pode oferecer diversos obstáculos durante sua trajetória. Para os estudantes que escolhem cursar uma universidade em uma cidade distante, um dos primeiros desafios é escolher a república para morar. Porém, é comum o caso de evasão de estudantes em repúblicas, muitas das vezes ocorrendo por incompatibilidade de personalidades. Assim, é interessante para o estudante conhecer a personalidade das pessoas que residem na república. Da mesma maneira, é interessante aos membros da fraternidade conhecerem o perfil do estudante que está interessado em conviver com eles. Assim, este trabalho propõe um método de recomendação que trate o problema da incompatibilidade de perfis, no qual será desenvolvido um sistema de recomendação que trabalhe com o perfil do estudante e da república - representado por traços de personalidade - e realize a recomendação adequada.

**Objetivo da Pesquisa:**

- Concepção de um sistema de recomendação de pessoas a grupos e vice-versa. A recomendação será baseada na personalidade dos usuários e dos grupos. Desta forma, se espera recomendar usuários à grupos de acordo com a similaridade entre suas personalidades.
- Aplicação de questionário Big Five Inventory, objetivando recomendação entre pessoas e grupos.

**Endereço:** Morro do Cruzeiro-Centro de Convergência

**Bairro:** Campus Universitário

**CEP:** 35.400-000

**UF:** MG

**Município:** OURO PRETO

**Telefone:** (31)3559-1368

**Fax:** (31)3559-1370

**E-mail:** cep.propp@ufop.edu.br

Continuação do Parecer: 3.657.800

- Avaliação de métricas para realização das recomendações

**Avaliação dos Riscos e Benefícios:**

Relação riscos-benefícios adequada.

**Comentários e Considerações sobre a Pesquisa:**

A pesquisa se mostra relevante e justificada teórica e metodologicamente, como Monografia do Bacharelado em Ciência da Computação.

**Considerações sobre os Termos de apresentação obrigatória:**

Pendências sanadas.

**Conclusões ou Pendências e Lista de Inadequações:**

As pendências apontadas foram sanadas.

**Considerações Finais a critério do CEP:**

O Comitê de Ética em Pesquisa CEP/UFOP, de acordo com as atribuições definidas na Res. CNS 466/12 e/ou Res. CNS 510/16, manifesta-se pela APROVAÇÃO deste protocolo de pesquisa. Ressalta-se ao pesquisador responsável pelo projeto o compromisso de envio ao CEP/UFOP, um ano após o início do projeto, do relatório final ou parcial de sua pesquisa, encaminhado por meio da Plataforma Brasil, informando, em qualquer tempo, o andamento da mesma, comunicando também eventos adversos e eventuais modificações no protocolo.

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1431726.pdf	22/10/2019 13:34:17		Aceito
Outros	CronogramaAtualizado.pdf	10/10/2019 12:27:48	MARCOS PAULO FERREIRA RODRIGUES	Aceito
Folha de Rosto	FolhaDeRosto.pdf	10/09/2019 17:01:09	MARCOS PAULO FERREIRA RODRIGUES	Aceito
Projeto Detalhado / Brochura Investigador	Detalhamento.pdf	09/09/2019 19:46:01	MARCOS PAULO FERREIRA RODRIGUES	Aceito
Outros	Submissao.pdf	09/09/2019 19:45:39	MARCOS PAULO FERREIRA RODRIGUES	Aceito

**Endereço:** Morro do Cruzeiro-Centro de Convergência

**Bairro:** Campus Universitário **CEP:** 35.400-000

**UF:** MG **Município:** OURO PRETO

**Telefone:** (31)3559-1368 **Fax:** (31)3559-1370 **E-mail:** cep.propp@ufop.edu.br



UNIVERSIDADE FEDERAL DE  
OURO PRETO



Continuação do Parecer: 3.657.800

TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.pdf	09/09/2019 19:42:10	MARCOS PAULO FERREIRA RODRIGUES	Aceito
---	----------	------------------------	---------------------------------	--------

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

OURO PRETO, 23 de Outubro de 2019

---

**Assinado por:**  
**EVANDRO MARQUES DE MENEZES MACHADO**  
**(Coordenador(a))**

**Endereço:** Morro do Cruzeiro-Centro de Convergência

**Bairro:** Campus Universitário

**CEP:** 35.400-000

**UF:** MG

**Município:** OURO PRETO

**Telefone:** (31)3559-1368

**Fax:** (31)3559-1370

**E-mail:** cep.propp@ufop.edu.br

# Referências Bibliográficas

- Burke, R. (2007). Hybrid web recommender systems. In *The adaptive web*, pp. 377–408. Springer.
- Cazella, S. C.; Nunes, M. e Reategui, E. (2010). A ciência da opinião: Estado da arte em sistemas de recomendação. *André Ponce de Leon F. de Carvalho; Tomasz Kowaltowski..(Org.). Jornada de Atualização de Informática-JAI*, pp. 161–216.
- de Souza, A. E. R. (2013). Um modelo para recomendação de cursos de especialização baseado no perfil profissional do candidato.
- Ferwerda, B.; Schedl, M. e Tkalcic, M. (2015). Predicting personality traits with instagram pictures. *Proceedings of the 3rd Workshop on Emotions and Personality in Personalized Systems 2015*, pp. 7–10.
- Hector Garcia-Molina, Jeffrey D Ullman, J. W. (2001). *Implementação de sistemas de bancos de dados*. Campus.
- John, O. E. (1991). The big five inventory—versions 4a and 54. *Handbook of personality: Theory and research*.
- Motta, C. L. d.; Garcia, A.; Vivacqua, A.; Santoro, F. e Sampaio, J. (2011). Sistemas de recomendação. *Pimentel, M.; Fuks, H. “Sistemas colaborativos”. Rio de Janeiro: Elsevier*.
- Nunes, M. A. S. N.; Cerri, S. A. e Blanc, N. (2008). Improving recommendations by using personality traits in user profiles. *International Conferences on Knowledge Management and New Media Technology*, pp. 92–100.
- Nuno Miranda, T. G. (2015). Sistemas de recomendação para grupos. *Jiue 2015 5th Workshop in informatics of the university of Évora*.
- Oliver E John, Laura P. Naumann, C. J. S. (2008). *Handbook of personality: Theory and research*. The Guilford Press.

- Oliver E John, V. B.-M. (1998). Los cinco grandes across cultures and ethnic groups: Multitrait multimethod analyses of the big five in spanish and english. *Journal of Personality and Social Psychology*, pp. 729–750.
- Quijano-Sánchez, L.; Recio-García, J. A. e Díaz-Agudo, B. (2010). Personality and social trust in groups recommendations. *International Conference on Tools with Artificial Intelligence*.
- Reategui, E. B. e Cazella, S. C. (2005). Sistemas de recomendação. *XXV Congresso da Sociedade Brasileira de Computação*, pp. 306–348.
- Sandy M. Porto, Wanderson S. Costa, M. A. S. N. N. L. N. (2011). Como a extração de personalidade através do teclado pode beneficiar a personalização na educação. *XXIII Sociedade Brasileira de Inteligência Emocional*.
- Sommerville, I. (2011). *Engenharia de Software*. Pearson.