

**UNIVERSIDADE DO ESTADO DE SANTA CATARINA – UDESC
CENTRO DE CIÊNCIAS TECNOLÓGICAS – CCT
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA – PPGCA**

NATHÁLIA LOCATELLI CEZAR

**ESTRATÉGIA PARA REDUZIR OS PROBLEMAS DE COLD START EM SISTEMAS
DE RECOMENDAÇÃO BASEADOS EM CONTEÚDO**

JOINVILLE

2024

NATHÁLIA LOCATELLI CEZAR

**ESTRATÉGIA PARA REDUZIR OS PROBLEMAS DE COLD START EM SISTEMAS
DE RECOMENDAÇÃO BASEADOS EM CONTEÚDO**

Dissertação submetida ao Programa de Pós-Graduação em Computação Aplicada do Centro de Ciências Tecnológicas da Universidade do Estado de Santa Catarina, para a obtenção do grau de Mestre em Computação Aplicada.

Orientador (a): Isabela Gasparini

Coorientador (a): Daniel Lichnow

JOINVILLE

2024

**Ficha catalográfica elaborada pelo programa de geração automática da
Biblioteca Universitária Udesc,
com os dados fornecidos pelo(a) autor(a)**

Cezar, Nathália Locatelli

Estratégia para reduzir os problemas de Cold Start em
Sistemas de Recomendação Baseados em Conteúdo /
Nathália Locatelli Cezar. -- 2024.

117 p.

Orientador: Isabela Gasparini

Coorientador: Daniel Lichtnow

Dissertação (mestrado) -- Universidade do Estado de
Santa Catarina, Centro de Ciências Tecnológicas, Programa
de Pós-Graduação em Computação Aplicada, Joinville, 2024.

1. Sistemas de Recomendação. 2. Cold Start User. 3.
Perfil inicial. 4. Abordagem Baseado em Conteúdo. I.
Gasparini, Isabela. II. Lichtnow, Daniel. III. Universidade do
Estado de Santa Catarina, Centro de Ciências Tecnológicas,
Programa de Pós-Graduação em Computação Aplicada. IV.
Título.

NATHÁLIA LOCATELLI CEZAR

**ESTRATÉGIA PARA REDUZIR OS PROBLEMAS DE COLD START EM SISTEMAS
DE RECOMENDAÇÃO BASEADOS EM CONTEÚDO**

Dissertação submetida ao Programa de Pós-Graduação em Computação Aplicada do Centro de Ciências Tecnológicas da Universidade do Estado de Santa Catarina, para a obtenção do grau de Mestre em Computação Aplicada.

Orientador (a): Isabela Gasparini
Coorientador (a): Daniel Lichnow

BANCA EXAMINADORA:

Prof. Dr. Daniel Lichnow
UFSM

Membros:

Prof. Dr. Daniel Lichnow
UFSM

Prof. Dr. José Palazzo Moreira de Oliveira
UFRGS

Prof. Dr. Gabriel Machado Lunardi
UFSM

Joinville, 28 de Junho, 2024

AGRADECIMENTOS

Durante alguns anos sonhei com a finalização deste ciclo. Sinto que superei diversos obstáculos ao longo desses últimos anos no mestrado, tanto pessoais quanto profissionais. Gostaria de agradecer aos meus orientadores, que além de tornarem meu trabalho possível, se tornaram muito mais que apenas orientadores, excelentes ao me orientarem e indicar sempre o melhor caminho. Terei sempre um carinho especial por vocês, professora Isabela e professor Daniel. São quase 8 anos trabalhando junto de vocês, que honra a minha. Obrigada pelas oportunidades que tive ao longo dessa jornada.

Dedico este trabalho aos meus pais. Acredito que sem o incentivo e apoio da minha mãe, eu nada seria. Obrigada por ser meu porto seguro sempre mãe. Ao meu pai, que já partiu, sinto que esse trabalho tem muito do seu incentivo que recebi ao longo dos anos que vivemos juntos.

Nos últimos anos do mestrado, eu também tive o privilégio de construir uma família ao lado do Vítor. Obrigada por ser a calmaria no caos, meu maior fã e meu também apoio.

Gostaria também de agradecer aos meus amigos, que acreditaram no meu potencial e continuam vibrando por minhas conquistas, assim como celebro as de vocês. Em especial à alguns amigos que fizeram esse projeto ser possível também, além de todos os participantes que testaram o sistema, obrigada Matheus, Karine e Cláudia, pelo apoio e ajuda de sempre.

Por fim, agradeço a todos que fizeram com que esse sonho pudesse ser realizado. "If you can dream it, you can do it."

RESUMO

Com o aumento de informações disponibilizadas usando Tecnologia da Informação (TI) nos últimos anos, os usuários passaram a encontrar na Internet assuntos que lhe interessam. Entretanto os usuários ainda tendem a ter dificuldades de encontrar trabalhos relevantes a sua pesquisa. Sistemas de Recomendação (SRs) são ferramentas de software e técnicas que auxiliam a lidar com a sobrecarga de informações encontradas atualmente. Todavia, mesmo com todo avanço na área de Sistemas de Recomendação, existe um problema bastante conhecido e ainda não solucionado, o *Cold Start User*. Este problema surge quando o sistema não consegue criar possíveis recomendações confiáveis devido à falta de dados suficientes sobre o perfil do usuário. Algumas técnicas são utilizadas para lidar com este problema no momento da criação do perfil inicial do usuário, porém grande parte utiliza histórico do usuário ou de outros usuários ou fontes externas. Nesse contexto, o objetivo deste trabalho é apresentar uma técnica para minimizar o *Cold Start User* a partir da criação de um perfil inicial do usuário em uma abordagem Baseada em Conteúdo. Para tal, foi desenvolvido um Sistema de Recomendação de artigos utilizando uma proposta de construção do perfil inicial sem nenhum tipo de informação prévia. Para a criação do perfil inicial, foi proposto uma estratégia que apresenta pares de itens para que o usuário selecione um, dois ou nenhum item do par, e a cada escolha, informações são adicionadas ao perfil do usuário sendo posteriormente geradas recomendações aos usuários. A avaliação da estratégia ocorreu em duas etapas, em que foram realizados, em sistema de recomendação que utiliza a estratégia, testes funcionais com estudantes da graduação em Computação e após, uma avaliação experimental com um grupo de estudantes de graduação e pós-graduação, além de especialistas. Com as avaliações foram obtidos feedbacks e sugestões de melhoria tanto nos aspectos de *design* quanto nas funcionalidades do sistema e nas recomendações geradas. Constatou-se que a estratégia proposta trouxe um resultado positivo em relação aos testes realizados e pode ser considerada para aplicação em trabalhos futuros que aprofundem a avaliação de sua contribuição para redução do *Cold Start User*.

Palavras-chave: Sistemas de Recomendação. *Cold Start User*. Perfil inicial. Abordagem Baseado em Conteúdo.

ABSTRACT

With the increase in information made available using Information Technology (IT) in recent years, users have begun to find subjects of interest to them on the Internet. However, users still tend to have difficulty finding relevant works for their research. Recommender Systems (RSs) are software tools and techniques that assist in dealing with the information overload currently encountered. Nevertheless, even with all the advancements in the field of Recommendation Systems, there is a well-known and as yet unsolved problem, the Cold Start User. This problem arises when the system cannot create reliable recommendations due to a lack of sufficient data about the user's profile. Some techniques are used to deal with this problem at the time of creating the user's initial profile, but a large part of them uses the user's or other users' history or external sources. In this context, the aim of this work is to present a technique to minimize the Cold Start User by creating an initial user profile in a Content-Based approach. To this end, a recommendation system for papers was developed using a proposal for constructing the initial profile without any prior information. For the creation of the initial profile, a strategy was proposed that presents pairs of items for the user to select one, two, or no item from the pair, and with each choice, information is added to the user's profile, which is then used to generate recommendations to the users. The evaluation of the strategy took place in two stages, in which functional tests were carried out with undergraduate Computing students using a recommendation system that employs the strategy, and after that, an experimental evaluation was conducted with a group of undergraduate and postgraduate students, as well as experts. The evaluations provided feedback and suggestions for improvement in both the design aspects and the functionalities of the system and the generated recommendations. It was found that the proposed strategy yielded positive results in relation to the tests conducted and could be considered for application in future work that further evaluates its contribution to reducing the Cold Start User.

Keywords: Recommender Systems. *User Cold Start*. Initial profile. Content-Based approach.

LISTA DE FIGURAS

Figura 1 – Trabalhos por abordagens de recomendação	32
Figura 2 – Arquitetura da proposta	45
Figura 3 – Exemplo de Matriz de Similaridade	46
Figura 4 – Criação do perfil inicial - pares	47
Figura 5 – Exemplo Interação 1 - pares	48
Figura 6 – Exemplo Interação 2 - pares	48
Figura 7 – Exemplo Interação 3 - pares	49
Figura 8 – Simulação de interação de pares	50
Figura 9 – Criação do perfil inicial - lista	52
Figura 10 – Simulação da lista de seleção de itens do perfil	54
Figura 11 – Experimentos de interface em lista	60
Figura 12 – Interface com 10 itens	61
Figura 13 – Interface com 2 itens	61
Figura 14 – Tela Inicial	62
Figura 15 – Fluxo de telas do sistema	63
Figura 16 – Tela de Cadastro	63
Figura 17 – Tela de Login	64
Figura 18 – Tela de Instruções	64
Figura 19 – Tela de seleção de itens em pares - item não selecionado	65
Figura 20 – Tela de seleção de itens em pares - item selecionado	65
Figura 21 – Tela de seleção de itens em pares - Botão Recomendar	66
Figura 22 – Adicionando um item ao aperfil do usuário	66
Figura 23 – Adicionando dois itens ao aperfil do usuário	67
Figura 24 – Tela de Recomendações	67
Figura 25 – Tela Finalizar Experimentos	68

LISTA DE TABELAS

Tabela 1 – Quantidade de trabalhos por Mecanismo de Busca	29
Tabela 2 – Quantidade de trabalhos por CI	30
Tabela 3 – Quantidade de trabalhos por CE	31
Tabela 4 – Quantidade de trabalhos no CE5	31
Tabela 5 – Comparativo dos Trabalhos Selecionados	41
Tabela 6 – Avaliações dos top 10 artigos recomendados.	72
Tabela 7 – Avaliações dos top 5 artigos recomendados.	73
Tabela 8 – Perfil dos participantes do experimento - Instituição e Área de atuação	74
Tabela 9 – Respostas dos participantes: Sobre os pares de itens apresentados, quanto você sentiu que tinham relação um com o outro?	75
Tabela 10 – Respostas dos participantes: O que lhe motivou a escolher um item no par, ou a não escolha?	76
Tabela 11 – Respostas dos participantes: Sugestão de como iniciar o perfil em um sistema de recomendação	77
Tabela 12 – Respostas Parte 1 dos participantes: Sugestões de melhorias para um sistema de seleção de perfil em pares:	79
Tabela 13 – Respostas Parte 2 dos participantes: Sugestões de melhorias para um sistema de seleção de perfil em pares:	80
Tabela 14 – Eventos científicos relacionados à IHC	91
Tabela 15 – Eventos científicos relacionados à IHC	92
Tabela 16 – Eventos científicos relacionados à IHC	93
Tabela 17 – Eventos científicos relacionados à IHC	94
Tabela 18 – Eventos científicos relacionados à IHC	95
Tabela 19 – Eventos científicos relacionados à IHC	96
Tabela 20 – Eventos científicos relacionados à IHC	97
Tabela 21 – Eventos científicos relacionados à IHC	98
Tabela 22 – Eventos científicos relacionados à IHC	99
Tabela 23 – Eventos científicos relacionados à IHC	100
Tabela 24 – Eventos científicos relacionados à IHC	101
Tabela 25 – Eventos científicos relacionados à IHC	102
Tabela 26 – Eventos científicos relacionados à IHC	103
Tabela 27 – Eventos científicos relacionados à IHC	104
Tabela 28 – Eventos científicos relacionados à IHC	105
Tabela 29 – Eventos científicos relacionados à IHC	106

LISTA DE ABREVIAÇÕES E ACRÔNIMOS

BC	Baseada em Conteúdo
FC	Filtragem Colaborativa
HB	Híbrida
IHC	Interação Humano-Computador
MBA	Mecanismos de Busca Acadêmica
RI	Recuperação de Informação
SR	Sistema de Recomendação
SRs	Sistemas de Recomendação

SUMÁRIO

1	INTRODUÇÃO	12
1.1	MOTIVAÇÃO E ESCOPO DESTE TRABALHO	13
1.1.1	Sistema de recomendação para sugerir artigos para a Comunidade Brasileira de IHC	13
1.1.2	Estratégia de Pós-processamento aplicada a um Sistema de Recomendação de Artigos visando refletir os vários interesses do usuário	14
1.2	OBJETIVOS	16
1.2.1	Objetivo Geral	16
1.2.2	Objetivos Específicos	16
1.3	METODOLOGIA	16
1.4	ESTRUTURA DO TEXTO	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	SISTEMAS DE RECOMENDAÇÃO	18
2.1.1	Recomendação Baseada em Conteúdo	19
2.1.2	Filtragem Colaborativa	21
2.1.3	Híbrida	21
2.1.4	Métricas de Avaliação de Sistemas Recomendação	22
2.1.4.1	<i>Precision</i>	23
2.1.4.2	<i>Recall</i>	24
2.1.4.3	<i>NDCG - Normalized Discounted Cumulative Gain</i>	24
2.1.4.4	<i>MRR- Mean Reciprocal Rank</i>	24
2.2	COLD START	25
2.3	CONSIDERAÇÕES DO CAPÍTULO	27
3	MAPEAMENTO SISTEMÁTICO	28
3.1	QUESTÕES DE PESQUISA	28
3.2	STRING DE PESQUISA	28
3.3	CRITÉRIOS DE SELEÇÃO E PROCESSO	29
3.4	ANÁLISES E RESULTADOS	31
3.5	TRABALHOS SELECIONADOS	33
3.6	AMEAÇAS À VALIDADE	41
3.7	CONSIDERAÇÕES DO CAPÍTULO	42
4	MODELO DE SISTEMA DE RECOMENDAÇÃO DE ARTIGOS PARA MINIMIZAR COLD START	43
4.1	MODELO PROPOSTO	44
4.1.1	Arquitetura do Modelo	44

4.1.2	Processo de criação do perfil inicial - Opção 1	45
4.1.3	Processo de criação do perfil inicial - Opção 2	51
4.2	COMPARAÇÃO DOS TRABALHOS SELECIONADOS	55
4.3	CONSIDERAÇÕES DO CAPÍTULO	56
5	SISTEMA DE RECOMENDAÇÃO DE ARTIGOS COM ESTRATÉGIA PARA REDUÇÃO DO COLD START USER	57
5.1	TECNOLOGIAS UTILIZADAS	57
5.2	BASE DE ITENS/ARTIGOS A SEREM RECOMENDADOS	58
5.3	DEFINIÇÕES SOBRE INTERFACE/INTERAÇÃO DOS USUÁRIOS	59
5.4	SISTEMA IMPLEMENTADO	62
5.4.1	Cenário de Uso - 1	68
5.4.2	Cenário de Uso - 2	69
5.5	AVALIAÇÃO DO SISTEMA	69
5.5.1	Testes Funcionais	70
5.5.2	Experimentos do sistema	70
5.5.3	Resultados da avaliação de experimentos	71
5.6	CONSIDERAÇÕES SOBRE O CAPÍTULO	81
6	CONSIDERAÇÕES FINAIS	82
	REFERÊNCIAS	84
	APÊNDICE A – LISTA DE EVENTOS	91
	APÊNDICE B – QUESTIONÁRIO - PREFERÊNCIA DE INTERFACES	107
	APÊNDICE C – QUESTIONÁRIO - TESTES FUNCIONAIS	110
	APÊNDICE D – CONVITE - EXPERIMENTOS	115
	APÊNDICE E – QUESTIONÁRIO - EXPERIMENTOS	117

1 INTRODUÇÃO

Com a grande quantidade de dados disponíveis na Internet a cada dia, é necessário que existam sistemas capazes de auxiliar na identificação do que é de maior interesse para um usuário. Existem bibliotecas que possuem mecanismos de busca baseadas muitas vezes em técnicas de Recuperação de Informação - Information Retrieval (SALTON; MCGILL, 1983). Assim, os usuários formulam pesquisas a partir de argumentos que são formados por termos/palavra, sendo retornados os documentos mais relevantes. Mesmo com estes mecanismos os usuários tendem a ter dificuldades de encontrar trabalhos relevantes a sua pesquisa.

Para que a tarefa de identificar trabalhos e itens relacionados ao interesse do usuário seja facilitada vem sendo desenvolvidos Sistemas de Recomendação. Sistemas de Recomendação (SRs) são ferramentas de *software* e técnicas que fornecem sugestões de itens para seus usuários (RICCI; ROKACH; SHAPIRA, 2010) e auxiliam a lidar com a sobrecarga de informações encontradas atualmente. Segundo Ricci, Rokach e Shapira (2015), o interesse cresceu visto que: 1) Os sistemas de recomendação desempenham um papel altamente importante em sites da Internet como Amazon.com, YouTube, Netflix, Spotify, LinkedIn, Facebook, Tripadvisor, Last.fm e IMDb; 2) Em instituições de ensino superior em todo o mundo, disciplinas de graduação e pós-graduação agora são dedicados inteiramente aos SRs, tutoriais sobre SRs são muito populares em conferências de Ciência da Computação; e 3) Existem várias edições especiais em revistas acadêmicas que cobrem pesquisas e desenvolvimentos na área de SRs. No contexto de Bibliotecas Digitais, Sistemas de Recomendação podem sugerir artigos para os pesquisadores e ajudá-los a encontrar mais rapidamente aquilo que for de seu interesse (BAI et al., 2019).

Mesmo com todo avanço na área de Sistemas de Recomendação, visto os diversos trabalhos publicados e a aplicação destes sistemas nos mais diferentes domínios, ainda há muito o que explorar, no que se refere, por exemplo, a alguns dos problemas conhecidos destes sistemas. Um dos problemas mais conhecidos em Sistemas de Recomendação é o problema de *Cold Start*.

O problema de *Cold Start* pode ser definido como a impossibilidade de criar possíveis recomendações confiáveis devido à falta de dados sobre um novo usuário ou um novo item (MONTI; RIZZO; MORISIO, 2021). Esse problema surge quando um novo usuário (*Cold-Start User* - *Cold Start* do usuário) ou item (*Cold-Start Item* - *Cold Start* do Item) é adicionado ao SR e não há informações históricas para identificar as preferências do usuário (CARVALHO; RODRIGUES; OLIVEIRA, 2020). O problema do novo usuário *Cold-Start User* está relacionado ao fato de que um usuário que interagiu pouco no sistema ou fez poucas avaliações não receberá inicialmente recomendações de qualidade (LICHTNOW, 2015). O problema de *Cold-Start* do usuário afeta negativamente o desempenho do SR para fornecer recomendações significativas para novos usuários (ALLIOUI, 2017). Na abordagem denominada Filtragem Colaborativa existe o problema do novo item (*Cold-Start Item*), itens que foram avaliados por poucos usuários ou que não foram avaliados não serão recomendados.

Em Bai et al. (2019), é indicado que Sistemas de Recomendação podem ser usados

para sugerir artigos presentes em uma biblioteca digital para os pesquisadores e ajudá-los rapidamente a encontrar os documentos que precisam. Os Sistemas de Recomendação de artigos visam ajudar os pesquisadores a mitigar a sobrecarga de informação, e também encontrar documentos relevantes e classificar registros de publicações. Borba (2015) destaca que Sistemas de Recomendação podem ajudar a evitar que parte do material disponibilizado nunca seja utilizado.

1.1 MOTIVAÇÃO E ESCOPO DESTE TRABALHO

Neste capítulo são apresentados alguns trabalhos desenvolvidos anteriormente pelo grupo de pesquisa GPIE (Grupo de Pesquisa em Informática na Educação) da UDESC que serviram de base para a identificação do problema desta presente pesquisa. Os trabalhos utilizam uma técnica de minimizar o *Cold-Start* a partir da criação do perfil inicial, que necessita de aperfeiçoamento.

Estes trabalhos são Cezar (2019) e Souza, Lichnow e Gasparini (2022a). Os três lidaram com a implementação de um Sistema de Recomendação de Artigos para a comunidade de Interação Humano-Computador (IHC) no Brasil. Os trabalhos possuem foco na lista de recomendação apresentada ao usuário. Inicialmente em Cezar (2019) foi criado um Sistemas de Recomendação que usa a abordagem Baseada em Conteúdo. Dois desses trabalhos (Borba (2020), Souza, Lichnow e Gasparini (2022b)) focam em estratégias de pós-processamento para a lista de recomendação, enquanto que o primeiro Cezar (2019) foca em criar o perfil inicial do usuário a partir de termos identificados a partir do processamento do conteúdo da biblioteca digital do Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais. A seguir, são apresentados os detalhes de cada trabalho.

1.1.1 Sistema de recomendação para sugerir artigos para a Comunidade Brasileira de IHC

Em Cezar (2019) e Cezar et al. (2021) aplicaam técnicas para minimizar o *Cold Start* a partir da construção do perfil inicial do usuário. O Sistema de Recomendação para artigos criado exige que o usuário selecione termos de interesse para que o sistema pudesse gerar o seu perfil inicial.

O Sistema de Recomendação proposto foi construído para recomendar artigos para a comunidade de IHC que publicados no evento do Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais. Esses artigos estavam armazenados numa base de dados já criada pelo grupo de pesquisa. Desta forma, o SR tinha o propósito de facilitar a pesquisa para os participantes e pesquisadores do evento auxiliando-os na identificação de artigos relevantes.

Neste sentido, o Sistema de Recomendação apresentava apenas 10 termos e o usuário selecionava 5 deles para compor seu perfil. Esses termos apresentados ao usuário são as palavras-chave mais significativas dos *abstracts* nos anais do Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC).

O processo de seleção dessas palavras baseou-se no modelo de *bag-of-words* e cálculo do IDF (ADOMAVICIUS; TUZHILIN, 2005), onde para um documento d , o conjunto de pesos determinado pelos pesos determinado por qualquer função de ponderação que mapeie o número de ocorrências do termo t em um documento d para um valor real positivo, pode ser visto como um resumo quantitativo desse documento. Isso significa que a ordem exata das palavras em um documento é ignorada, mas o número de ocorrências de cada termo importa (MANNING; RAGHAVAN; SCHÜTZ, 2009).

Inicialmente foram identificadas 4.865 palavras nos resumos. Desta forma, seguiram-se filtragens sucessivas para que as palavras fossem as mais adequadas para representação do perfil de um usuário. Alguns critérios para a exclusão das palavras foram: remoção de *stop-words*, além disso foram removidas palavras que tinham um valor de IDF próximo de zero (palavras que ocorriam em muitos documentos). A etapa final consistiu da análise manual das palavras, realizada em conjunto com pesquisadores da área de IHC. Ao final da seleção, 178 palavras foram definidas para a lista de palavras a serem apresentadas ao usuário.

Assim, de forma aleatória, 10 dos 178 palavras eram apresentados ao usuário no momento do seu cadastro, de forma que ele selecionasse os 5 termos mais relevantes que iriam compor o seu perfil inicial. Para cada termo o usuário podia indicar a relevância atribuindo ao termo um valor numérico de 1 até 5.

Usando as palavras indicadas pelo usuário, era realizada uma consulta a biblioteca que retornava as palavras mais relevantes. O armazenamento dos itens da biblioteca digital e a consulta eram realizados usando os recursos do *ElasticSearch*. Assim eram retornados os documentos mais similares ao perfil do usuário.

Embora esta abordagem tenha contribuído em reduzir o *Cold-Start*, existem desafios e talvez os principais fossem definir quais termos apresentar ao usuário e como realizar sua extração a partir dos documentos presentes na biblioteca. Além disso, outra questão refere-se ao fato de que como as palavras eram apresentadas de forma aleatória, não representava uma garantia de que a diversidade de termos discutidos nos artigos presentes na base estivesse coberta, o que poderia fazer com que apenas parte dos interesses de um usuário fossem identificados. Sobre diversidade cabe destacar que trabalhos posteriores procuraram abordar isto mas não na formação do perfil inicial e são descritos a seguir.

1.1.2 Estratégia de Pós-processamento aplicada a um Sistema de Recomendação de Artigos visando refletir os vários interesses do usuário

No trabalho de Borba (2020) foi criada a partir do trabalho de Cezar (2019), uma abordagem de pós-processamento para ser aplicada após o sistema retornar as recomendações para o usuário, usando a abordagem Baseada em Conteúdo desenvolvida usando os recursos do *ElasticSearch*.

A estratégia de pós-processamento foi desenvolvida para ser aplicada após o sistema retornar as top 10 recomendações para o usuário, usando a abordagem Baseada em Conteúdo de-

senvolvida usando os recursos do *ElasticSearch*. A aplicação da estratégia de pós-processamento processa a lista dos dez itens recomendados e executa uma nova consulta no *ElasticSearch* utilizando a API *Explain*, que permite recuperar informações sobre o cálculo da similaridade entre o perfil e os documentos. A partir da API *Explain* é possível saber quais termos da *query*, que representa o perfil do usuário, constam no artigo recomendado.

Utilizando esses recursos o sistema verifica se algum termo contido no perfil do usuário não está contemplado nos 10 artigos recomendados, ou seja, se algum termo está ausente na lista de recomendação. Caso isso ocorra, uma nova consulta é realizada no *ElasticSearch* para obter um novo artigo a ser inserido na lista dos artigos a serem recomendados que contemple o termo que está ausente.

Desse modo a lista de itens recomendados tende a refletir todas as preferências do usuário por considerar todos os termos de interesses, pois caso um deles não esteja sendo incluso nas recomendações, uma nova recomendação será incluída na lista para contemplar o termo. Esta inclusão, por sua vez, é realizada na última posição possível da lista dos itens recomendados. O pós-processamento visa então incluir apenas um novo registro na última posição possível da lista de recomendações para contemplar os termos ausentes na recomendação, mas que são de interesses do usuário, sem prejudicar os termos de maiores prioridades do usuário que estão contemplados nos artigos das posições iniciais da lista de itens recomendados.

Se a análise realizada na etapa de pós-processamento identificar que todos os termos de interesse do usuário estão contemplados na lista de recomendação, então nenhum artigo da lista de recomendação é alterado e ela permanece igual ao resultado do processamento feito pelo *ElasticSearch*. Já o trabalho de Souza, Lichnow e Gasparini (2022b) foi criado a partir do trabalho de Borba (2020) uma abordagem de pós-processamento que efetua troca de itens da lista de recomendação levando em consideração dois atributos que compõem os artigos: o ano de publicação dos artigos e os termos presentes nos textos dos artigos.

A partir do ano de publicação dos artigos, os autores impedem a formação de uma bolha sobre um determinado ano, isto é, um grande número de ocorrência de artigos apenas de um mesmo ano na lista de recomendação. Acerca da distribuição dos termos de interesse, a diversificação da lista busca evitar a formação de uma bolha sobre artigos que provavelmente descrevem a mesma temática, para isso, é feita a troca desses artigos por outros que estejam também relacionados a outros interesses do usuário.

A diferença aqui em relação a Borba (2020) é que a diversidade da lista de recomendação é medida usando a métrica ILD (*Intra List Diversity* ou diversidade intra-lista) e procura-se desta forma não apenas realizar trocas na lista inicial de itens recomendados mas ver se as trocas feita possuem impacto positivo.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Minimizar o *Cold Start User* em Sistemas de Recomendação de Artigos Científicos que usam a abordagem Baseada em Conteúdo.

1.2.2 Objetivos Específicos

- Entender as técnicas para minimizar o *Cold Start User* usadas em Sistemas de Recomendação apresentadas em trabalhos anteriores através da realização do Mapeamento Sistemático da Literatura;
- Propor e desenvolver uma técnica que reduza o *Cold Start User*;
- Aplicar a técnica definida com usuários de uma biblioteca digital de artigos científicos;
- Analisar os resultados da utilização da técnica.

1.3 METODOLOGIA

Quanto aos objetivos, a presente pesquisa classifica-se como exploratória (GIL, 2009). Este tipo de pesquisa tem como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a construir hipóteses (GIL, 2009). Este trabalho visa explorar e identificar formas de minimizar o *Cold Start* que foram aplicadas em trabalhos que lidam com este problema de forma a propor algo que possa auxiliar na redução do problema de *Cold Start*.

Quanto a maturidade este trabalho com o objetivo de minimizar o Cold-Start em SRs classifica-se no tipo: “Apresentação de Algo Diferente” (WASLAWICK, 2014). O *Cold-Start* de fato é um problema existente, mas a proposta do trabalho é solucionar o problema de maneira diferente das já propostas. Desta maneira, o trabalho trará uma comparação de técnicas já realizadas e o principal objetivo é propor uma maneira de construir um perfil inicial de usuário de maneira que minimize o *Cold Start User*.

Podemos classificar este trabalho sendo primeiramente quantitativo (FONTELLES et al., 2006), visto que o trabalho irá medir o valor de *Precision* das recomendações geradas a partir da comparação de similaridade dos documentos com o perfil inicial do usuário. Entretanto, o trabalho pode ser classificado como uma pesquisa qualitativa visto que a partir do *feedback* explícito dos itens e de um questionário de satisfação, será avaliado como o usuário se sentiu utilizando o sistema e sua satisfação com as recomendações recebidas.

Neste sentido, a classificação quanto aos procedimentos de pesquisa, neste trabalho será utilizada a pesquisa bibliográfica. A pesquisa bibliográfica segundo Waslawick (2014) implica o estudo de artigos, teses, livros e outras publicações usualmente disponibilizadas por editoras e indexadas. Este passo é o primeiro a ser realizado para o entendimento inicial sobre Sistemas de Recomendação e *Cold Start*.

Em Waslawick (2014) a pesquisa experimental caracteriza-se pela manipulação de um aspecto da realidade pelo pesquisador. A pesquisa experimental implica ter uma ou mais variáveis experimentais que podem ser controladas pelo pesquisador, e uma ou mais variáveis observadas. Neste caso, considera-se que este trabalho seja uma pesquisa experimental pois teremos o Sistema de Recomendação como a variável controlada mas será observado a partir da construção do perfil para cada novo usuário no sistema, a recomendação feita, se aumentou ou não o valor de precisão.

1.4 ESTRUTURA DO TEXTO

O restante deste texto está organizado da seguinte forma: O Capítulo 2 apresenta a fundamentação conceitual, compreendendo Sistemas de Recomendação e um dos principais problemas, o *Cold Start*. O Capítulo 3 apresenta o Mapeamento Sistemático realizado, evidenciando a oportunidade de pesquisa identificada e explorada nessa dissertação. O Capítulo 4 apresenta a proposta para minimizar o *Cold Start User* em Sistemas de Recomendação de abordagem Baseada em Conteúdo. O Capítulo 5 apresenta as tecnologias que serão utilizadas na implementação da proposta e experimentos realizados. Por fim, o Capítulo 6 apresenta as considerações finais, apontando os próximos passos da pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os temas centrais envolvidos neste trabalho. A Seção 2.1 conceitua e descreve Sistemas de Recomendação e suas principais abordagens. A Seção 2.2 descreve o problema de *Cold Start* e suas principais técnicas para amenizar o problema em Sistemas de Recomendação.

2.1 SISTEMAS DE RECOMENDAÇÃO

Sistemas de recomendação (SRs) são ferramentas e técnicas de software que fornecem sugestões de itens com maior probabilidade de interesse para um usuário específico (RICCI; ROKACH; SHAPIRA, 2015). O sistema de recomendação é uma tecnologia de filtro de informações, que objetiva ajudar a encontrar os itens de interesse do usuário com rapidez e precisão (HONG; JEON; JEON, 2012).

Sistemas de Recomendação coletam ativamente vários tipos de dados a fim de construir suas recomendações. Os dados são principalmente sobre os itens a serem sugeridos e os usuários que receberão essas recomendações.

Item é o termo que representa o que o sistema recomenda para os usuários, assim, os itens de um Sistema de Recomendação podem ser, por exemplo, filmes, produtos ou artigos científicos. Por fim, o SR recomenda os top N artigos associados aos interesses de pesquisa de um pesquisador ou foco de pesquisa (BAI et al., 2019).

Desde que os Sistemas de Recomendação foram introduzidos, muitos algoritmos/abordagens de recomendação surgiram. Os Sistemas de Recomendação são geralmente classificados em três grandes abordagens: (a) Recomendações Baseadas no Conteúdo, o usuário recebe recomendações de itens que são similares àqueles que ele preferiu no passado; (b) Recomendações baseadas na Filtragem Colaborativa, o sistema recomenda itens para o usuário em questão que foram avaliados por usuários com gostos semelhantes; e (c) Abordagens Híbridas, quando métodos de Recomendações Colaborativas e Baseadas no Conteúdo são combinados para recomendar itens (CONSONI, 2014).

Nos últimos anos, grandes tecnologias de inteligência artificial foram criadas e geraram muita expectativa em torno disso, como ChatGpt, Gemini, Copilot e Midjourney. Essas tecnologias possuem em comum a dependência de um LLM (Large-Language Model) para funcionar. Um LLM é um Modelo de Linguagem de Grande Escala, e este é um modelo de linguagem treinado com grandes quantidades de textos. Alguns dos impactos gerados através dessas tecnologias em Sistemas de Recomendação foi a criação de Sistemas Conversacionais. Um Sistema Conversacional geralmente tenta coletar as preferências do usuário fazendo perguntas. Depois que a preferência do usuário é coletada, ele faz recomendações personalizadas ao usuário (SUN; ZHANG, 2018).

2.1.1 Recomendação Baseada em Conteúdo

Na abordagem Baseada em Conteúdo, o sistema aprende a recomendar itens semelhantes aos que o usuário gostou no passado (RICCI; ROKACH; SHAPIRA, 2015). Nesta abordagem a representação dos itens pode ser feita usando pares atributo-valor. Com o perfil do usuário acontece o mesmo, ele é representado por termos que são os atributos, e valores numéricos associados a cada termo, que indicam o quanto aquele termo define os interesses do usuário. O processo de recomendação consiste basicamente em combinar os atributos do perfil do usuário em relação aos atributos de um objeto de conteúdo (RICCI; ROKACH; SHAPIRA, 2011). No caso de um documento, a importância de uma palavra-chave no documento é determinada pela medida dos pesos que pode ser definido de diferentes maneiras. A maneira mais comum na Recuperação de Informação (RI) de especificar palavras-chave é a da *frequência do termo-frequência do documento inversa* chamada de métrica *TF-IDF* (ADOMAVICIUS; TUZHILIN, 2005) explicado a seguir. Além disso, dentro da área de Recuperação da Informação, uma forma de medir a similaridade de itens em um SR é através da similaridade do Cosseno, onde $sim(d1, q)$ é o resultado da distância dos vetores dos documentos.

Na abordagem Baseada em Conteúdo a representação dos itens pode ser feita usando pares atributo-valor. O perfil do usuário é representado por termos que são os atributos, e valores numéricos associados a cada termo, que indicam o quanto aquele termo define os interesses do usuário. A utilidade $u(c, s)$ de um item s para um usuário c é estimada baseada nas utilidades $u(c, s_i)$ atribuídas pelo usuário c para os itens $s_i \in S$ que são similares ao item s . A importância da palavra-chave k_j no documento d_j é determinada pela medida dos pesos w_{ij} que pode ser definido de diferentes maneiras. A maneira mais comum na RI para especificar palavras-chave é a métrica *TF-IDF*, que pode ser calculada pela fórmula a seguir:

$$TF-IDF_{i,d} = TF_{i,d} \times IDF_i \quad (1)$$

O *TF-IDF* atribui ao termo i um peso no documento d para este termo. Onde, $tf_{i,d}$ é o número de ocorrências do termo i no documento d ; e o idf_i , chamado de *Inverse Document Frequency* de um termo i , para ressaltar o efeito de termos que ocorrem com muita frequência em um item são significativos para determinação de relevância deste item (MANNING; RAGHAVAN; SCHÜTZ, 2009).

Assumindo que o total de documentos é representado por N , uma palavra-chave i que aparece em n_i deles. E também, considerando $f_{i,j}$ o número de vezes que a palavra-chave i aparece no documento d_j , $max_i f_{i,j}$ o número de termos no documento d_j , e df_i representa o número de documentos que o termo i aparece. As equações da frequência da palavra-chave no documento, e a frequência inversa da palavra-chave no documento são respectivamente:

$$TF_{i,d} = \frac{f_{i,j}}{max_i f_{i,j}} \quad (2)$$

$$IDF_i = \log \frac{N}{df_i} \quad (3)$$

Um documento pode ser definido como um vetor da maneira (ADOMAVICIUS; TUZHILIN, 2005):

$$d_j = (w_{1,j}, w_{2,j}, \dots, w_{k,j}) \quad (4)$$

Sendo $w_{1,j}$ é o termo presente na posição j do item 1, $w_{2,j}$ é o termo presente na posição j do item 2 e assim por diante com os outros termos dos outros documentos, caso existirem. Alguns exemplos de vetores de palavras-chave de documentos que falam sobre gêneros de filmes, são apresentados a seguir:

$$\begin{aligned} \text{perfil do usuário} &= (\text{terror, comédia}) \\ \text{documento1} &= (\text{terror, comédia}) \\ \text{documento2} &= (\text{ficção, romance}) \\ \text{documento3} &= (\text{documentário, romance}) \\ \text{documento4} &= (\text{romance, comédia}) \\ \text{TF-IDF} &= (\text{terror, comédia, ficção, romance, documentário}) \end{aligned}$$

Um exemplo de vetor de *TF-IDF*:

$$\text{TF-IDF} = (1; 1; 0,5; 0,75; 0,8)$$

Dentro da área de Recuperação da Informação, uma forma de medir a similaridade de itens em um SR é o Cosseno (BORBA, 2015):

$$sim(d1, d2) = \frac{\sum_{i=1}^k w_{1,i} \times w_{2,i}}{\sqrt{\sum_{i=1}^k w_{1,i}^2 \sum_{i=1}^k w_{2,i}^2}} \quad (5)$$

Onde $sim(d1, d2)$ é o resultado da distância dos vetores, variando de [0,1]; $w_{1,i}$ é o termo presente na posição i do item 1; e $w_{2,i}$ é o termo presente na posição i do item 2.

Um exemplo de vetores representando o perfil do usuário e um item para o cálculo da similaridade seria $u=(1,1)$ o perfil do usuário e $d_1=(1,1)$, o item com os termos representados de forma vetorial. Esses vetores representam que o documento 1 possui tanto o primeiro como o segundo termo em comum com as preferências do usuário u . Então a equação ficaria:

$$sim(d1, d2) = \frac{1 \times 1 + 1 \times 1}{\sqrt{(1^2 + 1^2) \times (1^2 + 1^2)}} = 1 \quad (6)$$

Como o valor da similaridade resultou em 1, ou seja, valor máximo, o documento 1 é bem similar às preferências do usuário.

Segundo Ricci, Rokach e Shapira (2015), as principais vantagens desta abordagem são: (a) Independência do usuário: o usuário possui liberdade para construir o seu modelo de perfil da maneira que achar melhor, pois interage somente com o que gosta e não necesita da opinião dos vizinhos; e (b) Itens novos: a abordagem é capaz de recomendar itens ainda não classificados por nenhum usuário. As principais desvantagens são:(a) Partida Fria (*Cold Start*):

o novo usuário não irá receber boas recomendações até que tenha avaliado alguns itens; (b) Superespecialização: quando são recomendados sempre itens similares aqueles já avaliados pelo usuário. A superespecialização pode acontecer na abordagem Baseada em Conteúdo pois esse tipo de abordagem não adota nenhum método para recomendar itens inesperados e, assim, apresenta um nível limitado de novidade na coleção. Um usuário, por exemplo, que avaliou positivamente livros de um determinado autor possivelmente receberá mais e mais livros desse mesmo autor (PAZZANI; BILLSUS, 2007).

2.1.2 Filtragem Colaborativa

Sistemas de Recomendação de Filtragem Colaborativa recomendam para o usuário itens que foram bem avaliados por usuários com perfil similar ao seu. Assim, a utilidade de um item para um usuário é estimada baseada nas utilidades atribuídas para o item pelos usuários que são "similares" ao usuário (ADOMAVICIUS; TUZHILIN, 2005). Um dos algoritmos mais utilizados em sistemas de recomendação que utilizam a Filtragem Colaborativa é o Baseados em Vizinhança (*Neighborhood Based*), sendo a variação mais utilizada é a *User-User* se o sistema tiver mais itens que usuários, onde procura-se avaliar quais usuários têm gostos similares (avaliaram os mesmos itens, dando notas similares) em relação a um determinado usuário (LICHTNOW, 2015). Uma das etapas nesta abordagem, seria baseada no cálculo da similaridade usando o Coeficiente de Pearson, comparando um usuário com os demais através das notas que deram aos itens. O Coeficiente de Pearson mede a força do relacionamento de duas variáveis.

Além da abordagem *User-User*, é proposta a abordagem *Item-Item* na qual a comparação é feita entre as avaliações dos itens e não entre usuários. A ideia principal da estratégia *Item-Item* é prever a nota que o usuário daria para um item com base nas notas que ele deu para itens semelhantes àquele. Para essa estratégia, o cálculo da similaridade pelo Cosseno, semelhante ao já citado, é uma métrica padrão e a que apresenta os melhores resultados (RICCI; ROKACH; SHAPIRA, 2011). As vantagens da Filtragem Colaborativa é de não considerar a descrição dos itens, as recomendações realizadas podem ser bastante interessantes e inesperadas ao usuário. Porém, abordagens colaborativas possuem a desvantagem do *Cold-Start*, assim como na abordagem Baseada em Conteúdo. Existem dois tipos de *Cold-Start* nesta abordagem: o *Cold-Start User* e o *Cold-Start Item* detalhadas na Seção 2.2. Além disso, há algumas desvantagens de utilizar a Filtragem Colaborativa: (a) Ovelha Cinza: usuários com gostos distintos demais, tornam-se desafiantes ao Sistema de Recomendação; (b) Escalabilidade: pois determinar usuários com perfil semelhante tem um custo computacional alto e; (c) Relevância da opinião dos usuários e da confiabilidade das avaliações;

2.1.3 Híbrida

Os Sistemas de Recomendação Híbridos combinam duas ou mais técnicas de recomendação para obter melhor desempenho reduzindo as desvantagens de uma única abordagem (BURKE,

2002). Métodos de Filtragem Colaborativa e Baseados em Conteúdo podem ser combinados de várias maneiras, por exemplo, fundindo suas previsões individuais em uma previsão única e mais robusta ou adicionando informações de conteúdo em um modelo de filtragem colaborativa (RICCI; ROKACH; SHAPIRA, 2015).

Existem ainda outras abordagens de recomendação. Segundo Ricci, Rokach e Shapira (2011), aparecem com frequência mais duas classificações, Demográfica e Baseada em Conhecimento:

- Demográfica: recomendações são realizadas de acordo com o perfil demográfico do usuário. O sistema utiliza por exemplo informações como idade, gênero, profissão e educação (RICCI; ROKACH; SHAPIRA, 2011). Assim identifica usuários que gostam de um determinado item;
- Baseada em Conhecimento: recomenda itens aos usuários com base no conhecimento que o sistema possui sobre como características de um item se encaixam nas necessidades de um usuário, e o quanto útil esse item será.

2.1.4 Métricas de Avaliação de Sistemas Recomendação

Em um Sistema de Recomendação, é necessário identificar o conjunto de propriedades que influenciam o sucesso do Sistema de Recomendação no contexto de uma aplicação específica. Existem três tipos de experimentos para o processo de avaliação da recomendação: *offline*, estudo com o usuário e *online* (RICCI; ROKACH; SHAPIRA, 2015).

Os experimentos *offline* usam conjuntos de dados existentes junto de um protocolo que modela o comportamento do usuário para estimar medidas de desempenho de recomendação, como a *precision*. Um experimento *offline* é realizado utilizando dados pré-coletados de um conjunto de usuários escolhendo ou avaliando itens. Utilizando esse conjunto de dados é possível tentar simular o comportamento de usuários que interagem com um Sistema de Recomendação (RICCI; ROKACH; SHAPIRA, 2015). Ao realizar esse tipo de experimento, assume-se que o comportamento do usuário quando os dados são coletados, será suficientemente similar ao comportamento do usuário quando o Sistema de Recomendação é implantado, para que se possa tomar decisões confiáveis baseadas na simulação.

Porém, o experimento que fornece as evidências mais fortes sobre o verdadeiro valor do sistema é uma avaliação *online*, onde o sistema é usado por usuários reais que executam tarefas reais. O experimento *online* é quando são executados em grande escala experimentos em um sistema implantado. Ainda segundo o autor, é mais confiável comparar alguns sistemas *online*, obtendo um ranking de alternativas, em vez de números absolutos.

No experimento de estudo com o usuário um pequeno conjunto de usuários é solicitado a executar um conjunto de tarefas usando o sistema, normalmente respondendo perguntas depois sobre sua experiência. Para avaliar a reação do usuário diante dos itens é necessário armazenar informações sobre elas. Estas informações coletadas são então armazenadas, e podem

ser chamadas de *feedback*. Estas reações juntamente com as descrições dos itens são exploradas durante o processo de aprendizagem do modelo do sistema e deve ser capaz de prever a relevância de novos itens. Porém, o usuário também pode ser capaz de deixar explícito seus interesses, como um perfil inicial (RICCI; ROKACH; SHAPIRA, 2011).

Existem dois tipos de técnicas adotadas para se obter um *feedback* do usuário. O *feedback* pode ser explícito ou implícito. *Feedback* explícito é quando o sistema exige que o usuário avalie os itens.

O *feedback* explícito possui três tipos de abordagens. Estas abordagens podem ser do tipo: (a) *like/dislike*: quando os itens são classificados como relevantes ou não utilizando uma escala de sistema binário de avaliação; (b) avaliação por meio de uma escala numérica, os itens são julgados; e (c) comentários em textos: comentários feitos sobre um item são coletados e apresentados para os usuários como uma maneira de facilitar o processo de decisão.

Uma das maneiras de avaliação citada, é por meio de notas que os usuários podem dar para a recomendação recebida. O usuário pondera uma nota dentro de uma escala definida para indentificar o quanto gostou ou desgostou do sistema e da recomendação do item.

O *feedback* implícito é quando o sistema não exige nenhum envolvimento explícito do usuário, e deriva-se através do monitoramento das ações do usuário. Tipos de feedback implícito incluem histórico de compras, histórico de navegação, padrões de pesquisa ou até movimentos do mouse. Por exemplo, marcar uma página como favorito é interpretado como uma forte evidência de interesse do usuário nessa página. Os documentos visitados são processados como exemplos positivos de interesses do usuário e documentos não visitados como exemplos negativos (RICCI; ROKACH; SHAPIRA, 2011).

A seguir, algumas das principais métricas avaliação usadas em sistemas de recomendação são descritas. Para o presente trabalho a de maior relevância é a precisão.

2.1.4.1 Precision

A precisão, do inglês *precision*, é usada para medir a precisão de Sistemas de Recomendação recomendando artigos relevantes aos pesquisadores (BAI et al., 2019). Quanto maior o valor, mais precisa é a recomendação feita. A precisão pode ser definida como a taxa de itens relevantes no resultado retornado, ela mostra a capacidade da recomendação de apresentar apenas itens úteis (OLMO; GAUDIOSO, 2008).

$$Precision = \frac{\text{itens relevantes}}{\text{total de itens recomendados}} \quad (7)$$

Fonte: Bai et al. (2019)

A *precision* pode ser calculada para um top k itens, sendo k um valor positivo. Geralmente, quando utilizada dessa forma a métrica denomina-se *Precision@k* ou *Precision at k*. Desta forma, a fórmula ficaria (RICCI; ROKACH; SHAPIRA, 2015):

$$Precision@k = \frac{\text{@k itens relevantes}}{\text{total de @k itens recomendados}} \quad (8)$$

2.1.4.2 Recall

A revocação, do inglês *recall*, quantifica a fração de artigos relevantes recuperados em todo o conjunto de documentos que estão na lista de resultados de recomendação (BAI et al., 2019). O *Recall* pode ser definida como a taxa de itens relevantes que foram recomendados em relação a quantidade total de itens relevantes.

$$Recall = \frac{\text{itens relevantes recomendados}}{\text{total de itens relevantes}} \quad (9)$$

Fonte: Bai et al. (2019)

Existem ainda outras métricas para avaliar a recomendação como por exemplo NDCG e MRR. Ambas avaliam a qualidade dos itens de acordo com a posição que estão no ranking de uma lista de itens recomendados.

2.1.4.3 NDCG - Normalized Discounted Cumulative Gain

É usado para avaliar a qualidade de um determinado tipo de lista de recomendação. Se o usuário acha que o item é muito relevante para sua pesquisa, o ganho é alto, caso contrário, o ganho é 0. Quanto mais relevante o item, maior o ganho. É desejável que o mais relevante dos itens apareça no topo da lista recomendada (BAI et al., 2019).

$$DCG = \frac{1}{|U|} \sum_{u=1}^{|U|} \sum_{j=1}^J \frac{g_{uj}}{\max(1, \log_b(j))} \quad (10)$$

$$NDCG = \frac{DCG}{maxDCG} \quad (11)$$

Onde U é o conjunto de usuários que participaram neste Sistema de Recomendação, $|U|$ é o número de usuários em U , J é o número de artigos recomendados para os usuários, j é a posição do artigo na lista de recomendação, b é o valor constante, e g_{uj} representa o "ganho" que o usuário teve do artigo j . Sendo $maxDCG$ o DCG ideal, ou seja, o DCG máximo possível para uma posição j .

2.1.4.4 MRR- Mean Reciprocal Rank

A Média da Classificação Recíproca, do inglês *Mean Reciprocal Rank*, é usada para determinar a qualidade das listas de artigos recomendados. O MRR enfatiza a *precision* antecipada sobre o *recall*. O MRR para uma consulta é a classificação inversa do primeiro documento relevante encontrado numa lista de recomendação. Se preocupa apenas com a classificação dos

artigos relevantes na lista recomendada e dá uma média sobre todos os documentos relevantes (BAI et al., 2019).

$$MRR = \frac{1}{N} \sum_{i=1}^N \frac{1}{rank_i} \quad (12)$$

Fonte: Bai et al. (2019)

Onde: N representa o número de itens e $rank_i$ é a posição do i -ésimo item.

2.2 COLD START

Dentre as limitações de um Sistema de Recomendação citadas, um problema muito bem conhecido é chamado de *Cold Start*, ou Partida Fria, relacionado a performance do sistema em novos usuários e em novos itens (RICCI; ROKACH; SHAPIRA, 2015)

Existem dois tipos de *Cold-Start*: o *User Cold Start* e o *Item Cold Start*. O problema do novo usuário *New User Cold Start*, está relacionado ao fato de que um usuário que interagiu pouco no sistema ou fez poucas avaliações e não receberá inicialmente recomendações de qualidade (RICCI; ROKACH; SHAPIRA, 2015). Esse é um problema existente tanto na abordagem Baseada em Conteúdo quanto na Filtragem Colaborativa. E na Filtragem Colaborativa existe o problema do novo item *New Item Cold Start*, itens que foram avaliados por poucos usuários ou que não foram avaliados não serão recomendados (JANNACH et al., 2010). Abordar o problema de *Cold Start* pode ser importante para o engajamento e retenção de um novo usuário recém chegado no sistema.

Alyari e Navimipour (2018) apresentam uma revisão sistemática da literatura que teve como objetivo avaliar estudos individuais sobre Sistemas de Recomendação e realizar um estudo abrangente de pesquisa sobre suas técnicas. Os autores indicam que embora muitos esforços de pesquisa tenham sido feitos, ainda existem algumas questões que requerem serem aprofundados, como o problema de *Cold Start*.

Abdullah et al. (2021) apresentam um *survey* com o objetivo de apresentar várias abordagens e técnicas empregadas por pesquisadores e os desafios associados na obtenção de informações necessárias para recomendação de itens para usuários procurando minimizar o *Cold Start*.

No trabalho de Panda e Ray (2022) os autores indicam que pesquisas recentes sobre abordagens e algoritmos que tentam engajar novos usuários e promover novos itens. Os autores destacam que os problemas de *Cold Start* representam sérios desafios para o valor comercial dos SRs. Novos usuários que não achariam as recomendações úteis deixariam de usar o sistema afetando o engajamento do usuário.

Quando o sistema entra na fase de adquirir informações sobre as preferências do usuário, são realizados os métodos de elicitação de preferências. Existem quatro métodos mais comuns que os Sistemas de Recomendação utilizam para adquirir essas preferências do usuário sendo

escalas de avaliação dos itens, pesos nos atributos, críticas ou comportamento explícito e comportamento implícito (RICCI; ROKACH; SHAPIRA, 2015).

O método de escalas ou notas de avaliação, em inglês, *rating scales* é o método de avaliação explícita mais utilizado. Os usuários utilizam escalas (normalmente via escala de *Likert* de 5 pontos ou escalas de até 10 pontos para atribuir notas de acordo com suas preferências (CENA; VERNERO; GENA, 2010; COSLEY et al., 2003; DOOMS; PESSEMIER; MARTENS, 2011; GENA et al., 2011; SPARLING; SEN, 2011). Outro tipo de escalas de avaliação são usadas como “concordo fortemente, concordo, neutro, discordo, discordo fortemente” ou ainda avaliações binárias quando o usuário é questionado simplesmente se o item é bom ou ruim (GOLDBERG et al., 2001; SCHAFER et al., 2007).

A elicitação de preferências via pesos nos atributos se origina da área de análise de decisão, onde a utilidade de multi-atributos é utilizada como um padrão para a tomada de decisão (BETTMAN; LUCE; PAYNE, 1998). Esses atributos normalmente são um conjunto de palavras que descrevem os itens associados a um contexto numa escala indicando do melhor valor para o pior valor, como por exemplo, para o contexto de preço os atributos seriam “muito bom, bom, ruim, o pior, muito ruim” (BETTMAN; LUCE; PAYNE, 1998). Desta forma, cada atributo que o usuário interage possui um peso se baseando em gostos maiores que outros.

Outro método de eliciar as preferências do usuário é utilizar feedback detalhado nas recomendações, também conhecido como comportamento ou feedback explícito. Quando um usuário realiza uma avaliação por meio de uma crítica ou comentário no item, ele ajuda na melhora da confiança de decisão de outros usuários sobre aquele item.

E o comportamento implícito também pode ser utilizado. Na maioria das situações do mundo real, o feedback implícito está muito mais disponível do que avaliações e não requer nenhum esforço extra do lado do usuário (RICCI; ROKACH; SHAPIRA, 2015). Por exemplo, em uma página Web/site, você pode fazer com que os usuários visitem um URL ou cliquem em um anúncio como um feedback positivo, ou ainda ações de compra/consumo (RICCI; ROKACH; SHAPIRA, 2015). Estes tipos de comportamento podem ser usados para identificar as preferências dos usuários e gerar recomendações.

Algumas técnicas para minimizar o Cold Start são discutidas baseadas em domínios de aplicação específicos. O sistema do *MovieLens* (JELASSI; Ben Yahia; NGUIFO, 2013) por exemplo, constrói o perfil do usuário a partir de suas preferências de filmes, com base na seleção de categorias de filmes seguido de avaliações (*ratings*) dos filmes recomendados. Já o *FOAFing* é outro SR que descobre, explora e recomenda músicas com abordagem Baseada em Conteúdo a partir do perfil do usuário construído de informações vindas de mídia social (CELMA; RAMÍREZ; HERRERA, 2005; CELMA; HERRERA; SERRA, 2006).

Em Sistemas de Recomendação voltados para a recomendação de artigos, informações para construção do perfil do usuário são extraídas no sistema a partir de ações anteriores do usuário, assim são utilizados dados do histórico do usuário como por exemplo *feedback* implícito, que inclui cliques e visualizações anteriores do usuário dentro do sistema (RICCI; ROKACH;

SHAPIRA, 2015).

Alguns trabalhos utilizam métodos como solicitar a avaliação dos itens ou dados demográficos do usuário como idade, gênero e ocupação. No trabalho de Son (2016) as técnicas utilizadas em trabalhos que buscam minimizar o problema do Cold Start do novo usuário podem ser divididos em três principais categorias: a) utilizar fontes de dados adicionais, b) escolher os grupos mais proeminentes de usuários e, c) melhorar a previsão usando métodos híbridos. A ideia de utilizar fontes de dados adicionais está relacionada a dados demográficos do usuário ou as opiniões dos usuários extraídas de comentários nas redes sociais. A técnica de escolher os grupos mais proeminentes é relacionada a ideia de clusterizar os usuários por algum atributo, podendo ser inclusive pelos dados demográficos, e ter um conjunto de usuários ativos que representem boa parte dos outros usuários daquele grupo ou cluster. Por fim, as técnicas de utilizar métodos híbridos se refere a trabalhos que utilizam mais de uma técnica de recomendação.

Por outro lado, quando não existe nenhum histórico do usuário, nenhuma interação dentro do sistema, o problema de *Cold Start* também aparece, e é preciso definir como identificar as preferências e gostos do usuário, especialmente quando se trata do primeiro usuário no sistema ou um novo usuário sem depender do histórico de interação de outros. Existem muitos desafios também quanto a explorar como esses itens são mostrados ao usuário, quais caminhos são utilizados para definir quais itens serão avaliados e qual técnica é a mais apropriada. Desta maneira, apresenta-se uma possibilidade de exploração de pesquisa.

Por intermédio de uma pesquisa exploratória, constatou-se que poucos são os trabalhos secundários voltados ao problema de *Cold Start* e suas técnicas para minimizar o problema em Sistemas de Recomendação. Especialmente trabalhos relacionados ao problema de um novo usuário no sistema, chamado de *User Cold Start*, ou *Cold Start* do usuário que indiquem técnicas de construção do perfil inicial deste usuário, sem o uso de interações prévias com o sistema. Os trabalhos apresentados no início deste capítulo indicam que o problema é um tópico para pesquisas futuras e alguns destacam que é um problema ainda não resolvido. Desta maneira, com o intuito de verificar como *Cold Start* do usuário é minimizado, realizou-se um Mapeamento Sistemático da Literatura (MSL) para identificar técnicas e soluções para a construção inicial do perfil, detalhado no Capítulo 3.

2.3 CONSIDERAÇÕES DO CAPÍTULO

Neste capítulo foram apresentados os principais fundamentos das áreas relacionadas a abordagem desenvolvida nesta dissertação, sendo elas: Sistemas de Recomendação, suas principais abordagens, métricas de avaliação e considerações sobre técnicas para minimizar o *Cold Start*. A partir da construção do perfil inicial do usuário, o problema de *Cold Start* é minimizado, este problema ocorre em mais de uma das abordagens de recomendação e é um problema bastante conhecido e ainda não solucionado. É nesta direção que a abordagem desta dissertação é fundamentada e implementada.

3 MAPEAMENTO SISTEMÁTICO

Para identificar formas de minimizar o *Cold Start* de novos usuários de Sistemas de Recomendação (*new user problem/ cold start user*) foi realizado um Mapeamento Sistemático da Literatura (MSL) que será detalhado neste capítulo. O mapeamento conduzido foi fundamentado baseado no processo de Petersen et al. (2008).

O Mapeamento Sistemático da Literatura tem o intuito de identificar formas de minimizar o *Cold Start* de novos usuário de Sistemas de Recomendação disponíveis na literatura além disso, obter uma visão geral da área por meio da classificação e quantificação de trabalhos existentes na literatura.

3.1 QUESTÕES DE PESQUISA

Com a intenção de obter as contribuições disponíveis na literatura para minimizar o *Cold Start*, foi definida a questão principal de pesquisa [QP] e as questões secundárias para fornecer subsídios para entender as especificidades do tópico de pesquisa estudado (CABREJOS; VIANA; SANTOS, 2018).

QP1 : Quais soluções vem sendo utilizadas para reduzir o problema do *Cold Start* do usuário?

QP1.1 : Como o perfil inicial do usuário é modelado nestes trabalhos?

QP1.2 : Quais os domínios de aplicação destes trabalhos?

QP1.3 : Como vem sendo a evolução das soluções para minimizar o *Cold Start* ao longo do tempo?

3.2 STRING DE PESQUISA

Por intermédio de uma *string* de busca, é realizada a busca de artigos. Uma pesquisa prévia foi executada para a extração de palavras-chave. Para responder as questões de pesquisa foram executadas consultas em diferentes bibliotecas digitais. Com base nisso, as palavras-chave que melhor comtemplaram os trabalhos buscados para esse MSL, foram: (1) Sistemas de Recomendação e (2) *Cold-Start*. Foram testadas várias versões da *string*. Deste modo, na realização do MSL, foram obtidos os artigos que coicidem com a consulta:

ALL(recommender systems or recommendations systems) AND ALL(user cold start) AND ABS(cold start)

Esta consulta procura identificar trabalhos que envolvam Sistemas de Recomendação considerando os termos “*recommender systems*” or “*recommendation systems*” em qualquer lugar do artigo (“*ALL*”). Pelas avaliações realizadas estes dois termos são os mais utilizados em artigos que tratem de sistemas de recomendação. Durante os testes de string foi observado que

trabalhos que minimizavam o problema de *Cold-Start* relacionado ao problema do novo usuário eram relacionados a expressão *user cold start*. Ainda foi usado como argumento de consulta do termo “cold start”, termo amplamente consagrado na área de Sistemas de Recomendação, sendo que neste caso foi indicado que este termo deveria estar no resumo/abstract do artigo (“ABS”). Isto foi feito com o objetivo de identificar artigos que tratam do problema do *Cold Start* e não apenas menciona o termo. Neste sentido, considerou-se que artigos que buscam minimizar o *Cold Start* possuem grande probabilidade de mencionar o termo “cold start” no seu abstract.

Para o mapeamento deste trabalho, os Mecanismos de Busca Acadêmica (MBA) selecionados foram a ACM Digital Library, o IEEE e o Scopus. As bases da ACM e IEEE foram selecionadas devido à proximidade com a área da Computação e a Scopus por ser um dos mecanismos que mais indexa documentos, especificamente periódicos (BUCHINGER; CA-VALCANTI; HUNSELL, 2014). O MBA Scopus obteve o maior número de resultados como esperado, não empregaram-se filtros por categoria, e não se restringiu a busca por data. Na Tabela 1 é possível observar o número de artigos retornados por cada MBA e o número total de artigos para a realização do Mapeamento.

Tabela 1 – Quantidade de trabalhos por Mecanismo de Busca

Mecanismo de Busca	Quantidade
ACM	90
Scopus	374
IEEE	44
Total	508

3.3 CRITÉRIOS DE SELEÇÃO E PROCESSO

Sobre o conjunto de artigos obtidos, executou-se a análise das informações nos campos de título, resumo e palavras-chave após todos os metadados serem baixados e organizados em uma planilha. Desta forma, aplicando os critérios de inclusão e exclusão, de modo que os artigos retornados sejam incluídos ou excluídos do processo. Os trabalhos que satisfizeram todos os Critérios de Inclusão (CI), mostrados a seguir, foram aceitos.

CI1 . A publicação está escrita na língua inglesa.

CI2 . A publicação precisa ter mais de 4 páginas.

Durante a análise dos trabalhos, aqueles que atendiam ao menos um dos Critérios de Exclusão citados a seguir foram removidos:

CE1 A publicação não está disponível para leitura de modo aberto via portal CAPES;

CE2 A publicação não descreve técnica para minimizar *Cold Start* do usuário;

CE3 A publicação não é um artigo científico primário;

CE4 A publicação está duplicada, isto é, foi selecionada em outra biblioteca digital;

CE5 A publicação usa histórico do usuário, fonte externa (e.g. redes sociais do usuário), ou dados demográficos do usuário como forma de iniciar o perfil de um novo usuário.

Na análise dos artigos, ao não ser possível concluir a aplicação de determinado critério com base no título, resumo e palavras-chaves, outras partes do texto precisaram ser lidas e avaliadas, como a introdução, conclusão ou então todo o estudo.

Desta forma, dos 508 artigos retornados pela busca, 469 artigos atenderam aos critérios de inclusão de acordo com a Tabela 2. Após remover os artigos que não passaram pelos critérios de inclusão, foram criadas 4 colunas em uma planilha. As colunas correspondiam ao tipo de abordagem utilizada na recomendação, que tipo de itens estavam sendo recomendados, qual a fonte dos dados para criação do perfil do usuário e uma coluna extra de detalhes para uma breve explicação sobre a fonte usada para criar o perfil.

A leitura dos trabalhos foi realizada em três fases, e em cada uma tentando obter as informações das colunas citadas: 1) a leitura do título, resumo, palavras-chave; 2) leitura da introdução e conclusão; 3) leitura completa.

Tabela 2 – Quantidade de trabalhos por CI

Critério de Inclusão	ACM	SCOPUS	IEEE	Total
Total retornado	90	374	44	508
CI 1	90	356	44	490
CI 2	83	345	41	469

Para desconsiderar os artigos que não ajudariam a responder as questões de pesquisa, os artigos passaram pelos critérios de exclusão em cada uma das fases de leitura. Se não fosse possível identificar se o critério de exclusão era atendido e a informação para as colunas não era suficiente após a primeira fase, a segunda fase era realizada e, quando ainda insuficiente, a terceira fase garantia a correta filtragem dos trabalhos. Após o Critério de Exclusão quatro, 324 artigos estavam disponíveis para leitura e descreviam técnicas para minimizar o *Cold-Start* do usuário.

Porém, foi constatado que:

- Vários trabalhos criavam o perfil do usuário através de um histórico de interações (não apenas interações iniciais) do usuário;
- Vários trabalhos usavam dados demográficos;
- Vários trabalhos usavam fontes de dados externas como Redes Sociais de onde os dados eram extraídos.

Como o intuito deste trabalho é entender como criar este perfil antes de qualquer tipo de interação prévia, isto é no momento em que o usuário realiza seu cadastro o CE5 foi aplicado e a quantidade de trabalhos retornados por MBA pode ser vista na Tabela 4. Foram identificados na base da ACM 54 artigos que utilizavam histórico, 9 artigos que utilizavam dados demográficos e 3 artigos que utilizavam alguma fonte externa. Na base da IEEE, 22 artigos utilizavam histórico, 6 artigos que utilizavam dados demográficos e 5 artigos que utilizavam alguma fonte externa. Por fim, na base da SCOPUS, 168 artigos utilizavam histórico, 33 artigos que utilizavam dados demográficos e 5 artigos que utilizavam alguma fonte externa. Após a aplicação do CE5 e a remoção dos artigos duplicados, apenas 19 trabalhos foram selecionados para a leitura completa de acordo com a Tabela 3.

Tabela 3 – Quantidade de trabalhos por CE

Critério de Exclusão	ACM	SCOPUS	IEEE	Total
Total retornado	83	345	41	469
CE 1	83	249	41	373
CE 2	74	231	39	344
CE 3	73	216	39	326
CE 4	73	212	39	324
CE 5	7	6	6	19

Tabela 4 – Quantidade de trabalhos no CE5

	Histórico	Dados Demográficos	Externo	Total
ACM	54	9	3	66
SCOPUS	168	33	5	206
IEEE	22	6	5	33

3.4 ANÁLISES E RESULTADOS

A análise foi realizada entre os meses de Fevereiro e Março de 2023, a partir dos 19 artigos restantes após serem aplicados os critérios de exclusão, foi possível responder à questão de principal de pesquisa, constatando que há propostas visando minizar o problema do *Cold Start*.

Considerando a forma de criar o perfil inicial do usuário a partir da leitura dos trabalhos, é possível identificar as seguintes possibilidades (P):

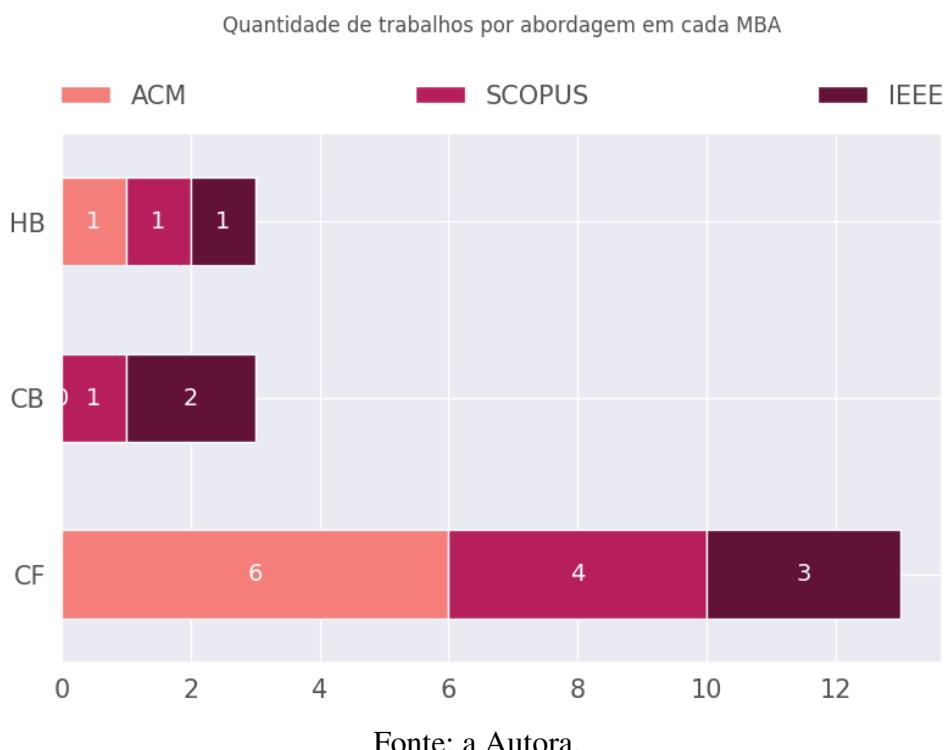
- P1 Solicitar que o usuário responda perguntas/questionário;
- P2 Solicitar a avaliação do usuário para um conjunto de itens,
- P3 Solicitar que o usuário selecione itens;
- P4 Solicitar para o usuário a seleção/cadastro de palavras-chave/tags;

P5 Usar itens populares entre usuários mais antigos (ou usuários considerados especialistas) do Sistema de Recomendação.

A técnica de perguntas/questionário (P1) consiste em realizar alguma(s) pergunta(s) ao usuário assim que este entra no sistema para identificar suas preferências. Quando é utilizada a avaliação, é solicitado ao usuário que avalie alguns itens no seu momento de entrada no sistema (P2) ou então que selecione itens de interesse dentre os apresentados (P3). A técnica de seleção/cadastro de palavras-chave/tags (P4) é quando o sistema solicita que o usuário selecione ou cadastre palavras que melhor descrevam seus interesses, e por fim, a técnica usando itens populares (P5), que são itens que foram melhor avaliados por usuários ativos ou chamados de especialistas.

A Figura 1 apresenta o gráfico com a categorização dos trabalhos de acordo com a abordagem de recomendação. Na base da ACM 6 dos artigos selecionados são Filtragem Colaborativa e 1 de Abordagem Híbrida. Já na base da IEEE, são encontradas as três principais abordagens de recomendação, sendo Filtragem Colaborativa (FC), Baseado em Conteúdo (BC) e Híbrida (HB). Na base da SCOPUS, também são encontradas as três principais abordagens de recomendação.

Figura 1 – Trabalhos por abordagens de recomendação



Fonte: a Autora.

Como dados estatísticos alcançados, têm-se a distribuição de artigos pelos mecanismos de busca utilizados, apresentados também na Figura 1. A maioria dos estudos são provenientes da ACM, cerca de 36,84% cada. Por último a IEEE e a Scopus com 31,57% cada do resultado

total de artigos por MBA. Acredita-se que a porcentagem maior da base ACM deva-se ao fato de ser o mecanismo de busca principal da área pelos eventos científicos na área de Recomendação.

Sobre os anos de publicação, há um crescimento dos trabalhos de 2013 até 2017, e após um decréscimo, temos uma constância de publicações entre 2020 a 2021. Percebe-se que há publicações em todos os anos sobre o assunto e que o problema de *Cold Start* ainda é um importante tópico a ser pesquisado.

A partir dos resultados gerais do Mapeamento foi possível relatar o número de trabalhos selecionados por abordagem de recomendação, por MBA consultado e por anos de publicação. Além disso, foram identificadas possibilidades de criação de perfil inicial nos trabalhos selecionados que serão detalhados na próxima seção.

3.5 TRABALHOS SELECIONADOS

Um estudo mais aprofundado nos 19 artigos voltados à construção do perfil inicial do usuário foi realizado. Com base na leitura destes artigos foram extraídas informações relevantes ao processo de minimizar o *Cold Start*.

No trabalho de Shi et al. (2021) um novo método de representação de conteúdo em texto é apresentado a partir da recomendação de notícias, nomeado WG4Rec. Os autores acreditam que as soluções já conhecidas principalmente de cross-domain podem ser melhoradas a partir das relações explícitas entre textos. Para estabelecer essas relações, os autores constroem um modelo de grafo de palavras utilizando similaridade semântica, co-ocorrência e a relação de cliques (histórico) entre as notícias. Estas relações conectam várias notícias e capturam vários tipos de relações e associações entre as palavras do texto. Para modelar as preferências do usuário e reduzir o *Cold-Start* são utilizadas as relações entre as palavras das notícias, e para isso é utilizado um algoritmo de uma rede neural para obter as interações do usuário com essas notícias. A partir de cada interação do usuário com os itens a rede neural identifica a relação entre as palavras e adiciona como preferência do usuário. A partir das interações do usuário com as notícias e páginas da Web clicadas, são aplicadas as estratégias de relações entre as palavras para obter as top-palavras relacionadas as notícias que o usuário interagiu, sendo mapeadas como seu perfil, e assim mostrando as palavras relacionadas as interações como recomendação ao usuário. Na avaliação de resultados o modelo desenvolvido é comparado a outros algoritmos *baseline* que utilizam outras abordagens de recomendação e obtêm os interesses do usuário de maneira semelhante. Por conta do modelo utilizar outras fontes de interação em outros domínios (cross-domain) é possível lidar com novos usuários. O modelo nos resultados para usuários *Cold Start* se destaca em relação a outros comparados que não utilizam ambos grafo de palavras e interações em outros domínio. Além disso, o trabalho utiliza a NDCG (JäRVELIN; KEKÄLÄINEN, 2002) que é uma métrica de avaliação de desempenho, para 5 e 10 itens recomendados apresentando um número superior no modelo proposto para a qualidade da lista de resultados retornada em relação aos itens que o usuário realmente indicou como relevantes comparado a todos os outros

algoritmos de baseline. O grafo de palavras e as interações entre cenários/domínios lidam melhor e ajudam a encontrar o conteúdo relacionado para usuários em situação de *Cold Start*.

Em Christakopoulou, Radlinski e Hofmann (2016), através de uma abordagem aplicada para recomendação de restaurantes, foi desenvolvido um *framework* para identificar quais perguntas fazer para que o sistema aprenda rapidamente as preferências de um novo usuário.

No momento que o usuário se cadastra no sistema, é utilizado o algoritmo de Thompson Sampling para entender que tipo de perguntas seriam feitas. Os autores analisaram as pesquisas relacionadas à restaurantes em um mecanismo de busca comercial na Web para entender que tipo de perguntas fazer no sistema. Dentre as palavras encontradas nas *queries* encontradas, os termos são relacionados a qual culinária (mexicana, tailandesa, etc), adjetivos (bom, famoso, etc) para descrever as comidas, e o tipo de alimento (peixes, churrasco, sushi, etc) além de combinações de termos de localização dos lugares. O sistema ao receber um novo usuário, seleciona algumas perguntas para fazer. A partir da resposta do usuário os parâmetros do modelo são atualizados, a pergunta realizada é removida e então a lista de recomendação é feita. É importante ressaltar que na fase de fazer as perguntas o sistema precisa fazer as perguntas certas para minimizar o número delas, pois quanto melhor a qualidade dela, melhor a resposta para entender as preferências. Para uma base de 200 usuários tinha-se 200 restaurantes. Após identificar as preferências do usuário, foram selecionados 10 itens baseados no que o usuário gostava associadas como *likes*/curtidas do usuário e 10 itens que não faziam parte das categorias selecionadas associadas como *dislikes*/descurtidas do usuário. Os resultados demonstram que o sistema é capaz de se adaptar as preferências do usuário e compara os algoritmos de seleção de perguntas, indicando qual obteve a melhor performance na questão de qualidade das perguntas para melhor traduzir as preferências. O sistema mostrou resultados utilizando uma métrica de média de precisão (MANNING; RAGHAVAN; SCHÜTZE, 2009; SALTON; MCGILL, 1983) e demonstrou sua melhor performance através de perguntas absolutas, que consistem em um termo absoluto, como “sempre”, “nunca”, “apenas”, “todos” etc.

No trabalho de Felício et al. (2017) é proposta uma abordagem chamada de “PdMS”, um Modelo de Previsão/Predição de Seleção para recomendar filmes aos usuários. A ideia do algoritmo proposto consiste em fazer a melhor seleção de itens para apresentar ao usuário que não possui nenhuma informação prévia quando entra no sistema. Cada vez que o usuário dá o *feedback* de um item recomendado, a próxima recomendação será realizada baseada nesta avaliação. Os algoritmos responsáveis pela classificação do modelo de seleção com o melhor desempenho são os algoritmos de UCB1 e o e-Greedy. Quando um novo usuário entra no sistema ele precisa fazer algumas avaliações de itens, atribuindo notas de 1 a 5 para o item. Através de avaliações similares dos mesmos itens de usuários já existentes no sistema, estes são agrupados em *clusters* de preferências. Assim que novas avaliações são feitas por estes usuários, são comparadas com as avaliações que foram previstas pelo algoritmo. A abordagem PdMS entra neste contexto selecionando quais itens previstos deve recomendar ao usuário baseado nas previsões e nas preferências reais. O objetivo é de recomendar os itens com as avaliações mais

altas. Desta maneira, o PdMS será aplicada aos usuários que são inicialmente novos usuários e permanecem em *Cold Start* até que tenham fornecido uma certa quantidade de avaliações ou *feedback* sobre os itens recomendados. O trabalho utiliza a NDCG como métrica de avaliação de desempenho para 5, 10, 15 e 20 itens recomendados e foram utilizados outros algoritmos como *baseline* de comparação de desempenho. Destacaram-se melhorias positivas nas recomendações para *Cold Start* de usuários feitas utilizando 4 datasets diferentes nos experimentos, houve um aumento na métrica indicando uma qualidade maior, em torno de 10% na lista de itens recomendados.

Já o trabalho Hristakeva et al. (2017) propõe uma estratégia de recomendar artigos para o caso de uso da ferramenta de referências Mendeley. As preferências do usuário podem ser adquiridas pelos métodos de *feedback* implícito, atividades recentes, indicação de disciplina de interesse e palavras-chave de interesses de pesquisa. Quando um usuário novo se cadastra na ferramenta e não possui nenhuma informação prévia, as preferências podem ser mapeadas através da seleção de 1 disciplina disponível em uma lista de 30 opções no momento do cadastro e que melhor represente seu interesse. Recomendações feitas a partir deste método de adquirir a preferência são baseadas em itens que são populares naquela disciplina e o que está em alta naquela disciplina. Outra forma de lidar com o usuário *Cold Start* que o trabalho desenvolve, é através de também no momento do cadastro na ferramenta, o usuário poder especificar um número de palavras-chave para representar seus interesses de pesquisa, e estas palavras são obtidas através da digitação do usuário. A lista de recomendação é criada a partir das palavras de interesse de pesquisa comparadas aos campos de título, resumo, palavras-chave do autor de um conjunto global de artigos populares pré-selecionados dentro de cada disciplina. Partindo dos métodos apresentados para adquirir as preferências do usuário, as recomendações baseadas na seleção da disciplina atingem os melhores valores de acurácia. A métrica de precisão é observada e segue o mesmo comportamento mas muda ao longo da quantidade de itens recomendados, quanto maior o número de itens menor o valor de itens relevantes recomendados.

Em Chao e Guangcai (2020) é desenvolvida uma abordagem chamada Leaders' Advice Model(LAM) para recomendação de filmes. Para minimizar o *Cold Start*, a partir de histórico de avaliações de usuários, o sistema seleciona usuários líderes. Estes usuários líderes são escolhidos baseado na quantidade de avaliações deste usuário. Para cada gênero tem pelo menos um usuário líder. Para cada usuário líder possível, são selecionados 6 gêneros que possuem o maior número de avaliações dos usuários. Quando um novo usuário entra no sistema, ele precisa selecionar um gênero de filmes de sua preferência, a partir do gênero selecionado o conjunto de itens recomendados é o conjunto itens do usuário líder daquele gênero que possuem a maior pontuação do conjunto de dados de avaliação do líder. Desta forma, a partir do gênero respondido do novo usuário, os itens populares avaliados pelos líderes aparecem na lista de recomendados do novo usuário. A avaliação foi realizada utilizando uma métrica para avaliar quantidade de líderes no sistema, e quanto maior o número melhor a performance. Quando comparado com outros modelos, o modelo proposto quando avaliada a performance para usuários *Cold Start*, teve uma

performance melhor.

No trabalho de Lin et al. (2012) é proposto um sistema de recomendação de notícias de abordagem Híbrida. Para minimizar o problema de *Cold Start*, o trabalho recomenda para novos usuários itens que foram selecionados a partir de opiniões implícitas de experts, ou usuários especialistas. O algoritmo assume a partir de uma rede de usuários em uma comunidade de leitura de notícias, que existe uma relação entre os interesses compartilhados entre os usuários experts e o novo usuário, a partir disso, o sistema recomenda os itens melhor avaliados dos experts para o interesse em comum com o novo usuário. Para avaliar o método proposto, compararam a métrica de RMSE do trabalho apresentado contra o algoritmo de LDA e o de Filtragem Colaborativa. Quando avaliado na recomendação de itens para novos usuários, o algoritmo proposto supera em todos os datasets de teste os algoritmos de comparação da métrica utilizada. Com isso, conclui-se que o trabalho conseguiu lidar com o *Cold-Start*.

Em Amatriain et al. (2009) é explorado como avaliadores experts ou especialistas em um determinado domínio podem prever o comportamento da população em geral. A partir da opinião de usuários especialistas, o sistema recomenda filmes para novos usuários que não possuem avaliações de itens ou ainda, pouquíssimas avaliações e o sistema precisa identificar suas preferências. Desta forma, quando um usuário entra no sistema e avalia um item, o algoritmo encontra as avaliações dos especialistas que possuem os mesmos interesses que o novo usuário. Neste trabalho os autores comentam sobre utilizar um modelo de detecção de especialistas e através de crawler web obter as avaliações destes em um determinado domínio. A partir de um algoritmo de clusterização, foram encontrados os usuários similares aos especialistas. A avaliação do modelo proposto foi realizado através da satisfação do usuário via feedback explícito de itens que agradaram ou não o usuário e qual o nível da qualidade da lista de recomendação, variando de muito ruim a muito boa. Em comparação com outros algoritmos, como por exemplo o novo usuário receber de maneira randômica os itens recomendados, ou apenas o algoritmo de KNN de clusterização, o Experts CF como foi denominado o trabalho, superou os resultados, indicando o melhor resultado de satisfação do usuário com o método proposto.

No trabalho de Crawford (2012) é apresentada um SR para recomendação de filmes. Para minimizar o *Cold Start* a partir da modelagem do perfil do usuário através de escolhas por pares de preferência. A ideia é que seja apresentado ao usuário dois itens e questionar qual item o usuário prefere. Normalmente, quando o usuário entra no sistema, este interage sucessivamente com pares de itens até que o sistema ou usuário julgue suficiente. O primeiro par de itens normalmente é selecionado de maneira randômica, e os próximos baseados nas seleção do usuário. Quando um novo usuário começa a preferir entre os pares de itens, o sistema nessas preferências encontra usuários correspondentes, i.e., que também avaliaram os mesmos itens. Quando os perfis dos usuários semelhantes são identificados, a média destes vetores de preferências é o perfil do novo usuário. Cada item dos pares apresentados para o usuário são na verdade um item que resume um conjunto de itens. Para isso, o trabalho clusteriza os itens e seleciona para mostrar ao usuário o item mais popular para representar o conjunto. Para uma

visualização melhor do item, cada item era representado pelo poster do filme. Baseado nessas interações, o usuário podendo finalizar a qualquer momento essa fase de inicialização de perfil, são realizadas as recomendações.

A avaliação do sistema foi conduzida com o objetivo de comparar a performance da abordagem com mais outros algoritmos de seleção. O algoritmo denominado PWDT teve uma performance melhor que um algoritmo que seleciona itens randomicamente pois o usuário pulou menos perguntas, e desta maneira o sistema consegue entender melhor suas preferências, além da métrica utilizada ser melhor no algoritmo proposto do que no algoritmo baseline.

Em Walunj et al. (2022) é proposta a criação de uma aplicação que permite que o usuário visitando uma cidade descubra locais de interesse e o que deseja encontrar na cidade. O sistema utiliza o modelo proposto para recomendar hotéis, restaurantes e pontos turísticos. Com o objetivo de entender as preferências de um novo usuário e minimizar o *Cold Start* do usuário, o trabalho propõe realizar algumas perguntas pra este novo usuário antes de realizar recomendações. O usuário precisa selecionar que tipo de locais deseja encontrar: hotéis, restaurantes ou atrações turísticas. O sistema então solicita algumas perguntas baseadas no tipo de item que o usuário selecionou. Se o usuário selecionou hotel, é solicitado que digite o país ou cidade, o tipo de viagem que está fazendo ou por que ele está visitando o local. Se escolheu restaurante o usuário precisa inserir o nome do restaurante que deseja visitar e baseado na filtragem colaborativa, o sistema retorna restaurantes similares aquele inserido. Por fim se o usuário selecionou atração turística, o modelo pergunta primeiramente que cidade o usuário esta visitando e que tipo de lugares ele gostaria de visitar. Os top lugares e sugestões de locais são realizados com base na resposta do usuário, após o cálculo de similaridade entre o perfil e os itens, retornando o top 5 locais recomendados. Os autores entendem que dessa maneira conseguiram lidar com o *Cold-Start* puro visto que não foi necessária a utilização de interações previstas com o sistema.

O trabalho de Okada, Tan e Kamioka (2021) é baseado em 5 fatores que englobam gêneros de música, este trabalho propõe uma maneira de lidar com usuários *Cold Start* através da idade e tipo de cérebro, recomendando músicas aos usuários. Quando um novo usuário entra no sistema, é solicitado que responda questionários sobre idade, Quociente de Empatia (EQ) e sobre Quociente de Sistematização Revisado (SQ-R). EQ é um Questionário de 60 itens projetado para medir a empatia do usuário nível. Por outro lado, o SQ-R é um questionário de 75 itens para medir o nível de sistematização do usuário. Os questionários são utilizados para definir o tipo de cérebro do usuário. Baseado no tipo de cérebro são associados os 5 fatores de música e a partir de cada um é possível mapear quais são os gêneros musicais de preferência. Com a preferência mapeada, a lista de recomendação baseada nos gêneros é mostrada ao usuário. Foi utilizado a métrica de RMSE para comparar a eficácia do modelo proposto utilizando idade e tipo de cérebro com modelos que utilizam apenas idade ou tipo de cérebro. A partir disso o resultado obtido foi de que utilizando as duas características, a performance das recomendações é melhor.

No trabalho de Felício et al. (2016) é realizada uma proposta para lidar com o perfil inicial do usuário sem um histórico através da seleção de itens em pares. A proposta foi criada

para recomendar artes ou quadros de pintura e roupas. O trabalho não explicita como chega nos pares a serem apresentados ao usuário, mas a ideia principal é mostrar os itens e entender quais o usuário prefere entre dois itens apresentados pelos pares e então agrupar usuários que selecionaram itens similares em clusters. A partir disso é possível recomendar itens para o usuário baseado nas percepções dos outros usuários deste cluster. Para a avaliação, foi comparado o modelo proposto VP-Rec com algoritmos baseline a partir da métrica utilizada para 5,10,15 e 20 itens recomendados resultando numa performance do modelo proposto superior a todos os outros algoritmos.

Em Zhang et al. (2020), é proposto um modelo que recomenda filmes. A partir dos itens melhor avaliados de usuários já presentes no sistema, o novo usuário recebe suas recomendações a partir dessa comunidade. O sistema identifica os vizinhos possíveis através de uma rede bipartida. A partir da interação do usuário com esses itens recomendados a partir das avaliações dos outros usuários, o usuário é alocado num cluster de preferências de usuários. A interação usuário-item é representada por uma rede bipartida é construída entre usuários e itens, onde cada aresta representa uma interação. A performance do sistema proposto foi avaliado através da métrica de precisão e acurácia. A precisão variou bastante com o aumento do número de recomendações, mas a acurácia aumentou comparada a outras abordagens.

O trabalho de Khan, Smyth e Coyle (2021) tem como objetivo criar um sistema de recomendação de gastronomia buscando ter um sistema personalizado e que explore os diferentes fatores dos alimentos que um usuário procura. Para isso, a proposta deste trabalho é através da criação de tópicos das receitas de comidas, representadas como documentos estruturados, o novo usuário selecionar suas preferências. Para cada conjunto de itens que formam um tópico, foram criadas tags que poderiam ser associadas a cada tópico: tipo de alimento, contexto, sabor e complexidade de preparo. A tag mais frequente em cada tópico era selecionada para ser o label descritivo do tópico. Na primeira interação do novo usuário com o FRS, o sistema mostra um conjunto de 288 itens em ordem de relevância e o usuário deve selecionar pelo menos 20 opções que ele gosta ou consome com frequência. O mesmo é feito para as opções que o usuário não gosta. Para resumir melhor as preferências e o que o usuário não gosta, os itens de gostei (*like*) e desgostei (*dislike*) recebem um peso em relação aos itens que não foram avaliados. O FRS então compara o perfil do usuário, com as receitas do sistema identificadas como os documentos do sistema, recomendando assim receitas similares a seus gostos. Além disso o sistema identifica outros gostos do usuário através da sua interação no sistema, identificando seus vizinhos semelhantes. Para a avaliação do FRS, foram comparadas as métricas de cobertura, acurácia e satisfação do usuário com um algoritmo baseline de abordagem baseada em conteúdo. A cobertura atingiu mais de 90% nas receitas, a acurácia do algoritmo performou melhor que o algoritmo de baseline e a satisfação do usuário foi de mais de 56%.

Em Lazemi e Ebrahimpour-Komleh (2017) para lidar com o problema de *Cold Start* do usuário recomendando itens de entretenimento de humor, o trabalho pede que o usuário realize algumas avaliações nas piadas que mais o agradam, antes de usá-lo com notas de 1 a 5.

Quando um usuário realiza uma avaliação ele é solicitado também que expresse seus sentimentos quanto aquele item através de categorias de emoções, como por exemplo felicidade. Os usuários vizinhos que são os semelhantes ao usuário em questão são encontrados através da categoria de emoção daquele item. O foco principal é na recomendação de itens baseado no perfil do usuário através da expressão de seu humor/sentimento combinado com a filtragem colaborativa tradicional. O trabalho não especifica como mostra os itens para os usuários. A avaliação é feita através da métrica de MAE (WILLMOTT, 1981). O método proposto só usa matriz de emoção para obtenção de similaridade de usuários, devido à alta qualidade da escolha de vizinhos , ressaltando boas escolhas de recomendação para os usuários.

O trabalho de Martins et al. (2013) tem como objetivo apresentar um método que utilize tags para associar os itens de vídeos e músicas. A partir de um conjunto inicial de tags e uma métrica selecionada, as tags mais relevantes deste conjunto são selecionadas para serem mostradas para o usuário. E a partir deste conjunto, gerar as tags concorrentes das tags selecionadas nas próximas iterações. O trabalho apresenta uma maneira de o usuário expressar por tags suas preferências nos itens, a ideia geral é de um conjunto tags ser apresentado ao usuário, e quando o usuário selecionar as que deseja, estas são utilizadas como entrada para o sistema de recomendação enquanto que as tags não selecionadas entram para uma espécie de lista negra para que não sejam recomendadas novamente para este usuário. O processo de iteração de seleção das tags é repetido até que um resultado satisfatório seja obtido, ou seja, até que o usuário decida parar de adicionar tags. Após esse processo espera-se que o usuário realize o feedback dos itens recomendados. Na avaliação de resultados os autores concluem que a partir do feedback realizado a proposta alcançou ganhos de até 39% de precisão em cenários *Cold-Start* comparado a outros métodos.

Procurando resolver o problema de *Cold-Start* do usuário, o trabalho de He, He e Liu (2020) propõe uma solução baseada em tags para recomendar filmes. Este trabalho utiliza tags associadas aos usuários para classificá-los quanto a suas preferências. As tags representam cada item, e são criadas a partir das características de cada item para representá-lo, por exemplo: título. As tags também podem ser criadas pelos próprios usuários que classificaram o item, para melhor descrevê-lo. Quando um novo usuário entra no sistema, ele precisa avaliar alguns itens, então o sistema interpreta esta informação como a tag associada ao item avaliada pelo usuário. Após a fase de seleção de tags, o sistema identifica quais usuários semelhantes que também possuem as mesmas tags que o usuário em questão para poder classificá-los como similares e recomendar itens ao usuário em questão baseado no que seus vizinhos gostaram. A avaliação foi realizada através da métrica de RMSE (WILLMOTT, 1982) para a predição das tags que seriam recomendadas ao novo usuário, resultados demonstram que a acurácia do trabalho com esta nova proposta de gerar tags melhorou significativamente em relação a outros algoritmos.

No trabalho de Darshna (2018) os autores propõem uma abordagem que utiliza tanto a abordagem baseada em conteúdo quanto a filtragem colaborativa para recomendar filmes e músicas. Quando um novo usuário entra no sistema ele possui botões para ambas as opções de

recomendação. As músicas recomendadas são agrupadas e clusterizadas por categorias como por exemplo tempo, acústica, instrumentais entre outros. Quando um usuário já possui uma lista de músicas em seu perfil, as opção de recomendação baseada em conteúdo encontra a similaridade das músicas da playlist do usuário com as músicas dos clusters. A filtragem colaborativa vai comparar a similaridade entre dois usuários que escutam as mesmas musicas. Porém para o caso do usuário *Cold Start*, ou seja, que não possui nenhuma lista de música, as músicas mais populares, ou seja, as que receberam as melhores avaliações de usuários já ativos, serão recomendadas a estes usuários. A avaliação do sistema foi realizada através da demonstração da performance do sistema quanto a quantidade de artistas, a quantidade de musicas, número de clusters e tempo de execução do algoritmo de clusterização conseguia fazer.

Já o trabalho de Velde et al. (2021) tem como objetivo recomendar objetos de aprendizagem para melhorar uma habilidade do aluno. Para isso, a proposta busca estimar por nível de dificuldade o perfil do novo usuário. O trabalho é conduzido baseado nos experimentos. Inicia-se com um tema aprendido pelos estudantes e então o sistema precisa medir o grau de dificuldade neste aprendizado para então classificar o usuário. O grau de dificuldade é medido através de tarefas no sistema que seriam perguntas para os usuários responderem. Os itens melhor avaliados por outros usuários através de notas, são recomendados para este novo usuário baseado na similaridade seu nível de dificuldade. O trabalho avalia a taxa de esquecimento dos alunos, a performance da sessão para aprendizado nesses exercícios e a acuracia do sistema entre o que foi de fato e o previsto pelo sistema. Os valores do que realmente aconteceram para os itens apresentados anteriormente de categoria de avaliação, tiveram uma performance superior ao previsto pelo sistema. Isso significa que o sistema conseguiu realizar seu objetivo com resultados satisfatórios.

Por fim, o trabalho de Fernández et al. (2020) utilizando a abordagem Baseada em Conteúdo, busca identificar os interesses do usuário sem exigir muito esforço do mesmo para indicá-los. Através da estratégia Given-2, 2 filmes aleatórios são mostrados para o novo usuário para identificar seus interesses e adicionados ao seu perfil inicial. A partir disso, é calculada a similaridade do perfil do usuário com outros itens onde usuários semelhantes são agrupados em clusters. O maior foco do trabalho é na evolução do perfil e não no perfil inicial, desta maneira o comportamento esperado é através da avaliação do comportamento do novo usuário com novos itens. A avaliação realizada é através do algoritmo baseline de clusterização comparado ao algoritmo proposto. Os resultados mostram que o algoritmo proposto possui uma performance melhor que a do baseline.

A Tabela 5 mostra todos os trabalhos selecionados com sua respectiva abordagem de recomendação, tipo de item recomendado e qual técnica para a construção do perfil inicial utilizou. Para cada abordagem é utilizada a seguinte nomenclatura “FC” para Filtragem Colaborativa, “BC” para Baseado em Conteúdo “HB” para Híbrida.

A ordem dos trabalhos apresentados na Tabela 5 é a mesma ordem da descrição de cada trabalho apresentado, além disso é também a ordem que os trabalhos foram selecionados ao

Tabela 5 – Comparativo dos Trabalhos Selecionados

Trabalho	Abordagem	Domínio	Técnica de perfil
Shi et al. (2021)	FC	Notícias	seleção/cadastro de palavras-chave/tags
Christakopoulou, Radlinski e Hofmann (2016)	FC	Restaurantes	solicitar que o usuário responda perguntas/questionário
Felício et al. (2017)	FC	Filmes	solicitar a avaliação do usuário para um conjunto de itens
Hristakeva et al. (2017)	FC	Artigos	solicitar que o usuário selecione itens e/ou seleção/cadastro de palavras-chave/tags
Chao e Guangcai (2020)	FC	Filmes	itens populares
Lin et al. (2012)	HB	Notícias	itens populares
Amatriain et al. (2009)	FC	Filmes	itens populares
Crawford (2012)	FC	Filmes	solicitar que o usuário selecione itens
Walunj et al. (2022)	BC	Hotéis, restaurantes e pontos turísticos	solicitar que o usuário responda perguntas/questionário
Okada, Tan e Kamioka (2021)	BC	Músicas	solicitar que o usuário responda perguntas/questionário
Felício et al. (2016)	FC	Quadros de arte e roupas	solicitar que o usuário selecione itens
Zhang et al. (2020)	FC	Filmes	itens populares
Khan, Smyth e Coyle (2021)	HB	Receitas gastronômicas	solicitar que o usuário selecione itens
Lazemi e Ebrahimpour-Komleh (2017)	FC	Piadas	solicitar a avaliação do usuário para um conjunto de itens
Martins et al. (2013)	FC	Vídeos e músicas	solicitar que o usuário selecione itens
He, He e Liu (2020)	FC	Filmes	solicitar a avaliação do usuário para um conjunto de itens
Darshna (2018)	HB	Músicas e filmes	itens populares
Velde et al. (2021)	FC	Objetos de aprendizagem	solicitar a avaliação do usuário para um conjunto de itens
Fernández et al. (2020)	BC	Filmes	solicitar que o usuário selecione itens

longo do processo da aplicação dos critérios do Mapeamento por ordem das bases na planilha de dados.

3.6 AMEAÇAS À VALIDADE

As questões de pesquisa, *string* de buscan e critérios de inclusão e exclusão foram estabelecidos antes da realização do Mapeamento Sistemático da Literatura. Além disso, o

processo foi realizado de forma independente pela autora e, quando a inclusão ou exclusão de um artigo era controversa, a decisão era tomada em conjunto. É importante considerar que é possível que estudos relevantes ainda não tenham sido contemplados. Os MBAs considerados foram os mais importantes para a área. Por fim, as classificações realizadas também podem ser uma ameaça à validade, pois pode ser um dos problemas apontados por Pretorius e Budgen (2008).

3.7 CONSIDERAÇÕES DO CAPÍTULO

O objetivo deste estudo foi identificar na literatura, através de um Mapeamento Sistemático da Literatura, artigos que minimizam o problema do *Cold Start*. Como resultados, foram identificados 19 artigos que apresentam técnicas para a redução do problema de *Cold Start* (**QP**). As técnicas são identificadas por solicitar que o usuário responda perguntas/questionário, solicitar a avaliação do usuário para um conjunto de itens, solicitar que o usuário selecione itens, seleção/cadastro de palavras-chave/tags e por fim técnicas de itens populares (**QP1.1**). Além disso, verificou-se que a maioria dos estudos são para o domínio de filmes, além disso, domínios de gastronomia, objetos de aprendizagem, itens de entretenimento de humor, vídeos, restaurantes, artigos, notícias, hotéis, música e itens de arte (**QP1.2**). Ao longo dos anos de publicação dos artigos selecionados, as técnicas de solicitar a avaliação do usuário para um conjunto de itens e solicitar que o usuário selecione itens estão presente em boa parte dos anos, a partir de 2016 a técnica de solicitar que o usuário responda perguntas/questionário começa a ser vista nas publicações e a partir de 2018 a técnica de itens populares torna-se presente nos artigos. Ao longo dos anos também nota-se uma crescente diversidade das técnicas ao longo dos anos, não apenas uma técnica presente em todos os artigos publicados no mesmo ano (**QP1.3**).

Por conseguinte, com base no mapeamento foi possível realizar uma análise das técnicas voltadas para a construção do perfil inicial do usuário, agrupando e classificando-as. Ao total, foram obtidas 5 técnicas. Boa parte dessas técnicas utilizam ainda avaliações de outros usuários, o novo usuário nem sempre é tratado como o primeiro usuário ao utilizar o sistema. Dessa forma, as técnicas que podem ser aplicadas a um novo usuário considerando que este não precise de avaliações de outros usuários, serão consideradas para a proposta desenvolvida neste trabalho.

Pode-se ainda constatar a quase ausência de trabalhos que lidem com o problema do novo usuário em Sistemas de Recomendação que usam a Filtragem Baseada em Conteúdo. Quando encontrados eles não são aplicados a Bibliotecas Digitais.

A realização do presente mapeamento resultou em uma publicação internacional, em 'Proceedings of the 26th International Conference on Enterprise Information Systems (ICEIS 2024)' (CEZAR et al., 2024).

4 MODELO DE SISTEMA DE RECOMENDAÇÃO DE ARTIGOS PARA MINIMIZAR COLD START

Neste capítulo será apresentado o Modelo de Sistema de Recomendação de Artigos proposto neste trabalho. O modelo procura usar estratégias para minimizar o *Cold Start* a partir da construção do perfil inicial do novo usuário.

Algumas das técnicas apresentadas no capítulo 3, podem ser consideradas para definição estratégias para minimizar o *Cold Start*:

- Solicitar que o usuário responda perguntas/questionário;
- Solicitar a avaliação do usuário para um conjunto de itens,
- Solicitar que o usuário selecione itens;
- Solcitar para o usuário a seleção/cadastro de palavras-chave/tags;
- Usar itens populares entre usuários mais antigos (ou usuários considerados especialistas) do Sistema de Recomendação.

Conforme já salientado no capítulo 3, para a técnica de solicitar a avaliação do usuário, foi considerada similar a técnica de solicitar que o usuário selecione itens, pois foi entendido que o usuário selecionando itens que deseja está realizando uma avaliação.

Essas técnicas não envolvem a interação prévia do usuário no sistema, i.e, um histórico do usuário ou o uso de fontes externas para contrução do perfil (e.g. dados do usuário presentes em redes sociais) para que o sistema construa um perfil inicial do usuário.

Dentre as quatro técnicas, duas são consideradas mais relevantes para a proposta do presente trabalho: solicitar a avaliação do usuário para um conjunto de itens e seleção/cadastro de palavras-chave/tags. Isto pelo fato de que recomendar itens populares exige que outros usuários já tenham interagido com o sistema e a elaboração de perguntas/questionários é algo que possui grande dependência do domínio e exigiria a contrução de questionário específico (e.g. se a biblioteca digital possuir artigos da área de banco de dados o questionário deverá permitir a identificação de interesses específicos dentro da área de banco de dados). A seleção/cadastro de palavras-chave/tags chegou a ser considerada, mas exigiria muito provavelmente o auxílio de especialistas da área ou técnicas mais sofisticadas como apresentado em Khan, Smyth e Coyle (2021) para identificação dos termos que resumiriam o conteúdo da biblioteca digital.

Assim, a opção acabou sendo feita pela técnica que envolve a avaliação de itens por parte do usuário. Algo similar é feito no MovieLens (CHANG; HARPER; TERVEEN, 2015) e alguns dos trabalhos identificados no Mapeamento Sistemático como (LAZEMI; EBRAHIMPOUR-KOMLEH, 2017) (HE; HE; LIU, 2020) e (VELDE et al., 2021). Esta opção envolve questões que são abordadas na apresentação do modelo que é feita na seção a seguir.

4.1 MODELO PROPOSTO

O principal objetivo deste trabalho é utilizar técnicas para minimizar o *Cold Start* a partir da criação do perfil inicial do usuário, fazendo isto sem depender de informações prévias ou externas.

A proposta é apresentar ao usuário um conjunto de itens que permita a construção de um perfil inicial. Questões envolvidas aqui dizem respeito: (1) *Quais itens solicitar que o usuário avalie?* e (2) *Quantos itens solicitar que o usuário avalie?*

Sobre a primeira questão, é importante apresentar um conjunto de itens que seja representativo, isto é diverso. Um dos aspectos envolvidos na qualidade da recomendação é a diversidade dos itens na lista de recomendação. Autores defendem que diversidade pode aumentar o interesse de um usuário, fazendo com que ele descubra novos itens e melhore sua experiência com o sistema (ZHOUA et al., 2010) (ADOMAVICIUS; KWON, 2012) (ABISHEVA; GARCIA; SCHWEITZER, 2016). Assim, o objetivo da diversificação é colocar o usuário frente a novos itens, que tenham um impacto positivo na sua satisfação, mantendo bons níveis de acurácia (LUNARDI, 2021). Isto deve ser considerado na formação do perfil inicial de um usuário, evitando apresentar ao usuário itens muito similares o que geraria um problema de superespecialização já no momento inicial de uso do sistema.

Neste trabalho, será aplicada a diversidade máxima de itens antes da geração da recomendação, na formação do perfil do usuário. Definir como diversificar a lista de itens recomendados para um usuário não é o objetivo principal deste trabalho, que poderia aplicar estratégia similar a utilizada em (CEZAR et al., 2021) e (SOUZA; LICHTNOW; GASPARINI, 2022b). Desta maneira, para gerar um perfil inicial que permita recomendar itens que incluam todos os gostos do usuário, o modelo proposto neste trabalho consiste em apresentar itens diversificados ao usuário, para que este possa selecionar o que gosta. Outro aspecto considerado é o de que é assumido que normalmente os usuário de bibliotecas digitais com artigos científicos terão interesse em trabalhos que estejam relacionados com os temas de pesquisa mais recentes.

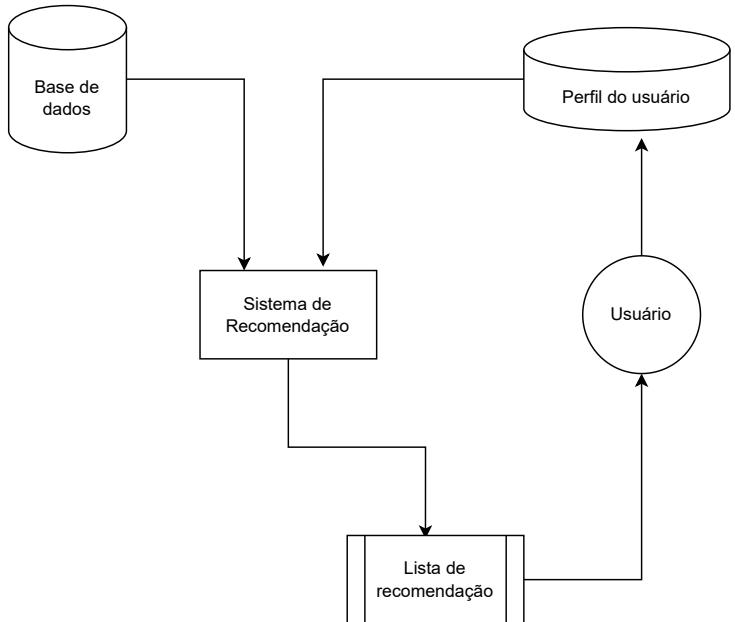
Assim, é proposto que os itens apresentados sejam de artigos diversificados (i.e. cubram diferentes assuntos) dentre os mais recentemente publicados. Esta estratégia reduz o universo de artigos a serem apresentados ao usuário inicialmente e de certa forma contribui com a resposta a questão (2) *Quantos itens solicitar que o usuário avalie?*

4.1.1 Arquitetura do Modelo

A arquitetura básica modelo proposto é ilustrado na Figura 2. Nesta figura é apresentada a base de itens a serem recomendados (artigos) e o perfil do usuário.

Na arquitetura proposta é necessário que a similaridade entre os itens seja previamente armazenada. Assim, através do cálculo de similaridade entre os documentos será gerada uma matriz com valores de similaridade onde consta a similaridade de cada um dos documentos com os demais. O cálculo da similaridade dos documentos pode ser realizado usando o TF-IDF e a

Figura 2 – Arquitetura da proposta



Fonte: a Autora.

similaridade do cosseno, métricas bastante utilizadas na área de Recuperação de Informação.

Na Figura 3 é exemplificada a matriz de similaridade dos documentos. Observa-se que a similaridade do documento 1 com ele mesmo é 1, a similaridade do documento 1 com o documento 2 é de 0,2, a similaridade do documento 1 com o documento 3 é de 0,8, e a similaridade do documento 1 com o documento 4 é 0,5 e assim em diante. A matriz segue da mesma maneira para os outros documentos entre si de acordo com a linha e coluna associada ao número do documento.

Já no perfil do usuário constarão a avaliação feita pelos usuários nos itens, sendo indicados a avaliação feita no momento da construção inicial e o conjunto de palavras-chave que identificam os interesses iniciais do usuário que foram extraídos a partir desta avaliação. O usuário irá avaliar/selecionar os itens apenas no momento do cadastro, por isso no momento da construção inicial. No momento que o usuário selecionar um item, as palavras-chaves e título do item mostrado, serão armazenados no perfil como termos de interesse.

As próximas seções apresentam duas opções individuais, não sendo conjuntas, de criação do perfil inicial para esta proposta. Cada uma é uma possível escolha para o desenvolvimento do modelo a ser definido.

4.1.2 Processo de criação do perfil inicial - Opção 1

O processo de criação do perfil inicial do usuário é ilustrado na Figura 4. Este inicia no momento que o usuário realiza o cadastro. Feito isto, ao usuário serão mostrados itens para avaliar, sendo mostrados sempre pares de itens.

Feito o cadastro inicial, o sistema recupera um conjunto de itens da base. Considera-se do

Figura 3 – Exemplo de Matriz de Similaridade

	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
d1	1	0,2	0,8	0,5	0,32	0,36	0,4	0,72	0,65	0,78
d2		1	0,6	0,7	0,51	0,65	0,7	0,68	0,2	0,76
d3	0,8	0,6	1	0,25	0,31	0,56	0,43	0,7	0,63	0,22
d4	0,5	0,7	0,25	1	0,67	0,3	0,66	0,78	0,6	0,74
d5	0,32	0,51	0,31	0,67	1	0,32	0,3	0,62	0,11	0,7
d6	0,36	0,65	0,56	0,3	0,32	1	0,15	0,82	0,62	0,44
d7	0,4	0,7	0,43	0,66	0,3	0,36	1	0,72	0,65	0,78
d8	0,72	0,68	0,7	0,78	0,62	0,82	0,72	1	0,6	0,3
d9	0,65	0,2	0,63	0,6	0,11	0,62	0,65	0,6	1	0,58
d10	0,78	0,76	0,22	0,74	0,7	0,44	0,78	0,3	0,58	1

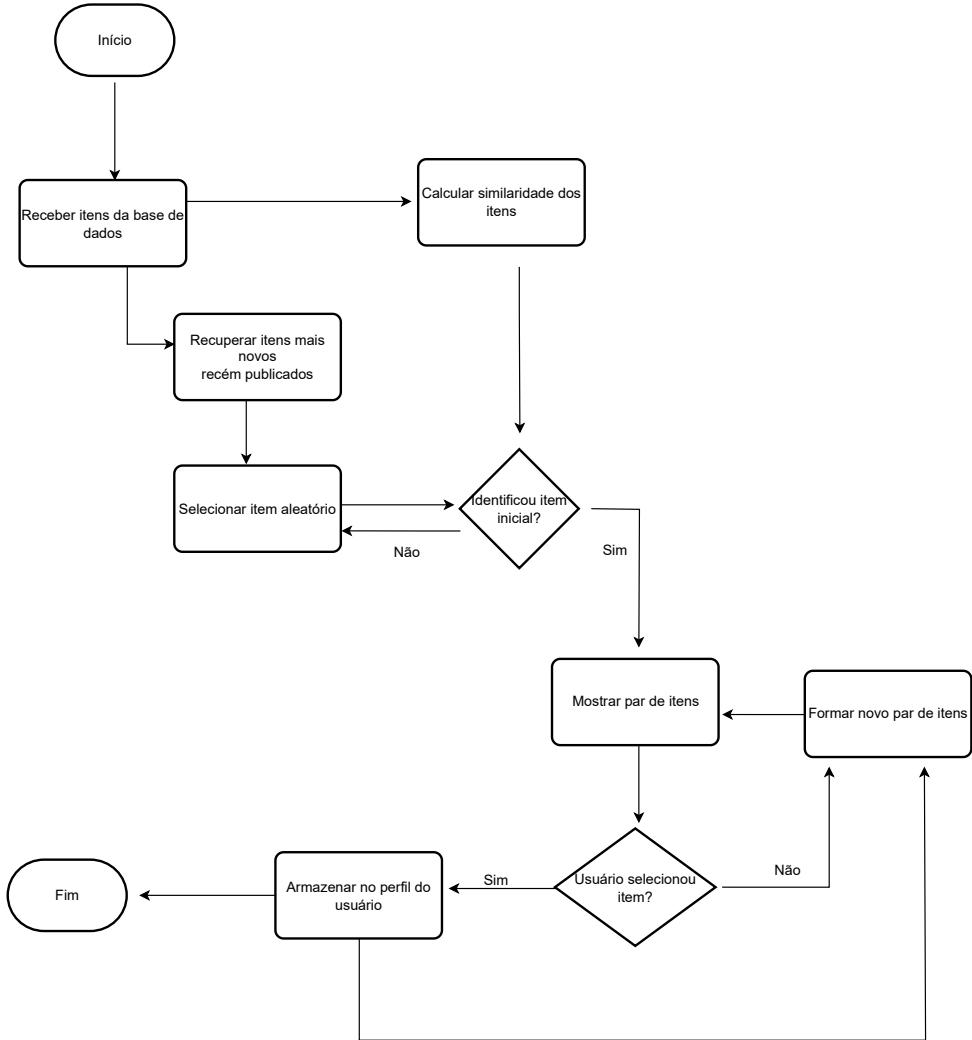
Fonte: a Autora.

conjunto total de itens, uma lista de itens mais novos, ou seja, a partir do maior ano de publicação dentro da base de dados, os itens que foram publicados nesse ano. O motivo da abordagem utilizar esses itens mais novos é devido a fato de que é assumido que o usuário normalmente estará interessado nos temas mais recentes de pesquisa dos eventos científicos.

Para a identificação do item inicial, item que formará o par inicial de escolha no sistema, a proposta no modelo consiste em selecionar de forma aleatória um item dentre os itens da lista de recém publicados. Desta maneira, no momento que o usuário entra para se cadastrar, o sistema identifica um item desta lista de recém-publicados de maneira aleatória. Este é o item inicial e a partir dele será formado o par.

O item que formará o par inicial além do item inicial, será o item menos similar ao item inicial. A identificação do item menos similar será feita usando a matriz de similaridade entre os itens previamente construída. Assim, conforme mostrado na Figura 3, suponha que item

Figura 4 – Criação do perfil inicial - pares



Fonte: a Autora.

inicial seja o documento 1. A matriz de similaridade do documento 1 é uma lista com todos os seus valores de similaridade, e é ordenada. Com o item inicial sendo o documento 1 o item mais similar à ele será o próprio item, neste caso, o documento 1 (1 de grau de similaridade). Conforme pode ser observado na Figura 5, é possível identificar então que o item menos similar ao documento 1 será o documento 2 (0,2 de grau de similaridade). Este será o par inicialmente mostrado ao usuário.

Neste momento o usuário pode selecionar um item dos pares, ou nenhum. Para ambas as opções o próximo par a ser escolhido será a partir do item intermediário da matriz de similaridade do documento 1. Se o usuário selecionar um dos itens, este será armazenado em sua lista de interesses, mais especificamente, o título e as palavras-chaves do documento serão armazenados no perfil do usuário. Conforme no exemplo da Figura 3, tem-se inicialmente uma matriz de 10x10, quando olhamos para a lista de similaridade do documento 1, é possível identificar que o item intermediário, que seria o 5º item dessa lista. Sendo assim, o próximo item a ser considerado para o novo par seria o documento 4 (0,5 de grau de similaridade), que ocupa a 5ª posição na

lista de similaridade do documento 1. Considere que o usuário selecionou o item 2 e este foi armazenado em seu perfil.

Na segunda interação observa-se que a lista de similaridade do documento 4 também será ordenada. O segundo item que formará o segundo par, será o item menos similar ao item alvo, que neste caso é o documento 4. Conforme observado na Figura 6, o documento menos similar ao documento 4, é o documento 3 (0,25 de grau de similaridade). Neste momento o usuário pode selecionar um item dos pares, ou nenhum, novamente. Se selecionar um dos itens, este será armazenado em seu perfil. Senão, o sistema gera um novo par à ele(a). Considere que o usuário não gostou de nenhum um item.

Na terceira interação de seleção de pares, a partir da Figura 6 é possível identificar que o 5º item da lista de similaridade do documento 3 é o documento 7 (0,66 de grau de similaridade), sendo ele o próximo item do novo par. Conforme a Figura 7, o outro item que formará o par com o documento 7 será o menos similar à ele, neste caso, o documento 5 (0,3 de grau de similaridade). A próxima interação de par seria a partir do documento 9, sendo o 5º item da lista de 10 itens de similaridade, neste exemplo.

Figura 5 – Exemplo Interação 1 - pares

		d1 d2 d3 d4 d5 d6 d7 d8 d9 d10										
		d1	1	0,2	0,8	0,5	0,32	0,36	0,4	0,72	0,65	0,78
↓ ORDENAR												
		d1	0,2	0,32	0,36	0,4	0,5	0,65	0,72	0,78	0,8	1

Fonte: a Autora.

Figura 6 – Exemplo Interação 2 - pares

		d1 d2 d3 d4 d5 d6 d7 d8 d9 d10										
		d4	0,5	0,7	0,25	1	0,67	0,3	0,66	0,78	0,6	0,74
↓ ORDENAR												
		d4	0,25	0,3	0,5	0,6	0,66	0,67	0,7	0,74	0,78	1

Fonte: a Autora.

Figura 7 – Exemplo Interação 3 - pares

	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
d4	0,4	0,7	0,43	0,66	0,3	0,36	1	0,72	0,65	0,78
	0,3	0,36	0,4	0,43	0,65	0,66	0,7	0,72	0,78	1

Fonte: a Autora.

É importante considerar os itens já selecionados pelo usuário, para que caso um desses itens escolhidos para formarem o novo par não sejam iguais aos já selecionados. Neste caso, considere na terceira interação que o usuário já tenha selecionado o documento 5. Como o documento 5 já está na lista de selecionados, o algoritmo selecionaria o item mais próximo ao documento 5, sendo neste caso o documento d6 (0,36 de grau de similaridade) sendo o próximo da lista de similaridade de menos similar ao documento 7, conforme Figura 3. Se ainda assim este item já esteja na lista, o algoritmo procuraria pelo próximo item sucessivamente, sendo neste caso o documento 1 (0,4 de grau de similaridade). O mesmo processo serviria para o caso do item intermediário caso já esteja selecionado, o algoritmo procuraria o 6º item da lista e não mais o 5º numa lista de 10 itens. A ideia de mostrar o item mais próximo do item, caso este já tenha sido escolhido, é que assim o algoritmo preserva os níveis de similaridades.

Os itens apresentados nos pares são representados pelo título, palavras-chaves e *abstract* de cada documento. O usuário tem a opção de escolher um dos itens dos pares ou nenhum, independente da escolha, o mesmo processo será executado para formar um novo par a partir do item intermediário do par anterior.

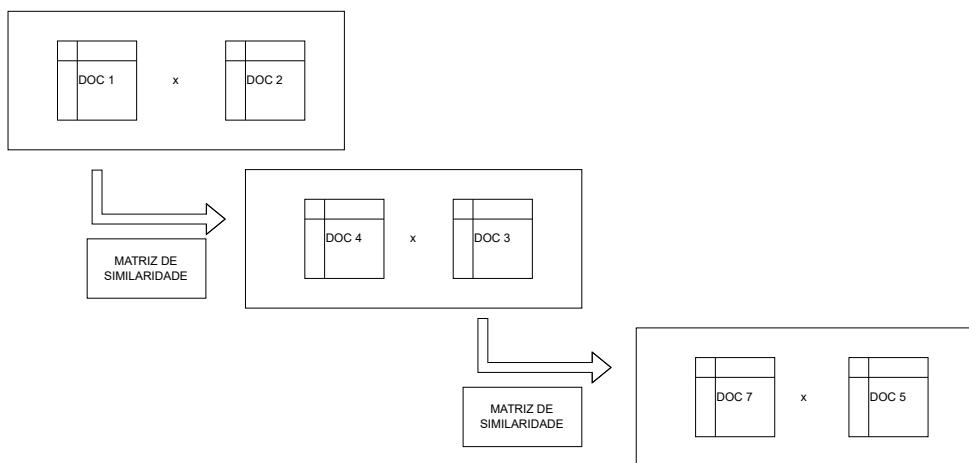
Outro cenário que pode acontecer no momento que o algoritmo seleciona um item é de que exista outro item com o mesmo peso de similaridade. Desta maneira, o sistema escolhe qualquer um dos dois itens de mesmo valor. Por exemplo, se na terceira interação tanto o documento 5 quanto o documento 6 tivessem o mesmo peso de similaridade de 0,3, o algoritmo poderia optar por qualquer um dos dois itens para formar par com o documento 7.

O intuito é que essas interações de pares sejam apresentadas ao usuário até que este selecione pelo menos 5 itens (regra 7 +/- 2 (MILLER, 1994)), sendo apresentado um conjunto de pares ilimitados para o usuário selecionar até que tenha o número mínimo de itens que gosta. Técnicas de comparação em pares são geralmente reconhecidas como uma ferramenta poderosa para apoio a tomada de decisão no processo de expressar preferências (MAIDA; OBWEGESER, 2012). Em Kalloori, Ricci e Gennari (2018), é identificado que em particular, as preferências pareadas ou em pares podem levar para uma maior satisfação do usuário para o procedimento

de seleção de termos de interesse. Além disso, preferências selecionadas em pares podem levar a uma maior precisão e qualidade de SRs em comparação com as avaliações (KALLOORI; RICCI; GENNARI, 2018). Dos trabalhos que utilizam a técnica de selecionar itens descritos no Capítulo 3, dois trabalhos utilizam a comparação via pares para explicitar os interesses do usuário (CRAWFORD, 2012; FELÍCIO et al., 2016). E serão comparados com esta proposta no final deste Capítulo.

Desta maneira, decidiu-se seguir por esta técnica neste presente trabalho. Foi identificado nos trabalhos Maida e Obwegen (2012) e Kalloori, Ricci e Gennari (2018) a preferências em pares que o número de comparações realizadas para obter o interesse do usuário, é em torno de 15 pares. Porém, a seleção de pares é de maneira aleatória nestes trabalhos e não existe um mínimo para escolha de itens. Já no presente trabalho foi definido um mínimo de itens a serem selecionados para permitir a geração da recomendação. Ainda optou-se nesse trabalho por não definir um máximo de interações de pares e que cada próxima seleção de item depende da anterior. O intuito de não limitar as interações de pares do usuário se deve ao fato de que caso o usuário desgoste de vários pares apresentados, e gere cada vez mais pares novos, pode chegar a um ponto em que não selecionou o número mínimo de itens que precisa e a quantidade de interações chegou ao fim. Além disso, a quantidade de interações pode variar com o número da base de itens.

Figura 8 – Simulação de interação de pares



Fonte: a Autora.

Na Figura 10 é possível identificar a formação dos pares conforme explicado anteriormente. A seguir é possível observar como o algoritmo de seleção de pares funciona:

Na linha 4 do algoritmo é possível identificar que é verificado se o item menos similar não está nos itens mostrados ao usuário e nem na lista de documentos já selecionados por este. Caso o item não se enquadre em nenhuma dessas condições, o par de itens é mostrado ao usuário na Linha 6. Na linha 7 do algoritmo, é possível identificar o momento de seleção de um artigo e será armazenado na lista de interesses na Linha 8. É possível identificar na linha 10 do algoritmo novamente verificando se o item não está nos itens mostrados ao usuário e nem na lista de

Algorithm 1 Cálculo de Seleção dos Pares

Require: Matriz de similaridade M , Itens mais recentes J

$\Delta_D \leftarrow$ Documentos selecionados pelo usuário

$\Delta_N \leftarrow$ Documentos mostrados ao usuário

$ItemA \leftarrow$ seleciona item aleatório a partir de J

while item menos similar não está em Δ_D e Δ_N **do**

$ItemB \leftarrow$ Recupera o menos similar de $ItemA$

 Mostra $ItemA$ e $ItemB$ para usuário

if usuário selecionou item **then**

$\Delta_D \leftarrow$ recebe item selecionado

end if ▷ se usuário selecionou ou não item, algoritmo continua da mesma forma

while item intermediário não está em Δ_D e Δ_N **do**

$ItemMed \leftarrow$ Recupera item intermediário

end while

end while

documentos já selecionados por este, mas desta vez é o item intermediário que será utilizado na formação do próximo novo par, independente do usuário ter selecionado algum item ou não na Linha 7. A partir das interações é definido o perfil do usuário. Desta forma, o perfil construído consistirá de um conjunto de palavras obtidos dos artigos que o usuário gostou a partir do título e palavras-chaves.

É possível observar na Figura 2 que após o usuário ter o seu perfil inicial criado, o usuário irá receber então a recomendação e o algoritmo de Cezar (2019) Baseado em Conteúdo, é executado. Para receber a recomendação será necessário que o usuário faça *login* no sistema com suas credenciais cadastradas no mesmo momento que escolheu os termos de interesse.

4.1.3 Processo de criação do perfil inicial - Opção 2

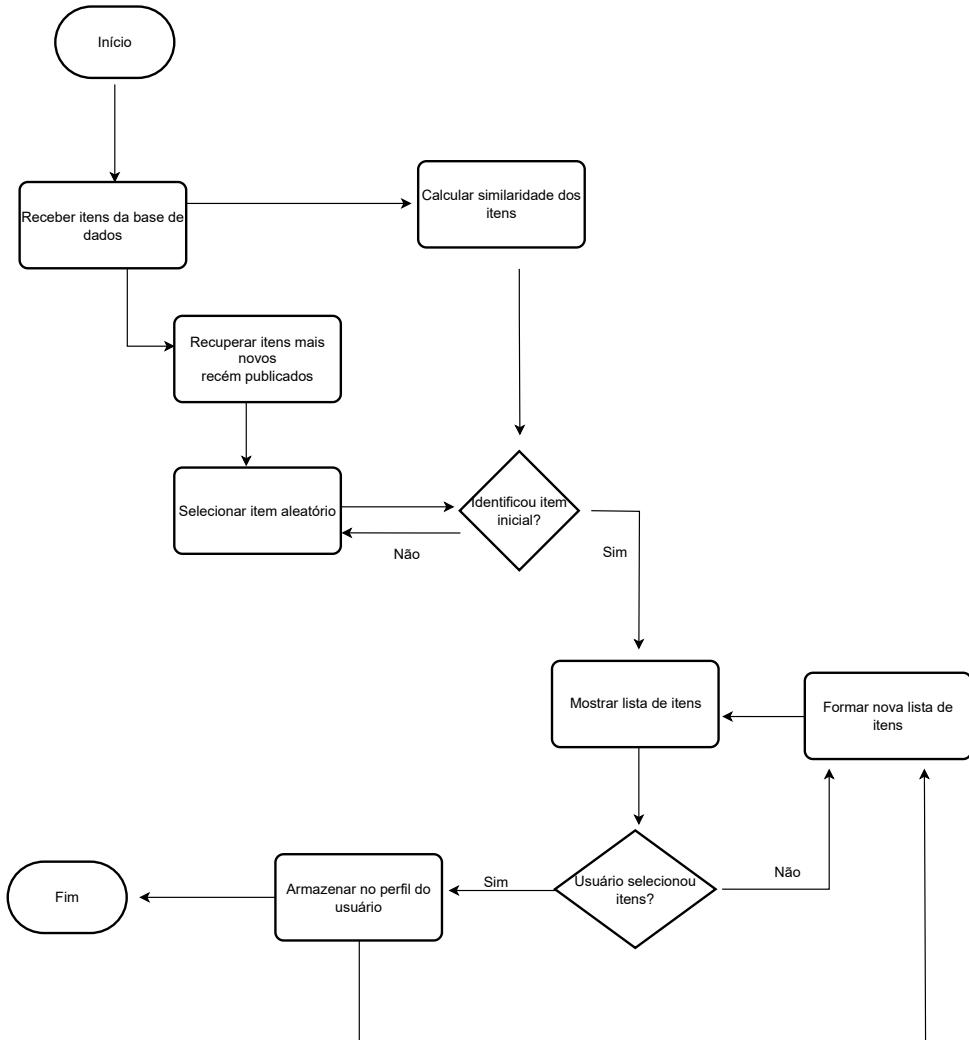
O processo de criação do perfil inicial do usuário é ilustrado na Figura 9. Este inicia no momento que o usuário realiza o cadastro. Feito isto, ao usuário serão mostradas duas listas de 10 itens (MILLER, 1994) cada para selecionar quais identifica serem de seu interesse, sendo 5 itens o mínimo de itens que o usuário precisa selecionar.

Para a identificação do item inicial, que irá gerar o restante dos itens da lista, a proposta no modelo consiste em selecionar de forma aleatória um item dentre os itens recém publicados. Desta maneira, no momento que o usuário entra para se cadastrar, o sistema identifica um item desta lista de recém-publicados de maneira aleatória. Este é o item inicial e a partir dele serão formados os outros itens da lista de interesses.

O algoritmo começa com o item inicial e então identifica o seu item menos similar, ambos os itens mostrados na lista de prováveis itens de interesse.

A identificação do item menos similar será feita usando a matriz de similaridade entre os itens previamente construída. Assim, conforme mostrado na Figura 3, suponha que item inicial seja o documento 1. A matriz de similaridade do documento 1 é uma lista com todos os seus

Figura 9 – Criação do perfil inicial - lista



Fonte: a Autora.

valores de similaridade, e é ordenada. Com o item inicial sendo o documento 1 o item mais similar à ele será o próprio item, neste caso, o documento 1 (valor 1 de grau de similaridade). Conforme pode ser observado na Figura 5, é possível identificar então que o item menos similar ao documento 1 será o documento 2 (0,2 de grau de similaridade). Conforme no exemplo da Figura 3, tem-se inicialmente uma matriz de 10x10, quando olhamos para a lista de similaridade do documento 1, é possível identificar que o item intermediário, ou seja, do meio, seria o 5º item dessa lista. Sendo assim, o próximo item a ser considerado para a lista de interesses seria o documento 4 (0,5 de grau de similaridade), que ocupa a 5ª posição na lista de similaridade do documento 1.

O próximo item que formará a lista de interesses será o item menos similar ao documento 4. Conforme observado na Figura 6, o documento menos similar ao documento 4, é o documento 3 (0,25 de grau de similaridade). Na terceira interação de seleção dos itens da lista de interesse, a partir da Figura 6 é possível identificar que o 5º item da lista de similaridade do documento 3 é o documento 7 (0,66 de grau de similaridade), sendo ele o próximo item. Conforme a Figura 7, o

item menos similar ao documento 7 é o documento 5 (0,3 de grau de similaridade). A próxima interação seria a partir do documento 9, sendo o 5º item da lista de 10 itens de similaridade, neste exemplo.

É importante considerar os itens já selecionados pelo usuário e itens já mostrados ao usuário, para que caso um desses itens escolhidos para estarem na lista de mostrados ao usuário não esteja novamente ou não seja selecionado novamente. Neste caso, considere na terceira interação que o usuário já tenha selecionado o documento 5. Como o documento 5 já está na lista de selecionados, o algoritmo selecionaria o item mais próximo ao documento 5, sendo neste caso o documento 6 (0,36 de grau de similaridade) sendo o próximo da lista de similaridade de menos similar ao documento 7, conforme Figura 3. Se ainda assim este item já esteja na lista, o algoritmo procuraria pelo próximo item sucessivamente, sendo neste caso o documento 1 (0,4 de grau de similaridade). O mesmo processo serviria para o caso do item intermediário caso já esteja selecionado, o algoritmo procuraria o 6º item da lista de similaridade e não mais o 5º considerando a lista de similaridade com 10 itens. A ideia de mostrar o item mais próximo do item, caso este já tenha sido escolhido, é que assim o algoritmo preserva os níveis de similaridades.

Os itens apresentados ao usuário são representados pelo título, palavras-chaves e *abstract* de cada documento. O usuário tem a opção de escolher pelo menos 5 itens da primeira lista contendo 10 no total ou, nenhum, independente da escolha, o mesmo processo será executado para formar uma nova lista a partir de outro item aleatório vindo dos itens recém-publicados. Neste caso, se o usuário não selecionou nenhum item na primeira lista, é requisitado que nesta segunda lista de 10 itens ele escolha o mínimo de 5 itens ou o número restante do que já foi selecionado.

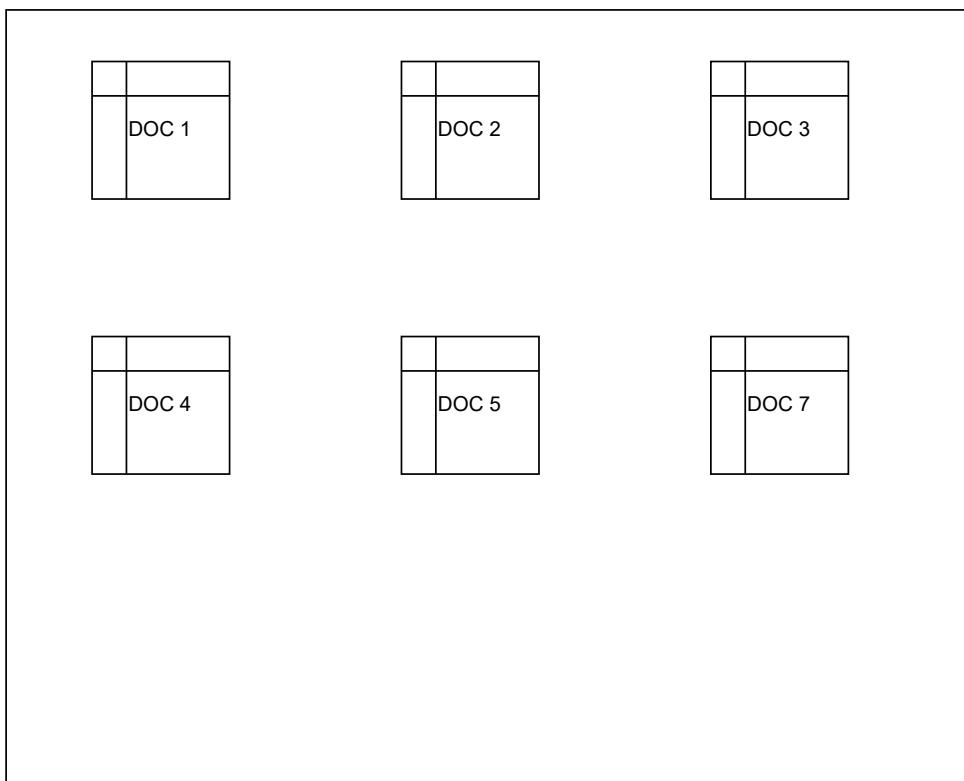
Outro cenário que pode acontecer no momento que o algoritmo seleciona um item é de que exista outro item com o mesmo peso de similaridade. Desta maneira, o sistema escolhe qualquer um dos dois itens de mesmo valor. Por exemplo, se na terceira interação tanto o documento 5 quanto o documento 6 tivessem o mesmo peso de similaridade de 0,3, o algoritmo poderia optar por qualquer um dos dois itens para formar par com o documento 7.

Os itens apresentados ao usuário são representados pelo título, palavras-chaves e *abstract* de cada documento.

Na Figura 10 é possível identificar a formação dos itens na lista de possíveis interesses conforme explicado anteriormente para a avaliação do usuário. A seguir é possível observar como o algoritmo funciona:

Na linha 6 do algoritmo é possível identificar que é verificado se o item menos similar não está nos itens mostrados ao usuário e nem na lista de documentos já selecionados por este. Caso o item não se enquadre em nenhuma dessas condições, o par de itens é adicionado a lista de mostrados ao usuário na Linha 8. É possível identificar na linha 9 do algoritmo novamente verificando se o item não está nos itens mostrados ao usuário e nem na lista de documentos já selecionados por este, mas desta vez é o item intermediário que será utilizado na formação dos

Figura 10 – Simulação da lista de seleção de itens do perfil



Fonte: a Autora.

próximos itens. Na linha 16 do algoritmo, é possível identificar o momento de seleção do(s) artigo(s) e será armazenado na lista de interesses na Linha 17. A partir das interações é definido o perfil do usuário. Desta forma, o perfil contruído consistirá de um conjunto de palavras obtidos dos artigos que o usuário gostou a partir do título e palavras-chaves.

É possível observar na Figura 2 que após o usuário ter o seu perfil inicial criado, o usuário irá receber então a recomendação e o algoritmo de Cezar (2019) Baseado em Conteúdo, é executado. Para receber a recomendação será necessário que o usuário faça login no sistema com suas credenciais cadastradas no mesmo momento que escolheu os termos de interesse.

Algorithm 2 Cálculo da Lista de Seleção dos Itens do perfil

Require: Matriz de similaridade M , Itens mais recentes J

- 1: $\Delta_D \leftarrow$ Documentos selecionados pelo usuário
- 2: $\Delta_N \leftarrow$ Documentos mostrados ao usuário
- 3: $\Delta_L \leftarrow$ Lista de documentos a serem mostrados para o usuário
- 4: Seleciona item aleatório a partir de J
- 5: **for** até 10 itens **do**
- 6: **while** item menos similar não está em Δ_D e Δ_N **do**
- 7: Recupera o menos similar
- 8: Adiciona em L
- 9: **while** item intermediário não está em Δ_D e Δ_N **do**
- 10: $ItemMed \leftarrow$ Recupera item intermediário
- 11: Adiciona L
- 12: **end while**
- 13: **end while**
- 14: **end for**
- 15: Mostra L
- 16: **if** usuário selecionou item **then**
- 17: $\Delta_D \leftarrow$ recebe itens selecionados
- 18: **end if**

4.2 COMPARAÇÃO DOS TRABALHOS SELECIONADOS

Comparando a proposta apresentada com os trabalhos apresentados no Capítulo 3, as técnicas de criação de perfil que foram consideradas para este trabalho foram as de solicitar a avaliação ou seleção do usuário para um conjunto de itens. Este trabalho apresenta uma proposta de seleção de itens, onde o primeiro item é randômico. A partir disso, sempre terá um item menos similar ao item alvo conforme descrito nas duas seções anteriores.

Os trabalhos que solicitam avaliação do usuário Felício et al. (2017), Lazemi e Ebrahimpour-Komleh (2017), He, He e Liu (2020), Velde et al. (2021), a partir de um conjunto de itens pré-selecionado o usuário normalmente deve avaliar os itens através de notas de 1 a 5. Detalhes sobre como estes itens são selecionados não são apresentados, mas normalmente os itens são selecionados de forma aleatória. Já Crawford (2012), Felício et al. (2016) utilizam a formação do primeiro par apresentados ao usuário para selecionar preferências normalmente de maneira aleatória e depois a partir de usuários que avaliaram o mesmo item, os próximos pares também avaliados por esses semelhantes são mostrados ao usuário. Estes trabalhos acabam fazendo uso de dados históricos de outros usuários, dados estes nem sempre disponíveis, algo que na abordagem proposta não é necessário. Por fim Fernández et al. (2020) utiliza uma técnica de escolher 2 itens aleatórios para o usuário selecionar e então serem adicionados ao seu perfil.

Outro ponto a ser destacado é que dos trabalhos citados com suas técnicas, apenas um trabalho solicita que o usuário selecione/avalie itens utilizando a abordagem Baseada em Conteúdo (FERNÁNDEZ et al., 2020), enquanto que os outros trabalhos que utilizam a abordagem Baseada em Conteúdo utilizam a técnica de perguntas/questionário (WALUNJ et al., 2022; OKADA;

TAN; KAMIOKA, 2021).

A partir disso, é possível observar as diferenças que o modelo proposto neste presente trabalho possui em relação aos que foi identificado a partir do mapeamento sistemático e da identificação de trabalhos relacionados:

1. A proposta não depende de itens pré-selecionados por algum critério externo (itens populares entre outros usuários, por exemplo);
2. Apenas um item é escolhido de forma aleatória, sendo os próximos apresentados para seleção de acordo com sua similaridade em relação ao item inicial ou a seleções feitas pelos usuários;
3. A proposta utiliza a abordagem Baseada em Conteúdo, onde técnicas para minimizar o *Cold Start* vem sendo pouco exploradas;
4. A proposta procura enfatizar questões de diversidade já na formação do perfil inicial.

4.3 CONSIDERAÇÕES DO CAPÍTULO

Neste capítulo foi apresentada a proposta para minimizar o *Cold-Start* em um Sistema de Recomendação de abordagem Baseada em Conteúdo, incluindo o algoritmo que foi implementado (ver Seção 5) e fluxo do sistema. O próximo capítulo descreve a implementação do modelo de criação do perfil do usuário proposto a partir do processo de criação do perfil inicial em pares.

5 SISTEMA DE RECOMENDAÇÃO DE ARTIGOS COM ESTRATÉGIA PARA REDUÇÃO DO COLD START USER

Este capítulo descreve a criação de um Sistema de Recomendação para avaliar a técnica de criação do perfil inicial do usuário descrito no capítulo 4. Inicialmente são descritas as tecnologias utilizadas na criação do sistema, após é feita a descrição da base de itens (artigos) que foram utilizados nos testes. São ainda descritos estudos sobre a definição de algumas das interfaces do sistema junto a usuários, isto foi feito especialmente em relação as interfaces relacionadas a criação do perfil inicial do usuário. E por fim, são apresentados a descrição dos experimentos realizados com participantes e os resultados da avaliação.

5.1 TECNOLOGIAS UTILIZADAS

A abordagem do sistema proposto escolhida é a baseada em Conteúdo. Para realizar este tipo de recomendação, como explicado no Capítulo 2, são utilizadas técnicas de Recuperação da Informação. O sistema foi implementado usando Python como principal linguagem de programação e o framework Web Flask, e a ferramenta ElasticSearch.

O *ElasticSearch* foi utilizado para indexar os artigos e recuperar os artigos da base a partir da similaridade com o perfil do usuário. O *ElasticSearch* é um mecanismo de pesquisa de código aberto construído baseado no Apache Lucene, que é uma biblioteca de mecanismo de pesquisa de texto. O Elasticsearch é escrito em Java e usa o Lucene internamente para toda a sua indexação e pesquisa. É orientado a documentos, isso significa que ele armazena objetos ou documentos inteiros. Ele não apenas os armazena, mas também indexa o conteúdo de cada documento para torná-los mais fáceis de pesquisar (GORMLEY; TONG, 2015).

O *ElasticSearch* permite interações na interface cliente através de uma camada *HTTP Rest*. O *ElasticSearch* disponibiliza diversas *API's* para indexação, administração, buscas e monitoramento de dados através dos comandos GET, PUT, POST e/ou DELETE. Uma das APIs mais utilizadas é a *API _search*. Essa API é utilizada para buscar os dados que estão dentro dos índices trazendo informações adicionais como a relevância dos resultados.

O *ElasticSearch* usa um índice invertido, que é designado para permitir pesquisas rápidas de textos completos bastante utilizado na área de Recuperação de Informação. Segundo Gormley e Tong (2015), um índice invertido consiste em uma lista de palavras-chave que aparecem em qualquer documento, e para cada um desses termos, uma lista de documentos, nos quais o termo aparece.

No retorno de uma consulta do *ElasticSearch*, os documentos vem ordenados pela sua similaridade em relação a um consulta, pois a ferramenta possui um algoritmo *default* de similaridade para calcular o *score*, a pontuação de relevância destes documentos (GORMLEY; TONG, 2015).

Existem três fatores principais para o *score* de um documento em uma determinada busca (PEIXOTO, 2019) (GORMLEY; TONG, 2015):

- *Term – Frequency*: Quanto mais um termo aparece em um único *field*, mais relevante ele é;
- *IDF*: Quanto mais documentos possuem o termo pesquisado, menos relevante ele é;
- *Field Length*: *Fields* menores, com menor volume de texto, são naturalmente mais relevantes do que campos maiores, pois quanto maior for o texto, menor é o seu peso.

O *ElasticSearch* utiliza uma variação da medida de similaridade *BM-25* para calcular o score (PEIXOTO, 2019; ROBERTSON; ZARAGOZA et al., 2009), baseado nos principais fatores citados anteriormente.

$$\sum_i^n IDF(q_i) \frac{f(q_i, D) \times (k_1 + 1)}{f(q_i, D) + k_1 \times (1 - b + b \times \frac{fieldLen}{avgFieldLen})} \quad (13)$$

Segundo (CO, 2018) (TURNBULL, 2015), na fórmula 13: q_i é o i-ésimo termo da *query*; $IDF(q_i)$ a frequência inversa do documento do i-ésimo termo da *query*; O tamanho do documento é dividido pelo comprimento médio do documento como ($fieldLen / avgFieldLen$), essa divisão representa o quanto um documento é relativo ao comprimento médio do documento. A constante b nos permite ajustar com quanta influência o valor de ($fieldLen / avgFieldLen$) tem no *score*. O valor de b é fixo sendo igual a 0.75 no *ElasticSearch*. k_1 determina características da saturação da frequência do termo. Ou seja, limita o quanto um único termo de consulta pode afetar a pontuação de um determinado documento. E, $f(q_i, D)$, significa “quantas vezes o i-ésimo termo da *query* ocorre no documento D ? ”, o *TF*. O raciocínio para esta variável é que, quanto mais vezes os termos da consulta ocorrerem em um documento, maior será sua pontuação.

No Capítulo 2, foram apresentadas as fórmulas de *TF – IDF* em Sistemas de Recomendação que utilizam a abordagem Baseada em Conteúdo. No *ElasticSearch* conforme a fórmula do *IDF*, que mede com que frequência um termo ocorre em todos os documentos, é calculado:

$$\ln \left(1 + \frac{docCount - f(q_i) + 0.5}{f(q_i) + 0.5} \right) \quad (14)$$

Na fórmula 14 temos que: *docCount* é o número total de documentos que possuem o termo no *field*; $f(q_i)$ é o número de documentos que contém o i-ésimo termo da *query*. Desta maneira o *ElasticSearch* realiza a relevância dos artigos.

5.2 BASE DE ITENS/ARTIGOS A SEREM RECOMENDADOS

Os itens a serem recomendados foram artigos publicados em eventos relacionados a área de Interação Humano-Computador publicados em eventos da SBC. Existia já uma base de dados modelada e estruturada com 466 artigos dos anais do Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC) desde 1998 a partir do trabalho de (CEZAR,

2019). Na base de dados estão armazenados os metadados sobre cada artigo como título, autores, ano de publicação, palavras-chave e resumo. Para aumentar a base de dados, foram adicionados periódicos e eventos relacionados a área de Interação Humano-Computador que são todos citados pela Comissão Especial de Interação Humano-Computador (IHC) da SBC (Sociedade Brasileira de Computação). A partir desses eventos, os artigos publicados nos anais foram extraídos e armazenados em uma base também. Foram extraídas as informações de títulos, palavras-chaves, autores e ano de publicação de cada artigo em formato de arquivo CSV, tanto da base já criada anteriormente quanto dos novos artigos a serem adicionados. Com o auxílio da biblioteca Pandas do Python, foram concatenados os dados dos artigos em um único arquivo CSV, para a então utilização dentro do sistema que foi implementado.

Desta forma, existem em torno de 7.000 artigos obtidos a partir de 682 eventos científicos apresentados no Apêndice A, onde cada um pode ter uma ou mais edições por ano na base de dados. Cabe ressaltar a ausência de artigos publicados recentemente, isto é, artigos publicados até o ano de 2024. Os artigos da base foram os publicados até 2020.

5.3 DEFINIÇÕES SOBRE INTERFACE/INTERAÇÃO DOS USUÁRIOS

A partir da proposta apresentada no Capítulo 4, duas opções de seleção de itens do perfil do usuário via interface foram consideradas para serem implementadas no sistema criado, devendo ser feita a escolha entre elas. As duas opções são: (1) apresentar pares de itens e (2) apresentar 10 itens simultaneamente.

Para embasar esta escolha, foram executados experimentos com alunos de graduação em Ciência da Computação e especialistas da área de Interação Humano-Computador para a definição da melhor interface, e do número ideal de itens de seleção. Nestes experimentos, foram criadas na ferramenta Figma uma simulação das interfaces de seleção de perfil no sistema de recomendação. O objetivo foi identificar se 10 itens sendo apresentados na lista seriam adequados ou se isto geraria uma sobrecarga no usuário. Por outro lado, também foi experimentado se a apresentação em pares até que o usuário selecione pelo menos 5 itens seria o mais adequado.

Em ambas as alternativas, na primeira tela era possível visualizar uma breve descrição do trabalho, explicando o objetivo da escolha de interface. A interface com 10 alternativas foi apresentada com itens em duas colunas e o usuário conseguia visualizar todos os itens ao descer a tela para baixo, podendo selecionar itens ou não. Quando acessava esta interface o usuário conseguia visualizar os primeiros dois itens da lista (ver Figura 11), porém ao passar a tela para baixo conseguia ver a lista de itens completa (Figura 12). Já na interface onde apenas 2 itens eram apresentados o usuário conseguia interagir visualizando um par por vez, selecionando um, dois ou nenhum item (ver Figura 13). Em ambas as interfaces, quando o usuário selecionava um item, este mudava de cor, identificando a seleção do item.

O experimento foi conduzido através de uma reunião online com um total de 6 participantes, os quais eram Graduandos do curso de Ciência da Computação e/ou alunos que realizaram

a disciplina de IHC na UDESC, e após a demonstração das interfaces, os usuários foram convidados a responder um questionário sobre qual interface preferiram, além de sugestões. Este questionário pode ser visualizado no Apêndice B. No experimento, 83,3% dos *testers* utilizaram o computador para participar da reunião e avaliar a interface, os outros 16,7% utilizaram um *smartphone*. Independentemente do dispositivo utilizado, todos preferiram a interface que apresentava dois itens por vez pela interface em pares, sendo a interface selecionada para ser criada no sistema proposto. As características e especificidades que os fizeram escolher esta interface foram “devido a organização, e a disposição ficar mais clara.”, “Mesmo que esse modelo seja mais lento que o outro, dá a impressão de que eu precisei pensar menos para efetuar as escolhas, acho é uma perca justa, perde no tempo e mas ganha na "facilidade" de uso.”, “Gostei mais da interface em pares, pois a pessoa pode escolher entre apenas 2 opções e entendo que a carga de trabalho dele fique menor”.

Figura 11 – Experimentos de interface em lista



Fonte: a Autora.

Figura 12 – Interface com 10 itens



Fonte: a Autora.

Figura 13 – Interface com 2 itens

Selecione o item de sua preferência.
Para isto clique sobre o item desejado.



Fonte: a Autora

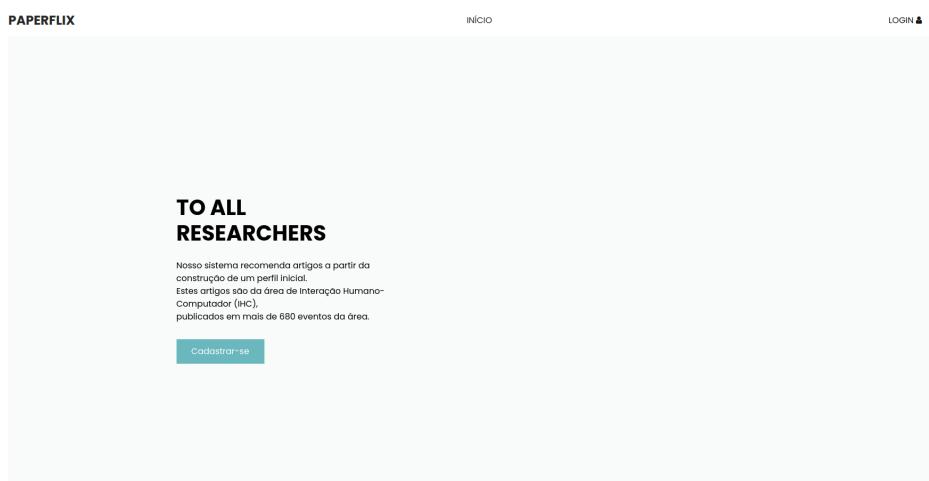
5.4 SISTEMA IMPLEMENTADO

A partir do experimento de interfaces (Seção 5.3) no qual a interface em pares foi escolhida, deu-se início a fase de implementação do sistema com as tecnologias principais utilizadas sendo Python e Flask, juntamente com o Elasticsearch (ver Seção 5.1).

O fluxo de telas do sistema pode ser visualizado na Figura 15, em que ao acessar com a tela inicial, o usuário se cadastra e em seguida é redirecionado para o *login*. Após, o usuário visualiza a tela de instruções do sistema, e então inicia-se a etapa de construção do perfil inicial através dos pares de itens. Caso o usuário já tenha selecionado no mínimo 5 itens, ele tem a opção de seguir para a etapa de ver suas recomendações geradas com base no seu perfil. Após avaliar suas recomendações através do feedback explícito, o usuário pode visualizar no final da lista de recomendações o botão de Finalizar o Experimento, onde é redirecionado para um questionário, e então finaliza o fluxo de telas do sistema.

Seguindo o passo a passo de fluxo do sistema apresentado, quando o usuário acessa o *link* para entrar no sistema, a tela inicial do sistema pode ser visualizada (Figura 14). Nesta tela inicial, são apresentados os botões de Cadastrar-se ao lado esquerdo da tela. Este botão redireciona o usuário para a tela onde realiza o seu cadastro no sistema. Na tela inicial também é possível visualizar o botão de *Login*, no lado superior direito.

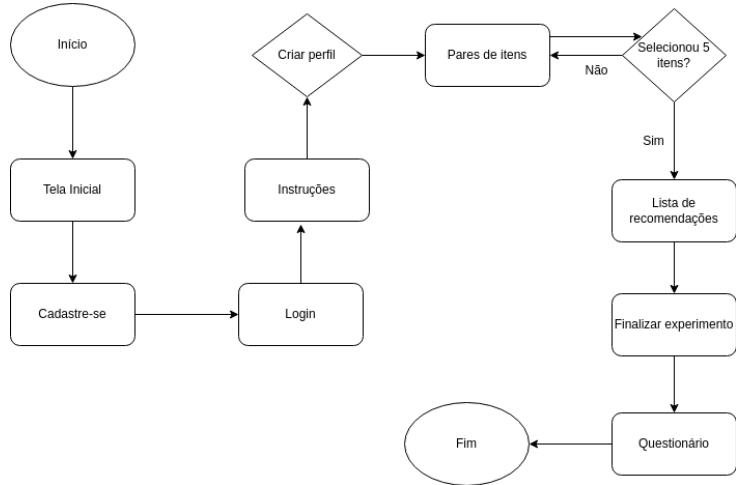
Figura 14 – Tela Inicial



Fonte: a Autora.

Quando o usuário clica em Iniciar (Figura 18), ele será redirecionado a tela de seleção de itens para o perfil. Conforme definido nos experimentos descritos na Seção 5.3, a interface selecionada foi a de pares. Desta maneira, o usuário visualizará pares de itens em que poderá optar por selecionar, um, dois ou nenhum item do par. O primeiro item do primeiro par é escolhido de maneira aleatória dentre os itens mais recentemente publicados, neste caso, últimos dois anos mais recentes da base de dados. A escolha de utilizar os itens dos dois últimos anos foi devido a quantidade de itens do ano mais recente ser considerada muito pequena, o que acarretou em repetir para usuários distintos (em testes feitos pela Autora) o mesmo item inicial. O segundo

Figura 15 – Fluxo de telas do sistema



Fonte: a Autora.

Figura 16 – Tela de Cadastro

Cadastro de Usuário

Após o cadastro, você será redirecionado para o login!

Fonte: a Autora.

item do primeiro par segue a lógica do algoritmo da proposta apresentado na Seção 4.1, que consistem em apresentar o item menos similar ao que já está sendo apresentado ao usuário. Cabe ainda destacar que a opção por selecionar um item/artigo do ano mais recente é fruto de que se assume que artigos mais recentes tratam de temas que são atualmente mais pesquisados e assim, tem talvez, maior chance de despertar o interesse de um usuário.

Quando o usuário deseja selecionar o item, basta clicar sob o *card* do item, e este será selecionado, identificando na tela com um símbolo de verificado em azul, conforme Figura 20.

A partir do momento que o usuário selecionou ao menos 5 itens, ele visualizará um botão de "Recomendar" (ver Figura 21). O usuário pode optar por seguir selecionando mais itens, ou ao clicar neste botão, receberá recomendações de artigos, com base no seu perfil criado a partir da seleção de itens.

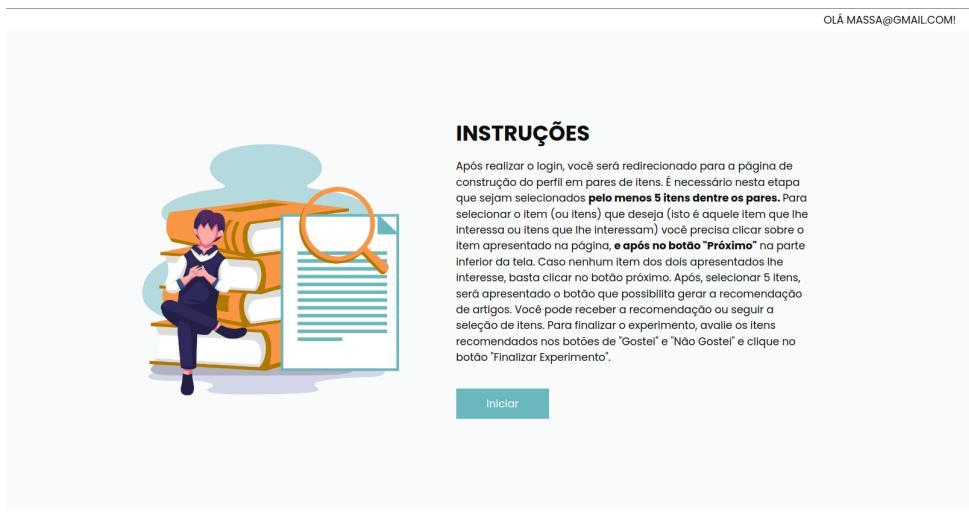
Figura 17 – Tela de Login

Por favor, insira seu email e senha:

Login/Email	Password	Login
-------------	----------	-------

Fonte: a Autora.

Figura 18 – Tela de Instruções

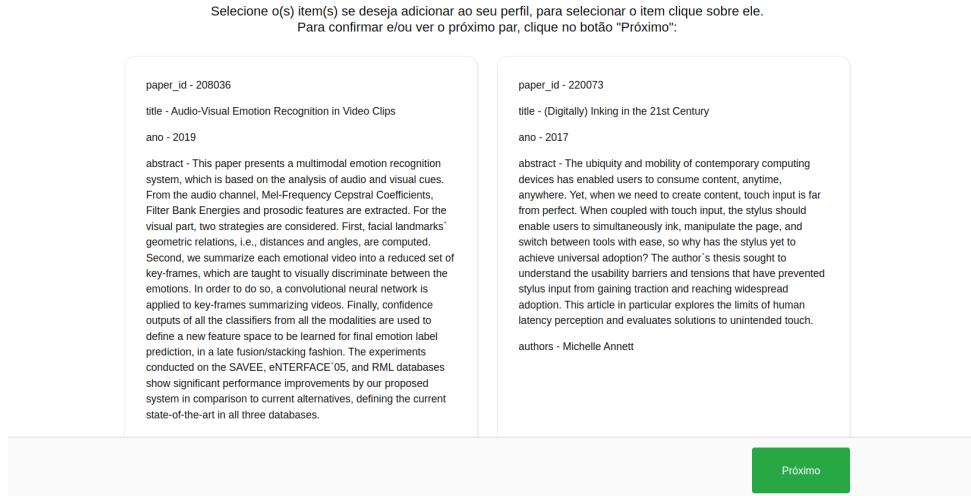


Fonte: a Autora.

Cada item que o usuário seleciona, estará sendo adicionado ao seu perfil o resumo ou *abstract* do artigo selecionado. Desta maneira, cada *abstract* adicionado ao seu perfil, será concatenado em um texto único completo com todas as palavras de todos os *abstracts* adicionados ao perfil do usuário anteriormente, e este texto é enviado para o ElasticSearch como uma *query*.

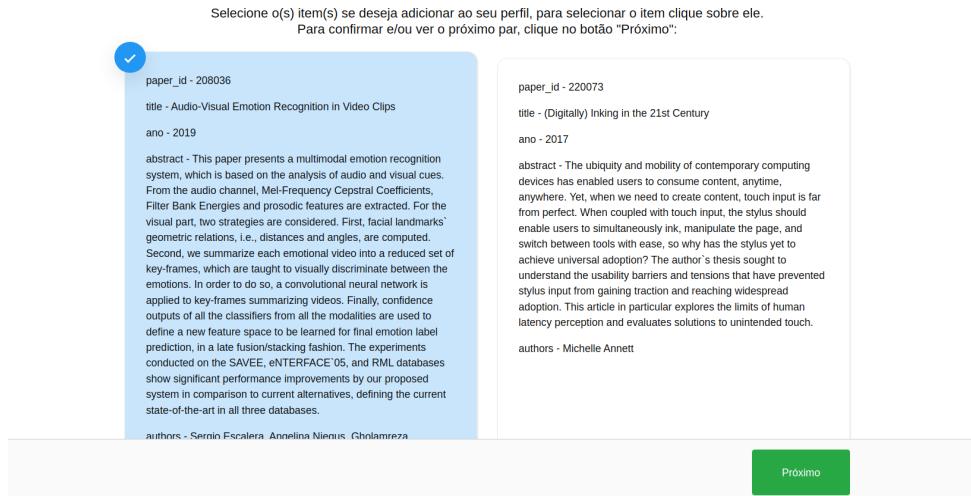
Por exemplo, ao se deparar com o par apresentado na Figura 20, o usuário selecionou o item da esquerda, de título “Audio-Visual Emotion Recognition in Video Clips”. Desta maneira em seu perfil, será adicionado o conteúdo do *abstract* após a remoção de *stop words* em seu perfil

Figura 19 – Tela de seleção de itens em pares - item não selecionado



Fonte: a Autora.

Figura 20 – Tela de seleção de itens em pares - item selecionado



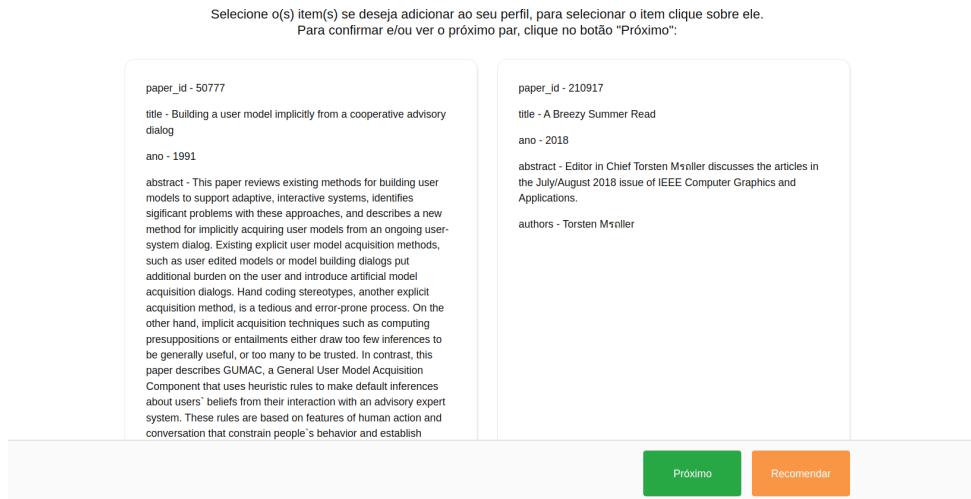
Fonte: a Autora.

(ver Figura 22). No caso do usuário selecionar dois itens, os dois resumos seriam adicionados ao perfil, e este ficaria como um texto completo a partir dos dois abstracts adicionados, conforme Figura 23.

O texto completo com todos os *abstracts* dos itens selecionados é representado como uma *query*, representando o perfil do usuário. Para que o perfil seja comparado com os termos dos artigos que estão indexados no ElasticSearch, é necessário que eles sejam comparados. Uma *query* de pesquisa, utiliza o operador *AND* ou o operador *OR* entre as palavras. No caso das palavras do perfil, elas serão enviadas com o operador *OR*.

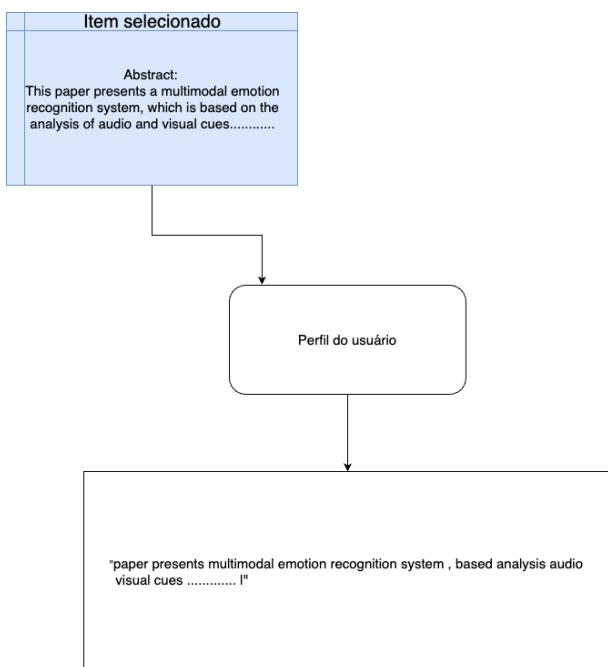
O processo de recomendação inicia-se a partir da comparação dos dados indexados com

Figura 21 – Tela de seleção de itens em pares - Botão Recomendar



Fonte: a Autora.

Figura 22 – Adicionando um item ao aperfil do usuário

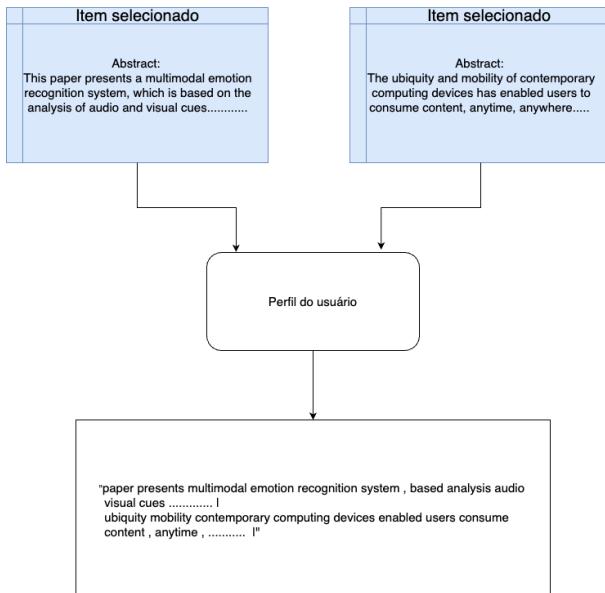


Fonte: a Autora.

os do perfil do usuário, ou seja, o processo de Recuperação da Informação. A *query* é enviada para o ElasticSearch via *JSON*. A partir desse momento, o ElasticSearch calcula a similaridade dos itens para que sejam mostrados os semelhantes aos que o usuário selecionou como os de seu interesse. Considerações sobre como a similaridade entre textos é calculada no ElasticSearch é descrita na seção 5.1.

A partir do momento que o usuário opta por receber recomendações, o sistema retorna uma lista de recomendação de artigos que possuem palavras em seus resumos semelhantes as

Figura 23 – Adicionando dois itens ao perfil do usuário



Fonte: a Autora.

palavras/resumos adicionados no perfil do usuário, como é possível visualizar na Figura 24.

Figura 24 – Tela de Recomendações

Olá massa@gmail.com!

Seu perfil

ID: 4832

Título: Therapy Game: uma nova plataforma para auxiliar nas áreas de fisioterapia e terapia ocupacional

Resumo: Video games are constantly evolving and new ideas for user interactions are being proposed. Within this context, innovative technological artifacts are being created in order to enable interactive flexibility, involvement in the game and rewarding experiences for the players. In recent years, the medical field has benefited from video games with new treatments in the area of physical and occupational therapy. This paper presents a low-cost technological infrastructure, capable of creating games for the area of body movement rehabilitation, utilizing an innovative wearable control that interacts with this new video game. In order to validate this proposal, we conducted initial experiments with potential users, pediatric neurologists and physical therapists. The user experience was a success and the results indicate promising usability in the medical area.

Ano: 2015

ID: 4946

Título: Analyzing User Experience in MobileWeb, Native and ProgressiveWeb Applications: A User and HCI Specialist Perspectives

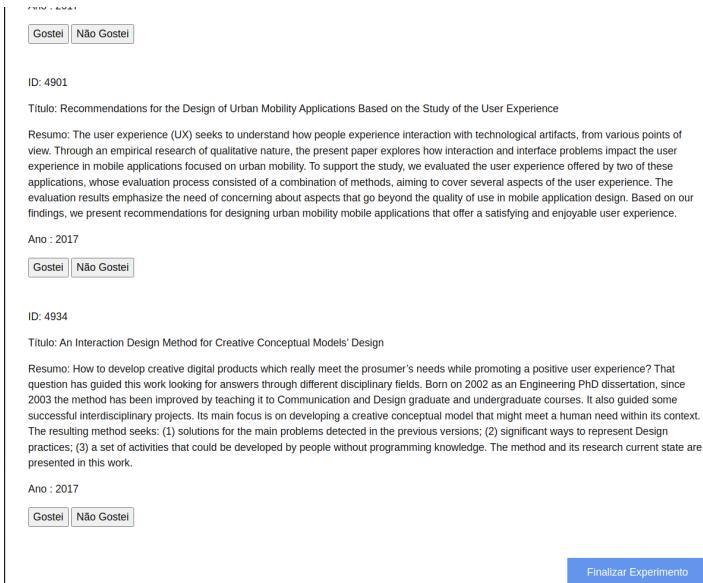
Resumo: Progressive Web App (PWA) is a new approach to the development of mobile application proposed by Google in 2015. It combines technology resources of both web and native applications. The challenges of designing interfaces for different applications platforms, such as web and native Android, has been discussed in recent years. However, PWAs are a recent technology and their impact regarding user experience have been little exploited. In this paper, we present the findings of an experimental study with 8 participants that explored the aspects of user experience on three different platforms. We carried out a qualitative analysis that focused on the comparison of the user experience during the participants' interaction with PWA, web mobile and native Android applications. Two distinct perspectives were defined to support our data analysis. First, the user perspective was considered and the participants' feedback was explored. After, focusing on the human-computer interaction specialist perspective we examined the users' facial expressions with the aims of identifying which emotions they sensed during interactions with each application. We gathered evidence that an overall positive user experience can be achieved even if the user had some interaction issues. There is no bias indicating that either a specific platform or interface element offer more enjoyable interactions.

Ano: 2018

Fonte: a Autora.

Ao receber as recomendações o usuário irá visualizar 10 itens, sendo os 10 documentos mais similares ao seu perfil retornado pelo *ElasticSearch*. Para cada item recomendado, o usuário pode dar um *feedback* sobre este a partir dos botões Gostei e Não Gostei. Ao clicar nos botões, o sistema computa automaticamente o *feedback* dado pelo usuário. Ao final da lista de recomendação, o botão de “Finalizar Experimento” (Figura 25) ao ser clicado, redireciona o usuário para um questionário, onde foram feitas perguntas para a avaliação do sistema (ver Seção 5.5).

Figura 25 – Tela Finalizar Experimentos



Fonte: a Autora.

A seguir são apresentados dois cenários de uso do sistema por usuários seguindo o fluxo apresentado.

5.4.1 Cenário de Uso - 1

O aluno José, 22 anos, estuda na Universidade Estadual de Santa Catarina. José está finalizando a graduação em Ciência da Computação e é um pesquisador júnior que irá submeter um artigo, pela primeira vez no próximo IHC sobre o tema usabilidade. Para isso, ele irá acessar o sistema Web criado com o modelo de recomendação. José então, entra na página e clica na opção login. Como ainda não se cadastrou, deverá clicar no botão que o redirecionada para a página de cadastro.

Para usar o sistema o usuário deverá efetuar o login. Caso ainda não esteja cadastrado, deverá realizar um cadastro na página de cadastro, fornecendo dados básicos (nome, email, senha). Estes dados serão armazenados no repositório de perfis do usuário. Cabe ressaltar que inicialmente nenhum interesse está definido para o usuário ocorrendo o problema de *Cold Start*.

No momento do cadastro, existe a etapa de seleção de itens de interesse. Dessa maneira, são apresentados pares de itens para o usuário escolher um item dentre eles de que mais gosta. Quando o usuário atinge 10 itens selecionados, o usuário então finaliza o cadastro, e realiza o login. Os documentos são apresentados em pares, representados pelo título e palavras-chave na visualização.

Conforme o usuário demonstra interesse em um documento dentre os pares, para cada um que seleciona são adicionadas as palavras-chave de cada documento em seu perfil. O sistema então apresenta a recomendação de artigos baseada nos interesses indicados por José comparados aos documentos e por fim, ele clica em "Finalizar Experimento" onde é redirecionado para um

questionário de avaliação.

5.4.2 Cenário de Uso - 2

O aluno José, 22 anos, estuda na Universidade Estadual de Santa Catarina. José está finalizando a graduação em Ciência da Computação e é um pesquisador júnior que irá submeter um artigo, pela primeira vez no próximo IHC sobre o tema usabilidade. Para isso, ele acessa o sistema Web criado com o modelo de recomendação. José então, entra na página e clica na opção login. Como ainda não se cadastrou, deverá clicar no botão que o redirecionada para a página de cadastro.

Para usar o sistema o usuário deverá efetuar o login. Caso ainda não esteja cadastrado, deverá realizar um cadastro na página de cadastro, fornecendo dados básicos (nome, email, senha). Estes dados serão armazenados no repositório de perfis do usuário. Cabe ressaltar que inicialmente nenhum interesse está definido para o usuário ocorrendo o problema de *Cold Start*.

No momento do cadastro, existe a etapa de seleção de itens de interesse. Desta maneira, é apresentada ao usuário uma lista de 10 itens para o usuário escolher 5 dentre eles de que mais gosta. Quando o usuário atinge 5 itens selecionados, o usuário então finaliza o cadastro, e realiza o login. Os documentos que são apresentados em pares e são representados pelo título e palavras-chave na visualização.

Conforme o usuário demonstra interesse em um documento, para cada um que seleciona são adicionadas as palavras-chave de cada documento em seu perfil. O sistema então apresenta a recomendação de 10 artigos baseada nos interesses indicados por José comparados aos documentos. Por fim, ele clica em "Finalizar Experimento" onde é redirecionado para um questionário de avaliação.

5.5 AVALIAÇÃO DO SISTEMA

Para avaliação do quanto a estratégia de construção do perfil inicial permitia a geração de boas recomendações, foram feitas duas avaliações: realização de testes funcionais com estudantes de Computação; avaliação com um grupo de usuários público-alvo do sistema. Para que pudessem ser feitos os testes, o sistema foi disponibilizado em um Website, a partir da plataforma de hospedagem de domínio em Nuvem *Heroku*¹. Para ambas as avaliações, foi feito o convite para participação no experimento via *email*.

A primeira avaliação consistiu de testes funcionais com Graduandos do curso de Ciência da Computação e/ou alunos que realizaram a disciplina de IHC na UDESC. Esta avaliação permitiu a validação de funcionalidades, dificuldades e percepções dos *testers* no sistema.

A segunda avaliação teve como objetivo avaliar efetivamente a precisão do sistema. Nesta avaliação foram convidados tanto de alunos de Bacharelado em Ciência da Computação,

¹ <<https://new-tcc.herokuapp.com/>>

quanto para pesquisadores na área de IHC. Por fim, nesta avaliação, na última tela os usuários visualizaram um questionário para avaliar as recomendações geradas aos usuários.

5.5.1 Testes Funcionais

A primeira avaliação realizada do sistema consistiu de testes funcionais para a validação das funcionalidades e elementos de interação do sistema para um grupo de pessoas convidadas por conveniência. Para esta etapa de avaliação, graduandos do curso de Ciência da Computação e/ou alunos que realizaram a disciplina de IHC na UDESC foram convidados a participar como *testers* do sistema de recomendação. Cada participante recebeu por email os *links* de acesso ao sistema, as instruções de uso que podem ser visualizadas no seguinte link² e a um questionário para ser respondido ao final do fluxo do sistema, para verificar as funcionalidades em cunho qualitativo. Esse questionário tinha como intuito entender as dificuldades e percepções dos testers encontrados no sistema. O questionário utilizado na avaliação está no Apêndice C. Ao total, foram 4 participantes, que receberam o email, utilizaram o sistema, e responderam ao questionário. Para cada tela do sistema foram realizados questionamentos sobre as percepções, dificuldades ou erros encontrados pelos *testers*. Dessa maneira, na tela inicial foram encontrados pontos positivos, que segundo os *testers* “Gostei bastante, tela simples que foca as explicações sobre o teste e esclarece como deve ser realizado o processo.”, “Achei bastante intuitivo as rotinas para cadastrar e também para fazer o login.”, foram encontrados alguns erros de *link* entre páginas no sistema em um dos ícones da tela. Um dos testers também relatou não ter a possibilidade de voltar para a tela de instruções, após fazer o login e voltar para a tela inicial. Somente se fizesse o cadastro novamente. Porém, o intuito é que o usuário seguisse um fluxo único, do início do experimento ao fim, conforme apresentado na Figura 15. A média de tempo de avaliação foi em torno de 18 minutos de uso do sistema e esse tempo foi medido a partir das respostas dos testers as perguntas sobre horários de início e fim da avaliação no sistema.

Com os testes funcionais, importantes *feedbacks* foram coletados para a realização de melhorias como questões de erros de páginas. Algumas sugestões de *design* foram sugeridas, porém como o intuito deste teste era entender se os usuários tinham alguma dificuldade ou se encontrariam algum erro de funcionamento no sistema, logo, entendeu-se que além das sugestões, os *feedbacks* foram positivos quanto ao funcionamento do sistema. Outro *feedback* coletado foi que os *testers* relataram que as instruções foram claras, o que possibilitou que a próxima fase de avaliação, pudesse ser realizada.

5.5.2 Experimentos do sistema

A segunda avaliação do sistema foi realizada através de um experimento com um grupo de usuários público-alvo do sistema. Para tal, foi realizado um convite via *email* para pesquisadores

² <https://docs.google.com/document/d/1f_KOp4kaE-VUIBuallajvD2oqW5JL8ddUc74gPs0Hng/edit?usp=sharing>

e estudantes da área de IHC convidadas por conveniência. O *email* enviado pode ser visualizado no Apêndice D e os feedbacks fornecidos por cada participante pode ser consultado neste link³.

Essa avaliação foi realizada de maneira a avaliar se a técnica utilizada para criação do perfil inicial do usuário, possibilitou a geração de recomendações adequadas aos usuários do sistema. As avaliações foram feitas usando a métrica de *Precision@k* sendo k igual à 10 e 5. A métrica *Recall* que normalmente acompanha a Precisão, não foi avaliada pois para isso, seria necessário conhecer o total de itens relevantes recomendados, ou seja, todas as recomendações possíveis para cada usuário. A avaliação também foi feita a partir de um questionário ao final do fluxo do sistema, quando o usuário clicava no botão "Finalizar o Experimento". Este questionário contém perguntas descritivas e de Escala de Likert, e pode ser visualizado no Apêndice E.

5.5.3 Resultados da avaliação de experimentos

Nesta seção são apresentados os resultados da avaliação. Os resultados são formados pelas avaliações feitas pelos usuários participantes durante suas interações no sistema desenvolvido. Os experimentos foram realizados no período de 16/04/2024 a 30/04/2024 e no total foram 24 participantes.

Do total de participantes, 41% dos participantes eram de áreas da Computação, e 54% eram da área específica de IHC (Tabela 8). Dos participantes 11 eram Docentes/Pesquisador(a) enquanto que 11 eram estudantes de pós-graduação e 2 eram estudantes de graduação. Estes mesmos participantes também escolheram mais de uma opção e 7 deles além destes perfis colocaram que também eram profissionais da indústria. Além disso, os participantes que responderam ao questionário eram de diversas regiões do país e suas universidades, como Mato Grosso, Pará, Paraná, São Paulo e Rio Grande do Sul. Apenas 5 participantes eram da UDESC.

A Tabela 6 mostra a relação dos 10 itens mostrados para cada um dos 24 usuários na lista de recomendação, quantos itens cada um deles Gostou e a taxa de *Precision@10*. A métrica de *Precision* é apresentada através da porcentagem e mede quantos itens relevantes foram mostrados no *top-k* de recomendações (HERLOCKER et al., 2004; RICCI; ROKACH; SHAPIRA, 2015). Dos 240 itens recomendados, 182 foram bem avaliados pelo usuário como "Gostei", enquanto que 58 itens foram avaliados como "Não Gostei". A taxa de itens classificados como "Gostei" é maior que 75,8% na média.

Os resultados mostrados na Tabela 6 refletem a quantidade individual de avaliações para os artigos recebidos, com a taxa de *Precision@10*. Os artigos recebidos na recomendação foram processados através do perfil de cada usuário. Do total de 24 participantes, 19 avaliaram bem mais de 5 dos 10 itens recomendados.

Além disso, foi feita a avaliação da Precisão até 5 (*Precision@5*). Através dessa filtragem de itens, é possível avaliar a quantidade de itens relevantes para o usuário diante de um top k

³ <<https://docs.google.com/spreadsheets/d/11jC9pp2l0Cfylg4ODcrxFzpo09r7VeaDJFr2SO5FyU/edit?usp=sharing>>

Participante	Gostou	<i>Precision@10 em %</i>
Participante 1	10	100
Participante 2	10	100
Participante 3	7	70
Participante 4	5	50
Participante 5	5	50
Participante 6	5	50
Participante 7	4	40
Participante 8	9	90
Participante 9	10	100
Participante 10	7	70
Participante 11	9	90
Participante 12	8	80
Participante 13	7	70
Participante 14	8	80
Participante 15	7	70
Participante 16	8	80
Participante 17	9	90
Participante 19	9	90
Participante 19	9	90
Participante 20	7	70
Participante 21	8	80
Participante 22	8	80
Participante 23	8	80
Participante 24	5	50

Fonte: a Autora.

Tabela 6 – Avaliações dos top 10 artigos recomendados.

artigos menor, para definir com mais clareza do que o usuário mais gostou na recomendação recebida.

Os resultados mostrados na Tabela 7 refletem a quantidade individual de avaliações para os artigos recebidos, com a taxa de *Precision@5*. Diante dos resultados, a taxa de itens classificados como "Gostei" é maior que 68%, considerando como 100% um total de 120 artigos.

Ao total 24 pessoas participaram do experimento de maneira voluntária acessando o sistema via URL enviada por *email*. No corpo do email os participantes foram instruídos a executar determinadas tarefas no protocolo de tarefas: cadastro no sistema, a realização do login, a avaliação das informações apresentadas para o usuário sobre como selecionar o perfil, o acesso a lista de recomendação com top 10 itens recomendados e a avaliação dos itens de acordo com o seu interesse, juntamente com o tempo de avaliação.

Após a avaliação da Precisão, foram observadas as respostas a partir do questionário de avaliação ao final do fluxo do sistema, quando os usuários optavam por "Finalizar o Experimento".

Quando questionados no questionário de avaliação sobre os pares de itens apresentados, o quanto os participantes sentiram que tinham relação um com o outro, estes deveriam responder

Participante	Gostou	Precision @5 em %
Participante 1	5	100
Participante 2	5	100
Participante 3	5	100
Participante 4	1	20
Participante 5	5	100
Participante 6	1	20
Participante 7	1	20
Participante 8	4	80
Participante 9	5	100
Participante 10	2	40
Participante 11	4	90
Participante 12	4	80
Participante 13	4	70
Participante 14	3	60
Participante 15	2	40
Participante 16	3	60
Participante 17	4	80
Participante 19	5	100
Participante 19	4	80
Participante 20	3	60
Participante 21	3	60
Participante 22	3	60
Participante 23	3	60
Participante 24	3	60

Fonte: a Autora.

Tabela 7 – Avaliações dos top 5 artigos recomendados.

em relação a quantos pontos na Escala Likert que variava de 1 a 10 pontos. É possível ver as respostas através da Tabela 9, onde 33% dos participantes votaram em 7 pontos (Escala Likert).

Do total de participantes 87% deles sentiram que os itens tinham relação um com o outro na Escala Likert de 1 a 7 pontos.

Do total de participantes, 12 deles avaliaram em até 5 pontos para o nível de quanto os pares tinham relação com o outro. Quando questionados sobre o que lhes motivou a escolher um item no par, ou a não escolha os *feedbacks* foram boa parte sobre itens de sua área de interesse, totalizando 8 participantes com essa escolha, sendo os Participantes 6,7,14,15,17,18,20 e 23 (ver Tabela 10). Além disso, destes participantes citados, nenhum gostou de menos que 5 itens recomendados. Com esses *feedbacks* é possível identificar que os participantes reconheceram que os pares não tinham tanta relação entre si, e procuraram seus interesses nos pares. Pode-se identificar que esse é um reflexo do algoritmo proposto, visto que para gerar o próximo par, o item inicial do próximo é um item que não é nem muito similar com o primeiro item anterior, e nem tão similar com o segundo item do par anterior.

Em contrapartida, a taxa de participantes que avaliou com um mínimo de 8 pontos

Participante	Instituição	Área de atuação
Participante 1	Controladoria Geral de Santa Catarina	Área da Computação
Participante 2	UDESC	Área da Computação
Participante 3	UDESC	Área da Computação
Participante 4	UFPel	Área da Computação, Área de IHC
Participante 5	UFACar	Área de IHC
Participante 6	IFRS	Área da Computação, Área de IHC
Participante 7	PUCRS	Área de IHC
Participante 8	UDESC	Área da Computação, Área de IHC, Informática na Educação
Participante 9	UDESC	Matemática
Participante 10	UFPA	Área da Computação
Participante 11	UFPR	Área da Computação
Participante 12	UFMG	Área da Computação
Participante 13	UFPR	Área da Computação, Área de IHC
Participante 14	UFPA	Área da Computação
Participante 15	UDESC	Área da Computação
Participante 16	Udesc	Área da Computação, Área de IHC
Participante 17	UFSM	Área da Computação
Participante 19	UFPA	Área da Computação, Área de IHC
Participante 19	IFSC	Área da Computação, Área de IHC
Participante 20	UFSCar	Área de IHC
Participante 21	Unisinos	Área da Computação, Área de IHC
Participante 22	UFC	Área da Computação, Área de IHC
Participante 23	Kwaba	Área de IHC
Participante 24	UDESC	Área da Computação

Fonte: a Autora.

Tabela 8 – Perfil dos participantes do experimento - Instituição e Área de atuação

Participante	Sobre os pares de itens apresentados, quanto você sentiu que tinham relação um com o outro?
Participante 1	9
Participante 2	6
Participante 3	4
Participante 4	4
Participante 5	5
Participante 6	3
Participante 7	5
Participante 8	7
Participante 9	7
Participante 10	7
Participante 11	8
Participante 12	7
Participante 13	3
Participante 14	3
Participante 15	7
Participante 16	4
Participante 17	4
Participante 19	7
Participante 19	4
Participante 20	4
Participante 21	8
Participante 22	4
Participante 23	7
Participante 24	7

Fonte: a Autora.

Tabela 9 – Respostas dos participantes: Sobre os pares de itens apresentados, quanto você sentiu que tinham relação um com o outro?

e máximo em 10 pontos na escala Likert sobre o quanto os itens tinham relação entre si, totalizou em 12%. Quando questionados sobre o que lhes motivou a escolher ou não o item, os participantes relataram que “Eu escolhi um item no par, ou optei por não escolher nenhum, baseado na minha afinidade e entendimento sobre o tema. Se nenhum dos itens apresentados me parece próximo ou comprehensível, preferi não selecionar nenhum.” segundo o Participante 2, e segundo o Participante 12 “Interesse pessoal com o tema, tamanho do abstract (muito curtos não retinham informação o suficiente para ganhar meu interesse a não ser que o título fosse bem interessante)”. Todos estes usuários gostaram de no mínimo 8 itens da lista de recomendação.

Na pergunta de sugestão de como iniciar o perfil em um sistema de recomendação, as opções pré-estabelecidas que os participantes poderiam escolher eram: “Digitar palavra”, “Digitar texto” e “Tópicos para seleção”. Conforme pode ser visualizado na Tabela 11, 14 do total de 24 participantes escolheram como primeira opção a de tópicos para seleção, enquanto que 8 participantes escolheram a opção “Digitar palavra” e 2 participantes escolheram “Digitar

Participante	O que lhe motivou a escolher um item no par, ou a não escolha?
Participante 1	Eu escolhi um item no par, ou optei por não escolher nenhum, baseado na minha afinidade e entendimento sobre o tema. Se nenhum dos itens apresentados me parece próximo ou compreensível, preferi não selecionar nenhum.
Participante 2	
Participante 3	afinidade com o tema e qualidade do resumo apresentado
Participante 4	O conteúdo apresentado nos textos de resumo e título
Participante 5	O que seria mais próximo do meu interesse
Participante 6	Minha principal motivação foi escolher itens que estavam relacionados com meus temas de pesquisa. Em segundo lugar, busquei identificar itens de maior qualidade/relevância através da análise do título/abstract.
Participante 7	Acho que não entendi o teste. Fiquei procurando, dentre os pares, aqueles que tinham relação com os meus interesses e não um com o outro. E não achei itens que estivessem relacionados a meus interesses de pesquisa. Apesar de ser um trabalho de IHC, me foram ofertados trabalhos mais na linha de computação gráfica. Acabei escolhendo alguns itens mais aleatórios, pois o sistema começou a repetir as alternativas.
Participante 8	Um bom resumo escrito, objetivo claro, método e resultados.
Participante 9	Itens voltados a área de atuação - Matemática, Educação, Ensino e Pesquisas Bibliográficas, itens ligados a temática de interesse.
Participante 10	não sei ao certo como responder, mas acho que foi a leitura do título
Participante 11	Interesse pessoal com o tema, tamanho do abstract (muito curtos não retinham informação o suficiente para ganhar meu interesse a não ser que o título fosse bem interessante).
Participante 12	Áreas acadêmicas que normalmente não tenho acesso no dia-a-dia, porém possuo interesse em saber mais.
Participante 13	Meus interesses de pesquisa
Participante 14	Itens que me fossem da minha área de interesse acadêmica e que possuíssem um resumo com mais linhas de descrição
Participante 15	Intuição, aquilo que mais faz parte de meus interesses, bater o olho e gerar interesse.
Participante 16	fui motivada pelos meus interesses em pesquisa acadêmica
Participante 17	A temática do artigo estar ou não relacionada aos tópicos de pesquisa com os quais já tive contato.
Participante 19	A descrição do abstract foi importante para decidir se escolheria o artigo ou não.
Participante 19	Observei os títulos dos artigos apresentados e aqueles que tinham similaridade com minhas áreas de interesse. Aquele artigo que eu gostaria de ler.
Participante 20	Acredito que o tema do item e a objetividade do resumo.
Participante 21	
Participante 22	O quanto eu achei que o artigo tinha a ver com minha área de pesquisa e que possivelmente eu leria
Participante 23	Selecionei ao menos 1 quando fazia sentido para minha área de interesse. E não selecionei nenhum quando ambos não estavam na minha área de interesse.
Participante 24	

Fonte: a Autora.

Tabela 10 – Respostas dos participantes: O que lhe motivou a escolher um item no par, ou a não escolha?

Participante	Sugestão de como iniciar o perfil em um sistema de recomendação:
Participante 1	Tópicos para seleção
Participante 2	Digitar palavra, Tópicos para seleção
Participante 3	Tópicos para seleção
Participante 4	Tópicos para seleção, Importar dados de outro sistema.
Participante 5	Digitar texto
Participante 6	Digitar texto, Tópicos para seleção
Participante 7	Tópicos para seleção
Participante 8	Tópicos para seleção
Participante 9	Digitar palavra, Tópicos para seleção
Participante 10	Digitar palavra
Participante 11	Digitar palavra, Tópicos para seleção
Participante 12	Tópicos para seleção, Seria legal, principalmente para professores, poder importar o perfil lattes ou uma solução semelhante. Dessa forma, o sistema teria os artigos publicados por eles como base para a recomendação.
Participante 13	Tópicos para seleção
Participante 14	Tópicos para seleção
Participante 15	Digitar palavra
Participante 16	Tópicos para seleção
Participante 17	Tópicos para seleção
Participante 19	Digitar palavra
Participante 19	Tópicos para seleção
Participante 20	Digitar palavra, Tópicos para seleção
Participante 21	Digitar palavra
Participante 22	Tópicos para seleção
Participante 23	Tópicos para seleção
Participante 24	Tópicos para seleção

Fonte: a Autora.

Tabela 11 – Respostas dos participantes: Sugestão de como iniciar o perfil em um sistema de recomendação

“texto”. Dos 14 participantes que escolheram “Tópicos para seleção” como primeira opção, 8 deles são da área específica de IHC.

Além disso, outras duas opções foram enviadas no questionário em acréscimo ao voto de “Tópicos para seleção”, sendo “Importar dados de outro sistema” pelo Participante 4, e a opção de “Seria legal, principalmente para professores, poder importar o perfil lattes ou uma solução semelhante. Dessa forma, o sistema teria os artigos publicados por eles como base para a recomendação.” pelo Participante 12. A questão de iniciar o perfil a partir de tópicos foi cogitada para esta proposta, porém foram encontradas dificuldades em como mostrar todo o conteúdo da base em tópicos, e que estes fossem diversos, para que o usuário pudesse escolher dentre seus diferentes interesses.

Quando questionados sobre impressões sobre o sistema no geral e sugestões de melhorias

para um sistema de seleção de perfil em pares, pelo menos 5 participantes comentaram sobre a possibilidade de um filtro de palavras-chaves, que a partir delas o perfil inicial pudesse ser definido, ou ainda indicar quais palavras ignorar para a construção do perfil, o Participante 19 relatou que “Neste cenário, as palavras-chaves são importantes para ajudar a escolher o artigo, bem como uma indicação prévia do intervalo de ano, pois apareceram alguns artigos muito antigos (mais de 10 anos). Após a primeira etapa do sistema, a lista de artigos estava bem coerente com as minhas escolhas iniciais.” enquanto que o Participante 12 comentou sobre “Possibilidade de filtro de palavras-chave que alguém queira ver ou já optar por ignorar.”.

Sobre impressões do sistema no geral, os participantes relataram suas percepções e podem ser visualizadas na Tabela 12 e Tabela 13.

A partir destes *feedbacks* coletados, percebe-se que os participantes identificaram que muitos itens não foram exatamente do seu interesse, esperavam encontrar mais itens do que lhes interessava mais, o que por um lado pode ser interessante por indicar a diversidade dos itens mostrados na formação do perfil, algo que a técnica aplicada busca atingir. Outra possibilidade é que esses resultados indicam que a técnica usada possa ser uma estratégia para minimizar a superespecialização. O lado negativo de trazer algo diferente dos interesses do usuário é o risco do usuário deixar de usar o Sistema de Recomendação, visto que o usuário possa achar que o sistema não irá reocmendar nada que lhe agrade.

Participante	Sugestões de melhorias para um sistema de seleção de perfil em pares:
Participante 1	Uma sugestão de melhoria poderia ser a comparação dos macrotemas dos artigos ao invés dos próprios artigos. Isso poderia ajudar a identificar interesses mais amplos e padrões de preferências mais facilmente, permitindo uma recomendação mais assertiva e abrangente. Por exemplo, em vez de escolher entre dois artigos específicos sobre Interação Humano-Computador, o usuário poderia escolher entre temas mais amplos como "Usabilidade", "Acessibilidade", "Realidade Virtual" ou "Interfaces Naturais do Usuário".
Participante 2	Assim como na escolha do perfil tinha o link dos artigos, seria legal se nas recomendações também tivesse - entendo que talvez não era o foco do trabalho nesse momento, mas seria legal se fosse possível mudar o perfil do usuário após já ter gerado as recomendações
Participante 3	Fazer um sistema que coloque um tema para seleção inicialmente, acabou sendo muito demorado para vir os artigos que eu tive algum interesse
Participante 4	Imagino que seja um protótipo e a interface final não será neste modelo. A minha sugestão seria a sua incorporação em outro sistema para que a recomendação se torne parte orgânica e não intrusiva.
Participante 5	Apareceram poucos itens nos pares que eram de meu interesse. Por isto ficou complicada a escolha. No final o recomendado foi adequado a escolha.
Participante 6	Os artigos que apareceram pra mim repetiam muito os mesmos assuntos (ex.: acessibilidade e 3D). Acabei não encontrando nenhum diretamente relacionado aos meus interesses: design intercultural (cross-cultural design). No fim, acabei selecionando itens de acessibilidade pois era o tema mais próximo que estava aparecendo. Sugiro tentar uma maior variedade de tópicos na hora de montagem dos pares de artigos. Também reparo que apareceram muitos itens que não eram revisados pelos pares, tais como introduções, anúncios, apresentações, etc. Esses itens, em geral, não me interessavam. Acredito que esses itens não deveriam ser exibidos.
Participante 7	Foi muito difícil para mim conseguir selecionar os itens, pois nenhum se relacionava aos meus interesses. Acabei gostando de algumas sugestões, mas elas não pareciam ter qualquer relação com o que selecionei. Acredito que devia haver alguma indicação inicial de tópicos de interesse, para se iniciar a apresentação de opções.
Participante 8	Achei os resultados da aplicação bons. O processo para definir o perfil foi rápido, e 70% da recomendação tinha correlação com o meu perfil.
Participante 9	Acredito que a apresentação das palavras-chaves seria interessante.
Participante 10	Eu achei o sistema com poucos elementos gráficos e de cores, acho que precisa chamar mais atenção nas partes que merecem mais destaque.
Participante 11	Possibilidade de filtro de palavras-chave que alguém queira ver ou já optar por ignorar.
Participante 12	Seria interessante poder escolher autores também, dessa forma o sistema recomendaria artigos de autores específicos que quero acompanhar.

Fonte: a Autora.

Tabela 12 – Respostas Parte 1 dos participantes: Sugestões de melhorias para um sistema de seleção de perfil em pares:

Participante	Sugestões de melhorias para um sistema de seleção de perfil em pares:
Participante 13	Demorou um pouco para começar a aparecer artigos relevantes para o meu interesse. Tenho um pouco de dúvidas sobre o quanto útil pode ser um sistema de recomendação que começa do zero, sem a escolha de tópicos de interesse a priori. Acho que pode ser muito bom para quem tem interesses de pesquisa mais gerais, porém para quem tem campos mais específicos, pode demorar um pouco para aparecer artigos relevantes. Além disso, notei muitos artigos com abstracts bem curtos, que provavelmente devem ser short paper ou resumos. Talvez a opção de filtrar entre essas categorias (resumo, short, full) também seja interessante.
Participante 14	Nada a declarar
Participante 15	Interessante, agiliza o ajuste do perfil se comparado com uma lista. Acredito que isto pode ser mais investigado e utilizado
Participante 16	No meu caso o artigo sugerido do lado direito repetiu muito o tema, muitos artigos sobre 3D, poderia apresentar temas mais aleatórios.
Participante 17	Um mecanismo de feedback para o usuário informar se gostou ou não do par.
Participante 19	Neste cenário, as palavras-chaves são importantes para ajudar a escolher o artigo, bem como uma indicação prévia do intervalo de ano, pois apareceram alguns artigos muito antigos (mais de 10 anos). Após a primeira etapa do sistema, a lista de artigos estava bem coerente com as minhas escolhas iniciais.
Participante 19	No momento que estava buscando, muitos títulos de repetiram. Fiquei em dúvida, se estava correto. Acabei selecionando mais de uma vez, um artigo de meu interesse. É isso mesmo?
Participante 20	Apresentar itens com tema e tamanho de resumo similares.
Participante 21	Acredito que seria interessante colocar as keywords.
Participante 22	Acho que um filtro inicial por área de interesse poderia ser muito útil, começou a sugerir alguns artigos fora de temas que geralmente eu pesquiso então fiquei só passando pro próximo. O design da interface também é algo que vale muito a pena investir, de forma a facilitar que referências seja exportadas, mais temas possam ser buscados, etc.
Participante 23	Adição da Opção: "Ambos não me agradam/interessam" para quando eu precisar passar para próxima oferta de pares.
Participante 24	Adição das palavras-chave do artigo antes do resumo dele. Talvez ter um pré-filtro de palavras-chave opcional antes da seleção para a sugestão dos pares já virem um pouco alinhada com o objetivo.

Fonte: a Autora.

Tabela 13 – Respostas Parte 2 dos participantes: Sugestões de melhorias para um sistema de seleção de perfil em pares:

5.6 CONSIDERAÇÕES SOBRE O CAPÍTULO

Este capítulo descreveu a avaliação do Sistema de Recomendação implementado, sistema esse que utiliza estratégia para criação de perfil descrita na Seção 4.1. Por meio de todas as etapas realizadas foi possível obter significativas sugestões de melhorias para o sistema.

Além da avaliação da precisão ter um resultado positivo, com a aplicação do questionário foi possível compreender melhor as reações dos usuários e obter sugestões. Apesar de 23 dos 24 participantes terem sugestões de melhorias, aproximadamente 80% dos participantes curtiram pelo menos metade dos itens recomendados.

6 CONSIDERAÇÕES FINAIS

O *Cold Start* é um dos problemas presentes e ainda não solucionado em Sistemas de Recomendação, mesmo sendo sempre citado na área. Desta forma, este trabalho teve como objetivo criar uma técnica que não utiliza nenhum tipo de histórico ou interação prévia do usuário, e considera apenas os dados existentes no sistema. Além disso, esta proposta utiliza a abordagem Baseada em Conteúdo de recomendação, uma abordagem pouco considerada nos trabalhos onde busca-se uma solução de minimizar o *Cold Start* a partir de um perfil inicial.

Para alcançar os objetivos neste trabalho e identificar as técnicas para minimizar o *Cold Start*, foi realizado um Mapeamento Sistemático sobre abordagens e técnicas para minimizar o *Cold Start* em Sistemas de Recomendação conforme publicado em (CEZAR et al., 2024). As técnicas identificadas para a construção do perfil inicial são solicitar que o usuário responda perguntas/questionário, solicitar a avaliação do usuário para um conjunto de itens, seleção/cadastro de palavras-chave/tags e técnicas de que consideram itens populares entre os usuários dos sistemas de recomendação. Dos 19 artigos identificados que apresentam técnicas para a redução do problema de *Cold-Start*, apenas 3 utilizam a abordagem Baseada em Conteúdo, e destes 3 apenas 1 utiliza técnica que não envolve um esforço para criar perguntas/questionário, utilizando a técnica de seleção/avaliação de itens.

A técnica definida neste trabalho consiste em apresentar ao usuário um conjunto de itens a serem por ele selecionados/avaliados de forma a permitir a construção de um perfil inicial. Os itens que serão mostrados para o usuário são alguns dos itens existentes na base de itens a serem recomendados. Esta estratégia de avaliar alguns itens é usada em outros sistemas, mas no trabalho o diferencial é o fato de que procura-se considerar itens diversos, isto é, são apresentados ao usuário para avaliação itens pouco similares entre si.

Para avaliar a técnica proposta para criação do perfil inicial do usuário, foi desenvolvido um Sistema de Recomendação de artigos que utiliza a seleção de pares de itens para a construção de um perfil inicial. O sistema desenvolvido apresenta no momento de criação do perfil do usuário, pares de artigos onde este pode selecionar um, dois ou nenhum item do par. Conforme o usuário seleciona os itens, informações são adicionadas ao seu perfil, sendo após recomendados 10 itens similares às informações adicionadas no perfil do usuário.

A avaliação do Sistema de Recomendação foi feita através de um estudo com o usuário de duas maneiras: a avaliação da taxa de *Precision* e além disso, através de um questionário. De acordo com a avaliação feita pelos usuários, conclui-se que os artigos foram bem recebidos pelos usuários, apesar de terem dificuldades na seleção dos itens, pois muitos indicaram que seria necessário a criação de mais alguns filtros para o início da criação do perfil baseado nos itens apresentados. Além disso, foi constatado que a maneira de apresentar os itens a serem avaliados para o usuário tem impacto.

Na avaliação realizada no sistema, foi possível identificar que 80% dos usuários participantes gostaram de pelo menos metade dos itens recomendados. Desta forma, a avaliação

realizada com a amostra de usuários coletada, indica que os resultados foram positivos quanto as recomendações feitas a partir do perfil inicial criado no sistema. Além disso, foi possível identificar algumas possíveis contribuições para questões como representação da diversidade de interesses no perfil inicial dos usuários.

Desse modo, acredita-se que a estratégia de construção do perfil inicial dos usuários proposta neste trabalho pode ser uma opção de criação de perfil inicial para minimizar o *Cold-Start* em Sistemas de Recomendação.

Os trabalhos futuros compreendem: a) Testar técnicas onde existam maior número de itens e também de áreas diversas; b) Aprofundar quais informações adicionar ao perfil do usuário a partir do item selecionado, melhorando a representação do perfil; c) Mostrar inicialmente artigos mais citados de uma área; d) Considerar a evolução do perfil do usuário; e) Mostrar tópicos/palavras para usuários para a criação do perfil inicial ou começar por uma pesquisa inicial e após aplicar a técnica; f) Avaliar quais informações mais relevantes para mostrar para o usuário do item - Permitir que os usuários tenham acesso aos artigos publicados na íntegra; g) Considerar itens que o usuário não selecionou para o perfil e não mostrar mais artigos deste tema/tópico (ou mais fortemente relacionados ao tópico); h) Avaliar a aplicação da técnica proposta neste trabalho em outros domínios e/ou abordagens de recomendação.

REFERÊNCIAS

- ABDULLAH, Nor Aniza et al. Eliciting auxiliary information for cold start user recommendation: A survey. **Applied Sciences (Switzerland)**, v. 11, n. 20, 2021. ISSN 20763417.
- ABISHEVA, Adiya; GARCIA, David; SCHWEITZER, Frank. When the filter bubble bursts: Collective evaluation dynamics in online communities. **WebSci 2016 - Proceedings of the 2016 ACM Web Science Conference**, p. 307–308, 2016.
- ADOMAVICIUS, Gediminas; KWON, Youngok. Diversity Using Ranking-Based Techniques. **IEEE Transactions on Knowledge and Data Engineering**, v. 24, n. 5, p. 896–911, 2012. ISSN 1041-4347. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5680904>>.
- ADOMAVICIUS, Gediminas; TUZHILIN, Alexander. Toward the Next Generation of Recommender Systems : A Survey of the State-of-the-Art and Possible Extensions. **IEEE Access**, IEEE, v. 17, n. 6, p. 734–749, 2005.
- ALLIOUI, Youssouf EL. A novel approach to solve the new user cold-start problem in recommender systems using collaborative filtering. **International Journal of Scientific Engineering Research Volume**, v. 8, n. 11, p. 273–281, 2017. Disponível em: <https://en.wikipedia.org/wiki/Collaborative_filtering>.
- ALYARI, Fatemeh; NAVIMIPOUR, Nima Jafari. Recommender systems: A systematic review of the state of the art literature and suggestions for future research. **Kybernetes**, v. 47, n. 5, p. 985–1017, 2018. ISSN 0368492X.
- AMATRIAIN, Xavier et al. The wisdom of the few: A collaborative filtering approach based on expert opinions from the web. **Proceedings - 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009**, p. 532–539, 2009.
- BAI, Xiaomei et al. Scientific Paper Recommendation : A Survey. **IEEE Access**, IEEE, v. 7, p. 9324–9339, 2019.
- BETTMAN, James R.; LUCE, Mary Frances; PAYNE, John W. Constructive Consumer Choice Processes. **Journal of Consumer Research**, v. 25, n. 3, p. 187–217, 12 1998. ISSN 0093-5301. Disponível em: <<https://doi.org/10.1086/209535>>.
- Borba, Caroline. Estratégia de Pós-processamento aplicada a um Sistema de Recomendação de Artigos visando refletir os vários interesses do usuário Isabela Gasparini Daniel Lichnow. 2020.
- BORBA, Eduardo. Sistema de Recomendação de Objetos de Aprendizagem no Ambiente Adaptweb. **Trabalho de Conclusão de Curso - Bacharelado em Ciência da Computação**, Universidade Estadual de Santa Catarina (UDESC), 2015.
- BUCHINGER, Diego; CAVALCANTI, Gustavo Andriolli de Siqueira; HOUNSELL, Marcelo Da Silva. Mecanismos de busca acadêmica: uma análise quantitativa. **Revista Brasileira de Computação Aplicada**, v. 6, n. 1, p. 108–120, 2014.

BURKE, Robin. Hybrid Recommender Systems : Survey and Experiments. **User Modeling and User-Adapted Interaction**, ACM, v. 12, n. 6, p. 331–370, 2002.

CABREJOS, Luis Jorge Enrique Rivero; VIANA, Davi; SANTOS, Rodrigo Pereira dos. Planejamento e Execução de Estudos Secundários em Informática na Educação: Um Guia Prático Baseado em Experiências. **Jornada de Atualização em Informática na Educação**, v. 7, n. 1, p. 21–52, 2018.

CARVALHO, Licínio; RODRIGUES, Fátima; OLIVEIRA, Pedro. A hybrid recommendation algorithm to address the cold start problem. In: _____. [S.l.: s.n.], 2020. p. 260–271. ISBN 978-3-030-14353-4.

CELMA, Oscar; HERRERA, Perfecto; SERRA, Xavier. Bridging the music semantic gap. **CEUR Workshop Proceedings**, v. 187, p. 927–934, 2006. ISSN 16130073.

CELMA, Óscar; RAMÍREZ, Miquel; HERRERA, Perfecto. Foafing the Music: A music recommendation system based on RSS feeds and user preferences. **ISMIR 2005 - 6th International Conference on Music Information Retrieval**, p. 464–467, 2005.

CENA, Federica; VERNERO, Fabiana; GENA, Cristina. Towards a Customization of Rating Scales in Adaptive Systems. In: De Bra, Paul; KOBZA, Alfred; CHIN, David (Ed.). **User Modeling, Adaptation, and Personalization**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 369–374. ISBN 978-3-642-13470-8.

CEZAR, Nathalia et al. Exploring strategies to mitigate cold start in recommender systems: A systematic literature mapping. In: INSTICC. **Proceedings of the 26th International Conference on Enterprise Information Systems - Volume 1: ICEIS**. [S.l.]: SciTePress, 2024. p. 965–972. ISBN 978-989-758-692-7.

CEZAR, Nathália Locatelli. **Sistema de recomendação para sugerir artigos para a Comunidade Brasileira de IHC**. 2019. Monografia (Bacharel em Ciência da Computação), UDESC (Universidade do Estado de Santa Catarina), Joinville, Brasil.

CEZAR, Nathália Locatelli et al. Applying a post-processing strategy to consider the multiple interests of users of a Paper Recommender System. **ACM International Conference Proceeding Series**, 2021.

CHANG, Shuo; HARPER, F. Maxwell; TERVEEN, Loren. Using Groups of Items for Preference Elicitation in Recommender Systems. p. 1258–1269, 2015.

CHAO, Xu; GUANGCAI, Cui. Collaborative Filtering and Leaders' Advice Based Recommendation System for Cold Start Users. In: **Proceedings of the 2020 6th International Conference on Computing and Artificial Intelligence**. New York, NY, USA: Association for Computing Machinery, 2020. (ICCAI '20), p. 158–164. ISBN 9781450377089. Disponível em: <<https://doi.org/10.1145/3404555.3404644>>.

CHRISTAKOPOULOU, Konstantina; RADLINSKI, Filip; HOFMANN, Katja. Towards conversational recommender systems. **Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, v. 13-17-August-2016, n. 3, p. 815–824, 2016.

CO, Elastic. **Practical BM25 - Part 2: The BM25 Algorithm and its Variables.** 2018. ElasticSearch Blog. Disponível em: <<https://www.elastic.co/blog/practical-bm25-part-2-the-bm25-algorithm-and-its-variables>>.

CONSONI, Gilberto. Recuperação de informação em sistemas de recomendação : análise da interação mediada por computador e dos efeitos da filtragem colaborativa na seleção de itens no website da Amazon.com. **Tese - Programa de Pós Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul**, Universidade Federal do Rio Grande do Sul (UFRGS), 2014.

COSLEY, Dan et al. Is seeing believing? how recommender system interfaces affect users' opinions. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY, USA: Association for Computing Machinery, 2003. (CHI '03), p. 585–592. ISBN 1581136307. Disponível em: <<https://doi.org/10.1145/642611.642713>>.

CRAWFORD, Diane. Initial Profile Generation in Recommender Systems Using Pairwise Comparison. **Communications of the ACM**, IEEE, v. 36, n. 11, p. 18, 2012. ISSN 15577317.

DARSHNA, Parmar. Music recommendation based on content and collaborative approach reducing cold start problem. **Proceedings of the 2nd International Conference on Inventive Systems and Control, ICISC 2018**, IEEE, n. Icisc, p. 1033–1037, 2018.

DOOMS, Simon; PESSEMIER, Toon De; MARTENS, Luc. An online evaluation of explicit feedback mechanisms for recommender systems. In: . [S.l.: s.n.], 2011. p. 391–394.

FELÍCIO, Crícia Z. et al. A multi-Armed bandit model selection for cold-start user recommendation. **UMAP 2017 - Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization**, p. 32–40, 2017.

FELÍCIO, Crícia Z. et al. Vp-rec: A hybrid image recommender using visual perception network. p. 70–77, 2016.

FERNÁNDEZ, Dlego et al. A Content-Based Approach to Profile Expansion. **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, v. 28, n. 6, p. 981–1002, 2020. ISSN 02184885.

FONTELLES, Mauro José et al. METODOLOGIA DA PESQUISA CIENTÍFICA: DIRETRIZES PARA A ELABORAÇÃO DE UM PROTOCOLO DE PESQUISA. **Australian Journal of Physiotherapy**, v. 52, n. 4, p. 237–239, 2006. ISSN 00049514.

GENA, Cristina et al. The Impact of Rating Scales on User's Rating Behavior. In: KONSTAN, Joseph A et al. (Ed.). **User Modeling, Adaption and Personalization**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 123–134. ISBN 978-3-642-22362-4.

GIL, Antonio Carlos. **Métodos e Técnicas de Pesquisa Social**. 6. ed. [S.l.: s.n.], 2009. v. 264. 27–28 p. ISBN 9788522451425.

GOLDBERG, Ken et al. Eigentaste: A Constant Time Collaborative Filtering Algorithm. **Information Retrieval**, v. 4, n. 2, p. 133–151, 2001. ISSN 13864564.

GORMLEY, Clinton; TONG, Zachary. **Elasticsearch: The Definitive Guide**. 1. ed. [S.l.]: O'Reilly Media, Inc., 2015.

HE, D. P.; HE, Z. L.; LIU, C. Recommendation algorithm combining tag data and naive bayes classification. **Proceedings - 2020 3rd International Conference on Electron Device and Mechanical Engineering, ICEDME 2020**, p. 662–666, 2020.

HERLOCKER, Jonathan L. et al. Evaluating collaborative filtering recommender systems. **ACM Trans. Inf. Syst.**, Association for Computing Machinery, New York, NY, USA, v. 22, n. 1, p. 5–53, jan 2004. ISSN 1046-8188. Disponível em: <<https://doi.org/10.1145/963770.963772>>.

HONG, Kwanghee; JEON, Hocheol; JEON, Changho. User Profile-Based Personalized Research Paper Recommendation System. **2012 8th International Conference on Computing and Networking Technology (INC, ICCIS and ICMIC)**, IEEE, p. 134–138, 2012.

HRISTAKEVA, Maya et al. Building Recommender systems for scholarly information. **ACM International Conference Proceeding Series**, Part F127853, p. 25–32, 2017.

JANNACH, Dietmar et al. **Recommender Systems: An Introduction**. [S.l.]: Cambridge University Press, 2010.

JÄRVELIN, Kalervo; KEKÄLÄINEN, Jaana. Cumulated gain-based evaluation of ir techniques. **ACM Trans. Inf. Syst.**, Association for Computing Machinery, New York, NY, USA, v. 20, n. 4, p. 422–446, oct 2002. ISSN 1046-8188. Disponível em: <<https://doi.org/10.1145/582415.582418>>.

JELASSI, Mohamed Nader; Ben Yahia, Sadok; NGUIFO, Engelbert Mephu. A personalized recommender system based on users' information in folksonomies. **WWW 2013 Companion - Proceedings of the 22nd International Conference on World Wide Web**, p. 1215–1223, 2013.

KALLOORI, Saikishore; RICCI, Francesco; GENNARI, Rosella. Eliciting Pairwise Preferences in Recommender Systems. p. 329–337, 2018.

KHAN, Mansura A.; SMYTH, Barry; COYLE, David. Addressing the complexity of personalized, context-aware and health-aware food recommendations: an ensemble topic modelling based approach. **Journal of Intelligent Information Systems**, Journal of Intelligent Information Systems, v. 57, n. 2, p. 229–269, 2021. ISSN 15737675.

LAZEMI, Soghra; EBRAHIMPOUR-KOMLEH, Hossein. Improving collaborative recommender systems via emotional features. **Application of Information and Communication Technologies, AICT 2016 - Conference Proceedings**, IEEE, p. 1–5, 2017.

LICHTNOW, Daniel. Sistemas de Recomendação: Breve Histórico e Perspectivas. Jornada Acadêmica Integrada - UFSM, p. 88–105, 2015.

LIN, Chen et al. PRemiSE: Personalized News Recommendation via Implicit Social Experts. In: **Proceedings of the 21st ACM International Conference on Information and Knowledge Management**. New York, NY, USA: Association for Computing Machinery, 2012. (CIKM '12), p. 1607–1611. ISBN 9781450311564. Disponível em: <<https://doi.org/10.1145/2396761.2398482>>.

LUNARDI, Gabriel Machado. Diversificação multi-atributo para diminuir o efeito “bolha de filtro” em sistemas de recomendação de notícias. 2021.

MAIDA, Martina; OBWEGESER, Nikolaus. Pairwise comparison techniques for preference elicitation : using test-retest reliability as a quality indicator. **International Conference on Information Resources Management (Conf-Irm)**, n. May, 2012.

MANNING, Christopher D.; RAGHAVAN, Prabhakar; SCHÜTZ, Hinrich. **An Introduction to Information Retrieval***An Introduction to Information Retrieval*. [S.l.]: Cambridge University Press, 2009.

MARTINS, Eder F. et al. Measuring and addressing the impact of cold start on associative tag recommenders. **WebMedia 2013 - Proceedings of the 19th Brazilian Symposium on Multimedia and the Web**, p. 325–332, 2013.

MILLER, G A. The magical number seven, plus or minus two: some limits on our capacity for processing information. v. 101,2, 1994.

MONTI, Diego; RIZZO, Giuseppe; MORISIO, Maurizio. **A systematic literature review of multicriteria recommender systems**. Springer Netherlands, 2021. v. 54. 427–468 p. ISSN 15737462. ISBN 0123456789. Disponível em: <<https://doi.org/10.1007/s10462-020-09851-4>>.

OKADA, Keisuke; TAN, Phan Xuan; KAMIOKA, Eiji. Five-Factor Musical Preference Prediction for Solving New User Cold-Start Problem in Content-Based Music Recommender System. **IISA 2021 - 12th International Conference on Information, Intelligence, Systems and Applications**, IEEE, 2021.

OLMO, Félix Hernández del; GAUDIOSO, Elena. Evaluation of recommender systems: A new approach. **Expert Syst. Appl.**, v. 35, p. 790–804, 10 2008.

PANDA, Deepak Kumar; RAY, Sanjog. Approaches and algorithms to mitigate cold start problems in recommender systems: a systematic literature review. **Journal of Intelligent Information Systems**, Journal of Intelligent Information Systems, p. 341–366, 2022. ISSN 15737675. Disponível em: <<http://dx.doi.org/10.1007/s10844-022-00698-5>>.

PAZZANI, Michael J; BILLSUS, Daniel. Content-Based Recommendation Systems. p. 325–326, 2007.

PEIXOTO, Jessé. **Como o elasticsearch decide o que é relevante e ordena o resultado de uma busca?** 2019. LinkedIn Article. Disponível em: <<https://www.linkedin.com/pulse/como-o-elasticsearch-decide-que-%C3%A9-relevante-e-ordena-de-jess%C3%A9-peixoto/>>.

PETERSEN, Kai et al. Systematic mapping studies in software engineering. **12th International Conference on Evaluation and Assessment in Software Engineering, EASE 2008**, n. February 2015, 2008.

PRETORIUS, Rialette; BUDGEN, David. A mapping study on empirical evidence related to the models and forms used in the uml. In: **Proceedings of the Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement**. New York, NY, USA: Association for Computing Machinery, 2008. (ESEM '08), p. 342–344. ISBN 9781595939715. Disponível em: <<https://doi.org/10.1145/1414004.1414076>>.

RICCI, Francesco; ROKACH, Lior; SHAPIRA, Bracha. Introduction to recommender systems handbook. In: **Recommender Systems Handbook**. [S.l.: s.n.], 2010. v. 22, p. 1–35.

- RICCI, Francesco; ROKACH, Lior; SHAPIRA, Bracha. **Recommender Systems: Handbook**. 2011. ed. [S.l.]: Springer, 2011. ISBN 9780387858197.
- RICCI, Francesco; ROKACH, Lior; SHAPIRA, Bracha. **Recommender Systems**. [S.l.: s.n.], 2015. ISBN 9781489976369.
- ROBERTSON, Stephen; ZARAGOZA, Hugo et al. The probabilistic relevance framework: Bm25 and beyond. **Foundations and Trends® in Information Retrieval**, Now Publishers, Inc., v. 3, n. 4, p. 333–389, 2009.
- SALTON, G.; MCGILL, M. J. Introduction to Modern Information. p. 375–384, 1983. Disponível em: <<http://portal.acm.org/citation.cfm?id=1893971.1894017>>.
- SCHAFFER, J. Ben et al. Collaborative filtering recommender systems. In: _____. **The Adaptive Web: Methods and Strategies of Web Personalization**. Berlin, Heidelberg: Springer-Verlag, 2007. p. 291–324. ISBN 9783540720782.
- SHI, Shaoyun et al. WG4Rec: Modeling Textual Content with Word Graph for News Recommendation. In: **Proceedings of the 30th ACM International Conference on Information Knowledge Management**. New York, NY, USA: Association for Computing Machinery, 2021. (CIKM '21), p. 1651–1660. ISBN 9781450384469. Disponível em: <<https://doi.org/10.1145/3459637.3482401>>.
- SON, Le Hoang. Dealing with the new user cold-start problem in recommender systems: A comparative review. **Information Systems**, Elsevier, v. 58, p. 87–104, 2016. ISSN 03064379. Disponível em: <<http://dx.doi.org/10.1016/j.is.2014.10.001>>.
- SOUZA, Ediana; LICHTNOW, Daniel; GASPARINI, Isabela. Estratégia de pós-processamento aplicada a um sistema de recomendação de artigos para a melhora da diversidade. In: . [S.l.: s.n.], 2022. p. 216–221.
- SOUZA, Ediana da Silva; LICHTNOW, Daniel; GASPARINI, Isabela. Estratégia de Pós-processamento Aplicada a um Sistema de Recomendação de Artigos para a Melhora da Diversidade. p. 216–221, 2022.
- SPARLING, E. Isaac; SEN, Shilad. Rating: How difficult is it? In: **Proceedings of the Fifth ACM Conference on Recommender Systems**. New York, NY, USA: Association for Computing Machinery, 2011. (RecSys '11), p. 149–156. ISBN 9781450306836. Disponível em: <<https://doi.org/10.1145/2043932.2043961>>.
- SUN, Yueming; ZHANG, Yi. Conversational recommender system. In: **The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval**. New York, NY, USA: Association for Computing Machinery, 2018. (SIGIR '18), p. 235–244. ISBN 9781450356572. Disponível em: <<https://doi.org/10.1145/3209978.3210002>>.
- TURNBULL, Doug. **BM25 The Next Generation of Lucene Relevance**. 2015. OpenSource Connections Blog. Disponível em: <<https://opensourceconnections.com/blog/2015/10/16/bm25-the-next-generation-of-lucene-relevance/>>.
- VELDE, Maarten van der et al. Alleviating the Cold Start Problem in Adaptive Learning using Data-Driven Difficulty Estimates. **Computational Brain and Behavior**, Computational Brain Behavior, v. 4, n. 2, p. 231–249, 2021. ISSN 2522087X.

- WALUNJ, V. et al. Smart tour advisor using machine learning and natural language processing. p. 53–57, 2022.
- WASLAWICK, Raul Sidnei. **Metodologia de Pesquisa para Ciência da Computação**. [S.l.: s.n.], 2014. v. 53. 160 p. ISSN 1098-6596. ISBN 9788578110796.
- WILLMOTT, Cort J. On The Validation of Models. **Physical Geography**, Taylor Francis, v. 2, n. 2, p. 184–194, 1981. Disponível em: <<https://doi.org/10.1080/02723646.1981.10642213>>.
- WILLMOTT, C. J. Some comments on the evaluation of model performance. **Bulletin - American Meteorological Society**, v. 63, n. 11, p. 1309–1313, 1982. ISSN 0003-0007.
- ZHANG, Zhipeng et al. Alleviating New User Cold-Start in User-Based Collaborative Filtering via Bipartite Network. **IEEE Transactions on Computational Social Systems**, IEEE, v. 7, n. 3, p. 672–685, 2020. ISSN 2329924X.
- ZHOUA, Tao et al. Solving the apparent diversity-accuracy dilemma of recommender systems. **Proceedings of the National Academy of Sciences of the United States of America**, v. 107, n. 10, p. 4511–4515, 2010. ISSN 00278424.

APÊNDICE A – LISTA DE EVENTOS

Tabela 14 – Eventos científicos relacionados à IHC

Nome do evento
Computer Human Interaction (CHI)
ACM Conference on Computer-Supported Cooperative Work
Social Computing
ACM Conference on Pervasive and Ubiquitous Computing (UbiComp)
ACM Symposium on User Interface Software and Technology
IEEE Transactions on Affective Computing
ACM/IEEE International Conference on Human Robot Interaction
International Journal of Human-Computer Studies
IEEE Transactions on Human-Machine Systems
Behaviour
Information Technology
Conference on Designing Interactive Systems
International Conference on Multimodal Interfaces (ICMI)
International Journal of Human-Computer Interaction
ACM Transactions on Computer-Human Interaction (TOCHI)
HCI International
Mobile HCI
IEEE Transactions on Haptics
International Conference on Intelligent User Interfaces (IUI)
International ACM Conference on Assistive Technologies (Assets)
International Conference on Tangible, Embedded, and Embodied Interaction
International Conference on User Modeling, Adaptation and Personalization (UMAP)
Foundations and Trends in Machine Learning
IEEE Transactions on Cybernetics
IEEE Transactions on Systems
International Journal of Computer-Supported Collaborative Learning
Science Technology and Human Values
Computers in Human Behavior
2015 IEEE Radar Conference - Proceedings
IET Control Theory and Applications
Topics in Cognitive Science
Cyberpsychology
International Journal of Intelligent Systems
ITCS 2016 - Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science
IEEE Transactions on Affective Computing
FPGA 2017 - Proceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays
Web Semantics
User Modelling and User-Adapted Interaction
MM 2015 - Proceedings of the 2015 ACM Multimedia Conference
Games and Culture
IEEE Transactions on Human-Machine Systems

Tabela 15 – Eventos científicos relacionados à IHC

Nome do evento
Proceedings - SCA 2015: 14th ACM SIGGRAPH / Eurographics Symposium on Computer Animation
Journal of Information Systems
IEEE Transactions on Haptics
Transactions on Interactive Intelligent Systems
Knowledge and Information Systems
AllThingsCellular 2015 - Proceedings of the 5th Workshop on All Things Cellular: Operations
International Journal of Human Computer Studies
Journal of Computational Design and Engineering
IEEE Transactions on Emerging Topics in Computing
SIGGRAPH Asia 2015 Courses
Proceedings of the Annual Conference of the International Speech Communication Association
Cognition
Eurographics Workshop on 3D Object Retrieval
Proceedings - 2016 2nd International Conference on Open and Big Data
ACM SoCC 2015 - Proceedings of the 6th ACM Symposium on Cloud Computing
First Monday
Technology
12th Workshop on Analytic Algorithmics and Combinatorics 2015
IEEE Haptics Symposium
IEEE Transactions on Computational Social Systems
Behaviour and Information Technology
ASIA CCS 2016 - Proceedings of the 11th ACM Asia Conference on Computer and Communications Security
Interaction Studies
Human-Computer Interaction
Computer Speech and Language
ACM Transactions on Computer-Human Interaction
Electronic Commerce Research
2015 International Symposium on Network Coding
IEEE-RAS International Conference on Humanoid Robots
International Journal of Human-Computer Interaction
ACM/IEEE International Conference on Human-Robot Interaction
MADiMa 2016 - Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management
SA 2016 - SIGGRAPH ASIA 2016 Courses
19th Power Systems Computation Conference
CPP 2017 - Proceedings of the 6th ACM SIGPLAN Conference on Certified Programs and Proofs
Computer Methods in Biomechanics and Biomedical Engineering
Proceedings of the 9th International Conference on Web and Social Media
International Journal of Child-Computer Interaction
Virtual Reality
2015 IEEE 16th Workshop on Control and Modeling for Power Electronics
Internet-QoE 2016 - Proceedings of the 2016 ACM SIGCOMM Workshop on QoE-Based Analysis and Management of Data Communication Networks
Proceedings - 2016 4th IEEE International Conference on Mobile Cloud Computing
HotMiddlebox 2015 - Proceedings of the 2015 ACM SIGCOMM Workshop on Hot Topics in Middleboxes and Network Function Virtualization
Proceedings - 2016 IEEE 20th International Enterprise Distributed Object Computing Conference
Proceedings of the Conference on Object-Oriented Programming Systems

Tabela 16 – Eventos científicos relacionados à IHC

Nome do evento
2015 IEEE Energy Conversion Congress and Exposition
Journal of Internet Commerce
IEEE Consumer Electronics Magazine
Foundations and Trends in Human-Computer Interaction
Computers and Graphics
ASE 2016 - Proceedings of the 31st IEEE/ACM International Conference on Automated Software Engineering
ACM Transactions on Accessible Computing
2015 13th International Symposium on Modeling and Optimization in Mobile Displays
ASM 2015 - Proceedings of the 1st International Workshop on Affect and Sentiment in Multimedia
2016 IEEE 7th International Symposium on Power Electronics for Distributed Generation Systems
Proceedings - International Workshop on Software Protection
2015 4th International Workshop on Optical Wireless Communications
Journal on Multimodal User Interfaces
HT 2015 - Proceedings of the 26th ACM Conference on Hypertext and Social Media
Proceedings of the 7th ACM SIGOPS Asia-Pacific Workshop on Systems
Universal Access in the Information Society
Advanced Robotics
International Journal of Intelligent Systems and Applications
Paladyn
2016 IEEE/ACM 24th International Symposium on Quality of Service
Proceedings of the International Conference on Compilers
2017 IEEE International Conference on Pervasive Computing and Communications
Entertainment Computing
Social Network Analysis and Mining
2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology
Eye Tracking Research and Applications Symposium (ETRA)
IEEE International Workshop on Machine Learning for Signal Processing
Proceedings of the ACM Conference on Computer Supported Cooperative Work
HotMobile 2017 - Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications
Conference on Human Factors in Computing Systems - Proceedings
Australasian Journal of Information Systems
ISWC 2015 - Proceedings of the 2015 ACM International Symposium on Wearable Computers
International Conference on Intelligent User Interfaces
Proceedings - 2015 3rd IEEE International Conference on Mobile Cloud Computing
Journal of Organizational and End User Computing
International Journal of Technology and Human Interaction
Presence: Teleoperators and Virtual Environments
2016 9th IAPR Workshop on Pattern Recognition in Remote Sensing
Proceedings - 2015 IEEE International Conference on Autonomous Robot Systems and Competitions
ISPASS 2016 - International Symposium on Performance Analysis of Systems and Software
Journal of Gaming
Virtual Worlds
2016 Science of Security for Cyber-Physical Systems Workshop
9th International Conference on Power Electronics - ECCE Asia: 'Green World with Power Electronics'
2015 IEEE International Workshop on Measurements and Networking

Tabela 17 – Eventos científicos relacionados à IHC

Nome do evento
International Conference on Energy Economics and Environment - 1st IEEE Uttar Pradesh Section Conference
IEEE Biomedical Circuits and Systems Conference: Engineering for Healthy Minds and Able Bodies
Proceedings of the International Workshop on Temporal Representation and Reasoning
International Journal of Computer Games Technology
L@S 2017 - Proceedings of the 4th (2017) ACM Conference on Learning at Scale
2016 Petroleum and Chemical Industry Technical Conference
2016 International Conference on Hardware/Software Codesign and System Synthesis
Proceedings of the ACM Conference on Computer Supported Cooperative Work
Interacting with Computers
Society of Petroleum Engineers - SPE Deepwater Drilling and Completions Conference
ACM International Conference Proceeding Series
Proceedings - IEEE International Workshop on Robot and Human Interactive Communication
Proceedings - 1st International Workshop on Bringing Architectural Design Thinking Into Developers' Daily Activities
CODASPY 2017 - Proceedings of the 7th ACM Conference on Data and Application Security and Privacy
Proceedings - 14th IEEE International Conference on Trust
SIAM Conference on Control and Its Applications 2015
Proceedings of the ACM Great Lakes Symposium on VLSI
ASME 2016 Internal Combustion Engine Fall Technical Conference
Interactions
Australasian Journal of Engineering Education
2015 IEEE International Ultrasonics Symposium
European Wireless Conference 2016
MobiWac 2016 - Proceedings of the 14th ACM International Symposium on Mobility Management and Wireless Access
Proceedings of 21st European Wireless Conference
Proceedings - 11th International Workshop on Automation of Software Test
2015 IEEE Conference on Control and Applications
Computational Social Networks
CONCREEP 2015: Mechanics and Physics of Creep
Journal of Visual Languages and Computing
ESAIR 2015 - Proceedings of the 2015 Workshop on Exploiting Semantic Annotations in Information Retrieval
Proceedings - 1st International Workshop on TEchnicaland LEgal Aspects of Data pRIVacy and Security
20th AIAA International Space Planes and Hypersonic Systems and Technologies Conference
BioWireleSS 2016 - Proceedings
24th Mediterranean Conference on Control and Automation
Intelligent Decision Technologies
2015 IEEE 2nd International Future Energy Electronics Conference
IEEE/ASME International Conference on Advanced Intelligent Mechatronics
Proceedings - 4th FME Workshop on Formal Methods in Software Engineering
ACM International Conference Proceeding Series
Journal of Interactive Learning Research
International Journal of Speech Technology
Proceedings - 23rd Euromicro International Conference on Parallel
2016 Australian Control Conference
Proceedings of the ASME Turbo Expo
ISARC 2016 - 33rd International Symposium on Automation and Robotics in Construction
ASME 2016 11th International Manufacturing Science and Engineering Conference

Tabela 18 – Eventos científicos relacionados à IHC

Nome do evento
Brazilian Symposium on Games and Digital Entertainment
Proceedings - 1st International Workshop on Complex Faults and Failures in Large Software Systems
2015 18th International Conference on Electrical Machines and Systems
2016 Picture Coding Symposium
OR58: The OR Society Annual Conference
Improving the Seismic Performance of Existing Buildings and other Structures 2015 - Proceedings of the 2nd ATC and SEI Conference on Improving the Seismic Performance of Existing Buildings and Other Structures
New Mathematics and Natural Computation
2016 12th International Conference on Network and Service Management
International Journal of Cognitive Informatics and Natural Intelligence
Proceedings of the Workshop on Advanced Visual Interfaces AVI
SYSTOR 2015 - Proceedings of the 8th ACM International Systems and Storage Conference
2016 IEEE International Conference on Renewable Energy Research and Applications
ACL-IJCNLP 2015 - 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing
IEEE International Symposium on Broadband Multimedia Systems and Broadcasting
Connection Science
Proceedings - Wireless Health 2015
2016 International Conference on Information Systems
2016 Principles
International Topical Meeting on Nuclear Reactor Thermal Hydraulics 2015
WearSys 2015 - Proceedings of the 2015 Workshop on Wearable Systems and Applications
MSCC 2016 - Proceedings of the 3rd MobiHoc International Workshop on Mobile Sensing
Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing
ENSsys 2015 - Proceedings of the 3rd International Workshop on Energy Harvesting and Energy Neutral Sensing Systems
2016 IEEE Symposium Series on Computational Intelligence
23rd European Conference on Information Systems
Proceedings - 2014 IEEE International Symposium on Multimedia
GECCO 2015 - Companion Publication of the 2015 Genetic and Evolutionary Computation Conference
Proceedings of the 13th IASTED International Conference on Biomedical Engineering
3DTV-Conference
CHIIR 2016 - Proceedings of the 2016 ACM Conference on Human Information Interaction and Retrieval
International Conference on Environmental Pollution and Public Health
Proceedings - 7th International Workshop on Modeling in Software Engineering
CCSW 2015 - Proceedings of the 7th ACM Cloud Computing Security Workshop
IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems
Journal of Information Science and Engineering
Proceedings of the 10th ACM Workshop on Programming Languages and Analysis for Security
Interaction Design and Architecture(s)
Proceedings - 2016 10th International Conference on Compatibility
Proceedings - 2015 International Conference on Cloud Computing and Big Data
24th European Conference on Information Systems
2015 IEEE International Symposium on Power Line Communications and Its Applications
Society of Petroleum Engineers - SPE Thermal Well Integrity and Design Symposium
2015 Texas Symposium on Wireless and Microwave Circuits and Systems
Proceedings - 9th International Workshop on Cooperative and Human Aspects of Software Engineering
Proceedings of the 2016 IEEE 1st International Conference on Data Stream Mining and Processing
River Flow - Proceedings of the International Conference on Fluvial Hydraulics

Tabela 19 – Eventos científicos relacionados à IHC

Nome do evento
2016 11th IEEE International Symposium on Industrial Embedded Systems
CHIIR 2017 - Proceedings of the 2017 Conference Human Information Interaction and Retrieval
Proceedings - 2015 Brazilian Symposium on Computing Systems Engineering
International Journal of Mobile Human Computer Interaction
2015 Asia-Pacific International Symposium on Electromagnetic Compatibility
Structures Congress 2017: Blast
3rd International Workshop on Exploratory Search in Databases and the Web
Journal of Advanced Computational Intelligence and Intelligent Informatics
Journal of Interactive Online Learning
ACM International Conference Proceeding Series
Proceedings - 12th International Conference on Intelligent Environments
Proceedings - IEEE 9th International Symposium on Embedded Multicore/Manycore SoCs
2015 International Conference on the Network of the Future
AI and Society
International Conference on Nuclear Engineering
1st IMEKO TC4 International Workshop on Metrology for Geotechnics
Advances in Human-Computer Interaction
2017 International Workshop on Antenna Technology: Small Antennas
Proceedings - 2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing
Webology
Proceedings - 2016 26th International Workshop on Power and Timing Modeling
WISCS 2016 - Proceedings of the 2016 ACM Workshop on Information Sharing and Collaborative Security
28th International Technical Meeting of the Satellite Division of the Institute of Navigation
Proceedings of 2016 International Conference on Electrical and Information Technologies
Proceedings of SEHPCCSE 2015: 3rd International Workshop on Software Engineering for High Performance Computing in Computational Science and Engineering - Held in conjunction with SC 2015: The International Conference for High Performance Computing
ACADEMICMINDTREK 2015 - Proceedings of the 19th International Academic Mindtrek Conference
ICISSP 2015 - 1st International Conference on Information Systems Security and Privacy
Journal of Internet Banking and Commerce
Digital Creativity
Proceedings of LLVM-HPC 2016: The 3rd Workshop on the LLVM Compiler Infrastructure in HPC - Held in conjunction with SC 2016: The International Conference for High Performance Computing
Proceedings of the 11th IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis
Proceedings of 2015 International Conference on Information and Communication Technology and Systems
Proceedings - 2015 10th Asia Joint Conference on Information Security
Proceedings of 2016 13th International Conference on Remote Engineering and Virtual Instrumentation
Proceedings of the European Conference on Games-based Learning
TVX 2015 - Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video
Human IT
13th HONET-ICT International Symposium on Smart MicroGrids for Sustainable Energy Sources Enabled by Photonics and IoT Sensors
2015 IEEE 7th International Conference on Intelligent Computing and Information Systems
Proceedings of the 3rd International Conference on Applied Robotics for the Power Industry
2015 IEEE/SICE International Symposium on System Integration
Human Technology
Proceeding - 2015 IEEE International Conference on Antenna Measurements and Applications

Tabela 20 – Eventos científicos relacionados à IHC

Nome do evento
Proceedings - 2015 IEEE/ACM 8th International Symposium on Software and Systems Traceability
2015 International Conference on IC Design and Technology
Conference Proceedings - 2014 IEEE Games
Proceedings - 2014 3rd International Conference on User Science and Engineering: Experience. Engineer. Engage
WCTE 2016 - World Conference on Timber Engineering
2015 36th IEEE Sarnoff Symposium
Proceedings of 2015 International Conference on Data and Software Engineering
Qwerty
2015 IEEE Online Conference on Green Communications
ACM IMCOM 2016: Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication
IEEE Proceedings of the 2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids
Proceedings - International Conference on Machine Learning and Cybernetics
Proceedings of IEEE International Conference on Technological Advancements in Power and Energy
78th EAGE Conference and Exhibition 2016: Efficient Use of Technology - Unlocking Potential
Proceedings - 2015 10th IEEE International Conference on Design and Technology of Integrated Systems in Nanoscale Era
2016 39th International Conference on Telecommunications and Signal Processing
Conference on Human Factors in Computing Systems - Proceedings
ICSPCC 2016 - IEEE International Conference on Signal Processing
International Symposium on Wireless Personal Multimedia Communications
Proceedings of the 2015 7th IEEE International Conference on Cybernetics and Intelligent Systems
IEEE Symposium on Large Data Analysis and Visualization 2015
3rd International Workshop on Software Development Lifecycle for Mobile
45th AIAA Thermophysics Conference
Proceedings of 2016 8th International Conference on Information Technology and Electrical Engineering: Empowering Technology for Better Future
South African Computer Journal
2016 IEEE International Conference on Mechatronics and Automation
Proceedings - 15th International Conference on Computational Science and Its Applications
Society of Petroleum Engineers - SPE Saudi Arabia Section Annual Technical Symposium and Exhibition
2016 International Conference on Advanced Computer Science and Information Systems
ACM International Conference Proceeding Series
DSM 2016 - Proceedings of the International Workshop on Domain-Specific Modeling
MA3HMI 2016 - Proceedings of the Workshop on Multimodal Analyses Enabling Artificial Agents in Human-Machine Interaction
Proceedings of the SPE/IADC Middle East Drilling Technology Conference and Exhibition
2015 IEEE International Conference on Peer-to-Peer Computing
2015 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision
IEEE SSCI 2014 - 2014 IEEE Symposium Series on Computational Intelligence - CIMSIVP 2014: 2014 IEEE Symposium on Computational Intelligence for Multimedia
Proceedings of the 2016 6th International Conference on System Engineering and Technology
2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence
Fujitsu Scientific and Technical Journal
Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct
Proceedings of the 6th International Congress on the Science and Technology of Steelmaking
3rd International Winter Conference on Brain-Computer Interface

Tabela 21 – Eventos científicos relacionados à IHC

Nome do evento
Maritime-Port Technology and Development - Proceedings of the International Conference on Maritime and Port Technology and Development
Near Surface Geoscience 2015 - 21st European Meeting of Environmental and Engineering Geophysics
Proceedings of the 6th International Workshop on Runtime and Operating Systems for Supercomputers
International Journal of Automation and Smart Technology
Proceedings - 2017 2nd International Workshop on Science of Smart City Operations and Platforms Engineering
Proceedings of the 2016 Workshop on Multimodal Virtual and Augmented Reality
LTA 2016 - Proceedings of the 1st Workshop on Lifelogging Tools and Applications
Proceedings of IAMG 2015 - 17th Annual Conference of the International Association for Mathematical Geosciences
Proceedings of the 1st ACM SIGMOD Workshop on Network Data Analytics
Near Surface Geoscience 2015 - 21st European Meeting of Environmental and Engineering Geophysics
Proceeding - 2016 2nd International Conference on Science in Information Technology
Proceedings - 2015 8th International Conference on BioMedical Engineering and Informatics
Proceedings - 2015 9th International Conference on Complex
2016 IEEE 11th Annual International Conference on Nano/Micro Engineered and Molecular Systems
Proceedings - 2015 3rd International Workshop on RFID and Adaptive Wireless Sensor Networks
Proceedings of 2016 Prognostics and System Health Management Conference
Smart Grid Conference
TEI 2015 - Proceedings of the 9th International Conference on Tangible
1st International Conference on Emerging Trends in Engineering
2016 IEEE-NPSS Real Time Conference
Proceedings - 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics
Proceedings of the 13th International Conference on Telecommunications
2016 IEEE Conference on Quality Management
IEEE Conference on Computational Intelligence and Games
Proceedings - 2016 IEEE 25th North Atlantic Test Workshop
Proceedings - 2017 International Conference on Communication
2016 8th International Conference on Knowledge and Smart Technology
2017 Seminar on Detection Systems Architectures and Technologies
3rd International Workshop on Simulation for Energy
ACM International Conference Proceeding Series
CSSE 2015 - 20th International Symposium on Computer Science and Software Engineering
EAN 2016 - 54th International Conference on Experimental Stress Analysis
Proceedings of the International Conference on Cloud Security Management
2015 21st International Conference on Automation and Computing: Automation
ASME 2015 Gas Turbine India Conference
Proceedings of IEEE International Conference on Teaching
UIST 2016 Adjunct - Proceedings of the 29th Annual Symposium on User Interface Software and Technology
Proceeding - 2016 International Conference on Radar
Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work
Signal Processing in Photonic Communications
2015 2nd Asia-Pacific World Congress on Computer Science and Engineering
2nd Workshop on Advancements in Social Signal Processing for Multimodal Interaction 2016

Tabela 22 – Eventos científicos relacionados à IHC

Nome do evento
6th Research in Engineering Education Symposium: Translating Research into Practice
Proceedings - 2015 5th International Conference on Advances in Computing and Communications
TMS Annual Meeting
2015 International Conference on Control
2016 13th International Conference on Synthesis
ISPTS 2015 - 2nd International Symposium on Physics and Technology of Sensors: Dive Deep Into Sensors
Proceedings - 2015 12th Web Information System and Application Conference
Proceedings of the 2016 IEEE North West Russia Section Young Researchers in Electrical and Electronic Engineering Conference
2015 3rd International Conference on Electric Power Equipment - Switching Technology
2015 International Conference on Logistics
32nd International Symposium on Automation and Robotics in Construction and Mining: Connected to the Future
33rd AIAA Applied Aerodynamics Conference
ENBENG 2017 - 5th Portuguese Meeting on Bioengineering
Proceedings - 2015 IEEE International Electric Machines and Drives Conference
Proceedings - 23rd Euromicro International Conference on Parallel
Proceedings of 2014 International Conference on Interactive Collaborative Learning
Proceedings of 2015 IEEE International Conference on Power
Sensors 2016 - Topical Conference at the 2016 AIChE Annual Meeting
VISIGRAPP 2016 - Proceedings of the 11th Joint Conference on Computer Vision
Proceedings of 2016 International Conference on ICT in Business
Proceedings of the 28th Australian Computer-Human Interaction Conference
2015 International Conference on 'Stability and Control Processes' in Memory of V.I. Zubov
2015 Power Generation Systems and Renewable Energy Technologies
2017 IEEE 37th International Conference on Electronics and Nanotechnology
International Conference on Research and Innovation in Information Systems
Proceedings - 2015 4th Mediterranean Conference on Embedded Computing
Recent Progress in Steel and Composite Structures - Proceedings of the 13th International Conference on Metal Structures
ACM International Conference Proceeding Series
ICOQN 2016 - 2016 15th International Conference on Optical Communications and Networks
MMHealth 2016 - Proceedings of the 2016 ACM Workshop on Multimedia for Personal Health and Health Care
Proceedings - 2015 IEEE International Conference on System Engineering and Technology
Proceedings - 2015 International Conference on Developments in eSystems Engineering
Proceedings of the 53rd Annual Rocky Mountain Bioengineering Symposium
ValueTools 2016 - 10th EAI International Conference on Performance Evaluation Methodologies and Tools
2015 IEEE 7th National Conference on Computing
2017 4th International Conference on Industrial Engineering and Applications
23rd International Conference in Central Europe on Computer Graphics
Proceedings - 2015 International Conference on Computers
TMP 2016 - 5th International Conference on ThermoMechanical Processing
CrossTalk
Distillation Symposium 2016 - Topical Conference at the 2016 AIChE Spring Meeting and 12th Global Congress on Process Safety
IEEE SSCI 2014 - 2014 IEEE Symposium Series on Computational Intelligence - CIHLI 2014: 2014 IEEE Symposium on Computational Intelligence for Human-Like Intelligence
SA 2016 - SIGGRAPH ASIA 2016 Symposium on Education

Tabela 23 – Eventos científicos relacionados à IHC

Nome do evento
13th International Industrial Simulation Conference 2015
19th International Symposium on VLSI Design and Test
ACM International Conference Proceeding Series
Proceedings - 2016 International Symposium on Educational Technology
Proceedings of the International Conference on Intellectual Capital
2015 International Symposium on Mathematical Sciences and Computing Research
41st International Computer Music Conference
Proceedings - 2015 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology
Proceedings - 2015 International Conference on Platform Technology and Service
Proceedings of 2015 International Conference on Asian Language Processing
Proceedings of the 2016 3rd International Conference on Information and Communication Technologies for Disaster Management
2015 12th International Conference on Electrical Engineering
2016 25th International Scientific Conference Electronics
2016 9th International Conference on Power Drives Systems
2016 International Conference on Communication Systems and Networks
2016 International Conference on VLSI Systems
International Journal of Imaging and Robotics
IPAS 2016 - 2nd International Image Processing
Proceedings - 2nd SBR Brazilian Robotics Symposium
Proceedings of PAW 2016: 1st PGAS Applications Workshop - Held in conjunction with SC 2016: The International Conference for High Performance Computing
2014 International Conference on Energy Systems and Policies
2016 IEEE Colombian Conference on Communications and Computing
New Review of Information Networking
WTMC 2016 - Proceedings of the 2016 ACM International Workshop on Traffic Measurements for Cybersecurity
Proceedings - 2015 7th International Conference on Information Technology in Medicine and Education
5th International Conference on Manipulation
Advanced Materials
Information and Control
Proceedings - 2015 3rd International Conference on Computer
Proceedings of INDIS 2015
TEI 2017 - Proceedings of the 11th International Conference on Tangible
2016 5th International Symposium on Next-Generation Electronics
Archiving 2015 - Final Program and Proceedings
Proceedings of the 2017 3rd Conference on Mobile and Secure Services
Workshop on Specialty Optical Fibers and their Applications
Proceeding - 2016 International Seminar on Sensors
Proceedings of the 2016 International Symposium on Antennas and Propagation
Proceedings of the International Display Workshops
SKY 2016 - 7th International Workshop on Software Knowledge
2015 International Conference on Computing
57th AIAA/ASCE/AHS/ASC Structures
Annual Workshop on Network and Systems Support for Games
Electrical Transmission and Substation Structures 2015: Technical Challenges and Innovative Solutions in Grid Modernization - Proceedings of the 2015 Electrical Transmission and Substation Structures Conference

Tabela 24 – Eventos científicos relacionados à IHC

Nome do evento
Proceedings - 2015 IEEE Conference on System
Proceedings - 2016 International Conference on Robots and Intelligent System
2015 13th International Conference on Engineering of Modern Electric Systems
2016 11th France-Japan and 9th Europe-Asia Congress on Mechatronics
2016 International Conference on Robotics and Automation Engineering
2016 Open Conference of Electrical
2017 IEEE International Symposium on Power Line Communications and its Applications
CS MANTECH 2015 - 2015 International Conference on Compound Semiconductor Manufacturing Technology
SA 2016 - SIGGRAPH ASIA 2016 Crafting Spirituality: A Pedagogic Project for Digital Heritage: Digital Sculpting
13th International Conference on Industrial Logistics
20th International Conference on Miniaturized Systems for Chemistry and Life Sciences
9th International Conference on Simulation and Modelling in the Food and Bio-Industry 2016
Computational Science and Engineering - Proceedings of the International Conference on Computational Science and Engineering
Foundations and Frontiers in Computer
SA 2016 - SIGGRAPH ASIA 2016 Symposium on Education: Talks
SIGGRAPH Asia 2015 Mobile Graphics and Interactive Applications
The Path Forward for Wood Products: A Global Perspective - Proceedings of Scientific Papers
2015 IEEE Colombian Conference on Communications and Computing
2016 2nd International Conference on Open Source Software Computing
2016 6th International Workshop on Computer Science and Engineering
Geophysics 2015 - 11th EAGE International Scientific and Practical Conference and Exhibition on Engineering and Mining Geophysics
Middleware 2016 Doctoral Symposium - Colocated with ACM/IFIP/USENIX Middleware 2016
Moscow University Computational Mathematics and Cybernetics
Proceedings - 2016 IEEE International Conference on Cloud Computing in Emerging Markets
Proceedings of the 1st Workshop on Middleware for Edge Clouds and Cloudlets
6th International Defense and Homeland Security Simulation Workshop
IEEE SSCI 2014: 2014 IEEE Symposium Series on Computational Intelligence - RiiSS 2014: 2014 IEEE Symposium on Robotic Intelligence in Informationally Structured Space
Proceedings of the 18th International Conference on Engineering and Product Design Education: Design Education: Collaboration and Cross-Disciplinarity
2016 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications
ACM International Conference Proceeding Series
ACM SIGGRAPH Art Papers
Bioinformatics and Biomedical Engineering - Proceedings of the 9th International Conference on Bioinformatics and Biomedical Engineering
Proceedings of the 7th ACM International Workshop on Massively Multiuser Virtual Environments
Proceedings of the 8th International Workshop on Context-Oriented Programming
14th IMEKO TC10 Workshop on Technical Diagnostics 2016: New Perspectives in Measurements
2016 IEEE Global Electromagnetic Compatibility Conference
21st Offshore Symposium 2016: Emerging Offshore Technology and Deepwater Trends
ACM SIGGRAPH 2016 Appy Hour
American Fuel and Petrochemical Manufacturers
Innovations of Green Process Engineering for Sustainable Energy and Environment 2016 - Topical Conference at the 2016 AIChE Annual Meeting
International Conference on Concept Lattices and Their Applications
Proceedings of the 11th Workshop on Implementation
2015 IEEE Brasil RFID

Tabela 25 – Eventos científicos relacionados à IHC

Nome do evento
24th International Conference in Central Europe on Computer Graphics
American Fuel and Petrochemical Manufacturers
IEEE Symposium on Low-Power and High-Speed Chips
IMAPS 12th International Conference and Exhibition on Device Packaging
Proceedings of the 9th International Conference on Standardization and Innovation in Information Technology
Pulping
WINSYS 2013 - 10th International Conference on Wireless Information Networks and Systems
30th Center for Chemical Process Safety International Conference 2015 - Topical Conference at the 2015 AIChE Spring Meeting and 11th Global Congress on Process Safety
4th International Workshop on Innovative Simulation for Health Care
ACM International Conference Proceeding Series
Energy Transport Processes 2016 - Core Programming Area at the 2016 AIChE Spring Meeting and 12th Global Congress on Process Safety
ICGA Journal
Innovations in Process Research and Development 2016 - Topical Conference at the 2016 AIChE Spring Meeting and 12th Global Congress on Process Safety
Proceedings ACM SIGUCCS User Services Conference
Proceedings of the 2016 7th International SpaceWire Conference
Proceedings of the 6th IASTED International Conference on Modelling and Simulation
Proceedings of the Technical Association of the Graphic Arts
29th International Conference on Computer Applications in Industry and Engineering
AIMCAL Web Coating and Handling Conference 2015
American Fuel and Petrochemical Manufacturers
Computer Science and Applications - Proceedings of the Asia-Pacific Conference on Computer Science and Applications
Conference on Nuclear Training and Education 2015
EI Electrical Insulation Conference Materials and Application
Environmental Protection and Sustainable Ecological Development - Proceedings of the Environmental Protection and Sustainable Ecological Development
imPACT 2016 - Internet
Liaison Functions 2016 - Core Programming Area at the 2016 AIChE Spring Meeting and 12th Global Congress on Process Safety
Proceedings of the 14th International Conference on Nuclear Reaction Mechanisms
Technoetic Arts: a journal of speculative research
Tissue World New Orleans Conference 2016: Beyond Sustainability - Leading a Culture of Innovation and Responsible Production
ACM Conference on Human Factors in Computing Systems
ACM Conference on Computer-Supported Cooperative Work and Social Computing
IFIP International Conference on Human-Computer Interaction
International ACM SIGACCESS Conference on Computers and Accessibility
ACM Intelligent User Interfaces
ACM MobileHCI
ACM Symposium on User Interface Software and Technology
Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais
IEEE Symposium on 3D User Interfaces
IEEE Symposium on Visual Languages and Human-Centric Computing
ACM Designing Interactive Systems
ACM International Conference on Tangible, Embedded and Embodied Interactions
ACM International Conference on Human Robot Interaction
ACM International Conference on Multimodal Interfaces

Tabela 26 – Eventos científicos relacionados à IHC

Nome do evento
ACM Conference on Pervasive and Ubiquitous Computing (UbiComp)
ACM Interaction Design and Children
IEEE International Symposium on Robot and Human Interactive Communication
Engineering Interactive Computing Systems
ACM Conference on Recommender Systems
Computer-Human Interaction of Australia
IEEE International Conference on Computer Supported Cooperative Work in Design
International Conference on Human-Computer Interaction
Interactive Tabletops and Surfaces
Conference of the British Computer Society Human Computer Interaction Specialist Group
International Conference on Collaboration and Technology
Nordic forum for Human-Computer Interaction
Participatory Design Conference
IHM - Conférence Francophone sur l'Interaction Homme-Machin
ACM International Conference on Design of Communication
CHI PLAY
Association for the Advancement of Assistive Technology in Europe Conference
User Modeling, Adaptation and Personalization Conference
IARIA International Conference on Advances in Computer-Human Interactions
ACM Symposium on Applied Computing
Symposium on Spatial User Interaction
Graphics Interface
IEEE International Conference on Virtual Environments Human-Computer Interfaces and Measurement Systems
ACM International Conference on Interactive Surfaces and Spaces
International Working Conference on Advanced Visual Interfaces
Creativity
Cognition
Hawaii International Conference on System Sciences
ACM SIGGRAPH
ACM International Conference on Interactive Experiences for TV and Online Video
International Conference on Electronic Government
ACM Symposium on Virtual Reality Software and Technology
SBC Simpósio Brasileiro de Sistemas Multimídia e Web
ACM International Conference on PErvasive Technologies Related to Assistive Environments
Simpósio Brasileiro de Informática na Educação
SBC Symposium on Virtual and Augmented Reality
IEEE International Symposium on Mixed and Augmented Reality
Annual International Conference on Digital Government Research
International Conference on Computer Supported Education
IEEE Working Conference on Software Visualization
IEEE International Conference on Advanced Learning Technologies
International Conference on Software Development for Enhancing Accessibility and Fighting Info-exclusion
Annual International Computers, Software Applications Conference

Tabela 27 – Eventos científicos relacionados à IHC

Nome do evento
ACM Conference on Hypertext and Social Media
Workshop de Informática na Escola
ACM International Conference on Multimodal Interaction
Latin American Conference on Human Computer Interaction
International Conference on Advances in Computer Entertainment Technology
International Conference on Information Visualisation
Augmented Human International Conference
Simpósio Brasileiro de Sistemas de Informação
International Conference on Enterprise Information Systems
International Conference on Intelligent Tutoring Systems
Seminário Integrado de Software e Hardware
ACM Multimedia Systems Conference
Workshop sobre Aspectos da Interação Humano-Computador para a Web Social
IFIP International Conference on Informatics and Semiotics in Organisations
International Conference on Web Information Systems and Technologies
International Cross-Disciplinary Conference on Web Accessibility
AAAI Conference on Human Computation and Crowdsourcing
Eurohaptics
Conferencia Latinoamericana En Informatica
Congresso Internacional de Ergonomia e Usabilidade de Interfaces Humano Computador
Congresso da Sociedade Brasileira de Computação
Affective Computing and Intelligent Interaction
International Conference on Computers Helping People with Special Needs
IEEE Haptics Symposium
Simpósio Brasileiro de Jogos e Entretenimento Digital
International Conference Universal Access in Human-Computer Interaction
International Conference on Intelligent Virtual Agents
ACM Symposium on Document Engineering
Artificial Intelligence in Education
Conference on Sustainable Internet and ICT for Sustainability
Workshop on Internet of Things for Ambient Assisted Living
Electronic Government and the Information Systems Perspective
Latin American Web Congress
International Working Conference on Human-Centered Software Engineering
Symposium on Interactive 3D Graphics and Games
Interaction Latin America
International Symposium on Electronic Arts
Congresso Brasileiro de Pesquisadores em Cibercultura
Congresso Internacional de Design da Informação

Tabela 28 – Eventos científicos relacionados à IHC

Nome do evento
International Journal of Human-Computer Studies
ACM Transactions on Computer-Human Interaction
Personal and Ubiquitous Computing
ACM Interactions
Behavior
Information Technology
Human–Computer Interaction
Computer Supported Cooperative Work
Universal Access in the Information Society
Journal on Multimodal User Interfaces
Interacting with Computers
Advances in Human-Computer Interaction
SBC Journal on Interactive Systems
ACM Transactions on Accessible Computing
International Journal of Mobile Human Computer Interaction
Foundations and Trends in Human-Computer Interaction
International Journal of Child-Computer Interaction
Proceedings of the ACM on Human-Computer Interaction
Computers and Education
Computers in Human Behavior
Journal of Systems and Software
APPLIED ERGONOMICS
Design Studies
Journal of Visual Languages and Computing
IEEE Transactions on Systems, Man and Cybernetics
IEEE Transactions on Human-Machine Systems
Elsevier COMPUTER STANDARDS INTERFACES
Entertainment Computing
AI SOCIETY
User Modeling and User-Adapted Interaction
Revista Brasileira de Informática na Educação
Leonardo - MIT press
Digital Creativity
Scientometrics
Elsevier Computers Graphics
WORLD WIDE WEB (BUSSUM)
Online Information Review
IEEE TRANSACTIONS ON EDUCATION
IEEE Transactions on Learning Technologies
Simulation
Gaming
Revista Eletrônica de Comunicação, Informação e Inovação em Saúde
iSys - Revista Brasileira de Sistemas de Informação
JOURNAL OF THE BRAZILIAN COMPUTER SOCIETY (ONLINE)
REVISTA DE INFORMÁTICA TEÓRICA E APLICADA: RITA

Tabela 29 – Eventos científicos relacionados à IHC

Nome do evento
Revista de Sistemas e Computação
Computer Communications
Elsevier Cognition
ACM Transactions on Graphics
HUMAN MUTATION
Computer Graphics Forum
IEEE Transactions on Visualization and Computer Graphics
IEEE Transactions on Mobile Computing
IEEE Transactions on Affective Computing
Ecological Informatics
HUMAN FACTORS
The Visual Computer: International Journal of Computer Graphics
IEEE Computer Graphics and Applications
JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING
Springer Cognition, Technology & Work
IEEE Transactions on Haptics
Transactions on Interactive Intelligent Systems
Games for Health: Research, Development, and Clinical Applications
ACM Transactions on Autonomous and Adaptive Systems
International Journal of Computer Applications in Technology
JOURNAL OF USABILITY STUDIES
JMIR Mhealth Uhealth
JMIR Serious Games
HUMAN COMPUTATION
International Journal of Cultural and Creative Industries
JMIR HUMAN FACTORS
ACM Transactions on Computer Systems
International Journal of Computer Games Technology
INTERNATIONAL JOURNAL OF PERVASIVE COMPUTING AND COMMUNICATIONS (PRINT)
Presence - MIT press

APÊNDICE B – QUESTIONÁRIO - PREFERÊNCIA DE INTERFACES

Questionário de Perfil

X ...

B I U ↵ X

Esta reunião é destinada para uma sessão de brainstorming e grupo focal para discutirmos sobre as interfaces apresentadas, observamos suas diferenças e avaliarmos as preferências e a satisfação do usuário. As duas interfaces apresentadas são **em pares** ou **em lista**. As interfaces apresentadas são de baixa fidelidade, proposta para o trabalho de mestrado da aluna Nathália Locatelli Cesar, estudante do PPGCAP da UDESC.

Pedimos que preencham seus dados de perfil.

Idade: *

- Até 18 anos
- De 19 até 22 anos
- De 23 até 26 anos
- De 27 anos ou mais

Gênero: *

- Feminino
- Masculino
- Prefiro não dizer

Minha área de atuação: *

Texto de resposta curta

Já fez a disciplina de IHC? Na graduação ou pós-graduação? *

Texto de resposta longa

Sobre a reunião. Você estava na sessão via qual dispositivo? Viu as telas com quais dispositivo?

- Celular/smarthphone
- Computador
- Tablet

Sobre as interfaces apresentadas na reunião com o grupo, qual você mais gostou? Quais as características e especificidades que fazem você ter escolhido tal interface?

Texto de resposta longa

Diga suas sugestões e comentários sobre as interfaces: *

Texto de resposta longa

Diga suas sugestões e comentários sobre a reunião:

Texto de resposta longa

APÊNDICE C – QUESTIONÁRIO - TESTES FUNCIONAIS

Testes funcionais - Paperflix

B **I** **U**  

Convidamos você para avaliar as funcionalidades do Sistema de Recomendação de artigos - Paperflix. Este sistema tem como intuito criar o perfil do usuário sem nenhum tipo de informação prévia e avaliar a criação do perfil através de pares de itens. O sistema faz parte do trabalho da mestrandia em Computação Aplicada Nathália Locatelli Cezar da Universidade do Estado de Santa Catarina (UDESC) e orientada pela professora Isabela Gasparini (UDESC) e pelo professor Daniel Lichnow (UFSM).

Em anexo ao e-mail que você recebeu, há um arquivo Instruções , em que estão descritas as principais funcionalidades do sistema. Você deverá realizá-las e detalhar o seu parecer reportando as suas percepções, dificuldades ou erros encontrados.

E-mail *

E-mail válido

Informe a data da avaliação:

Mês, dia, ano



Informe o horário de início da avaliação:

Horário



Informe o horário final da avaliação:

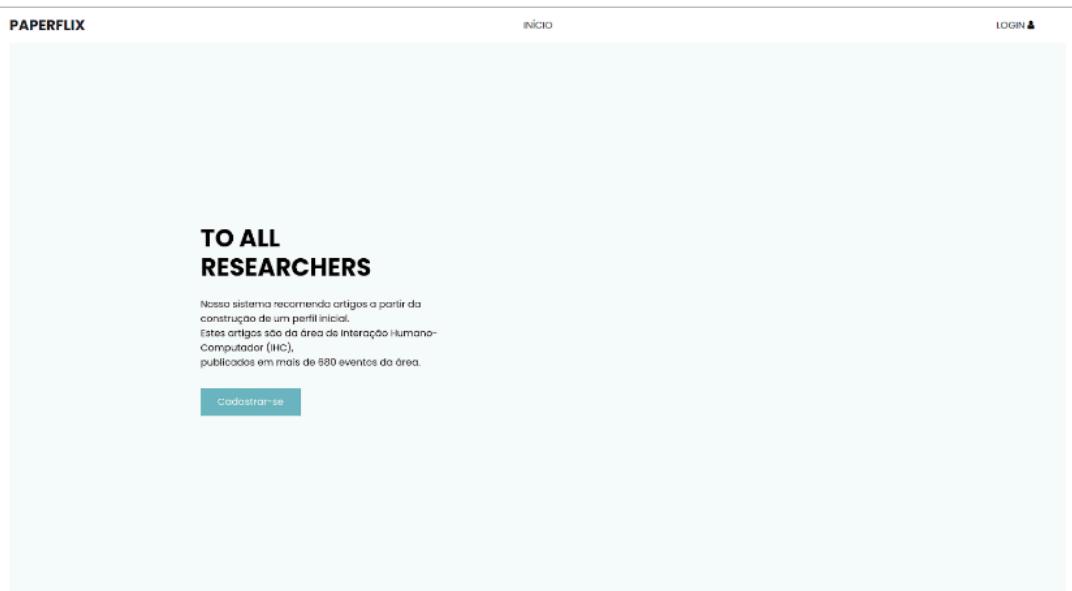
Horário



Descreva as configurações do seu dispositivo para o teste (Ex: PC x smartphone, sistema operacional, resolução de tela, especificações da máquina, etc.)

Texto de resposta longa

Tela inicial



TO ALL RESEARCHERS

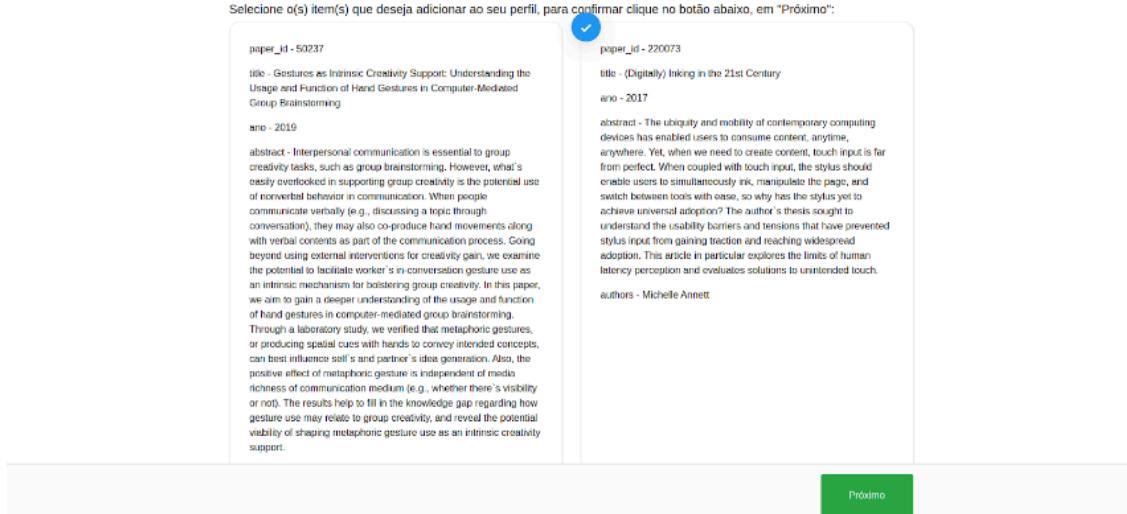
Nossa sistema recomenda artigos a partir da construção de um perfil inicial. Estes artigos são da área de Interação Humano-Computador (IHC), publicados em mais de 680 eventos da área.

[Cadastrarse](#)

Sobre a tela inicial: qual o seu parecer em relação as suas percepções, dificuldades ou erros encontrados? *

Texto de resposta longa

Tela seleção de pares



Selecione o(s) item(s) que deseja adicionar ao seu perfil, para confirmar clique no botão abaixo, em "Próximo":

<input checked="" type="checkbox"/> paper_id - 50237	<input type="checkbox"/> paper_id - 220073
title - Gestures as Intrinsic Creativity Support: Understanding the Usage and Function of Hand Gestures in Computer-Mediated Group Brainstorming	title - (Digitally) Inking in the 21st Century
ano - 2019	ano - 2017
abstract - Interpersonal communication is essential to group creativity tasks, such as group brainstorming. However, what's easily overlooked in supporting group creativity is the potential use of nonverbal behavior in communication. When people communicate verbally (e.g., discussing a topic through conversation), they may also co-produce hand movements along with verbal contents as part of the communication process. Going beyond using external interventions for creativity gain, we examine the potential to facilitate worker's in-conversation gesture use as an intrinsic mechanism for bolstering group creativity. In this paper, we aim to gain a deeper understanding of the usage and function of hand gestures in computer-mediated group brainstorming. Through a laboratory study, we verified that metaphoric gestures, or producing spatial cues with hands to convey intended concepts, can best influence self's and partner's idea generation. Also, the positive effect of metaphoric gesture is independent of media richness of communication medium (e.g., whether there's visibility or not). The results help to fill in the knowledge gap regarding how gesture use may relate to group creativity, and reveal the potential viability of shaping metaphoric gesture use as an intrinsic creativity support.	abstract - The ubiquity and mobility of contemporary computing devices has enabled users to consume content, anytime, anywhere. Yet, when we need to create content, touch input is far from perfect. When coupled with touch input, the stylus should enable users to simultaneously ink, manipulate the page, and switch between tools with ease, so why has the stylus yet to achieve universal adoption? The author's thesis sought to understand the usability barriers and tensions that have prevented stylus input from gaining traction and reaching widespread adoption. This article in particular explores the limits of human latency perception and evaluates solutions to unintended touch.
authors - Michelle Annett	

[Próximo](#)

Sobre a tela de seleção em pares para criação de perfil: qual o seu parecer em relação as suas percepções, dificuldades ou erros encontrados? *

Texto de resposta longa

Tela de Recomendações

Seu perfil

ID: 4788

Título: Design of a Brazilian Portuguese Virtual Keyboard for People with Severe Motor Disability

Resumo: Virtual keyboards represent an important tool in the Assistive Technology area as they can assist in computer interaction by people with severe motor disability. Depending on the interaction method used, a differentiated arrangement of keys can speed up the task of typing a text. This work proposes a virtual keyboard layout which allows people with disabilities to write texts in Brazilian Portuguese language efficiently. The method of human-computer interaction used is a webcam-based cursor controller limited by continuous vertical and horizontal movements. Keys arrangement is based on digraphs frequency of occurrence in a corpus of Portuguese language. The performance of the virtual keyboard was evaluated by two tests. Firstly, a function which estimates the movement efficiency when typing a key is used to compare the layout with other solutions. Proposed layout presented typing speed faster than QWERTY (41%), OPTI (9%) and FITALY (9%). In user evaluation, users have performed the tasks significantly faster with the proposed keyboard layout.

Ano: 2014

ID: 4522

Título: Virtual Bicycle: An User Interface Tool for Navigation in Virtual Worlds

Resumo: One of the most complicated tasks when working with three-dimensional virtual worlds is the navigation process. Usually, this process requires the use of buttons and key-sequences and the development of interaction metaphors which frequently makes the interaction process difficult and inefficient. In these environments, very simple tasks, like look upward and downward can become extremely complicated. To overcome these obstacles, this work presents an interaction model for three-dimensional virtual worlds, based on the interpretation of the natural gestures of a real user while he/she is walking in a real world. This model is an example of a non-WIMP (Window, Icon, Menu, Pointer) interface. To test this model we created a device named virtual-bike. With this device, the user can navigate through the virtual environment exactly as if he were riding a real bike.

Ano: 1999

ID: 4938

Título: A Participatory Method for Multidimensional Analysis and Proposing Improvements in Collaborative Systems

Tela de Recomendações - fim da lista

ID: 4857

Título: Acessibilidade de Dispositivos Móveis: Mapeamento de Problemas e Estudo de Gestos de Toque

Resumo: Mobile devices can be a valuable in improving the quality of life of visually impaired people, enabling greater independence to perform daily tasks. A review was carried out to map the existing accessibility problems in the interaction of people with visual impairment with mobile devices. Seven groups of problems were found, which allowed establishing a set of recommendations and future studies to improve accessibility on mobile devices. A case study was planned and conducted addressing the group "gesture-based interaction", which had the highest number of problems in the literature review. The empirical study aimed to evaluate touch gestures made by 12 visually impaired people by using an application developed for this purpose. Gestures represented by the letters "V" and "Z" had performance similar to those already available on the Android Platform. Other factors have influenced the recognition with speed being the main factor in reducing the performance of gestures.

Ano: 2016

ID: 4995

Título: Intermediated Semiotic Inspection Method

Resumo: Semiotic Engineering is a consolidated HCI theory that perceives an interactive system's interface as a designer-to-user communication. In this context, the relevant quality property to be considered is communicability. Currently, there are two consolidated methods to evaluate the communicability of a system, Communicability Evaluation Method (CEM) and Semiotic Inspection Method (SIM). In this work, our purpose is to formally present the Intermediated Semiotic Inspection Method - ISIM, as a method to evaluate communicability. The ISIM was proposed to allow for the communicability evaluation from the perspective of indirect users. It was described within the context of educational systems, in a case study to analyze the professor's perspective on systems aimed at their students. However, ISIM has been more broadly used in different contexts, and with different participants. In this article our goal is to formally present ISIM, describing the necessary steps to its application. We also present an analysis of the studies that have applied the method. Based on our analysis we discuss how consolidated the ISIM is, and the different aspects it can evaluate in comparison to SIM and CEM.

Ano: 2018

Em relação a tela de Recomendações: qual o seu parecer em relação as suas percepções, dificuldades ou erros encontrados? *

Texto de resposta longa

Sobre a forma como as instruções são fornecidas: qual o seu parecer em relação as suas percepções, dificuldades ou erros encontrados? *

Link PDF

intruções: https://drive.google.com/file/d/1C8ahUgNkoCmi940kfthpr_3Pa_eFtbAS/view?usp=sharing

Texto de resposta longa

Relate aqui seus comentários gerais e sugestões de melhoria. *

Texto de resposta longa

APÊNDICE D – CONVITE - EXPERIMENTOS

Olá pessoal, espero que vocês estejam bem.

Me chamo Nathália Locatelli, sou mestrande em Ciência da Computação Aplicada, na UDESC. Estou finalizando meu projeto de pesquisa e preciso de sua ajuda. O meu trabalho é um Sistema de Recomendação de artigos publicados em anais de eventos relacionados a área de IHC (Interação Humano-Computador). Nossa objetivo neste trabalho foi desenvolver um método de criação de perfil inicial do usuário, sem nenhum tipo de informação prévia no sistema, através da seleção de itens em pares. E depois disso, recomendamos artigos que são similares ao perfil do usuário criado.

Agora precisamos testar este sistema, principalmente entender o resultado dessa criação de perfil inicial. O trabalho está sendo orientado pela professora Dra Isabela Gasparini (UDESC), e co-orientado pelo prof Dr Daniel Lichnow, da UFSM.

Utilizamos como base os artigos de mais de 680 eventos científicos e todas as sugestões apresentadas são de artigos dessa base.

Desta forma, gostaria de contar com a sua participação voluntária.

O período de avaliação será dos dias **16/04/2024 até 30/04/2024**, e o tempo de avaliação é de **10 a 15 minutos**.

Link do sistema: <https://new-tcc.herokuapp.com/>

A ideia é que você faça o cadastro, realize o login, leia com atenção as instruções e selecione pelo menos 5 itens dentre os pares. Observe que você pode selecionar um, dois ou nenhum item de cada par. Quando tiver esse mínimo de itens selecionados, você poderá visualizar um botão “Recomendar” na tela, e ao clicar nele poderá receber suas recomendações. Você pode optar por clicar neste botão a qualquer momento após ter o número mínimo de itens selecionados, optando por selecionar mais itens ou já receber suas recomendações.

Depois disso, você terá sua lista de recomendações, e em cada item poderá avaliar se Gostou ou Não Gostou da recomendação recebida, através dos botões de “Gostei” e “Não Gostei” abaixo de cada item. Indicamos que você avalie todos os itens, se possível.

Para finalizar, no fim da lista de recomendação existe o botão “Finalizar o Experimento”, e ao clicar nele você será redirecionado para um questionário de avaliação do método da criação dos pares.

Caso queira experimentar mais vezes é só fazer outro cadastro, e veja as diferentes sugestões de itens.

Qualquer dúvida, podem entrar em contato comigo no email: czr.nathalia@gmail.com

Atenciosamente
Isabela, Daniel e Nathália

APÊNDICE E – QUESTIONÁRIO - EXPERIMENTOS

Questionário de Perfil

...

B **I** **U**  

Este questionário é destinado a compreender a experiência com a criação de perfil dentro do sistema de recomendação (<https://new-tcc.herokuapp.com/>) proposta de trabalho da aluna Nathália Locatelli Cezar, mestrandra em Computação Aplicada Nathália Locatelli Cezar da Universidade do Estado de Santa Catarina (UDESC) e orientada pela professora Isabela Gasparini (UDESC) e pelo professor Daniel Lichnow (UFSM).

Perguntas sobre a seleção em pares

Email (insira o mesmo cadastrado no sistema): *

Texto de resposta curta

Gênero: *

- Feminino
- Masculino
- Prefiro não dizer

Minha área de atuação: *

- Área da Computação
- Área de IHC

...
Meu perfil: *

- Estudante de graduação
- Estudante de pós-graduação
- Docente/Pesquisador
- Profissional da indústria

Minha instituição: *

Coloque o nome da instituição em que está vinculado (universidade, nome da empresa, etc.)

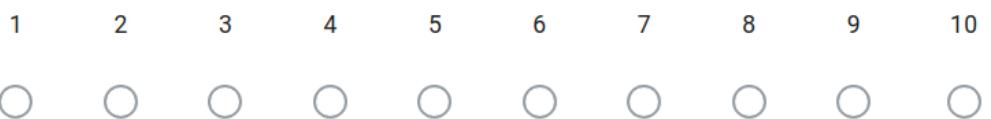
Texto de resposta curta

Sobre a criação de perfil, qual(is) dispositivo(s) você utilizou para acessar o sistema? *

- Celular/smarthphone
- Computador/notebook
- Tablet

Sobre os pares de itens apresentados, quanto você sentiu que tinham relação um com o outro? *

Sendo 1 indica menor grau de relacionamento e 10 indica maior grau de relacionamento



O que lhe motivou a escolher um item no par, ou a não escolha?

Texto de resposta longa

Sugestão de como iniciar o perfil em um sistema de recomendação:

- Digitar palavra
- Digitar texto
- Tópicos para seleção

Quais são suas impressões sobre o sistema no geral:

Descrição (opcional)

Sugestões de melhorias para um sistema de seleção de perfil em pares: *

Texto de resposta longa



Código para verificação: **H6WKZ834**

Este documento foi assinado digitalmente pelos seguintes signatários nas datas indicadas:



ISABELA GASPARINI (CPF: 024.XXX.689-XX) em 19/08/2024 às 15:25:14

Emitido por: "SGP-e", emitido em 30/03/2018 - 12:38:40 e válido até 30/03/2118 - 12:38:40.

(Assinatura do sistema)

Para verificar a autenticidade desta cópia, acesse o link <https://portal.sgpe.sea.sc.gov.br/portal-externo/conferencia-documento/VURFU0NfMTIwMjJfMDAwMzUzMThfMzUzNjFfMjAyNF9INldLWjgzNA==> ou o site

<https://portal.sgpe.sea.sc.gov.br/portal-externo> e informe o processo **UDESC 00035318/2024** e o código **H6WKZ834** ou aponte a câmera para o QR Code presente nesta página para realizar a conferência.