

Tudo é Ranking! Adaptações Cruzadas entre Busca e Recomendação

André Lima^{1*}, Davi Greco^{1*}, Lucas Eduardo Bernardes^{1*}, Messias Feres^{1*}

Marcos André Gonçalves², Leonardo Rocha¹, Felipe Viegas¹

¹Universidade Federal de São João del-Rei, ²Universidade Federal de Minas Gerais,
andrechalima@aluno.ufsj.edu.br;davigreco21@aluno.ufsj.edu.br;xxlucas0404xx@aluno.ufsj.edu.br;
messiasferes127@aluno.ufsj.edu.br;mgoncalv@dcc.ufmg.br;lcrocha@ufsj.edu.br;frviegas@dcc.ufmg.br

ABSTRACT

Information Retrieval (IR) and Recommender Systems (RS) have historically evolved as distinct research areas, with different data representations, algorithms, and evaluation practices. Yet, both are fundamentally concerned with ranking items according to user needs, preferences, or intent. This paper explores this shared foundation by investigating whether native algorithms from one paradigm can operate in the other through representation-level adaptations, without modifying their internal architectures. We propose two cross-task strategies: *Tag-Query*, which converts user interaction histories into textual queries for recommendation using retrieval models, and *Query-as-User*, which converts search queries into pseudo-user profiles for retrieval using recommender models. We evaluate these strategies using a robust experimental framework that encompasses multiple datasets, strong native baselines from both IR and RS, and complementary ranking metrics. The results show that these simple adaptations can achieve competitive performance relative to specialized models and, under certain conditions, outperform native baselines on specific evaluation metrics. More broadly, our findings suggest that meaningful transfer between search and recommendation can emerge through representation-level adaptations, indicating that the two paradigms may be more interoperable than traditionally assumed and opening a research path toward lightweight, reusable, and hybrid ranking pipelines.

KEYWORDS

information retrieval, recommender systems, ranking, cross-task adaptation, hybrid systems

1 INTRODUÇÃO

Recuperação de Informação (*Information Retrieval* – IR) e Sistemas de Recomendação (*Recommender Systems* – RS) foram historicamente desenvolvidos como áreas distintas: enquanto IR se consolidou em torno da correspondência entre consultas explícitas e documentos, RS enfatizou a modelagem

de preferências a partir de históricos de interação entre usuários e itens. Apesar dessas diferenças, ambas compartilham uma operação fundamental: ranquear itens com base em algum sinal de necessidade, preferência ou intenção do usuário. Essa proximidade conceitual já havia sido descrita por Belkin e Croft ao caracterizarem recuperação e filtragem de informação como “duas faces da mesma moeda” [2], e voltou a ganhar força com avanços recentes em representação neural, modelos de linguagem e arquiteturas generativas para sistemas unificados ou híbridos [10, 11, 17, 21, 22, 31, 37, 38, 41, 43–45]. Ainda assim, a integração efetiva entre IR e RS permanece limitada por uma barreira prática central: algoritmos de busca esperam consultas e documentos, enquanto algoritmos de recomendação esperam usuários, itens e históricos de interação.

Grande parte das propostas recentes tenta superar essa barreira por meio de modelos conjuntos, arquiteturas unificadas ou treinamento compartilhado entre tarefas [21, 24, 40, 42–44]. Embora essas abordagens sejam importantes, frequentemente exigem modificar arquiteturas, treinar modelos específicos ou construir pipelines próprios para cada cenário. Além disso, evidências recentes mostram que modelos unificados nem sempre superam métodos especializados em condições experimentais realistas [13, 24]. Comparações entre paradigmas também são frequentemente dificultadas por diferenças de pré-processamento, particionamento de dados, formulação de tarefas, orçamento de ajuste e métricas de avaliação [24, 27]. Assim, permanece aberta uma questão complementar e ainda pouco explorada: *algoritmos nativos de IR e RS podem ser reaproveitados entre domínios apenas pela transformação das representações de entrada?*

Neste artigo, partimos da premissa de que “tudo é ranking” para investigar essa questão. Em vez de propor uma nova arquitetura unificada, investigamos até que ponto a fronteira entre busca e recomendação pode ser atravessada por meio de adaptações relativamente simples nos sinais de entrada. A hipótese central é que, se IR e RS são ambos problemas de ranqueamento, então consultas podem ser reinterpretadas como perfis colaborativos, e históricos de usuários podem ser reinterpretados como intenções de busca. Essa mudança desloca a integração entre as áreas do nível arquitetural para o nível representacional: os algoritmos permanecem intactos, mas passam a operar sobre entradas traduzidas para o paradigma esperado.

Para explorar essa hipótese, propomos e avaliamos duas estratégias de adaptação cruzada. A primeira, denominada *Tag-Query* (IR → RS), transforma o histórico implícito de um

*Estes autores contribuíram igualmente como autores principais.

usuário em uma consulta textual explícita, permitindo que um modelo de recuperação lexical, como o BM25, seja utilizado como recomendador. A segunda, denominada *Query-as-User* (RS $\rightarrow$ IR), realiza o movimento inverso: consultas de busca são tratadas como pseudo-usuários em uma estrutura colaborativa, permitindo que modelos de recomendação sejam usados para ranquear documentos. Em ambos os casos, a adaptação ocorre sem alterar a arquitetura interna dos modelos, o que evidencia que parte da interoperabilidade entre IR e RS pode emergir apenas da reescrita dos sinais de entrada.

Avaliamos essas estratégias em um ambiente experimental completo, comparando-as a linhas de base sólidas e recentes de IR e RS. Os experimentos foram conduzidos em três conjuntos de dados com características distintas — MovieLens-25M, LastFM-360k e Amazon Electronics —, permitindo observar as adaptações em cenários com diferentes níveis de densidade, domínios, metadados textuais e padrões de interação. Todas as abordagens foram submetidas ao mesmo protocolo de pré-processamento, particionamento temporal, múltiplas execuções independentes e validação estatística pareada. No lado de IR, consideramos BM25 [25], Bi-Encoders [35] e ColBERT [15], cobrindo métodos lexicais, densos e de interação tardia; no lado de RS, utilizamos LightFM [16], BERT4Rec [19, 30] e ItemKNN [1, 23], contemplando filtragem colaborativa, recomendação sequencial e recomendação baseada em conteúdo. A avaliação emprega NDCG, Precision e Recall, permitindo comparar modelos nativos e adaptações híbridas tanto pela qualidade do ranking quanto pela capacidade de recuperar itens relevantes em diferentes profundidades da lista.

Os resultados mostram que o *Tag-Query*, embora não supere os recomendadores nativos mais fortes, aproxima-se deles em cenários favoráveis mesmo utilizando BM25, um modelo simples e sem treinamento supervisionado. Isso sugere que históricos de usuários podem fornecer sinais textuais suficientes para motores de busca quando adequadamente convertidos em consultas. Já o *Query-as-User* supera linhas de base de busca em algumas métricas, indicando que sinais colaborativos também podem beneficiar tarefas de recuperação. Em conjunto, esses achados mostram que adaptações representacionais podem viabilizar transferência significativa entre busca e recomendação, preservando parte do desempenho dos modelos originais.

A principal contribuição deste trabalho é demonstrar que adaptações simples de representação permitem aplicar algoritmos nativos a tarefas cruzadas, favorecendo o reaproveitamento de modelos e a construção de pipelines híbridos entre IR e RS. Em vez de depender exclusivamente de arquiteturas unificadas, mostramos que parte dessa integração pode emergir da reformulação das entradas que alimentam algoritmos já consolidados.

Este trabalho apresenta as seguintes contribuições:

- Uma formulação de adaptação cruzada entre IR e RS baseada na transformação das representações de entrada, preservando as arquiteturas dos algoritmos;

- Duas estratégias de adaptação representacional, *Tag-Query* e *Query-as-User*, que permitem, respectivamente, usar modelos de busca em cenários de recomendação e modelos de recomendação em tarefas de busca;
- Um arcabouço experimental para avaliar adaptações híbridas em múltiplos conjuntos de dados e métricas de avaliação com validação estatística;
- A comparação das adaptações propostas com linhas de base nativas fortes e recentes de IR e RS, demonstrando que adaptações simples podem alcançar desempenho competitivo, evidenciando o potencial de transferência entre paradigmas de ranking.

2 TRABALHOS RELACIONADOS

A aproximação entre Recuperação de Informação (IR) e Sistemas de Recomendação (RS) decorre do fato de que ambos tratam, fundamentalmente, de problemas de ranqueamento. Apesar de partirem de representações distintas — consultas e documentos em IR, usuários, itens e interações em RS —, os dois paradigmas buscam ordenar itens segundo sinais de necessidade, preferência ou intenção. Essa convergência tem impulsionado esforços de integração entre busca e recomendação, mas também evidencia um desafio central: os algoritmos de cada área foram concebidos para diferentes formas de representação dos dados.

2.1 Modelos Nativos de Busca e Recomendação

Em IR, a tarefa clássica é ranquear documentos a partir de uma consulta textual. Modelos lexicais esparsos, como BM25 [25], permanecem como referências importantes por sua eficiência, robustez e ausência de treinamento supervisionado. Além do ranqueamento tradicional, trabalhos em IR têm explorado diferentes formas de representar e organizar o conteúdo, incluindo a estimativa da importância de blocos estruturais em páginas Web [9] e a organização semântica e hierárquica de coleções textuais [33]. Recentemente, recuperadores densos passaram a representar consultas e documentos em espaços vetoriais semânticos, como em arquiteturas de *bi-encoders* [35]. Modelos de interação tardia, como ColBERT [15], preservam representações em nível de *token* até a etapa de ranqueamento, permitindo capturar correspondências mais finas entre consulta e documento.

Em RS, por outro lado, o ranqueamento é usualmente derivado de padrões de interação entre usuários e itens. Modelos colaborativos, como LightFM [16], aprendem fatores latentes a partir da matriz usuário-item e podem incorporar metadados para mitigar problemas como *cold-start*. Recomendadores sequenciais, como BERT4Rec [30], exploram a ordem temporal das interações por meio de mecanismos de atenção, enquanto abordagens baseadas em conteúdo, como ItemKNN [23], ranqueiam itens com base na similaridade em relação ao histórico do usuário.

Embora busca e recomendação resolvam problemas de ranqueamento, seus algoritmos operam sobre representações distintas. Este trabalho investiga se a transformação dessas representações é suficiente para reutilizar algoritmos nativos em tarefas cruzadas.

2.2 Modelagem Unificada entre IR e RS

A ideia de integrar busca e recomendação não é recente. Belkin e Croft [2] já caracterizavam a recuperação e a filtragem de informação como “duas faces da mesma moeda”, antecipando a visão de que ambas poderiam ser compreendidas como formas relacionadas de correspondência entre usuários e conteúdo. Trabalhos posteriores exploraram essa convergência por meio de objetivos conjuntos e da transferência de sinais entre tarefas, incluindo modelos como JSR [42] e SESRec [28]. Recentemente, arquiteturas unificadas e generativas, como UnifiedSSR [39] e GenSAR [27], passaram a integrar comportamentos heterogêneos de usuários e modelos de linguagem em pipelines que lidam com busca e recomendação.

Essas abordagens ampliam o espaço de modelagem na fronteira entre IR e RS, mas frequentemente dependem de arquiteturas específicas, treinamento conjunto ou formulações generativas projetadas para operar em múltiplas tarefas. Além disso, evidências empíricas indicam que modelos unificados nem sempre superam, de forma consistente, sistemas especializados sob condições experimentais realistas [24]. Seus resultados também podem ser sensíveis a escolhas de pré-processamento, geração de candidatos, particionamento de dados, orçamento de ajuste e configuração experimental [24, 27].

Esse cenário sugere que a integração entre IR e RS pode ser investigada não apenas como um problema de arquitetura, mas também como um problema de representação. Há uma questão complementar, menos explorada: até que ponto algoritmos já consolidados em um domínio podem ser reutilizados em outro se os sinais de entrada forem reescritos no formato esperado pelo modelo? É nessa direção que este trabalho se posiciona. Em vez de propor um novo modelo unificado, avaliamos adaptações cruzadas que preservam as arquiteturas originais e deslocam a integração para o nível de representação de dados.

2.3 Infraestrutura e Avaliação Cruzada

Outro desafio importante na aproximação entre IR e RS é a avaliação. Estudos sobre modelos híbridos ou unificados frequentemente utilizam conjuntos de dados, formulações de tarefa e protocolos de validação distintos, o que dificulta comparações diretas e reproduzíveis. Essa limitação é particularmente relevante quando se deseja avaliar se um algoritmo nativo de um domínio pode competir com métodos especializados do outro, pois diferenças de pré-processamento, métricas e seleção de candidatos podem influenciar as conclusões.

Diversas ferramentas apoiam a experimentação reproduzível em comunidades específicas. Em IR, Pyserini [18] e USearch [32] padronizam componentes de indexação e recuperação. Em RS, frameworks como iRec [29], iRev [3] e DeepCTR [26] facilitam a implementação e avaliação de modelos de recomendação. No entanto, esses arcabouços foram concebidos principalmente para fluxos de uma única tarefa, oferecendo suporte limitado à avaliação de adaptações cruzadas entre busca e recomendação.

Este trabalho avalia duas estratégias de adaptação entre IR e RS sob um protocolo comum. *Tag-Query* converte históricos de usuários em consultas textuais, enquanto *Query-as-User*

transforma consultas em perfis colaborativos. Ao compará-las com fortes linhas de base nativas em múltiplos conjuntos de dados, investigamos se a integração IR-RS pode emergir da reformulação das representações de entrada, sem modificar os algoritmos subjacentes.

Diferentemente da maior parte dos trabalhos anteriores, não propomos uma arquitetura unificada. Em vez disso, investigamos se a transferência entre os dois paradigmas pode ser obtida por meio de adaptações das representações de entrada, preservando modelos e procedimentos de treinamento originais. Essa perspectiva enfatiza interoperabilidade, reaproveitamento de modelos e simplicidade de implantação.

3 ADAPTAÇÕES CRUZADAS ENTRE BUSCA E RECOMENDAÇÃO

Nesta seção, apresentamos duas estratégias para aplicar algoritmos de IR e RS além de seus domínios originais. Partimos da premissa de que a principal diferença entre busca e recomendação está menos nos algoritmos do que na forma como os dados são representados. Assim, transformamos as entradas de um paradigma no formato esperado pelo outro, sem alterar os modelos subjacentes.

Investigamos duas direções complementares. Na $IR \rightarrow RS$, o histórico do usuário é convertido em uma consulta textual (*Tag-Query*), permitindo que um modelo de busca produza recomendações. Na $RS \rightarrow IR$, consultas são convertidas em perfis colaborativos (*Query-as-User*), permitindo que um modelo de recomendação ranquee documentos.

3.1 Busca-para-Recomendação: *Tag-Query*

A estratégia *Tag-Query* adapta um modelo de busca para resolver uma tarefa de recomendação. Em um cenário tradicional de RS, o sistema recebe o histórico de itens consumidos por um usuário e deve recomendar novos itens relevantes. Em vez de treinar um recomendador diretamente a partir desse histórico, propomos converter o perfil do usuário em uma consulta textual. Essa consulta sintetiza os principais interesses observados no histórico e é submetida a um modelo de recuperação, como o BM25, que passa a ranquear os itens do catálogo como se fossem documentos. A Figura 1 ilustra o fluxo de funcionamento dessa estratégia, desde o histórico do usuário até a geração da consulta submetida ao modelo de busca.

O processo ocorre em três etapas principais.

1. Enriquecimento dos itens. Inicialmente, cada item do catálogo é representado como um documento textual. No caso do MovieLens, por exemplo, esse documento é construído a partir de gêneros e *tags* associados aos filmes. Para reduzir ruído, mantemos apenas metadados cuja relevância ultrapassa um limiar previamente definido. Assim, cada item passa a ter uma representação textual indexável por um modelo de IR.

2. Construção do perfil textual do usuário. Para cada usuário, recuperamos os itens presentes em seu histórico de interações e coletamos as *tags* associadas a esses itens. Em seguida, agregamos os pesos das *tags* ao longo do histórico, de modo que conceitos recorrentes ou altamente relevantes ganhem maior importância. Essa agregação transforma um conjunto

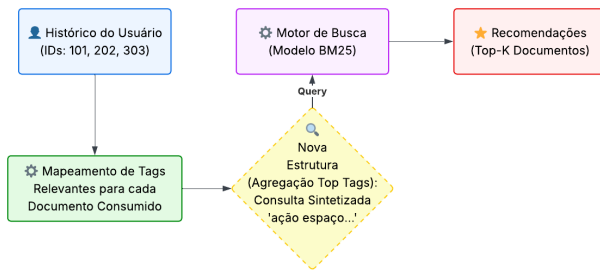


Figura 1: Fluxo de funcionamento da estratégia *Tag-Query*, em que o histórico de itens do usuário é convertido em uma consulta textual enviada a um modelo de busca.

de interações implícitas em uma representação textual dos interesses do usuário.

3. Síntese da consulta e ranqueamento. As *tags* agregadas são ordenadas por relevância acumulada, e os conceitos mais representativos são concatenados em uma pseudoconsulta. Essa consulta é submetida ao modelo BM25, que ranqueia os itens/documentos do catálogo com base na aderência deles ao perfil textual do usuário. O resultado final é uma lista recomendada, produzida por um algoritmo de busca a partir de uma entrada formulada como problema de recomendação.

A principal vantagem dessa adaptação é sua simplicidade. O modelo de IR não precisa ser modificado, treinado ou especializado para recomendação. Todo o processo de adaptação ocorre antes da chamada ao recuperador, pela transformação do histórico do usuário em uma intenção textual explícita. Dessa forma, *Tag-Query* testa até que ponto sinais de preferência podem ser expressos como consultas e explorados por modelos tradicionais de busca.

3.2 Recomendação-para-Busca: *Query-as-User*

A estratégia *Query-as-User* realiza o movimento inverso: adapta um modelo de recomendação para resolver uma tarefa de busca. Em IR, o sistema recebe uma consulta e deve ranquear documentos relevantes. Em vez de tratar a consulta como uma sequência textual a ser comparada diretamente com documentos, propomos interpretá-la como um pseudo-usuário. Os documentos associados à consulta passam, então, a compor um perfil de interação, permitindo que algoritmos colaborativos sejam usados para ranquear novos documentos.

A Figura 2 apresenta essa transposição. A intuição é que uma consulta pode ser vista como uma entidade que manifesta preferência por determinados documentos. Se várias consultas se relacionam com documentos em diferentes graus de relevância, é possível construir uma matriz semelhante à matriz usuário-item de recomendação. Nessa matriz, consultas ocupam o papel de usuários, documentos ocupam o papel de itens, e os escores de relevância assumem o papel de interações.

O processo também pode ser dividido em três etapas.

1. Transposição das entidades. Durante a preparação dos dados, cada consulta ou *tag* de busca é mapeada para uma entidade do tipo usuário. Cada documento ou item do catálogo é mapeado para uma entidade do tipo item. A relação de

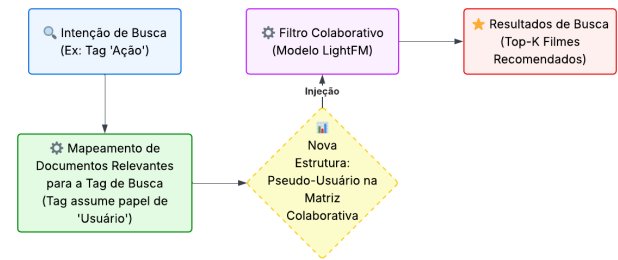


Figura 2: Fluxo de funcionamento da estratégia *Query-as-User*, em que consultas são tratadas como pseudo-usuários em uma matriz colaborativa.

relevância entre consulta e documento é utilizada como sinal de interação. Relações de baixa relevância são descartadas para reduzir o ruído e preservar apenas as associações mais informativas.

2. Construção da matriz colaborativa. A partir dessa transposição, construímos uma matriz esparsa em que linhas representam consultas e colunas representam documentos. Essa matriz tem a mesma estrutura esperada por algoritmos colaborativos de recomendação. Assim, um modelo como o LightFM pode ser treinado sem alterações arquiteturais, aprendendo representações latentes para consultas-pseudo-usuários e documentos-itens.

3. Ranqueamento de documentos. Após o treinamento, o modelo colaborativo estima a afinidade entre uma consulta e os documentos candidatos. O ranking final é produzido a partir desses pesos de afinidade, de modo análogo ao ranqueamento de itens para um usuário em RS. Com isso, uma tarefa originalmente formulada como busca passa a ser resolvida por um recomendador.

Essa adaptação permite explorar sinais colaborativos em tarefas de recuperação. Em vez de depender apenas da correspondência lexical ou semântica entre consulta e documento, *Query-as-User* aprende padrões de coocorrência entre consultas e itens relevantes. Isso pode revelar relações latentes entre intenções de busca e documentos que não são capturadas diretamente por modelos textuais tradicionais.

3.3 Síntese das Adaptações

As duas estratégias propostas compartilham o mesmo princípio: preservar os algoritmos e transformar apenas suas entradas. A *Tag-Query* responde à pergunta: *um histórico de usuário pode ser reescrito como uma consulta?* Já a *Query-as-User* responde à pergunta complementar: *uma consulta pode ser reescrita como um perfil de usuário?* Em conjunto, essas adaptações operacionalizam a ideia de que busca e recomendação podem ser aproximadas por uma camada comum de representação e ranqueamento.

Essa formulação tem implicações práticas e metodológicas. Do ponto de vista prático, ela permite reaproveitar algoritmos já consolidados em tarefas para as quais não foram originalmente projetados. Do ponto de vista metodológico, ela desloca parte da integração entre IR e RS do desenho

de novas arquiteturas para a construção de mapeamentos entre representações. Assim, as adaptações propostas não devem ser vistas apenas como dois métodos específicos, mas como exemplos de uma agenda mais ampla: investigar como diferentes sinais de usuário, conteúdo e interação podem ser traduzidos entre paradigmas de ranking.

4 CONFIGURAÇÃO EXPERIMENTAL

Esta seção descreve o protocolo experimental utilizado para avaliar se algoritmos nativos de busca e recomendação podem operar de forma competitiva em tarefas cruzadas por meio das adaptações propostas. O objetivo não é apenas comparar modelos isoladamente, mas testar a hipótese central deste trabalho em um ambiente controlado: se IR e RS compartilham uma base comum de ranqueamento, então a transformação das representações de entrada deve permitir que algoritmos de um domínio sejam avaliados no outro sem alterações arquiteturais.

Para isso, construímos uma configuração experimental abrangente, envolvendo três conjuntos de dados de domínios distintos, linhas de base fortes de IR e RS, particionamento temporal, métricas clássicas de ranqueamento em múltiplos cortes, execuções independentes e validação estatística. Todas as etapas, do pré-processamento à avaliação, foram executadas de forma padronizada para modelos nativos e adaptações híbridas, garantindo a comparabilidade dos resultados.

4.1 Conjuntos de Dados

Utilizamos três coleções amplamente empregadas na literatura de IR e RS que representam cenários com diferentes níveis de densidade, domínios de aplicação, disponibilidade de metadados textuais e padrões de interação. Essa diversidade é importante para avaliar se as adaptações propostas dependem de condições específicas dos dados ou se apresentam comportamento consistente em diferentes regimes experimentais.

MovieLens-25M [12] reúne avaliações de filmes com alta densidade relativa de interações. Seus metadados incluem gêneros e *tags* atribuídas pelos usuários [34], que são utilizados para compor representações textuais dos itens no cenário de IR e para sintetizar pseudoconsultas na adaptação IR $\rightarrow$ RS.

LastFM-360k [4] contém históricos de escuta musical em larga escala, com forte presença de cauda longa e de padrões de consumo recorrentes. As consultas de IR são derivadas dos atributos textuais de artistas e faixas, permitindo avaliar as adaptações em um domínio diferente de filmes (MovieLens-25M), com maior diversidade de itens.

Amazon Electronics [36] compreende interações de compra na categoria de eletrônicos e representa um cenário mais esparsos. As descrições textuais dos produtos são utilizadas diretamente na indexação para IR, tornando esse conjunto particularmente relevante para avaliar os limites das adaptações em um domínio com menor densidade colaborativa.

4.2 Pré-processamento e Particionamento

O pré-processamento foi definido de modo a preservar a comparabilidade entre modelos nativos e híbridos. Estratégias de seleção de instâncias têm mostrado que reduções seletivas

dos dados podem diminuir redundância e ruído, preservando a efetividade e reduzindo custos computacionais [7, 8]. No MovieLens-25M e no LastFM-360k, usuários e itens fora dos percentis 25 e 95 da contagem de interações foram descartados, reduzindo extremos de esparsidade e de popularidade excessiva. No Amazon Electronics, aplicou-se apenas o corte no percentil inferior, acrescido de um filtro *10-core*, que exige ao menos dez interações por usuário e por item. A Tabela 1 apresenta as estatísticas dos conjuntos de dados utilizados após pré-processamento.

Em todos os conjuntos, adotamos particionamento temporal. As interações mais recentes de cada usuário foram reservadas para teste, preservando a ordem cronológica dos eventos e aproximando-se de uma avaliação de um cenário realista, no qual o sistema deve ranquear itens futuros com base em sinais observados no passado. Para o MovieLens-25M, os últimos 20 itens relevantes por usuário foram reservados para teste; para o LastFM-360k, os últimos 50. O Amazon Electronics segue o mesmo critério temporal, sem corte fixo de itens por usuário.

Conjunto	#Usuários	#Itens	#Interações	#Consultas
MovieLens-25M	129.611	31.759	12.938.564	1.105
LastFM-360k	280.000	247.058	13.657.139	384
Amazon Electr.	147.529	51.700	2.473.355	986

Tabela 1: Estatísticas dos conjuntos de dados utilizados.

Em IR, cada item foi indexado como um documento textual. No MovieLens-25M, os documentos foram compostos pela concatenação de gêneros e *tags* do *Genome Tags* com relevância superior a 0,8, mesmo limiar utilizado na construção das pseudoconsultas em recomendação. No LastFM-360k, os atributos textuais de artistas e faixas compõem os documentos. No Amazon Electronics, as descrições dos produtos são utilizadas diretamente como conteúdo textual indexável.

4.3 Modelos Avaliados

Avaliamos modelos nativos de IR, modelos nativos de RS e as duas adaptações cruzadas propostas. Essa escolha permite comparar as estratégias híbridas com algoritmos fortes e representativos de cada domínio.

Modelos nativos de IR.

- **BM25** [25]: modelo lexical esparsos amplamente utilizado como referência em IR, baseado na frequência e raridade dos termos no corpus.
- **Bi-Encoder** [35]: recuperador denso que codifica consultas e documentos de forma independente em um espaço vetorial semântico, permitindo ranqueamento por similaridade.
- **ColBERT** [15]: modelo de recuperação densa com interação tardia, que preserva representações em nível de *token* até a etapa de *scoring*.

Modelos nativos de RS.

- **LightFM** [16]: modelo de filtragem colaborativa baseado em fatores latentes, capaz de incorporar metadados de usuários e de itens.
- **BERT4Rec** [30]: recomendador sequencial baseado em Transformers, que modela o histórico do usuário por meio de atenção bidirecional.

- **ItemKNN** [23]: abordagem baseada em conteúdo que ranqueia candidatos com base na similaridade em relação aos itens previamente consumidos pelo usuário.

4.4 Métricas de Avaliação

As abordagens foram avaliadas com métricas clássicas de ranqueamento [6, 14, 20], calculadas nos cortes $k \in 5, 10, 20, 50$. A avaliação em múltiplas profundidades permite observar tanto a qualidade do topo da lista quanto a capacidade de recuperar itens relevantes em posições mais altas do ranking.

- **NDCG@ k** (*Normalized Discounted Cumulative Gain*): mede a qualidade do ranqueamento considerando a posição dos itens relevantes, penalizando acertos localizados em posições mais baixas.
- **Precision@ k** : mede a proporção de itens relevantes entre os k primeiros resultados, capturando a exatidão do topo da lista.
- **Recall@ k** : mede a fração dos itens relevantes recuperados nos k primeiros resultados, capturando a cobertura do ranking.

Os resultados reportados correspondem à média de seis execuções independentes por configuração. Essa escolha reduz a influência de variações aleatórias e permite uma comparação mais estável entre modelos nativos e adaptações híbridas.

4.5 Validação Estatística

As diferenças de desempenho entre as abordagens foram avaliadas por meio do teste t **pareado** com correção de *Holm* ($\alpha = 0,05$). Resultados marcados com † nas tabelas indicam empate estatístico em relação ao melhor modelo do grupo.

5 RESULTADOS E DISCUSSÃO

Avaliamos as duas adaptações propostas, *Tag-Query* (IR $\rightarrow$ RS) e *Query-as-User* (RS $\rightarrow$ IR), sob o mesmo protocolo experimental aplicado aos modelos nativos de busca e de recomendação. A apresentação dos resultados está organizada por conjunto de dados, permitindo observar o comportamento das estratégias em domínios com diferentes níveis de densidade, disponibilidade de metadados textuais e padrões de interação. Em cada cenário, analisamos separadamente as direções IR $\rightarrow$ RS e RS $\rightarrow$ IR, comparando as adaptações com linhas de base representativas de seus respectivos paradigmas.

MovieLens-25M. Os resultados da Tabela 2 mostram que o MovieLens-25M é um cenário favorável para avaliar a hipótese de adaptação entre paradigmas, pois combina alta densidade de interações com metadados textuais ricos. Na direção IR $\rightarrow$ RS, o *Tag-Query* atinge NDCG@5 = 0.5413, superando o ItemKNN nesse corte e aproximando-se de recomendadores nativos mais fortes, embora permaneça abaixo de LightFM e BERT4Rec. Esse resultado é relevante porque o método usa uma formulação simples baseada em BM25, sem treinamento supervisionado nem alterações arquiteturais, indicando que históricos de consumo podem ser reescritos como consultas textuais suficientemente informativas para produzir recomendações competitivas. Na direção RS $\rightarrow$ IR, o *Query-as-User* apresenta o comportamento mais expressivo: obtém o

melhor Recall em todos os cortes e o melhor NDCG@50, superando modelos nativos nessas métricas. Isso sugere que tratar consultas como pseudo-usuários permite explorar padrões colaborativos que ampliam a cobertura do ranking, enquanto modelos como ColBERT permanecem mais fortes nos cortes menores, em que a precisão semântica no topo é decisiva.

LastFM-360k. Podemos observar, pela Tabela 3, um padrão semelhante, mas com diferenças importantes. O *Tag-Query* alcança NDCG@5 = 0.6100, valor próximo ao de LightFM (0.6455) e BERT4Rec (0.6205), reforçando que a transformação de históricos em consultas pode produzir rankings competitivos quando os itens dispõem de descrições textuais úteis. Entretanto, sua Precision e seu Recall ficam abaixo dos recomendadores nativos, especialmente em comparação ao LightFM, indicando que o histórico colaborativo de escutas ainda oferece um sinal mais forte para ordenar o topo da lista. Na direção RS $\rightarrow$ IR, o *Query-as-User* novamente obtém o melhor Recall em todos os cortes, superando BM25, Bi-Encoders e ColBERT nessa dimensão. Esse resultado reforça a principal promessa da adaptação: modelos colaborativos podem recuperar itens relevantes que métodos puramente lexicais ou semânticos não priorizam. Por outro lado, ColBERT mantém vantagem em NDCG e Precision, mostrando que sinais colaborativos favorecem cobertura, mas não substituem integralmente a correspondência textual no topo do ranking.

Amazon Electronics. A Tabela 4 evidencia os limites das adaptações em um cenário mais esparsos. Na direção IR $\rightarrow$ RS, o *Tag-Query* apresenta queda acentuada em relação aos recomendadores nativos, com NDCG@5 = 0.0653 em comparação a 0.1981 do ItemKNN. Esse comportamento sugere que, quando há poucas interações por usuário e metadados menos alinhados ao histórico de consumo, a pseudoconsulta sintetizada carrega sinal insuficiente para que o BM25 produza recomendações competitivas. Na direção RS $\rightarrow$ IR, o *Query-as-User* mantém desempenho razoável em relação ao BM25 em alguns pontos, mas fica abaixo dos recuperadores semânticos, especialmente Bi-Encoders e ColBERT, que se beneficiam do contexto das descrições textuais dos produtos. Esse resultado mostra que a adaptação cruzada não elimina a dependência do regime de dados: quando o sinal textual é forte e o sinal colaborativo é esparsos, modelos nativos de IR tendem a prevalecer.

Esses resultados, em conjunto, mostram que as adaptações propostas são mais do que *baselines* alternativas: revelam como diferentes tipos de sinal podem atravessar a fronteira entre busca e recomendação. O *Tag-Query* é mais efetivo quando o histórico do usuário pode ser convertido em uma intenção textual representativa, como ocorre em domínios com metadados ricos. Já o *Query-as-User* se beneficia de cenários em que há coocorrência suficiente entre consultas, itens e padrões de relevância para que um modelo colaborativo aprenda relações latentes úteis. Assim, os resultados não apontam para uma equivalência irrestrita entre IR e RS, mas para uma intercambialidade condicionada: algoritmos de um domínio podem ser reaproveitados no outro quando a transformação de entrada preserva o sinal necessário ao ranqueamento.

Avaliamos também a eficiência dos algoritmos sob a perspectiva do tempo de execução. A Tabela 5 apresenta o tempo

(a) RS e IR $\rightarrow$ RS					(b) IR e RS $\rightarrow$ IR				
Modelo	@k	NDCG	Precision	Recall	Modelo	@k	NDCG	Precision	Recall
LightFM	5	0.6467 [†]	0.7646 [†]	0.2719 [†]	BM25	5	0.0641	0.0597	0.0066
	10	0.6481 [†]	0.6887 [†]	0.4774		10	0.0555	0.0460	0.0099
	20	0.6701 [†]	0.5273	0.6937		20	0.0463	0.0352	0.0142
	50	0.7471	0.2643	0.8491		50	0.0390	0.0246	0.0202
BERT4Rec	5	0.6753	0.7969	0.2890	Bi-Encoders	5	0.0845	0.0780	0.0097
	10	0.6955	0.7426	0.5253		10	0.0771	0.0696	0.0160
	20	0.7441	0.5957	0.7937		20	0.0700	0.0602	0.0255
	50	0.8297	0.2992	0.9654		50	0.0671	0.0498	0.0464
ItemKNN	5	0.4375	0.5018	0.2602	ColBERT	5	0.1248	0.1123	0.0142
	10	0.5420	0.5311	0.5257		10	0.1183	0.1033	0.0210
	20	0.6422	0.4582	0.8062		20	0.1093	0.0913	0.0351
	50	0.7159	0.2368	0.9702		50	0.0983	0.0753	0.0610
Tag Query ^a	5	0.5413	0.5209	0.2157	Query-as-User ^b	5	0.0834	0.0806	0.0265
	10	0.5071	0.4336	0.3440		10	0.0828	0.0761	0.0392
	20	0.4790	0.3080	0.4672		20	0.0845	0.0706	0.0609
	50	0.5475	0.1643	0.6054		50	0.1027	0.0617	0.1037

Tabela 2: Resultados MovieLens-25M. Negrito marca o melhor valor por @k/métrica em cada subtabela; [†] marca empate estatístico com o melhor (teste t , Holm, $\alpha=0,05$). Híbridos: ^a Tag Query utiliza BM25; ^b Query-as-User utiliza LightFM.

(a) RS e IR $\rightarrow$ RS					(b) IR e RS $\rightarrow$ IR				
Modelo	@k	NDCG	Precision	Recall	Modelo	@k	NDCG	Precision	Recall
LightFM	5	0.6455	0.9541	0.3805	BM25	5	0.0194	0.0159	0.0023 [†]
	10	0.7184	0.8765	0.6978		10	0.0157	0.0109	0.0028
	20	0.7627	0.5603 [†]	0.8897		20	0.0122	0.0103	0.0043
	50	0.7801 [†]	0.2357	0.9356		50	0.0109	0.0078	0.0078 [†]
BERT4Rec	5	0.6205 [†]	0.9296 [†]	0.3707 [†]	Bi-Encoders	5	0.0245	0.0240	0.0042 [†]
	10	0.7029	0.8676	0.6910		10	0.0216	0.0191	0.0057
	20	0.7604	0.5707	0.9065		20	0.0189	0.0147	0.0092
	50	0.7821	0.2431	0.9652		50	0.0191	0.0121	0.0165 [†]
ItemKNN	5	0.4353	0.6323	0.3643	ColBERT	5	0.1156	0.1087	0.0122
	10	0.6000	0.6744	0.7675		10	0.1084	0.0975	0.0196
	20	0.7303	0.4857	0.9313		20	0.0987	0.0864	0.0317
	50	0.7468	0.2107	0.9701		50	0.0912	0.0727	0.0559
Tag Query ^a	5	0.6100	0.5436	0.3425	Query-as-User ^b	5	0.0669	0.0511	0.0335
	10	0.4862	0.3060	0.3855		10	0.0667	0.0422	0.0463
	20	0.5018	0.1652	0.4165		20	0.0699	0.0336	0.0737
	50	0.5383	0.0808	0.5097		50	0.0795	0.0256	0.1168

Tabela 3: Resultados LastFM-360k. Negrito marca o melhor valor por @k/métrica em cada subtabela; [†] marca empate estatístico com o melhor (teste t , Holm, $\alpha=0,05$). Híbridos: ^a Tag Query usa BM25; ^b Query-as-User usa LightFM.

médio das seis execuções de cada algoritmo analisado. Por limitação de espaço, restringimos a apresentação à coleção MovieLens-25M, embora o mesmo padrão geral tenha sido observado nas demais coleções. Na direção IR $\rightarrow$ RS, o *Tag-Query* executa em 952,72 segundos, abaixo de LightFM (1737, 56 s), ItemKNN (2150, 89 s) e, sobretudo, BERT4Rec (17598,63 s). Esse resultado reforça que, embora a adaptação baseada em BM25 não supere os recomendadores mais fortes em todas as métricas, ela oferece desempenho competitivo com menor custo computacional, tornando-se uma alternativa relevante em cenários em que simplicidade, reuso

e eficiência são desejáveis. Na direção RS $\rightarrow$ IR, o *Query-as-User* executa em 126,41 segundos, tempo muito próximo ao BM25 (101, 20 s) e substancialmente inferior ao dos modelos densos Bi-Encoders (1206, 53 s) e ColBERT (3646, 53 s). Considerando que o *Query-as-User* também obteve ganhos em métricas de cobertura nos resultados de efetividade, esse comportamento sugere que adaptações colaborativas podem oferecer uma alternativa competitiva para busca quando o objetivo é ampliar a recuperação de itens relevantes com custo inferior ao de recuperadores neurais mais complexos.

De modo geral, os resultados corroboram a mensagem central deste trabalho: a integração entre busca e recomendação

(a) RS e IR $\rightarrow$ RS					(b) IR e RS $\rightarrow$ IR				
Modelo	@k	NDCG	Precision	Recall	Modelo	@k	NDCG	Precision	Recall
LightFM	5	0.1618	0.1284	0.1466	BM25	5	0.6419	0.5851	0.2615
	10	0.1873	0.0948	0.2130		10	0.6330	0.5129	0.3724
	20	0.2193	0.0673	0.2984		20	0.6232	0.4170	0.4872
	50	0.2664	0.0412	0.4534		50	0.5934	0.2583	0.6091
BERT4Rec	5	0.1852	0.1491	0.1748	Bi-Encoders	5	0.6852	0.6307	0.2764
	10	0.2193	0.1117	0.2602		10	0.6841	0.5637	0.4040
	20	0.2580	0.0786	0.3641		20	0.6911	0.4804	0.5534
	50	0.3071	0.0457	0.5274		50	0.7028	0.3421	0.7671
ItemKNN	5	0.1981	0.1362	0.2052	ColBERT	5	0.7651	0.7551	0.1249
	10	0.2367	0.0999	0.2981		10	0.7481	0.7281	0.2150
	20	0.2776	0.0699	0.4151		20	0.7122	0.6651	0.3580
	50	0.3276	0.0403	0.5953		50	0.5891	0.4351	0.5120
Tag Query ^a	5	0.0653	0.0240	0.0684	Query-as-User ^b	5	0.5485	0.5700	0.1400
	10	0.0742	0.0167	0.0919		10	0.5081	0.5352	0.2558
	20	0.0883	0.0130	0.1386		20	0.4331	0.4574	0.3977
	50	0.1216	0.0109	0.2808		50	0.4847	0.3422	0.5408

Tabela 4: Resultados Amazon Electronics. Negrito marca o melhor valor por @k/métrica em cada subtabela; [†] marca empate estatístico com o melhor (teste t , Holm, $\alpha=0,05$). Híbridos: ^a Tag Query usa BM25; ^b Query-as-User usa LightFM.

(a) RS e IR $\rightarrow$ RS		(b) IR e RS $\rightarrow$ IR	
Modelo	Tempo (s)	Modelo	Tempo (s)
RS		IR	
LightFM	1737,56	BM25	101,20
BERT4Rec	17598,63	Bi-Encoders	1206,53
ItemKNN	2150,89	ColBERT	3646,53
IR $\rightarrow$ RS		RS $\rightarrow$ IR	
Tag Query ^a	952,72	Query-as-User ^b	126,41

Tabela 5: Tempo de execução MovieLens-25M (inferência/ranking sobre o conjunto de teste).

não precisa depender exclusivamente de novas arquiteturas unificadas. Estratégias simples de adaptação representacional já permitem que algoritmos nativos sejam reutilizados em tarefas cruzadas, alcançando desempenho competitivo e, em métricas específicas, superando métodos especializados. Mais importante, esses achados abrem uma agenda de pesquisa sobre como sinais lexicais, semânticos e colaborativos podem ser traduzidos entre paradigmas de ranking, permitindo construir pipelines híbridos mais leves, reutilizáveis e ajustáveis ao regime de dados disponível.

6 CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho investigou a aproximação entre Recuperação de Informação (IR) e Sistemas de Recomendação (RS) a partir da hipótese de que ambos podem ser tratados como problemas de ranqueamento. Em vez de propor uma arquitetura unificada, avaliamos se algoritmos nativos de um paradigma poderiam operar no outro por meio da transformação das representações de entrada. Para isso, propusemos duas adaptações cruzadas: *Tag-Query* (IR $\rightarrow$ RS), que reescreve históricos de usuários como consultas textuais, e *Query-as-User*

(RS $\rightarrow$ IR), que reescreve consultas como pseudo-usuários em uma estrutura colaborativa.

Os resultados mostram que adaptações representacionais simples podem viabilizar transferência significativa entre busca e recomendação. O *Tag-Query* aproxima-se de recomendadores nativos em cenários ricos em metadados textuais, enquanto o *Query-as-User* obtém ganhos de cobertura ao explorar sinais colaborativos em tarefas de busca. Ao mesmo tempo, os experimentos indicam que essa transferência depende da disponibilidade e da qualidade dos sinais textuais e colaborativos presentes em cada domínio.

Mais do que propor métodos específicos, este trabalho evidencia que algoritmos de busca e recomendação podem ser reutilizados em pipelines híbridos por meio da reformulação de suas entradas, oferecendo uma alternativa complementar às abordagens baseadas em arquiteturas unificadas. Como trabalhos futuros, pretendemos explorar novas estratégias de adaptação, incorporar modelos de linguagem na síntese de consultas e ampliar a avaliação para outros domínios, regimes de esparsidade e cenários de custo computacional.

Limitações

Este estudo possui três limitações principais. Primeiro, avaliamos apenas uma instanciação de cada direção de adaptação, utilizando BM25 no *Tag-Query* e LightFM no *Query-as-User*, o que restringe a generalização dos resultados para outros modelos de IR e RS. Segundo, o desempenho das adaptações depende da disponibilidade relativa de sinais textuais e colaborativos, como observado nos diferentes comportamentos entre os conjuntos de dados avaliados. Por fim, não realizamos análises sistemáticas de sensibilidade de parâmetros nem investigamos outras combinações de modelos, aspectos que permanecem como direções para trabalhos futuros.

CUIDADOS ÉTICOS

Utilizamos neste trabalho apenas *datasets* públicos e abertos para fins de pesquisa. Tais bases de dados já se encontram devidamente pseudoanonimizadas por seus respectivos criadores originais, não permitindo a reidentificação direta dos usuários. Portanto, esta pesquisa está em conformidade com as diretrizes da Lei Geral de Proteção de Dados Pessoais (LGPD) (Lei n.º 13.709/2018) no que tange ao tratamento de dados para fins acadêmicos. Este estudo não envolveu nova coleta de dados com seres humanos, atendendo, portanto à Resolução CNS n.º 510/2016).

AGRADECIMENTOS

Esta pesquisa foi financiada por: CNPq, Capes, Fapemig, Fapesp, AWS, NVIDIA, CIIA-Saúde e Instituto Nacional de Ciência e Tecnologia em Inteligência Artificial Responsável para Linguística Computacional, Tratamento e Disseminação de Informação (INCT-TILD-IAR; 408490/2024-1).

Ferramentas de Inteligência Artificial Generativa foram utilizadas apenas como suporte à revisão gramatical de todo conteúdo. Os autores realizaram uma revisão completa do texto e assumem total responsabilidade pela integridade das informações apresentadas.

REFERÊNCIAS

- [1] Yan Andrade, Naan Vasconcelos, Elisa Tuler, Adriano Pereira, and Leonardo Rocha. 2025. Uma Abordagem Híbrida para Sistemas de Recomendação com Base em Avaliações Textuais e Filtragem Colaborativa. In *Proceedings of the 31st Brazilian Symposium on Multimedia and the Web* (Rio de Janeiro/RJ). SBC, Porto Alegre, RS, Brasil, 455–463. <https://doi.org/10.5753/webmedia.2025.14831>
- [2] Nicholas J. Belkin and W. Bruce Croft. 1992. Information filtering and information retrieval: two sides of the same coin? *Commun. ACM* 35 (1992), 29–38. <https://api.semanticscholar.org/CorpusID:14679722>
- [3] Guilherme Bittencourt, Naan Vasconcelos, Yan Andrade, Nicollas Silva, Washington Cunha, Diego Roberto Colombo Dias, Marcos André Gonçalves, and Leonardo Rocha. 2026. Review-Aware Recommender Systems (RARSS): Recent Advances, Experimental Comparative Analysis, Discussions, and New Directions. *ACM Comput. Surv.* 58, 1 (2026), 21:1–21:49. <https://doi.org/10.1145/3744661>
- [4] Oscar Celma and Pedro Cano. 2008. From hits to niches? or how popular artists can bias music recommendation and discovery. In *Proceedings of the 2nd KDD workshop on large-scale recommender systems and the netflix prize competition*.
- [5] H.-H. Chen, W.-J. Duh, H.-H. Huang, M. P. Kato, J. Mothe, and B. Poblete (Eds.). 2023. *Proc. 46th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2023)*. ACM Press, New York, USA.
- [6] Murilo Cunha, Marília Silveira, Brenda Santana, Larissa Freitas, and Ulisses Corrêa. 2025. Optimizing and Evaluating a Retrieval-Augmented Generation System for Normative Document Retrieval in Hospital Settings. In *Proceedings of the 31st Brazilian Symposium on Multimedia and the Web* (Rio de Janeiro/RJ). SBC, Porto Alegre, RS, Brasil, 385–393. <https://doi.org/10.5753/webmedia.2025.16029>
- [7] Washington Cunha, Celso França, Guilherme Fonseca, Leonardo Rocha, and Marcos André Gonçalves. 2023. An Effective, Efficient, and Scalable Confidence-Based Instance Selection Framework for Transformer-Based Text Classification. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 665–674. <https://doi.org/10.1145/3539618.3591638>
- [8] Washington Cunha, Felipe Viegas, Celso França, Guilherme Fonseca, Leonardo Rocha, and Marcos André Gonçalves. 2025. A Noise-Oriented and Redundancy-Aware Instance Selection Framework. *ACM Transactions on Information Systems* 43, 1 (2025). <https://doi.org/10.1145/3701769>
- [9] David Fernandes, Edleno Silva de Moura, Berthier A. Ribeiro-Neto, Altigran Soares da Silva, and Marcos André Gonçalves. 2007. Computing Block Importance for Searching on Web Sites. In *Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management*. ACM, 165–174. <https://doi.org/10.1145/1321440.1321466>
- [10] Shijie Geng, Shuchang Liu, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang. 2022. Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5). In *Proceedings of the 16th ACM Conference on Recommender Systems* (Seattle, WA, USA) (*RecSys '22*). Association for Computing Machinery, New York, NY, USA, 299–315. <https://doi.org/10.1145/3523227.3546767>
- [11] Yuqi Gong, Xichen Ding, Yehui Su, Kaiming Shen, Zhongyi Liu, and Guannan Zhang. 2023. An Unified Search and Recommendation Foundation Model for Cold-Start Scenario. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, CIKM 2023, Birmingham, United Kingdom, October 21–25, 2023*, Ingo Frommholz, Frank Hopfgartner, Mark Lee, Michael Oakes, Mounia Lalmas, Min Zhang, and Rodrygo L. T. Santos (Eds.). ACM, 4595–4601. <https://doi.org/10.1145/3583780.3614657>
- [12] F Maxwell Harper and Joseph A Konstan. 2015. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)* 5, 4 (2015).
- [13] Zhankui He, Zhouhang Xie, Harald Steck, Dawen Liang, Rahul Jha, Nathan Kallus, and Julian J. McAuley. 2025. Reindex-Then-Adapt: Improving Large Language Models for Conversational Recommendation. In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining, WSDM 2025*. 866–875. <https://doi.org/10.1145/3701551.3703573>
- [14] K. Järvelin and J. Kekäläinen. 2002. Cumulated Gain-Based Evaluation of IR Techniques. *ACM Transactions on Information Systems (TOIS)* 20, 4 (October 2002), 422–446.
- [15] Omar Khattab and Matei Zaharia. 2020. ColBERT: Efficient and Effective Passage Search via Contextualized Late Interaction over BERT. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 39–48. <https://doi.org/10.1145/3397271.3401075>
- [16] Maciej Kula. 2015. Metadata Embeddings for User and Item Cold-start Recommendations. In *Proceedings of the 2nd Workshop on New Trends on Content-Based Recommender Systems (CBRecSys 2015) (CEUR Workshop Proceedings, Vol. 1448)*. 14–21.
- [17] Yongqi Li, Xinyu Lin, Wenjie Wang, Fuli Feng, Liang Pang, Wenjie Li, Liqiang Nie, Xiangnan He, and Tat-Seng Chua. 2024. A Survey of Generative Search and Recommendation in the Era of Large Language Models. *ArXiv abs/2404.16924* (2024). <https://api.semanticscholar.org/CorpusID:269430314>
- [18] Jimmy Lin, Xueguang Ma, Sheng-Chieh Lin, Jheng-Hong Yang, Ronak Pradeep, and Rodrigo Nogueira. 2021. Pyserini: An Easy-to-Use Python Toolkit to Support Replicable IR Research with Sparse and Dense Representations. *CoRR abs/2102.10073* (2021). [arXiv:2102.10073](https://arxiv.org/abs/2102.10073) <https://arxiv.org/abs/2102.10073>
- [19] Vinicius Machado, Murilo Schmitt, and Eduardo Spínosa. 2025. Extrapolation-Based Data Augmentation for Sequential Recommendation. In *Proceedings of the 31st Brazilian Symposium on Multimedia and the Web* (Rio de Janeiro/RJ). SBC, Porto Alegre, RS, Brasil, 238–247. <https://doi.org/10.5753/webmedia.2025.15127>
- [20] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press.
- [21] S. Merlo, G. Faggioli, and N. Ferro. 2025. Joint Information Retrieval and Recommendation: a Reproducibility Study. In *Proc. 33rd Italian Symposium on Advanced Database Systems (SEBD 2025)*, I. Bartolini, G. Cima, D. Firmani, D. Lembo, D. Martignghi, M. Mecella, and A. Poggi (Eds.). CEUR Workshop Proceedings (CEUR-WS.org), ISSN 1613-0073, <https://ceur-ws.org/Vol-4182/>, 165–175.
- [22] Marc Najork. 2023. Generative Information Retrieval. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023, Taipei, Taiwan, July 23–27, 2023*, Hsin-Hsi Chen, Wei-Jou (Edward) Duh, Hen-Hsen Huang, Makoto P. Kato, Josiane Mothe, and Barbara Poblete (Eds.). ACM, 1. <https://doi.org/10.1145/3701769>

- [//doi.org/10.1145/3539618.3591871](https://doi.org/10.1145/3539618.3591871)
- [23] Michael J. Pazzani and Daniel Billsus. 2007. Content-Based Recommendation Systems. In *The Adaptive Web*. Springer, 325–341.
 - [24] Gustavo Penha, Ali Vardasbi, Enrico Palumbo, Marco de Nadai, and Hugues Bouchard. 2024. Bridging Search and Recommendation in Generative Retrieval: Does One Task Help the Other?. In *Proceedings of the 18th ACM Conference on Recommender Systems*. <https://doi.org/10.1145/3640457.3688123>
 - [25] Stephen Robertson and Hugo Zaragoza. 2009. The Probabilistic Relevance Framework: BM25 and Beyond. *Foundations and Trends in Information Retrieval* 3, 4 (2009), 333–389. <https://doi.org/10.1561/15000000019>
 - [26] Weichen Shen. 2017. DeepCTR: Easy-to-use, Modular and Extensible package of deep-learning based CTR models. <https://github.com/shenweichen/deepctr>.
 - [27] Teng Shi, Jun Xu, Xiao Zhang, Xiaoxue Zang, Kai Zheng, Yang Song, and Enyun Yu. 2025. Unified Generative Search and Recommendation. *arXiv preprint arXiv:2504.05730* (2025).
 - [28] Zihua Si, Zhongxiang Sun, Xiao Zhang, Jun Xu, Xiaoxue Zang, Yang Song, Kun Gai, and Ji-Rong Wen. 2023. When Search Meets Recommendation: Learning Disentangled Search Representation for Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (Taipei, Taiwan) (SIGIR '23)*. Association for Computing Machinery, New York, NY, USA, 1313–1323. <https://doi.org/10.1145/3539618.3591786>
 - [29] Thiago Silva, Nicollas Silva, Heitor Werneck, Carlos Mito, Adriano CM Pereira, and Leonardo Rocha. 2022. Irec: An interactive recommendation framework. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 3165–3175.
 - [30] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM international conference on information and knowledge management*. 1441–1450.
 - [31] Juntao Tan, Shuyuan Xu, Wenyue Hua, Yingqiang Ge, Zelong Li, and Yongfeng Zhang. 2024. IDGenRec: LLM-RecSys Alignment with Textual ID Learning. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2024, Washington DC, USA, July 14-18, 2024*, Grace Hui Yang, Hongning Wang, Sam Han, Claudia Hauff, Guido Zuccon, and Yi Zhang (Eds.). ACM, 355–364. <https://doi.org/10.1145/3626772.3657821>
 - [32] Ash Vardanian. 2023. *USearch by Unum Cloud*. <https://doi.org/10.5281/zenodo.7949416>
 - [33] Felipe Viegas, Washington Cunha, Christian Gomes, Antônio Pereira, Leonardo Rocha, and Marcos Goncalves. 2020. CluHTM: Semantic Hierarchical Topic Modeling Based on CluWords. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 8138–8150. <https://doi.org/10.18653/v1/2020.acl-main.724>
 - [34] Jesse Vig, Shilad Sen, and John Riedl. 2012. The Tag Genome: Encoding Community Knowledge to Support Novel Interaction. *ACM Trans. Interact. Intell. Syst.* 2, 3, Article 13 (Sept. 2012), 44 pages. <https://doi.org/10.1145/2362394.2362395>
 - [35] Liang Wang, Nan Yang, Xiaolong Huang, Linjun Yang, Rangan Majumder, and Furu Wei. 2024. Multilingual E5 Text Embeddings: A Technical Report. *arXiv preprint arXiv:2402.05672* (2024).
 - [36] Qianyi Wang. 2025. Comparison of Different Recommendation Models based on Amazon Platform Dataset. In *Proceedings of the 2024 5th International Conference on Computer Science and Management Technology (ICCSMT '24)*. Association for Computing Machinery, New York, NY, USA, 74–79. <https://doi.org/10.1145/3708036.3708048>
 - [37] Wei Wei, Xubin Ren, Jiabin Tang, Qinyong Wang, Lixin Su, Suqi Cheng, Junfeng Wang, Dawei Yin, and Chao Huang. 2024. LLMRec: Large Language Models with Graph Augmentation for Recommendation. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining, WSDM 2024, Merida, Mexico, March 4-8, 2024*, Luz Angelica Caudillo-Mata, Silvio Lattanzi, Andrés Muñoz Medina, Leman Akoglu, Aristides Gionis, and Sergei Vassilvitskii (Eds.). ACM, 806–815. <https://doi.org/10.1145/3616855.3635853>
 - [38] Shiguang Wu, Wenda Wei, Mengqi Zhang, Zhumin Chen, Jun Ma, Zhaochun Ren, Maarten de Rijke, and Pengjie Ren. 2024. Generative Retrieval as Multi-Vector Dense Retrieval. In *Proceedings of the 47th International ACM Conference on Research and Development in Information Retrieval, SIGIR, Grace Hui Yang, Hongning Wang, Sam Han, Claudia Hauff, Guido Zuccon, and Yi Zhang (Eds.)*. ACM, 1828–1838. <https://doi.org/10.1145/3626772.3657697>
 - [39] J. Xie, S. Liu, G. Cong, and Z. Chen. 2024. UnifiedSSR: A unified framework of sequential search and recommendation. In *Proceedings of the ACM Web Conference 2024*. 3410–3419.
 - [40] Hamed Zamani and W. Bruce Croft. 2018. Joint Modeling and Optimization of Search and Recommendation. <https://arxiv.org/abs/1807.05631>
 - [41] Hamed Zamani and W. Bruce Croft. 2020. Learning a Joint Search and Recommendation Model from User-Item Interactions. In *Proceedings of the 13th International Conference on Web Search and Data Mining (WSDM'20)*. 717–725. <https://doi.org/10.1145/3336191.3371818>
 - [42] Hamed Zamani and W. Bruce Croft. 2020. Learning a Joint Search and Recommendation Model from User-Item Interactions. In *Proceedings of the 13th International Conference on Web Search and Data Mining (Houston, TX, USA) (WSDM '20)*. Association for Computing Machinery, New York, NY, USA, 717–725. <https://doi.org/10.1145/3336191.3371818>
 - [43] H. Zeng, S. Kallumadi, Z. Alibadi, R. Nogueira, and H. Zamani. 2023. A Personalized Dense Retrieval Framework for Unified Information Access, See [5], 121–130.
 - [44] K. Zhao, Y. Zheng, T. Zhuang, X. Li, and X. Zeng. 2023. Joint Learning of E-commerce Search and Recommendation with a Unified Graph Neural Network, See [5], 1461–1469.
 - [45] Yuyue Zhao, Jiancan Wu, Xiang Wang, Wei Tang, Dingxian Wang, and Maarten de Rijke. 2024. Let Me Do It For You: Towards LLM Empowered Recommendation via Tool Learning. In *Proceedings of the 47th International ACM Conference on Research and Development in Information Retrieval, SIGIR, Grace Hui Yang, Hongning Wang, Sam Han, Claudia Hauff, Guido Zuccon, and Yi Zhang (Eds.)*. ACM, 1796–1806. <https://doi.org/10.1145/3626772.3657828>