



Sistemas de recomendação em bibliotecas: iniciativas e proposta de um modelo teórico híbrido

Recommender systems in libraries: initiatives and proposal of a theoretical hybrid model

Bruna Beatriz de Moura Vieira ¹

<anurbvieira1993@gmail.com>

Ketry Gorete Farias dos Passos ²

<ketry.passos@udesc.br>

Vanessa Marie Salm ³

<vanessa.salm@udesc.br>

Submetido em: 08-05-2023	Aceito em: 08-07-2023
--------------------------	-----------------------

Resumo: Sistemas de recomendação são ferramentas de software que têm como objetivo gerar uma estimativa da utilidade que um item pode ter para um usuário, bem como analisar interações entre usuários e itens. São divididos entre sistemas de recomendação clássicos, que utilizam apenas um tipo de técnica de filtragem, e híbridos, que utilizam dois ou mais tipos de técnicas de filtragem aliados a uma técnica de recomendação híbrida. O objetivo deste artigo consiste em apresentar os sistemas de recomendação em bibliotecas e propor um modelo de recomendação híbrido teórico, adaptado ao contexto das bibliotecas. Para isto, foi utilizada uma abordagem qualitativa, exploratória, descritiva e bibliográfica para explorar o tema

1 Graduada em Biblioteconomia pela Universidade do Estado de Santa Catarina (2022). Tem experiência na área de Ciência da Computação, com ênfase em Sistemas de Computação.

2 Doutora e mestre em Ciência da Informação (2012-2019) pelo Programa de Pós-Graduação em Ciência da Informação - PGCIN/UFSC. Especialista em Gestão de Bibliotecas Escolares pela Universidade Federal de Santa Catarina - UFSC (2015). Graduada em Biblioteconomia - Hab. Gestão da Informação pela Universidade do Estado de Santa Catarina - UDESC (2009).

3 Doutora em Engenharia e Gestão do Conhecimento - Universidade Federal de Santa Catarina. Mestre em Sistemas de Informação - Arizona State University (2010). Graduada em Administração com Hab. em Gestão da Informação pelo Centro de Educação Superior (2006).



de sistemas de recomendação e sua aplicação em bibliotecas como uma resposta às crescentes demandas informacionais dos usuários. Foi percebido que a maioria das iniciativas de sistemas de recomendação em uso atualmente são do modelo clássico, e utilizam apenas dois tipos de filtragem – por conteúdo e colaborativa. Portanto, é necessário que pesquisas futuras tentem entender o motivo por trás da adoção dos sistemas de recomendação do modelo clássico em detrimento dos sistemas de recomendação híbridos, e que é necessário um maior enfoque no desenvolvimento e aplicação de sistemas de recomendação híbridos em bibliotecas, dado às suas vantagens em relação ao modelo clássico. Também é sugerida a realização de estudos sobre as outras diferentes técnicas de filtragem, pelo fato de possuírem o potencial de gerar novos tipos de estratégias de recomendação em bibliotecas.

Palavras-chave: Sistemas de recomendação; sistemas de recomendação híbridos; filtragem por conteúdo; filtragem colaborativa; recuperação da informação.

1 INTRODUÇÃO

A International Federation of Library Associations and Institutions - IFLA afirma no *Statement on Libraries and Intellectual Freedom* (1999) que os seres humanos têm o direito fundamental de acessar informações e que a liberdade intelectual é responsabilidade central de bibliotecas e de profissionais da área de informação, que devem visar a garantia do acesso à informação, o apoio ao aprendizado durante todo o ciclo de vida do usuário e equidade.

Observa-se que as bibliotecas vêm sofrendo com um aumento de demanda à medida que as necessidades informacionais dos usuários se tornam mais complexas em vista à crescente produção de informação, e problemas como a infoxicação estão se tornando mais presentes. Segundo Kumari, Raja e Sundaravalli (2021, p. 4456), “a sobrecarga de informações leva à infoxicação, também conhecida como poluição informacional, onde a contaminação de informações úteis com informações incorretas ocorre” (tradução nossa).



Embora o uso da inteligência artificial, também conhecida como IA, e automação tenham se tornado tópicos amplamente discutidos no âmbito da Biblioteconomia nos anos recentes, nota-se que suas histórias são intimamente envolvidas com a história das bibliotecas em si. Segundo Fachin (2009, p. 260): “A necessidade de recuperar a informação de forma rápida, precisa e relevante remonta à própria história da humanidade. Vannevar Bush, já em 1945, lançava a ideia da automatização da informação [...]”.

A International Federation of Library Associations and Institutions – IFLA, no relatório *IFLA Statement on Libraries and Artificial Intelligence* (2020), recomenda a interação de bibliotecários com pesquisadores da área de IA, automação e demais ciências computacionais para a criação de aplicações para bibliotecas que respondam à crescente necessidade informacional de seus usuários, assim tornando-os agentes centrais do processo de RI, capaz de advogarem por si próprios de maneira esclarecida.

Levando em consideração os argumentos expostos, entende-se que uma das maneiras de resolver tal dilema seria pela aplicação de um sistema de recomendação híbrido, ou SR híbrido. Segundo Amato e Straccia (1999, p.184):

Prover buscas de informação personalizadas e serviços de entrega de informações, como serviços adicionais na busca uniforme e genérica de informação oferecida hoje, é provavelmente o primeiro passo para fazer com que informações relevantes estejam disponíveis para as pessoas nos formatos e quantidades adequados, com o nível certo de detalhe, e no tempo certo através dos canais corretos, e com esforço mínimo por parte do usuário (tradução nossa).

Assim surgiu a seguinte lacuna de pesquisa: como um SR híbrido pode ser aplicado em bibliotecas, e quais componentes são necessários para sua utilização?

Para isso, o presente artigo tem como objetivo geral apresentar os sistemas de recomendação em bibliotecas e propor um modelo de recomendação híbrido teórico, adaptado ao contexto das bibliotecas.

Para este fim, o artigo é composto por uma revisão de literatura dividida em quatro partes: na primeira parte apresentam-se conceitos relacionados a sistemas de recomendação - SRs; em seguida descrevem-se as principais



técnicas utilizadas em SRs; após apresentam-se os SRs baseados em filtragem colaborativa, ou *FC*, e em filtragem por conteúdo, ou *FCO*, aplicados no contexto das bibliotecas; depois apresentam-se os SRs híbridos. Após, apresenta-se uma análise de iniciativas de SRs em uso em bibliotecas nos dias de hoje.

2 SISTEMAS DE RECOMENDAÇÃO

Sistemas de recomendação, ou SRs, são softwares compostos por ferramentas e técnicas que têm como finalidade auxiliar usuários na recuperação de informação por meio da sugestão de itens que vão ao encontro de suas necessidades.

Existem tipos diferentes de SRs, cada qual com foco na recomendação de um tipo específico de item, como filmes, livros, música, entre outros. Eles são formulados para fornecer sugestões personalizadas a um ou mais usuários no contexto em que estão inseridos (RICCI; ROKACH; SHAPIRA, 2022). Alguns dos SRs mais populares atualmente são o do Spotify, o da Netflix e o da Amazon.

Embora os SRs sejam associados às dificuldades associadas à recuperação de informação e tomada de decisão no contexto da sociedade da informação, é possível verificar que eles são fruto de um processo histórico cujo início data do século XVIII, onde a preocupação com a explosão informacional já era registrada por autores como Denis Diderot (1751, p.314) em sua obra *Encyclopédie*:

Enquanto os séculos continuarem a se desdobrar, o número de livros irá crescer continuamente, e nenhuma pessoa poderá prever quando chegará o tempo em que será quase tão difícil aprender qualquer coisa através de livros quanto através do estudo do universo inteiro (tradução nossa).

Nota-se que não é incomum que os usuários se baseiem nas opiniões de pessoas em seu círculo social para escolher produtos, livros, filmes, entre outros itens.



Os SRs representam uma maneira de tentar captar a essência destas interações sociais, apontando os usuários para novos itens, ou para itens com os quais os usuários já tenham tido contato anterior, mas que gostariam de reconsumir (RICCI; ROKACH; SHAPIRA, 2022).

Para entender o funcionamento e as vantagens dos SRs e sua ligação com as bibliotecas, é necessário fazer um exame de suas partes, principalmente no que diz respeito aos atores do sistema - conhecidos como stakeholders -, seus itens e suas interações.

Stakeholders são grupos ou indivíduos que podem afetar a entrega de recomendações para um usuário, ou serem afetados por ela (ABDOLLAHPOURI; BURKE, 2022). Podem ser separados em três grupos: consumidores/usuários (consomem recomendações), provedores (fornecem os itens que o usuário está procurando) e o sistema em si, representado por uma plataforma associada a um SR cujo objetivo é fazer uma ponte entre os provedores e os usuários, e os guiar até os itens.

Nota-se que os usuários não são meros atores passivos no que diz respeito à constituição do sistema. Enquanto os sistemas influenciam os resultados, o perfil do usuário também influencia no tipo de SR escolhido.

Já os itens são, junto dos usuários, um dos elementos mais importantes dos SRs, pois sem eles não existe o ímpeto inicial para a realização da recomendação em si. Eles representam todos os objetos que serão recomendados aos usuários, acompanhados de metadados de identificação. Em uma biblioteca, os itens são representados por livros, periódicos e demais recursos dos quais a biblioteca dispõe.

Outro elemento relevante são as interações dos SRs. Elas consistem no diálogo entre o usuário e um item, que gera informações que serão armazenadas em históricos para alimentar o algoritmo de recomendações do sistema. Podemos citar como um exemplo popular de interação as avaliações de clientes da Amazon: ao comprar um produto, um usuário pode realizar uma avaliação dele, na forma de notas em uma escala de uma a cinco estrelas.

As interações podem ser de vários tipos, como numéricas (quando o usuário pode dar notas ao item utilizando uma escala numérica), ordinais (avalia



opiniões e atitudes (BERNARDINO, 2022), e ocorrem quando o usuário é encorajado a escolher uma opção que determine o quanto ele concorda ou discorda de algo) e binárias (escolha entre duas alternativas, como sim ou não, ou positivo versus negativo).

No contexto das bibliotecas, os usuários são representados pelas pessoas que usufruem dos serviços das bibliotecas e que buscam itens, sejam eles livros, periódicos, dentre outros, oferecidos pela biblioteca, que por sua vez atua como o provedor. As interações entre os usuários e os itens presentes na biblioteca alimentariam o algoritmo do SR, que pode tomar formas diferentes, se adequando ao tipo específico de biblioteca em que está inserido, como bibliotecas universitárias, escolares, hospitalares, entre outras.

Partindo dos pontos anteriores, torna-se claro que existe uma sinergia entre as vantagens apresentadas pelos SRs e os desafios presentes nas bibliotecas atualmente, e que a aplicação de SRs em bibliotecas representa uma evolução natural no processo de RI e letramento informacional. Portanto, serão examinadas a seguir as técnicas utilizadas para a construção de SRs.

2.1 Técnicas utilizadas em SRs

Os SRs, além das partes citadas na seção anterior, são constituídos por diversas técnicas de filtragem, que têm como objetivo filtrar itens nos quais o usuário pode se interessar, colaborando com a recomendação final. Entre as técnicas de filtragem mais utilizadas atualmente temos a filtragem colaborativa, *FC*, e a filtragem por conteúdo, *FCO*.

Em relação a *FC*, pode-se considerar que ela faz parte de um grupo de técnicas de abordagens personalizadas, e que é um dos tipos mais populares de filtragem na atualidade (RICCI; ROKACH; SHAPIRA, 2022) e que tem como objetivo tentar prever quais itens são mais relevantes para um usuário a partir de uma análise de feedbacks de outros usuários com interesses similares. A lógica deste método considera que se um usuário A possui as mesmas características de um usuário B, que mostrou um feedback positivo para um



certo item, o usuário A também receberá a recomendação de maneira positiva, assim como outros usuários .

Pode-se utilizar o seguinte exemplo: uma biblioteca possui um usuário A, de 15 a 22 anos de idade e com um interesse em livros de romance. Para saber se ele(a) gostaria do livro “Orgulho e Preconceito”, de Jane Austen, analisam-se feedbacks que outros usuários semelhantes deram ao livro. Se os feedbacks forem positivos, o SR recomendará o livro. Se não, ele procurará outras alternativas mais adequadas.

A FC tem como uma de suas vantagens a maior facilidade na classificação de itens em comparação com outros tipos de filtragem, o que a torna atrativa para aplicação em bibliotecas. Nota-se que é necessária apenas a estruturação de um questionário para alimentar um histórico que servirá como base para recomendações futuras.

Como desvantagens podemos citar a hiperespecialização, que ocorre quando a variedade das recomendações é perdida pela sugestão de apenas uma categoria de itens (NIKOLAKOPOULOS *et al.*, 2022), e a dificuldade em lidar com a escalabilidade no SR, como explicado por Chen e McLeod (2006, p.118): “Dados de usuários são dinâmicos, o que significa que eles variam em um curto período de tempo. Usuários atuais podem mudar seus padrões de comportamento, e novos usuários podem entrar no sistema a qualquer momento” (tradução nossa).

A FC também pode ser classificada pelos tipos de abordagens em que pode ser utilizada, como abordagens baseadas em métodos de vizinhança e abordagens baseadas em modelos.

Na abordagem associada aos métodos de vizinhança, os feedbacks dados pelos usuários em outros itens são utilizados na predição de feedbacks para itens com quais o usuário alvo ainda não tenha interações. Esta predição pode ser baseada em usuários, onde feedbacks de *vizinhos* - usuários que possuem um padrão de feedbacks similar ao do usuário a quem o item está sendo recomendado - são utilizados para gerar recomendações (KONSTAN *et al.*, 1997), ou itens, que utilizam os feedbacks do próprio usuário em itens similares ao que está sendo recomendado (DESHPANDE; KARYPIS, 2004). Para os itens



serem considerados similares, diversos usuários do SR devem realizar feedbacks parecidos no mesmo item (NIKOLAKOPOULOS *et al.*, 2022). Por exemplo, para realizar a recomendação do livro “Eu robô”, do autor Isaac Asimov, para um usuário, o SR analisaria feedbacks do usuário a outros livros do mesmo gênero literário e os feedbacks de outros usuários do sistema para estabelecer a relação de similaridade e realizar a recomendação.

Já a abordagem de *FC* baseada em modelos tem como característica a utilização dos feedbacks dos usuários em um processo contínuo de aprendizagem de um modelo preditivo, ao invés de utilizá-los diretamente na previsão final. Para entender melhor esta abordagem, pode-se imaginar o seguinte cenário: um SE tem que gerar uma recomendação de um livro para um usuário A, de 20 a 35 anos e interessado em suspense. O SR vem sendo alimentado por feedbacks de diversos usuários com o mesmo perfil do usuário e ao invés de indicar, por exemplo, o livro de autoajuda “Quem mexeu no meu queijo?”, de Spencer Johnson, ele indicaria o livro de suspense “O talentoso Ripley”, de Patrícia Highsmith, pois ele já aprendeu, a partir dos feedbacks dos outros usuários do sistema com perfis similares, que este livro tem maior chance de ser bem recebido pelo usuário A. Após a recomendação, o sistema recebe o feedback do usuário A sobre o livro indicado e o armazena para aprimorar seu modelo preditivo.

A segunda técnica de filtragem mais utilizada atualmente nos SRs é a *FCO*. Ela tem como objetivo aprender as preferências dos usuários e gerar recomendações personalizadas baseando-se tanto nas informações disponíveis sobre os itens em si, quanto nos feedbacks dos usuários (BURKE, 2007). Pode ser classificada pelos tipos de abordagens em que pode ser utilizada, como a abordagem de modelo clássico e a abordagem de modelo aliada a abordagens semânticas.

No modelo clássico, as características do usuário são determinadas a partir das características dos itens com os quais ele tenha interagido. Estas características dos itens são armazenadas em forma de palavras-chave no sistema e são utilizadas pelo sistema para calcular a similaridade entre diferentes itens (RICCI; ROKACH; SHAPIRA, 2022). A recomendação é feita



através de uma comparação das características do usuário e das dos itens com os quais tenha interagido, assim resultando em uma representação do nível de interesse do usuário naquele objeto (LOPS; GEMMIS; SEMEMARO, 2011).

Para entender este conceito, podemos analisar o seguinte exemplo: se um usuário A de uma biblioteca deu um feedback positivo sobre um livro de romance, o SR entende que uma das características do perfil do usuário A é a de que ele tem interesse por livros deste gênero, então ele sugeriria outros livros similares.

Já na abordagem de *FCO* aliada a abordagens semânticas, os itens e usuários são representados por conceitos ao invés de palavras-chave, diferentemente da abordagem do modelo clássico. Para este fim, podem ser utilizadas técnicas exógenas ou endógenas de indexação semântica.

A indexação semântica endógena é baseada na hipótese de distribuição dos modelos distribucionais semânticos, que diz que o significado de uma palavra em um documento pode ser inferido por outros documentos que estejam inseridos no mesmo contexto. Já a indexação exógena tenta determinar o significado de uma palavra por meio de bases externas de conhecimento, como dicionários, enciclopédias etc. (MUSTO *et al.*, 2022).

Um SR baseado em *FCO* geralmente é composto por três componentes, cada qual realiza um passo da recomendação (MUSTO *et al.*, 2022): o analisador de conteúdo, ou *content analyzer* (representa itens através da extração de palavras-chave ou conceitos de suas descrições e alimenta os demais componentes), componente de aprendizagem de perfil, ou *profile learner* (constrói o perfil do usuário através dos itens com os quais ele interage e seus feedbacks, com a ajuda de algoritmos de aprendizagem de máquina (MITCHELL, 1997) e o componente de filtragem, ou *filtering component* (prediz se um item pode ser de interesse do usuário comparando as características de seu perfil com as características dos itens).

A seguir, serão examinadas as vantagens e desvantagens das técnicas citadas acima no contexto das bibliotecas.



2.2 SRs baseados em *FC* e *FCO* aplicados a bibliotecas

Baseado nas informações expostas anteriormente sobre os componentes necessários para a construção de um SR, é possível perceber que as bibliotecas se encontram numa posição privilegiada no que diz respeito à adoção e aplicação de um SR, pois já possuem bancos de dados compostos por dados não só de usuários, mas também dos itens que se encontram em seu acervo. Assim, se torna necessária somente a elaboração de um método pelo qual os usuários possam dar feedbacks nos materiais com os quais eles interagem.

O feedback nas bibliotecas pode ser considerado como explícito, mas, segundo Ricci, Rokach e Shapira (2022, p.10), “[...] o feedback explícito é muitas vezes indisponível, ou muito esparsa, já que vários usuários não se importam em dar o feedback”. Portanto, para tentar contornar este problema, é sugerido que o feedback seja realizado das seguintes maneiras no contexto das bibliotecas: *questionários*, aplicados durante a devolução dos itens que o usuário retirou para empréstimo, por meio de uma *ferramenta* no site da biblioteca que permita aos usuários darem feedbacks numéricos, binários ou ordinais aos itens, ou por *inferência*, através dos dados estatísticos de empréstimos aliados aos dados estatísticos de usuários.

Pode-se também afirmar que os SRs baseados em *FC* e em *FCO* são ambos interessantes para a aplicação em bibliotecas. Destacam-se os SRs baseados em *FC* associados aos métodos de vizinhança, devido à sua eficiência e facilidade de aplicação em comparação a outros métodos de filtragem (NIKOLAPOULOS *et al.*, 2022).

Embora os SRs apresentados anteriormente possuam vantagens em suas aplicações em bibliotecas, podemos citar como consequências não desejadas a hiperespecialização do sistema, problemas de escalabilidade devido ao alto número de itens e usuários das bibliotecas e a dificuldade em incentivar a interação entre os usuários e os itens por meio de feedbacks, o que pode levar a um baixo nível de confiabilidade.

Tem-se como uma possível solução para estes empecilhos: a aplicação de um SR híbrido, que possibilita a união de dois ou mais métodos de filtragem



com o objetivo de aumentar a eficiência e eficácia do sistema. A seguir, será feita uma análise deste tópico.

2.3 SRs híbridos

Os SRs híbridos surgiram como uma resposta aos problemas associados às diferentes técnicas de recomendação. Entende-se que juntando duas ou mais técnicas de recomendação e adicionando outros componentes a elas, como IA e machine learning, por exemplo, estas dificuldades podem ser supridas (RICCI; ROKACH; SHAPIRA, 2022).

Um fator importante para a construção e aplicação de SRs híbridos é a definição da estratégia de recomendação.

As estratégias de recomendação híbridas são compostas por, no mínimo, duas técnicas de filtragem aliadas a uma técnica de hibridização. São, neste caso, chamadas de componentes do sistema. A técnica de hibridização, por sua vez, pode ser dividida em sete tipos (BURKE, 2007): por peso, ou *weighted* (combinação da pontuação dada aos itens pelos componentes do sistema), por troca, ou *switching* (escolha feita pelo sistema entre seus componentes e sua aplicação no processo de recomendação), método misto, ou *mixed* (utiliza recomendações de diferentes componentes ao mesmo tempo), por combinação de características, ou *feature combination* (combina características de itens e sua utilização em um único algoritmo de recomendação), por aumento de características, ou *feature augmentation* (computação das características de um item por um componente, que alimentam um segundo componente no sistema para que ele possa realizar a recomendação final), por cascata, ou *cascade* (priorização de um dos componentes do sistema e uso do componente com menor prioridade para realizar desempates) e por meta-nível, ou *meta-level* (produção de um modelo por um dos componentes de recomendação do sistema, que então é utilizado pelo outro componente do sistema para gerar as recomendações finais).

Após estudos extensos, percebeu-se que as três seguintes estratégias de recomendações híbridas têm a maior média de recomendações corretas



(BURKE, 2007): aumento de características aliado a *FCO* e filtragem por conhecimento, cascata aliada a *FCO* e *FC*, e cascata aliada a filtragem por conhecimento e *FC*.

Sabe-se que as técnicas de *FCO* e *FC* são consideradas as mais utilizadas nos dias atuais, além de serem as de mais fácil aplicação em bibliotecas. Quando estas técnicas de filtragem são aliadas às técnicas de hibridização, seus processos são refinados e seus problemas em potencial, como a hiperespecialização, podem ser resolvidos.

Levando em consideração a revisão de literatura feita até aqui e a relação das estratégias de recomendação híbridas, em conjunto com o contexto do ambiente das bibliotecas, será examinada a seguir a pesquisa de iniciativas de SRs em uso atualmente em bibliotecas, como base para a proposta de um SR híbrido teórico em bibliotecas.

3 METODOLOGIA

A pesquisa caracteriza-se por uma abordagem qualitativa, exploratória, descritiva e bibliográfica com o objetivo principal de explorar o tema de SRs e sua aplicação em bibliotecas como uma resposta às crescentes demandas informacionais dos usuários. Segundo Bhandari (2020, p.1): “A pesquisa qualitativa envolve a coleta e análise de dados não numéricos (ex. texto, vídeo ou áudio) para entender conceitos, opiniões ou experiências (tradução nossa)”.

Como procedimentos metodológicos, empreendeu-se:

- Busca pelas palavras-chave "sistemas de recomendação", "sistemas de recomendação híbridos", "sistemas de recomendação em bibliotecas"; primariamente utilizadas no idioma português, mas também traduzidas para o inglês nas formas "recommender systems", " hybrid recommender systems", "recommender systems used in libraries";
- O período de busca compreendeu os meses entre 01/08/2022 e 01/11/2022;



- O levantamento bibliográfico foi feito no Google Scholar (<https://scholar.google.com.br/>), Science Direct (<https://www.sciencedirect.com/>) e BRAPCI (<https://www.brapci.inf.br/>);
- Revisão do levantamento bibliográfico;
- Pesquisa e seleção de iniciativas de SRs em uso atualmente, com base na sua popularidade, ampla adoção e relevância;
- Descrição de um modelo teórico de SR híbrido aplicado em uma biblioteca.

4 PRINCIPAIS INICIATIVAS DE SRs EM USO ATUALMENTE

A presente pesquisa teve como resultado o levantamento de três SRs utilizados em bibliotecas atualmente: *Related book links in Aleph OPAC*, *bx Recommender* e *Sciencedirect*. Os sistemas foram examinados em relação a seus componentes, para determinação do tipo de modelo que utilizam (clássico ou híbrido), em relação a suas funções e em relação ao contexto em que são utilizados. Com o objetivo de contextualizar a pesquisa, foram também levantados dados históricos sobre suas construções.

4.1 SRs utilizados em bibliotecas

Entende-se que SRs são softwares que têm como objetivo realizar recomendações de itens aos seus usuários que vão de encontro a suas necessidades específicas. Para alcançar este objetivo, os SRs seguem uma lógica interna pré-programada baseada em técnicas de filtragem e recomendação específicas. Eles podem realizar recomendações de inúmeros tipos de itens e são altamente personalizáveis, podendo funcionar em vários contextos.

Segundo Wolff *et al.* (2016, p. 9), “A sociedade tem se tornado cada vez mais dependente de dados, fazendo com que seja necessário garantir que todos os cidadãos sejam equipados com as habilidades necessárias para serem



alfabetizados em dados” (tradução nossa). Tal guinada à uma sociedade cada vez mais direcionada aos dados é um dos fatores que contribui para o número cada vez maior de SRs em uso atualmente.

Em bibliotecas, especialmente, foram identificados três SRs utilizados para contribuir com a alfabetização informacional de seus usuários, como o Related books links in Aleph OPAC, bX Recommender e Sciencedirect. É importante ressaltar que existem outros SRs em uso em bibliotecas atualmente, mas iremos focar nestes três exemplos pela sua popularidade, ampla adoção e relevância nos dias de hoje.

O **Related books links in Aleph OPAC** é uma extensão criada em 2010 para o software de gerenciamento de bibliotecas Aleph. Baseado em Perl e com o objetivo de realizar recomendações de itens de bibliotecas universitárias, ele se destaca pela facilidade na instalação e manutenção, sendo necessária somente a criação de duas tabelas e índices na base de dados Oracle, além da instalação de módulos Perl, que já estão inclusos no pacote de instalação (PHILLIPS, 2010).

Ele possui os seguintes métodos de recomendação, classificados por tipo de filtragem (KREBS, 2013): *peças que emprestaram esse item também emprestaram* (filtragem colaborativa por vizinhança, calcula recomendações com base no histórico de empréstimos de usuários em comum), trabalhos relacionados (filtragem por conteúdo, relaciona itens através dos cabeçalhos de assunto e áreas de classificação em registros MARC), outras edições incluem (filtragem por conteúdo, relaciona diferentes edições do mesmo item através da equivalência do ISBN).

Este SR é utilizado por várias universidades, como a Universidade de Dundee, na Escócia, e a PUCRS, no Brasil. Nota-se que a técnica de recomendação utilizada não é híbrida, pois cada método de recomendação utiliza apenas um único tipo de filtragem. O sistema resultante é de fácil utilização e compreensão, e se destaca pela sua eficiência, já que os cálculos de recomendações são feitos em segundo plano para manter a rapidez do sistema.



O **bX Recommender system** é um SR que tem suas origens em uma pesquisa conduzida por Herbert Van de Sompel e Johann Bolen, em conjunto com o Laboratório Nacional de Los Alamos, nos Estados Unidos, e com a empresa de automação para bibliotecas ExLibris (bX ARTICLE, 2022).

Seu objetivo é realizar recomendações de artigos a partir da mineração de dados de usuários de diferentes instituições acadêmicas ao redor do globo. Se o sistema detectar que um usuário visualizou dois ou mais artigos na mesma sessão de uso, ele realiza uma análise da conexão entre estes artigos e guarda esse registro em um histórico, que vai ser utilizado futuramente para gerar recomendações para outros usuários que visualizem os mesmos artigos, se assemelhando ao sistema de recomendações de serviços populares como a Netflix e Amazon (EX LIBRIS bX..., 2021).

Pelo fato deste recomendador possuir alto nível de granularidade e utilizar artigos provenientes de diversas revistas, editoras e plataformas (KREBS, 2013), considera-se que suas recomendações finais são verdadeiramente neutras no que diz respeito ao conteúdo dos artigos e à plataforma onde estão hospedados, assim tornando possível para o usuário tanto expandir o escopo de seus conhecimentos através de recomendações com alto nível de diversidade, como também atingir um nível mais alto de especificidade, se assim for desejado (bX ARTICLE, 2022).

É importante notar que, por utilizar técnicas de *FC* aliadas à *FCO*, o bX Recommender system pode ser considerado como um SR verdadeiramente híbrido. Atualmente, ele é utilizado na Universidade DePaul e na Universidade Lewis, ambas do estado de Illinois, nos Estados Unidos, entre outras.

Por fim, temos o **ScienceDirect**, que é um mecanismo de pesquisa de artigos científicos da Elsevier, baseado em um motor de pesquisas federado (HARNEGIE, 2013). Seu objetivo principal é combinar “publicações de conteúdo científico, técnico e médico confiáveis e completas, com funcionalidade inteligente e intuitiva para que os usuários possam permanecer informados em suas áreas e trabalhar de forma mais eficaz e eficiente” (tradução nossa) (ELSEVIER, 2021). O ScienceDirect é utilizado por mais de 15 milhões de



peessoas (L'HUILLIER, 2022) em todo o mundo, e possui mais de 18 milhões de artigos (SCIENCEDIRECT..., 2022).

Para realizar as recomendações aos usuários, o ScienceDirect conta com uma seção chamada Related Articles, cujo objetivo é gerar recomendações de artigos similares ao que o usuário está visualizando durante sua sessão de uso. Para este fim, é realizada uma análise de conteúdo e extração de palavras-chave do texto completo do artigo que está sendo visualizado pelo usuário (KREBS, 2013). Assim, nota-se que a técnica de recomendação utilizada por este sistema não é híbrida, pois utiliza somente a *FCO*.

Foram analisadas, então, as principais iniciativas de SRs que estão sendo utilizadas atualmente nas bibliotecas. Das três, nota-se que apenas uma utiliza técnicas de recomendação híbrida, e que todas utilizam as técnicas de *FCO* e *FC*.

Tendo em vista estas conclusões e os demais assuntos discutidos ao longo deste artigo, segue-se agora para a apresentação dos resultados, que consiste na proposta do modelo teórico de SR híbrido para aplicação em bibliotecas.

5 RESULTADOS

Após o levantamento das iniciativas de SRs utilizadas hoje em dia em bibliotecas, nota-se que existe uma lacuna no que diz respeito à utilização de SRs híbridos, vendo-se que as bibliotecas mostram predileção para o uso do modelo clássico. Sabemos, através da revisão de literatura feita anteriormente, que embora o modelo clássico de SRs funcionem bem, eles são sujeitos a alguns problemas como hiperespecialização, falta de feedbacks o suficiente para dar partida no sistema, problemas de escalabilidade, entre outros, e que os SRs híbridos surgiram como uma resposta a esses problemas.

Também foi possível determinar, através da pesquisa de Burke (2007) em parceria com a Universidade de Chicago, a Universidade da Califórnia e com o Escritório de Pesquisa Naval dos Estados Unidos, que três técnicas de hibridização apresentam os melhores resultados: a estratégia de aumento de



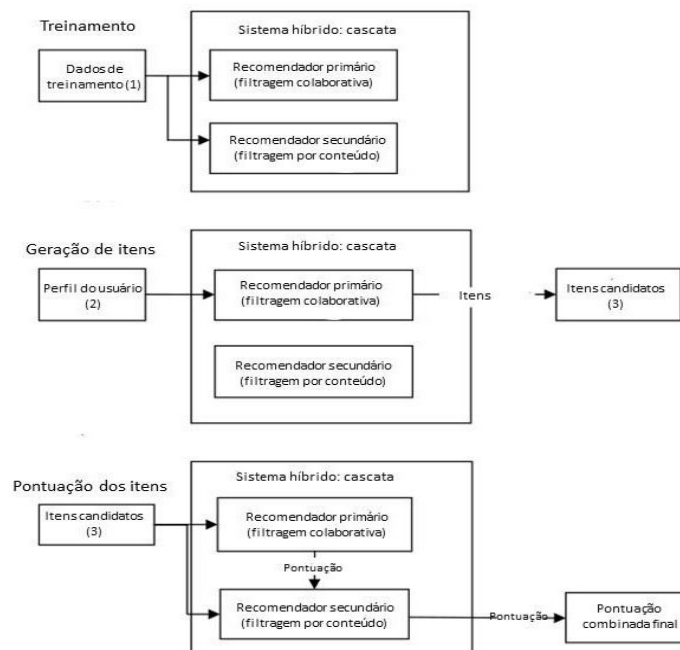
características aliada a *FCO* e filtragem por conhecimento, a estratégia de cascata aliada a *FCO* e *FC*, e, por fim, a estratégia de cascata aliada a filtragem por conhecimento e *FC*.

Para fazer uma ponte entre os SRs utilizados atualmente e o modelo de recomendação híbrido teórico proposto a seguir, a estratégia de cascata aliada a *FCO* e *FC* foi escolhida para ser utilizada no modelo proposto, assim utilizando as mesmas técnicas de filtragem já conhecidas e aplicadas nas iniciativas disponíveis de SRs em uso nos dias atuais vistas anteriormente, e as refinando com a ajuda da técnica de hibridização de cascata.

A técnica de cascata, como visto anteriormente, é baseada em um processo de refinamento iterativo que consiste na priorização de componentes de filtragem com o objetivo de construir uma ordem de preferência entre itens diferentes (ISINKAYE; FOLAJIMI; OJOKOH, 2015). Nela, o componente de filtragem que possui a maior prioridade - chamado de recomendador primário - realiza as recomendações, enquanto o componente de filtragem que possui a menor prioridade - chamado de recomendador secundário - é utilizado somente em casos de empate de recomendações (BURKE, 2007).

A Figura 1 apresenta um exemplo de funcionamento de um SR que utiliza esta estratégia de hibridização, utilizando o componente de *FC* como o recomendador primário e o componente de *FCO* como recomendador secundário. Podemos, a partir deste exemplo, analisar sua possível aplicação em bibliotecas.

Figura 1 – SR híbrido baseado na técnica de cascata aliado à FC e FCO



Fonte: Adaptado de Burke (2007).

O funcionamento deste sistema por ser dividido em três momentos:

- **Treinamento (1):** neste momento é realizada a etapa de treinamento do modelo, a partir de dados dos usuários - como seu histórico de empréstimo, idade, entre outros - e dos itens da biblioteca - como seus códigos de CDD/CDU e metadados - com o objetivo de refinar o sistema e o “ensinar” a realizar recomendações;
- **Geração de itens (2):** neste momento o sistema envia dados do perfil do usuário, como sua idade e histórico de empréstimo, para o recomendador primário com o objetivo de gerar uma lista de livros nos quais o usuário pode se interessar;
- **pontuação dos itens (3):** neste momento a lista de livros que foi gerada no passo anterior é enviada para o recomendador primário, que tem como objetivo pontuar os livros recomendados para construir uma ordem de preferência. Caso ocorra um empate entre as pontuações de dois ou mais livros, o recomendador secundário analisa a lista de livros gerada no passo anterior e a pontuação feita pelo recomendador primário e atua como o decisor final entre os



itens, aplicando um critério de desempate. Temos, então as pontuações e recomendações finais do sistema.

Observa-se que a ordem e as técnicas de filtragem utilizadas nos SRs híbridos têm impacto direto nas recomendações finais realizadas pelo sistema.

O fato de o recomendador primário utilizar a técnica de *FC*, que é baseada nos feedbacks de diferentes usuários do sistema, resulta em um sistema com recomendações mais diversas e redução do risco de hiperespecialização. Já a quebra dos empates da pontuação feita pelo recomendador secundário baseado em *FCO* ajuda a refinar ainda mais a recomendação. Invertendo a ordem das técnicas de filtragem, correríamos um risco maior de ter problemas com hiperespecialização, pois os itens seriam recomendados somente pelo seu conteúdo e a *FC* seria aplicada apenas a casos de empates na pontuação, o que resultaria em uma possível perda de variedade nas recomendações e de eficiência.

Vemos, então, como os métodos de hibridização funcionam e como um modelo teórico baseado no método de cascata aliado à *FC* e à *FCO* pode ser aplicado em bibliotecas.

A seguir, partindo da revisão de literatura feita ao longo do artigo e do modelo de recomendação híbrido teórico proposto, serão feitas as considerações finais.

6 CONSIDERAÇÕES FINAIS

O artigo propôs um modelo de recomendação híbrido adaptado para bibliotecas. Para este fim, foram apresentados conceitos relacionados a SRs clássicos e híbridos, bem como uma pesquisa sobre SRs utilizados em bibliotecas atualmente.

Percebe-se que as técnicas de filtragem mais populares permanecem sendo a filtragem por conteúdo e a filtragem colaborativa, tanto pela sua eficiência quanto pela sua eficácia, e que as iniciativas de SRs aplicadas a bibliotecas dão preferência a estas duas técnicas de filtragem. Também pode-se



perceber que embora seja amplamente aceito que os SRs híbridos apresentam melhorias nas recomendações finais, as bibliotecas ainda mostram preferência pela adoção de SRs do modelo clássico, como visto nas iniciativas de SRs em uso atualmente.

Entende-se que os sistemas de recomendação podem contribuir positivamente para o letramento informacional dos usuários, e que as bibliotecas têm apresentado interesse na sua adoção. Porém, é necessário tempo e uma maior aproximação das bibliotecas com este assunto para que a ideia se popularize em maior escala e para que modelos mais sofisticados, como os híbridos, sejam mais utilizados para gerar recomendações ainda mais refinadas.

Com base nos resultados obtidos, sugere-se que as próximas pesquisas sobre o assunto busquem compreender o motivo das bibliotecas terem preferência pelo uso de SRs do modelo clássico em detrimento do modelo híbrido. Também sugere-se um maior enfoque no desenvolvimento e aplicação de SRs híbridos, pois eles são uma forte tendência no campo das recomendações e sua aplicação é mais vantajosa quanto ao nível de refinamento. Por fim, também entende-se que é necessário que sejam realizados estudos sobre outros diferentes métodos de filtragem, como a filtragem por conhecimento e a filtragem demográfica, que não foram abordados neste artigo, mas possuem o potencial de gerar novos tipos de estratégias de recomendação em bibliotecas.

Abstract: Recommender systems are software tools that have as a goal the generation of an estimate of usefulness of items for a certain user, as well as analyzing interactions between users and items. They are divided between classic recommender systems, which use only one type of filtering technique, and hybrid, which use two or more types of filtering techniques allied to a hybrid recommendation technique. The goal of this paper is to introduce the recommender systems used in libraries and propose a theoretical hybrid recommender system model, adapted to the context and reality of libraries. For this purpose, a qualitative, exploratory, descriptive and bibliographic approach was used to explore the theme of recommender system and their use in libraries as an answer to users' growing



informational demand. A conclusion was made that most of the recommender system initiatives that are in use in libraries today are based on the classic recommender model and use only two types of filtering techniques – by content and collaborative. Therefore it is suggested that future research needs to try to comprehend the reason why the classic recommender model is used more, in detriment to hybrid recommender system, and that a bigger focus on the development and application of hybrid recommender system for libraries is needed, given its advantages when compared to the classic model. It is also suggested that further studies on other types of filtering techniques be made, given their potential to generate new recommendation strategies in libraries.

Keywords: Recommender systems; hybrid recommender systems; content filtering; collaborative filtering; information retrieval.

REFERÊNCIAS

ABDOLLAHPOURI, H.; BURKE, R. Multistakeholder recommender systems. *In*: RICCI, F.; ROKACH, L.; SHAPIRA, B. (ed.). **Recommender systems handbook**. 3. ed. New York: Springer Natura, 2022. p. 647-677.

AMATO, G.; STRACCIA, U. User profile modeling and applications to digital libraries. *In*: ABITEBOUL, S.; VERCOUSTRE, A. M. (ed.). Research and Advanced technology for digital libraries. **Lecture Notes in Computer Science**, v. 1696, p. 184-197, 1999. Disponível em: https://link.springer.com/chapter/10.1007/3-540-48155-9_13. Acesso em: 5 out. 2022.

BERNARDINO, D. 5 exemplos de escalas Likert para sua próxima pesquisa. s. d. Disponível em: <https://www.questionpro.com/blog/pt-br/exemplos-de-escalas-likert/>. Acesso em: 20 out. 2022.

BHANDARI, P. **What is qualitative research?** Methods & examples. 2020. Disponível em: <https://www.scribbr.com/methodology/qualitative-research/>. Acesso em: 5 out. 2022.

BURKE, R. Hybrid web recommender systems. *In*: BRUSILOVSKY, P.; KOBASA, A.; NEJDAL, W. (ed.). The adaptive web: methods and strategies of web personalization. **Lecture Notes in Computer Science**, v. 432, p. 377-408, 2007. Disponível em:



https://link.springer.com/chapter/10.1007/978-3-540-72079-9_12. Acesso em: 5 out. 2022.

bX ARTICLE Recommender: Enrich your discovery experience. **ExLibris**. 2022. Disponível em: <https://exlibrisgroup.com/products/bx-recommender/>. Acesso em: 8 out. 2022.

CHEN, A.; MCLEOD, D. Collaborative filtering for information recommendation systems. In: KHOSROW-POUR, M. (ed.). **Encyclopedia of E- Commerce, E-Government, and Mobile Commerce**. [S. l.]: IGI Global, 2006. p. 118-123.

DESHPANDE, M.; KARYPIS, G. Item-based top-N recommendation algorithms. **ACM Transactions on Information Systems**, v. 22, n. 1, p. 143-177, jan. 2004.

DIDEROT, D. **Encyclopédie**. Paris: Briasson, 1751. p. 314.

ELSEVIER. **ScienceDirect**: Overview & Factsheet. Amsterdam: Elsevier, 2021. Disponível em: https://www.elsevier.com/__data/assets/pdf_file/0004/1134931/factsheet-sciencedirect.pdf. Acesso em: 8 out. 2022.

EX LIBRIS bX Article Recommender for I-Share Institutions. **CARLI**. 2021. Disponível em: <https://www.carli.illinois.edu/products-services/i-share/discovery-interface/bxrecommender>. Acesso em: 5 out. 2022.

FACHIN, G. R. B. Recuperação inteligente da informação e ontologias: um levantamento na área da Ciência da Informação. **BIBLOS**, v. 23, n. 1, p. 259-283, 2009.

HARNEGIE, M. P. SciVerse science direct. **Journal of the Medical Library Association**: JMLA, v. 101, n. 2, p. 165, 2013.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). **IFLA statement on libraries and artificial intelligence**. 2020. Disponível em:

http://repository.ifla.org/bitstream/123456789/1646/1/ifla_statement_on_libraries_and_artificial_intelligence-full-text.pdf. Acesso em: 8 out. 2022.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). **IFLA statement on libraries and intellectual freedom**. 1999. Disponível em:

<http://repository.ifla.org/bitstream/123456789/1424/1/ifla-statement-on-libraries-and-intellectual-freedom-en.pdf>. Acesso em: 8. Out. 2022.

ISINKAYE, F. O.; FOLAJIMI, Y. O.; OJOKOH, B. A. Recommendation systems: Principles, methods and evaluation. **Egyptian Informatics Journal**, v. 16, n. 3, p. 261-273, 2015.



KONSTAN, J. A. *et al.* Grouplens: applying collaborative filtering to usenet news. **Communications of the ACM**, v. 40, n. 3, p. 77-87, 1997. Disponível em: <https://dl.acm.org/doi/pdf/10.1145/245108.245126>. Acesso em: 15 out. 2022.

KREBS, L. M. **Sistema de recomendação para bibliotecas universitárias**. Orientador: Rafael Port da Rocha. 2013. 95 f. Monografia (Bacharelado em Biblioteconomia) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/78367/000899325.pdf?sequence=1&isAllowed=y>. Acesso em: 15 out. 2022.

KUMARI, S. S.; RAJA, B. W. D.; SUNDARAVALLI, S. R. Analysis paralysis – the product of information explosion. **Annals of the Romanian Society for Cell Biology**, p. 4456–4458, 2021. Disponível em: <https://www.annalsofrscb.ro/index.php/journal/article/view/2993/2495>. Acesso em: 15 out. 2022.

L'HUILLIER, R. **ScienceDirect pilot aims to improve research discovery and access**. 2022. Disponível em: <https://www.elsevier.com/connect/science-direct-third-party-content>. Acesso em: 15 out. 2022.

LOPS, P.; GEMMIS, M.; SEMERARO, G. Content-based recommender systems: state of the art and trends. *In*: RICCI, F.; ROKACH, L.; SHAPIRA, B. (ed.) **Recommender systems handbook**. New York: Springer Nature, 2011. p. 73-105.

MITCHELL, T. M. **Machine learning**. Portland.: McGraw-Hill Science/Engineering/Math, 1997.

MUSTO, C. *et al.* Semantics and content-based recommendations. *In*: RICCI, F.; ROKACH, L.; SHAPIRA, B. (ed.) **Recommender systems handbook**. New York: Springer Nature, 2022. p. 251-301.

NIKOLAKOPOULOS, A. N. *et al.* Trust your neighbors: a comprehensive survey of neighborhood-based methods for recommender systems. *In*: RICCI, F.; ROKACH, L.; SHAPIRA, B. (ed.) **Recommender systems handbook**. 3. ed. New York: Springer Nature, 2022. p. 39-90.

PHILLIPS, M. **Related books links in Aleph OPAC**. 2010. Disponível em: <https://developers.exlibrisgroup.com/blog/related-books-links-in-aleph-opac/>. Acesso em: 20 out. 2022.

RICCI, F.; ROKACH, L.; SHAPIRA, B. (ed.). **Recommender systems handbook**. 3. ed. New York: Springer Nature, 2022.

SCIENCEDIRECT/Elsevier. **ITC Library**. 2022. Disponível em: <https://www.itc.nl/library/guides/information-literacy-guide/databases/sciencedirect-elsevier/>. Acesso em: 5 out. 2022.



BiblioCanto



66

WOLFF, A. *et al.* Creating an understanding of data literacy for a data-driven society.
The Journal of Community Informatics, v. 12, n. 3, p. 9-26, 2016.