



Texto licenciado sob a forma de uma licença *Creative Commons* Atribuição 4.0 Internacional



ID 898

Algoritmos de recomendação: um estudo sobre como funciona e como é descrito o sistema de recomendações da Netflix

Algoritmos de recomendación: un estudio sobre cómo funciona y cómo se describe el sistema de recomendaciones de Netflix

Recommendation Algorithms: a study on how Netflix's recommender system works, and it is described

Tiago Franklin Rodrigues Lucena

Doutor em Artes pela Universidade de Brasília (UnB). Professor Adjunto do Curso de Comunicação e Mídias na Universidade Estadual de Maringá (UEM) e do Mestrado em Comunicação na Universidade Estadual de Londrina (UEL). Maringá (PR). Brasil.

E-mail: tfllucena2@uem.br

ORCID: <https://orcid.org/0000-0002-0154-7417>

Eduarda Carretero Garcia

Bacharel do Curso de Comunicação e Mídias da Universidade Estadual de Maringá (UEM). Maringá (PR). Brasil. Foi bolsista PIBIC/CNPq.

E-mail: ra108552@uem.br

ORCID: <https://orcid.org/0009-0005-0592-3685>

Mariana Maronezzi Brezovsky

Graduanda do Curso de Matemática da Universidade Estadual de Maringá. Foi aluna de Iniciação Científica PIC/UEM. Maringá (PR). Brasil.

E-mail: marimaronezzi@gmail.com

ORCID: <https://orcid.org/0009-0008-5977-4234>

Thiago Fanelli Ferraiol

Doutor em Matemática pela Universidade Estadual de Campinas (UNICAMP). Professor Adjunto da Universidade Estadual de Maringá (UEM). Maringá (PR). Brasil.

E-mail: tferraiol@uem.br

ORCID: <https://orcid.org/0000-0001-6573-8311>

Resumo: As produções audiovisuais visualizadas por cada usuário na plataforma de *streaming* Netflix são baseadas, em parte, nos dados coletados, tratados e arquivados sobre como e o que foi consumido anteriormente por ele e por outros usuários. As sugestões de novos conteúdos são efetuadas por sistemas de recomendação e são operacionalizadas por um conjunto de algoritmos, que por muitas vezes são mantidos em segredo comercial. A Netflix, em seu site, propõe uma “uma descrição de alto nível” sobre o sistema de recomendação “em uma linguagem para leigos”. Este artigo analisa como esse texto explicita o funcionamento dessas ferramentas, articulando-o com autores que já fizeram parte do grupo de programadores da plataforma, outros críticos, e especialistas em algoritmos de recomendação. A análise demonstrou que, a partir da coleta de poucos dados do usuário, especialmente se comprado com o volume geralmente extraído de sites de redes sociais, é possível efetivar seu elaborado sistema de recomendação de forma personalizada. Os dados coletados se comportam como um “padrão de inclusão” e se constituem em matéria prima de um banco de dados que alimenta o sistema, criando um complexo perfil personalizado para cada indivíduo. Esse perfil é o que recomenda novos títulos no sistema de busca e orienta, principalmente, a posição do item nas fileiras na interface inicial. Por fim, a posição do título na interface e a fileira da qual faz parte influenciam significativamente na escolha da produção, o que tem consequências no contato com a diversidade de produtos audiovisuais, na manutenção da assinatura, e na experiência de consumo na plataforma.

Palavras-chave: Sistemas de Recomendação; Dados; Plataforma de streaming; Algoritmos.

Resumen: Las producciones audiovisuales visualizadas por cada usuario en la plataforma de *streaming* Netflix se basan, en parte, en los datos recopilados, procesados y archivados sobre lo que fue consumido previamente por él y otros usuarios. Las sugerencias de nuevos contenidos son realizadas por sistemas de recomendación y son operacionalizadas por un conjunto de algoritmos, que a menudo se mantienen en secreto comercial. Netflix, en su sitio web, propone una “descripción de alto nivel” sobre el sistema de recomendación “en un lenguaje para personas no expertas”. Este artículo analiza cómo ese texto explicita el funcionamiento de esas herramientas, relacionándolo con autores que han formado parte del grupo de programadores de la plataforma, otros críticos y expertos en algoritmos de recomendación. El análisis demostró que, a partir de la recopilación de pocos datos del usuario, especialmente en comparación con el volumen generalmente extraído de sitios de redes sociales, es posible efectuar su elaborado sistema de recomendación de forma personalizada. Los datos recopilados actúan como un “patrón de inclusión” y constituyen la materia prima de una base de datos que alimenta el sistema, creando un perfil personalizado complejo para cada individuo. Este perfil es el que recomienda nuevos títulos en el sistema de búsqueda y orienta, principalmente, la posición del elemento en las filas en la interfaz inicial. Finalmente, la posición del título en la interfaz y la fila a la que pertenece influyen significativamente en la elección de la producción, lo que tiene consecuencias en el contacto con la diversidad de productos audiovisuales, en el mantenimiento de la suscripción y en la experiencia de consumo en la plataforma.

Palabras clave: Sistemas de Recomendación; Datos; Plataforma de *streaming*; Algoritmos.

Abstract: The audiovisual productions watched by each user on Netflix – a streaming platform - are based, in part, on data collection, computing, and archiving about how and what was previously consumed by each user and by others. Suggestions for new content are made by recommendation systems and operated by a set of algorithms, which are often kept in commercial secret. Netflix, on its website, proposes a “high-level description” of its recommendation system “in layman’s language.” This article analyzes how this text explains the functioning of these tools, articulating it with authors who were already part of the group of platform’s developers, other critics, and specialists in algorithms. The analysis demonstrated that from the collection of little user data, especially if compared with the quantity of data usually extracted from social networking sites, it is possible to implement the recommendation system in an elaborated and customized way. The collected data behave as an inclusion pattern and constitute the raw material of a database that feeds the system, creating a complex personalized profile for each user. This profile is what recommends new titles in the search system and mainly guides the item’s position in the ranks in the initial interface. Finally, the position of the title in the interface and the row in which it is displayed significantly influence the choice of production. This, in turn, has consequences in the user’s contact with the diversity of audiovisual productions, the maintenance of the subscription, and the consumption experience on the platform.

Keywords: Recommender Systems; Data; Streaming Platform; Algorithms.

Introdução

O padrão de consumo de filmes e séries ganhou, nos últimos anos, novas possibilidades com a disseminação de serviços de *streaming*. Essas plataformas permitem o acesso a produções audiovisuais por meio de conexão à internet e várias empresas oferecem essa possibilidade: Amazon Prime, Globoplay, HBO Max, Apple TV, Paramount+, Disney+ e Mubi. Dentre elas, a Netflix tornou-se uma das mais conhecidas.

A plataforma de *streaming* de produções audiovisuais Netflix, fundada em 1997 por Reed Hastings e Marc Randolph, foi lançada inicialmente como locadora de vídeo pela internet. Em 1999, passou a promover locação ilimitada de longas metragens por meio de uma assinatura periódica, assim o assinante recebia as produções que solicitava via sistema postal (NETFLIX, 2022). O serviço passou a ser entregue de forma online em 2007 e a empresa se consolidou como uma das líderes de serviço digital, ultrapassando em 2021 a marca de 200 milhões de assinantes no mundo (*Ibid.*). A plataforma também passou a incluir séries e documentários de produção própria (JENNER, 2016), e diversos pesquisadores se interessaram por compreender as características audiovisuais e linguagem dessas produções no ambiente *streaming* (BRAGHINI; MONTAÑO LA CRUZ, 2019; CARLAN DA SILVEIRA; IUVA, 2020). Nos anos de pandemia, observou-se um aumento de assinaturas, além do tempo consumindo audiovisual (LOZIC, 2021). O crescimento permitiu que a Netflix aprimorasse ainda mais os algoritmos envolvidos no sistema de recomendação (GAW, 2022).

Quando um usuário acessa sua conta na Netflix, consegue visualizar cartazes de alguns filmes e séries na sua interface inicial. Essas produções que estão em destaque variam para cada usuário, uma vez que as indicações são baseadas, em parte, nos dados coletados, tratados e arquivados sobre como e o que foi consumido anteriormente. Também é levado em consideração dados de outros usuários com perfil semelhante. Essas sugestões são efetuadas por sistemas de recomendação e são operacionalizadas por um conjunto de algoritmos que garantem boa parte do sucesso comercial da empresa. Eles buscam, em última instância, a eficiência na oferta de conteúdo para o usuário para que este mantenha a assinatura do serviço (GOMEZ-URIBE; HUNT, 2016).

Para Goffey (2008), os algoritmos são, basicamente, a descrição de um método pelo qual uma tarefa deve ser resolvida. Segundo ele, os “algoritmos fazem coisas, e sua sintaxe incorpora uma estrutura de comando para permitir que isso aconteça” (GOFFEY, 2008, p.17). De acordo com Gillespie, os algoritmos “[...] não são



necessariamente *softwares*: em seu sentido mais amplo, são procedimentos codificados que, com base em cálculos específicos, transformam dados em resultados desejados” (2018, p.97). Ou seja, para este autor, os algoritmos são instruções e passos de como proceder em relação aos dados recebidos.

A plataforma de *streaming* não está sozinha no uso dos algoritmos, mas compartilha estratégias que também encontramos em serviços online de empresas como Alphabet (Google), Meta, Amazon e Apple. As consequências dos usos desses algoritmos no consumo cultural e social são ainda incertas, pois eles são “opacos”, de difícil acesso e de grande complexidade (O’NEIL, 2016). Cabe destacar que as empresas utilizam de códigos ocultos para manutenção da assimetria do poder, em defesa dos segredos de negócio, e para evitar que os usuários driblem os algoritmos caso sejam de código aberto (PASQUALE, 2016). Se por um lado as políticas públicas pedem transparência para que usuários saibam exatamente como e o que é coletado (como orienta a LGPD - Lei Geral de Proteção de Dados Pessoais - nº 13.709, de 14 de agosto de 2018), por outro se reconhece que essas são estratégias que se configuram como verdadeiras “caixas-pretas” do mercado, incorporadas dentro de uma lógica do capitalismo de vigilância (ZUBOFF, 2020).

As pesquisas que analisaram o sistema de recomendação da Netflix – *Netflix recommender system* (NRS) – assim o fizeram de perspectivas diferentes. Niko Pajkovic (2022) identificou como o NRS funciona e como ele impacta na orientação de um gosto. Para fazer isso, o autor se valeu da “engenharia reversa” com um experimento que simulou a experiência na página da Netflix com três tipos de usuários fictícios criados com perfis diferentes. Compreendendo que o algoritmo teve um papel essencial na experiência audiovisual, o autor abraçou uma perspectiva mais “materialista relacional” dos algoritmos para entendê-los não como objetos técnicos estáticos, mas como parte de uma cultura, respondendo e dando forma a ela.

Já Burroughs (2019) descreveu o NRS como um algoritmo colocado pela empresa como um “colaborador” para ajudar a “decidir em um momento de incerteza”. Para os assinantes, é dito que os algoritmos “garantem” que eles vão continuar gostando da programação da Netflix, e como uma solução para as complicações que vêm da fragmentação da indústria da mídia com centenas de produções.

Gaw (2022) combina a análise de documentos da Netflix com matérias na mídia e tweets de usuários sobre a plataforma para identificar a lógica por traz do conceito de gosto presente no NRS, bem como as suas contradições. Considerando que tudo na página advém da recomendação, a autora questiona as noções de datificação e da ideia de se traduzir gosto como sendo algo quantificável e individualizável. Cohn (2019), por sua vez, analisa criticamente diversos sistemas de recomendação, incluindo o NRS, e



argumenta que esses sistemas estão longe de ser neutros, naturais ou benevolentes, mas que eles dão forma e são formados pelas concepções de gênero, sexualidade, raça e classe de quem os cria. Aqui lemos mais uma vez que o NRS é posicionado pela empresa como uma tecnologia “necessária” para “guiar os assinantes da Netflix” em suas dificuldades de fazer escolha quando são confrontados com um número avassalador de possibilidades (COHN, 2019).

Por fim, Hallinan e Striplas (2016) analisam o prêmio Netflix (*Netflix prize*) criado pela empresa para premiar o desenvolvedor que aprimorar o NRS. No trabalho, destacam que por trás desse desafio técnico para engenheiros se escondeu uma oportunidade para se identificar as noções de cultura audiovisual da empresa. Desconhece-se, no entanto – e até o presente momento –, artigos que analisam diretamente o texto “Como funciona o sistema de recomendações da Netflix” (NETFLIX, 2019), disponível para os usuários, e como a partir dele podemos inferir sobre a cultura algorítmica por trás do NRS.

Metodologia

Levando em consideração essa disputa entre interesse comercial, proteção, e esclarecimento do usuário sobre o tratamento dos dados, este artigo, qualitativo e exploratório (GIL, 2008), articula o pensamento de Gillespie (2014), sua tradução (2018), Gomez-Uribe e Hunt (2016) e Aggarwal (2016) com o texto “Como funciona o sistema de recomendações da Netflix” (NETFLIX, 2019), presente no site da empresa. Trata-se de um documento que é disponibilizado para acesso e leitura em português no site da empresa, na aba “centro de ajuda” e foi acessado em 11 de fevereiro de 2022. Ele é dividido em quatro partes: “O básico” (1470 caracteres), “‘Forçando’ o sistema de recomendações” (671 caracteres), “Fileiras, classificações e representação de títulos” (945 caracteres) e “Como melhorar o nosso sistema de recomendações” (383 caracteres). Segundo a página, o texto oferece “uma descrição de alto nível”, ou seja, é voltado para usuário final com informações sobre o sistema de recomendação “em uma linguagem para leigos”.

Após leitura na íntegra do texto por duas autoras, foram identificados e destacados trechos no documento que apresentaram descrições nítidas sobre o possível funcionamento dos algoritmos incorporados no sistema de recomendação da plataforma. Os trechos foram recortados, confrontados e articulados com os pensamentos sobre o algoritmo de autores já mencionados. Cabe destacar que os autores Gomez-Uribe e Hunt (2016) estavam vinculados à empresa Netflix e forneceram

informações relevantes sobre como foi composto o sistema de recomendação da plataforma, além de descreverem seu propósito comercial. A proposta deste artigo, então, foi a de se utilizar do texto presente no site da plataforma para analisar os recursos que a Netflix utiliza para recomendar novos conteúdos.

Os critérios para a escolha dos três autores que foram articulados na análise do texto se deram pela diversidade representada. Como uma das vozes temos os autores Gomez-Uribe e Hunt (2016), que publicaram artigo com uma visão interna da Netflix. Trata-se de profissionais que atuaram diretamente na construção dos algoritmos da empresa. Gillespie (2014), e o seu texto em versão traduzida em 2018, foi selecionado por representar um dos principais autores a relacionar o algoritmo à cultura (BEER, 2013; GAW, 2022; MORRIS, 2015; PAJKOVIC, 2022; SEAVER, 2017), fazendo reflexões sobre as consequências e características dessas tecnologias na vida cotidiana. Por fim, a seleção de Aggarwal (2016) se deu por representar um dos principais autores sobre algoritmos e sistemas de recomendação por uma perspectiva técnico-matemática. Todos os autores são constantemente citados em outros artigos da área, conforme métricas do Google Scholar: 1) Gomez-Uribe e Hunt, citados por mais de 1.700 autores; 2) Gillespie, citado por mais de 2.700 outros artigos; e 3) Aggarwal, citado por mais de 1.600 autores.

Discussão

Com o objetivo de analisar como o texto “Como funciona o sistema de recomendações da Netflix” articula e dialoga com os pensamentos dos autores Gillespie (2014), Aggarwal (2016) e Gomez-Uribe e Hunt (2016), este artigo buscou esclarecer como os algoritmos são utilizados pela plataforma para destinar ao usuário um conteúdo personalizado e com maior potencial de ser consumido. O discurso é de se colocar como uma ferramenta que colabora na difícil tarefa de escolher por si só um filme/série diante de uma infinidade de produções do catálogo (COHN, 2019). No texto na página da Netflix, lemos: “Sempre que você acessa o serviço Netflix, nosso sistema de recomendações tenta ajudar você a encontrar uma série ou filme de forma fácil” (NETFLIX, 2019).

Conforme destaca David Sumpter (2019), acessar exatamente o que faz um algoritmo é algo difícil, uma vez que se constituem em segredos comerciais das empresas. Muitas delas sugerem que esses segredos não devam “ser divulgados em espaços públicos” (GILLESPIE, 2018, p.112). Os profissionais de redes sociais, de SEO – *Search Engine Optimization* –, e *spammers* também têm acesso restringido aos

algoritmos de funcionamento das plataformas, mas já se desenvolveu habilidades técnicas “para deduzir os critérios do algoritmo a partir de testes e engenharia reversa” (GILLESPIE, 2018, p.111). Sumpter (2019) usa o pensamento lógico-matemático para inferir sobre o funcionamento de diversos algoritmos de plataformas comerciais, pois lembra que existem princípios básicos por trás deles, e assim pode-se construir os modelos matemáticos e seu modo de funcionamento. Esse foi o caminho adotado por nós.

Analisando, então, o texto disponível no site da plataforma, reconhece-se que ele se difere dos documentos do tipo “Termos de Uso e Políticas de Privacidade”. Não se trata de um texto com esse teor, pois os documentos chamados de “Termos e Condições de Uso”, geralmente, estabelecem regras e diretrizes que os usuários devem concordar e seguir para usar o serviço disponibilizado. Esse aspecto de contrato está ausente na redação do texto publicado no site da Netflix. Já textos do tipo “Política de Privacidade de Dados” se referem aos direitos e deveres relacionados aos dados dos usuários, especialmente sobre como os dados são coletados pelo sistema em uso são utilizados (BRASIL, 2018). Sobre esses documentos, cabe lembrar que diversos artigos já se empenharam em analisá-los (TALIB *et al.*, 2014; WILSON *et al.*, 2016; WINKLER; ZEADALLY, 2016), mas muitas dessas análises recaem em observações quanto aos aspectos jurídicos da redação e desrespeito as políticas de privacidade. Além disso, ainda que sejam documentos que atendam às legislações vigentes sobre a relação entre consumidor e o provedor de serviço contratado, viu-se que eles não são claros, são ignorados pelos usuários (TALIB *et al.*, 2014; WILSON *et al.*, 2016) ou, mesmo quando lidos, apresentam problemas quanto à sua transparência (SIMÕES; CUNHA, 2018). De forma resumida, as políticas de privacidade da Netflix, que não são o objeto de estudo deste artigo, focalizam-se na relação contratual entre usuário e plataforma. Ritter (2021) propõe uma avaliação das políticas de privacidade da empresa sob a ótica da LGPD e parte da sua conclusão sobre a eficiência do NRS em recomendar conteúdo com os dados extraídos do assinante se deu por também acessar o texto objeto de nossa análise nesse artigo – sem, no entanto, se aprofundar em seus aspectos.

Embora o texto no site da Netflix não qualifique o que eles compreendem como “de forma fácil”, percebemos uma tentativa de explicar como funciona o sistema de recomendação. Simbolicamente, trata-se de uma forma de a empresa se posicionar como “colaboradora” do assinante e explicar como o sistema funciona para “ajudar” ou “guiar” o usuário na difícil tarefa de escolher algo do catálogo (COHN, 2019). O aceite ou a recusa de uma recomendação de filme ou série na interface do usuário retorna como um dado ao NRS, o que por sua vez é utilizado para aprimorar ainda mais o próprio NRS.



Sistemas de recomendação são parte mediadora de diversas mídias contemporâneas e são filtros de informação que visam, a partir de um conjunto de dados, inferir os interesses dos usuários. O princípio de funcionamento das recomendações está na forte relação entre as preferências dos usuários e os itens sugeridos (AGGARWAL, 2016). Os primeiros sistemas de recomendação foram criados no início dos anos 90, como uma resposta ao crescimento do volume de conteúdo disponível na internet. Segundo Aggarwal (2016), um dos primeiros artigos publicados sobre o assunto analisava os problemas decorrentes do uso crescente de comunicações eletrônicas, que terminavam por sobrecarregar o usuário com documentos desnecessários. Já em um dos sistemas de recomendação pioneiros, chamando de *Tapestry*, os usuários podiam listar critérios de filtragem de conteúdo, como palavras-chave e outros usuários com os quais compartilhavam interesses. Nesse sistema, por exemplo, um usuário poderia solicitar algo como “todos os artigos com a palavra *fintech* que contêm comentários do Ricardo Silva” (GOLDBERG *et al.*, 1992).

O texto da Netflix fornece algumas explicações sobre o direcionamento de conteúdo realizado pela empresa, bem como quais dados dos usuários são coletados e tratados. Cabe lembrar que os algoritmos de recomendação da Netflix são, na verdade, um conjunto de algoritmos responsáveis pelo direcionamento do conteúdo presente no catálogo da plataforma para o usuário. O artigo de Beer (2013) apresenta um pouco a complexidade do NRS ao mencionar o prêmio Netflix, que levou um grupo de desenvolvedores a melhorar em 8.34% a acurácia do sistema. O grupo levou 2.000 horas para combinar os 107 algoritmos responsáveis por articular as recomendações aos usuários. Assim, Beer (2013) destaca que não só o NRS é central para o sucesso comercial da Netflix, como também que o gosto fílmico (a taxa de acerto na recomendação) é mensurado-metrificado pela empresa. Obviamente, a noção e os conceitos do que é “gosto” podem ser questionados (GAW, 2022).

Gomez-Uribe e Hunt (2016), criadores iniciais da NRS, não escondem o interesse comercial por trás do uso de tal sistema, uma vez que lembram que a renda da empresa é vinculada a assinaturas e que, quanto mais tempo o usuário passa consumindo o conteúdo, maior é a probabilidade de que continue assinando o serviço de *streaming*. Trata-se da confirmação de um tipo de sistema onde a coleta de dados dos assinantes garante maior fidelidade dos usuários, num capitalismo de plataforma (SRNICEK, 2016) e de vigilância (ZUBOFF, 2020). A lógica do “extrativismo” dos dados se confirma como essencial para a lógica desse sistema comercial (MOROZOV, 2018).

Para melhor encaminhar a discussão, foram identificadas três grandes categorias de análise que emergiram da leitura do texto no site da Netflix. São elas: 1) identificação e explicação sobre como são coletados os dados do usuário; 2) explicação



de como são classificados os títulos dos filmes e séries; e 3) como são endereçadas as recomendações para os usuários. Embora esses elementos estejam interligados (identificação e coleta, classificação, e tratamento dos dados), para fins de análise, vamos considerar cada elemento separadamente. Entendemos, à luz da definição do termo “tratamento” presente na LGPD, que essas três instâncias se constituem em processos de tratamento dos dados. Ou seja, o artigo 5 em seu parágrafo X define tratamento como sendo toda operação realizada com dados pessoais, como as que se referem a “coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração” (BRASIL, 2018).

Da relação entre coleta e inclusão dos dados

Dentre as diversas características levantadas no texto do site da Netflix, identificamos alguns trechos que apontam para como a coleta de dados do usuário é operacionalizada. Segundo o texto, faz-se um cálculo de probabilidade de um usuário assistir a um título do catálogo. Esse cálculo tem como base um número de fatores, que são: “suas interações com o nosso serviço (como o que você assistiu e que nota deu a outros títulos), outros assinantes com gostos similares e preferências sobre nossos serviços e informações sobre os títulos, como gênero, categorias, atores, ano de lançamento, etc.” (NETFLIX, 2019). Além disso, a Netflix ainda acrescenta que além de “saber ao que você assistiu” também observa “o horário em que você assiste; os aparelhos nos quais você assiste à Netflix; e por quanto tempo você assiste” (*Ibid.*). Nessa descrição, percebe-se que são coletados os seguintes dados do comportamento do usuário na plataforma:

- a. tempo de interação do usuário com o conteúdo;
- b. avaliação do conteúdo consumido por parte do usuário (nota emitida para uma produção audiovisual);
- c. tempo de interação do usuário com cada tipo de conteúdo (filme, série, gênero, atores, ano de lançamento);
- d. horário de consumo (manhã, tarde ou noite);
- e. tipo de aparelho utilizado para consumo dos conteúdos (TV, smartphone, PC).



Também inferimos, porque não foi explicitamente mencionado, que os termos de busca dos usuários quando procuram títulos de filmes e séries também sejam coletados, uma vez que, conforme lemos no site Netflix:

Tentamos ter um sistema de busca fácil e rápido de usar. Quando você inserir algo na busca, os primeiros resultados que retornaremos terão como base as ações de outros assinantes que tenham inserido os mesmos termos ou termos similares (NETFLIX, 2019).

A seleção de dados que a Netflix coleta para direcionar conteúdo é fundamental para o funcionamento do sistema. Trata-se de uma estratégia de alimentação de um banco de dados atrelado ao perfil do usuário. Sobre isso, Gillespie (2014) destaca que mais importante do que apenas programar o funcionamento lógico do algoritmo é pensar com qual banco de dados ele deve dialogar. Para Aggarwal (2016), esses sistemas precisam lidar com grandes bancos de dados, a fim de extrair apenas o que é admitido como relevante a partir das preferências e interesses dos usuários. Dessa forma, uma atividade fundamental para que uma recomendação seja realizada é a coleta e a sistematização dos dados. Nesses sistemas, há uma seleção de quais dados são mais relevantes de serem incluídos para que a personalização possa acontecer com maior efetividade. Essa lógica por trás da seleção dos dados relevantes é chamada por Gillespie de *Padrões de Inclusão*. Isso não descarta, por exemplo, que outros dados sejam coletados pela Netflix. Mas, de acordo com a descrição no site da Netflix,, esses dados não são tratados no contexto do sistema de recomendação, conforme entendemos com a leitura do trecho: “o sistema de recomendações não inclui dados demográficos (como idade ou sexo) para tomar as decisões” (NETFLIX, 2019).

Esses padrões de inclusão não são exclusivos da Netflix. Os usuários, ao navegarem pela internet, deixam rastros que são coletados e armazenados por um conjunto de sistemas e empresas. Cada ação online é registrada e constitui uma base que alimenta um ou outro algoritmo. Quanto mais extensa a quantidade de registros, mais atividades são catalogadas e arquivadas (GILLESPIE, 2014). Do ponto de vista da quantidade de dados coletados de apenas um usuário (quando comparados àqueles que são coletados por empresas de redes sociais online, por exemplo), percebe-se que a Netflix não indica coletar dados sensíveis e pessoais do usuário para o funcionamento do sistema de recomendação. Num primeiro momento, tínhamos a expectativa de que a plataforma listaria dados coletados e correlacionados, tais como: informações de dados de localização do usuário (advindos do IP do dispositivo ou da *smart TV*), a etnia ou grupo social, e a faixa-etária. No entanto, esses dados do perfil do assinante não são declarados como coletados e tratados pelo sistema de recomendação, pois, conforme



trecho no site da Netflix, são informações vinculadas apenas ao contrato de serviço entre o usuário e a empresa.

Apesar disso, Ritter (2021) nos lembra que, embora a Netflix colete poucos dados cadastrais do seu assinante, a utilização de fontes agregadas de informação e a constante vigilância sistêmica das ações do usuário na plataforma potencializa o uso de técnicas de *profiling*. Essa técnica permite inferir sobre dados que não foram inicialmente coletados. Essa característica aumenta a assimetria de informação sobre de que forma e para que os dados pessoais são utilizados (RITTER, 2021). Assim, não é possível a um usuário da plataforma saber e ter controle total sobre como seus dados pessoais serão utilizados.

Para ilustrar essa dimensão do *profiling* sendo efetivada na Netflix, é possível citar matéria publicada no site Wired em 2018. Um debate acendeu, quando usuários negros começaram a perceber que nas recomendações de filmes/séries havia uma presença destacada de atores negros nos pôsteres dos filmes. Isso acontecia mesmo quando esses eram uma minoria dentro do elenco da produção assistida. Houve a suspeita de que havia, sim, a coleta de dados relacionados à etnia ou grupo social dos assinantes. Em sua defesa, a empresa esclareceu que esse elemento surgiu após a implementação de um algoritmo de *machine learning* em 2017, quando um sistema para pôsteres personalizados foi implementado. Especialistas destacaram que, quando se personaliza um conteúdo baseado no histórico de visualização, direcionar conteúdo baseado na raça, etnia e gênero seria um efeito emergente natural. Assim, reconheceu-se que mesmo não coletando os dados relacionados a essas características, os dados de histórico de visualização permitiram sua utilização para inferir sobre a raça/gênero dos usuários (TIKU, 2018).

Ao que parece, o sucesso do sistema de recomendação viria, também, do cruzamento desses dados com outros provenientes de milhares de usuários. A coleta de dados de muitos usuários é uma estratégia primordial nos negócios de mídia (FERNÁNDEZ-MANZANO *et al.*, 2016). Um elemento importante para que isso aconteça é a associação entre o consumo do usuário com informações sobre o tipo de conteúdo que ele assiste. É sobre essa classificação e etiquetagem do conteúdo que comentaremos a seguir.

Sistema de classificação das produções audiovisuais

Conforme lemos no site Netflix (2022), “informações sobre os títulos, como gênero, categorias, atores, ano de lançamento, etc.” são utilizadas e associadas para

que o sistema de recomendação seja efetivo. Cada item do catálogo está classificado dentro de um gênero (terror, suspense, drama), bem como no ano de lançamento (recente, filme antigo), quanto ao seu formato (série, filme, documentário), e com listagem dos atores presentes. Essas classificações são fundamentais para que os conteúdos sejam direcionados com a maior efetividade possível para o usuário. Nesse sentido, Gillespie (2018) comenta que a classificação realizada pelas plataformas para determinar quais conteúdos são mais relevantes a serem compartilhados dialoga com o que chama de *Avaliação de Relevância*. Segundo o autor, as empresas avaliam e classificam os dados coletados para direcionar conteúdos considerados mais adequados para um determinado usuário, dentre a gama de conteúdos e serviços oferecidos. Considerando um catálogo com milhares de títulos, a plataforma deve avaliar as informações e dispor numa interface os que devem ser exibidos.

Além disso, Gillespie (2018) lembra que esses dados podem ser incorporados dentro de uma lógica de *Ciclos de Antecipação*. Para o autor, após a coleta de múltiplos dados (suas ações nas plataformas, bem como dados de outros usuários), empresas como a Netflix possuem recursos para tentar antecipar os gostos, desejos e necessidades de seus assinantes. Essa antecipação entra na lógica da predição, tentar antecipar com certa assertividade um comportamento ou interesse do usuário e, possivelmente, vender essa informação a um mercado (ZUBOFF, 2020). A predição acontece quando, munido dos dados sobre o gênero ou formato dos filmes e séries que o usuário mais assistiu anteriormente, o sistema tenta antecipar os gostos do usuário e começa a colocar bem-posicionados outros filmes/séries com o mesmo gênero ou formato. Considerando que esse sistema oferece mais títulos com o mesmo gênero ou formato, é aqui que recai uma observação sobre a criação de “filtros-bolha”. Segundo David Sumpter (2019), existem duas principais expressões – que estão relacionadas, mas são ligeiramente diferentes – dadas quando os usuários se mantêm restritos a determinadas informações e conteúdos online. A primeira é chamada pelo autor de *Câmaras de Eco*. São criadas pelas próprias pessoas e consistem em selecionar conteúdos, sites, perfis para se manter conectadas online, ou seja, como sua fonte de informação, “criando seu próprio mundo dentro do qual suas opiniões reverberam” (*op. cit.*, p.149). A segunda são as filtros-bolha, sistemas determinados por algoritmos que mantêm uma pessoa cercada por uma bolha de informações restritas. Para Sumpter, esse fenômeno é observado na medida em que se percebe que a diversidade cultural pode estar sendo afetada pelo sistema que intensifica a sugestão de apenas “o mesmo” tipo de produção/conteúdo. Esse aspecto não é exclusivo da plataforma Netflix, mas já foi observado e analisado por autores como Pariser (2012) e Alexander (2016).

Por fim, o NRS funciona “recomendando” essa nova produção/conteúdo na interface principal. São os algoritmos que direcionam a presença de um título e sua localização nas chamadas fileiras. A posição desse título nessas fileiras pode garantir uma maior visibilidade e probabilidade do usuário acessá-lo. Sobre isso, Gomez-Uribe e Hunt (2016) reforçam que o aprimoramento desse serviço se deu com muitos testes. São esses testes replicados diversas vezes que permitem que os sistemas recomendem as produções audiovisuais para cada usuário. Será sobre esses testes e recomendações provenientes dos seus resultados que falaremos a seguir.

Como são feitas as recomendações para os usuários

O site da Netflix parece indicar duas formas como os títulos são recomendados na interface para o usuário: 1) os títulos que aparecem quando o usuário realiza a busca (pela barra de pesquisa); e 2) os títulos que estão na interface inicial com o usuário.

Sobre os títulos que aparecem quando o usuário digita na barra de pesquisa, lemos que os resultados apresentados levam em consideração “ações de outros assinantes que tenham inserido os mesmos termos ou termos similares” (NETFLIX, 2019). Sobre esse aspecto, Gomez-Uribe e Hunt afirmam que, quando um usuário faz uma pesquisa por um título, um sistema inteligente faz sugestões relacionadas mais à popularidade de outros vídeos do que por uma relação com o título pesquisado. Uma vez que nem todos os títulos pesquisados estão disponíveis no catálogo da Netflix, a recomendação é feita por algoritmos distintos. Como mesmo disse Ted Sarandos – codiretor executivo e diretor de conteúdo da Netflix –, os “algoritmos conduzem todo o nosso site [menu] – não há um centímetro de espaço editorial não calculado” (CURTIN; HOLT; SANSON, 2014, p.144). Um dos algoritmos sugere filmes a partir de consultas parciais (ao pesquisar “fren” é recomendado “*Frenemies*” – um filme de 2012). Outro, identifica um conceito a partir da consulta parcial (ao pesquisar “fren”, é recomendado “*French Movies*”). Para isso, são combinados tanto os dados de reprodução e de pesquisa, como metadados.

Esse recurso avançou após conhecimento acumulado sobre sistemas de busca, inclusive de páginas web. Gillespie (2018) lembra que os algoritmos de pesquisa eram baseados apenas em dizer qual a frequência com que determinados termos apareciam nas páginas indexadas online. Agora, os algoritmos incorporam mais informações contextuais sobre as páginas, incluindo onde estão hospedadas e “a frequência e como o site é relacionado por outros; e convocam técnicas de processamento de linguagem natural para melhor ‘entender’ tanto a consulta, quanto os recursos que o algoritmo pode

oferecer como resposta” (GILLESPIE, 2018, p.103). Para avaliar se os resultados apresentados foram os mais relevantes, as plataformas carecem de uma métrica independente e, assim, os engenheiros precisam decidir quais resultados parecem “corretos” e ajustar seu algoritmo para atingir esse resultado. Conforme continua o autor, fazer “alterações com base em evidências dos usuários, considerando cliques rápidos e pesquisas sem continuação como uma aproximação, não de relevância exatamente, mas de satisfação com o conteúdo oferecido” (*op. cit.*, p.104). São essas estratégias que são aplicadas a diferentes sistemas de recomendação.

Há, então, diversas categorias de sistemas de recomendação, como as de *filtragem baseada em conteúdo*, baseada em *conhecimento* e de *filtragem colaborativa* (AGGARWAL, 2016). As que nos interessam, visto o que foi sinalizado por Gomez-Uribe e Hunt (2016) sobre a Netflix, são as de filtragem colaborativa, que se referem à utilização da similaridade entre múltiplos usuários para computar a recomendação de itens ainda não avaliados (AGGARWAL, 2016). Várias contribuições no contexto de pesquisa em filtragem colaborativa foram impulsionadas pela competição aberta chamada de *Netflix Prize*, realizada em 2006, que convocou especialistas e interessados no mundo todo para aprimorar os sistemas de recomendações da Netflix mediante pagamento de um prêmio (MORRIS, 2015). Para Aggarwal (2016, p. 6), essa competição foi notável, pois entregou diversas lições úteis para a plataforma, como o uso de modelos baseados em fatores latentes. Esses modelos consistem em identificar um padrão nos dados e, a partir disso, reduzir a dimensionalidade da matriz de dados.

Porém, a maior efetividade do sistema de recomendação não se dá pelo recurso da busca do título, mas sim pela posição do título na interface inicial do usuário. Para isso, vemos que 80% dos títulos que são assistidos na plataforma Netflix são acessados pelas recomendações que a própria plataforma faz, por meio da indicação do filme/série na interface inicial, enquanto 20% são advindos do acesso pela barra de pesquisa (GOMEZ-URIBE; HUNT, 2016). Levando em consideração essa relevância da interface inicial para se acessar o conteúdo, lemos no texto no site: “Essa descrição indica que a posição do filme/série na interface da Netflix é importante para que o sistema de recomendação se efetive”.

Gomez-Uribe e Hunt (2016) sinalizam essa importância ao falar a respeito do *Personalized Video Ranker* (classificador/ranqueador de vídeo personalizado), ou PVR, na sua forma abreviada. Trata-se de um algoritmo usado para classificar e ordenar filmes e séries que serão apresentados na interface inicial do usuário (*homepage*). Essa interface apresenta 40 linhas com sugestões de conteúdos personalizados por temas, os quais vão de gêneros de filmes e vídeos mais assistidos da plataforma, a recomendações com base no que o usuário viu anteriormente (Porque você assistiu



<filme>). Aggarwal (2016, p. 43) aponta que, apesar de não haver acesso explícito aos métodos utilizados pela Netflix, essa recomendação possivelmente é realizada por métodos baseados em itens (*item-based*) de filtragem colaborativa, a partir de itens da vizinhança. Existe, segundo ele, dois tipos de métodos de algoritmos baseados em vizinhança, a filtragem colaborativa baseada em itens (*item-based*) e a baseada em usuários (*user-based*). Com relação à abordagem baseada em itens, é escolhido o conjunto daqueles mais similares aos do usuário. Por exemplo, as notas de um usuário A com relação aos filmes de ficção científica *Alien* e *Predador* “podem ser utilizadas para prever a nota do filme *Exterminador do Futuro*”.

Nesse sistema, cada fileira na interface possui 75 vídeos classificados e ordenados pelo PVR, o que determina tanto a ordem em que os filmes serão apresentados numa mesma fileira, quanto a ordem das fileiras na página inicial (GOMEZ-URIBE, HUNT, 2016). Ou seja, quais ficarão no topo com maior probabilidade de serem vistas e quais ficarão mais abaixo. Desse modo, o algoritmo PVR analisa a relevância da posição que cada título nessa interface deve tomar. Podemos relacionar esse aspecto retomando Gillespie (2018, p.106), quando nos fala sobre a *Avaliação de Relevância*. Cada algoritmo tem como premissa tanto um pressuposto sobre a análise adequada de relevância, quanto em uma representação desse pressuposto em uma técnica de avaliação (computacional). Implícitas nessas análises podem existir premissas incorporadas na ideia de relevância, bem como “existir atalhos incorporados em sua representação técnica dessa ideia, e podem existir fricções entre elas duas” (*Ibid.*).

A posição do título da interface é, então, um recurso essencial para que se aumente a probabilidade do usuário em clicar e acessar o título. Segundo o texto, “[e]m cada fileira, existem três camadas de personalização: a escolha da fileira (por exemplo, “Continuar assistindo”, “Em alta”, “Comédias premiadas” etc.), quais títulos aparecem na fileira e a classificação desses títulos” (NETFLIX, 2019). Vemos também que há uma organização da disposição dos títulos nessas fileiras, respeitando a forma de leitura ocidental:

As fileiras mais recomendadas aparecem na parte superior. Os títulos mais recomendados aparecem da esquerda para a direita de cada fileira, exceto se você selecionar os idiomas árabe ou hebraico. Nesses casos, os títulos recomendados aparecerão na ordem inversa (*Ibid.*).

Assim, a mesma fileira pode apresentar vídeos completamente diferentes para cada usuário. Há também um algoritmo específico para ordenar e sugerir os títulos mais assistidos, porém de forma personalizada: os conteúdos, dentre os mais populares, são



sugeridos de acordo com a maior probabilidade de serem vistos. Ambos os algoritmos citados combinam personalização com popularidade (GOMEZ-URIBE; HUNT, 2016).

Esse aspecto é atestado ao usuário no texto da plataforma:

Além de escolher quais títulos serão incluídos nas fileiras da sua página inicial da Netflix, nosso sistema também classifica cada título da fileira e as próprias fileiras usando algoritmos e sistemas complexos para oferecer uma experiência personalizada. Colocando de outra forma, quando você acessa a página inicial da Netflix, os nossos sistemas classificam os títulos de forma que eles sejam apresentados na melhor ordem possível para você desfrutar do serviço (NETFLIX, 2019).

Para descobrir quais algoritmos são mais eficientes e para quais situações seu uso é mais apropriado, a Netflix aplica testes de modo online e offline. Para fazer os testes A/B, os membros são submetidos a células distintas, em que cada célula representa uma combinação de filmes e organização diferentes. Deixa-se uma célula de controle para o experimento e células de teste, em média de 5 a 10 variações. A partir disso, após um período de 2 a 6 meses, é possível tirar conclusões estatísticas da influência de cada uma das células, como se os filmes recomendados foram assistidos, qual o tempo de visualização e o impacto na fidelização dos membros. Assim, os testes A/B são cruciais nas decisões a serem tomadas.

Segundo Gillespie (2014), os algoritmos são constantemente alterados e ajustados. Isso ocorre porque podem ser facilmente, instantaneamente e invisivelmente modificados, quando necessário, para executarem suas atividades com mais precisão. “A maioria dos usuários não se debruça sobre os critérios algorítmicos e tende a tratá-los como ferramentas não problemáticas a serviço de uma atividade maior: encontrar uma resposta, resolver um problema, entreter” (GILLESPIE, 2018, p.106). Ou seja, essas tecnologias são “encaixapretadas” pelos criadores (LATOURE, 2000; PINCH; BIJKER, 1984), o que nos leva a crer que são estáveis.

Existem outras métricas que também colaboram para a tomada de decisões, como o tempo de espera para acessar algum filme do catálogo, os dias assistidos, o número de filmes que não foram assistidos até o fim. Para auxiliar os testes online que medem o impacto direto do Sistema de Recomendações no usuário final (AGGARWAL, 2016, p. 228), como os testes A/B, os algoritmos são submetidos a testes offline, ou seja, sem ligação direta com o consumidor final. Nesse sentido, são utilizadas métricas estatísticas para compreender se as variações de um algoritmo foram efetivas para criar uma boa página inicial. Há uma abordagem baseada em regras, a qual organiza uma ordem genérica para posição de linhas principais, como “Continuar assistindo”, “Popular



na Netflix” e “Porque você assistiu <filme>” (ALVINO; BASILICO, 2015). Outra possibilidade é uma abordagem de classificação de linha, a partir de um sistema de pontuação para cada linha de forma independente, que organiza a ordem dos filmes e quais deles pertencerão à linha em questão. Por fim, como a classificação de linha é restrita à diversidade, é utilizada uma abordagem em estágio, que permite a relação dos filmes entre as linhas para gerar a página.

Finalmente, é importante reforçar que o texto em questão destaca quais dados são coletados e como eles interferem na recomendação das produções para o usuário. Não é mencionado que os dados coletados são interpretados numa lógica de mercado para guiar a produção de novos filmes e séries que levam em conta a análise combinada de diversos eventos pelo usuário (BURROUGHS, 2019). Esses eventos incluem também momentos em que o usuário deu pausa, abandonou um filme/série ou acelerou e adiantou o tempo de uma produção. Também é relevante lembrar que a Netflix cria recomendações baseadas cada vez mais em metadados gerados automaticamente dos eventos dos usuários combinados com “espectadores pagos” (pessoas contratadas pela empresa para assistir e taguear os conteúdos) que ajudam a taguear as produções com diversas palavras-chave (COHN, 2019).

Considerações finais

O texto da Netflix aponta que, com poucos dados coletados do usuário, o serviço consegue efetivar seu sistema elaborado de recomendação de títulos de forma personalizada. Quando comparado a outras plataformas e sites de rede sociais, por exemplo, de fato são poucas as informações extraídas pela plataforma de *streaming*. São essas informações que se comportam como um Padrão de Inclusão e que se constituem matéria prima de um banco de dados que alimenta o sistema, criando um perfil personalizado para cada conta/indivíduo. Esse perfil é o que direciona os próximos títulos recomendados com a sugestão no sistema de busca, bem como – e principalmente – a posição desses títulos nas fileiras e ordens na interface inicial.

Embora o texto analisado seja curto, acreditamos que ele expõe de forma simples o modo de funcionamento do sistema. Não esperávamos que a plataforma apontasse possíveis problemas advindos desses recursos e nem foi o foco deste artigo analisar as consequências culturais desse tipo de sistema de recomendação. Essa discussão pode ser ampliada com outras literaturas que destacam a criação de um efeito de “filtro-bolha” no consumo cultural, ou até mesmo de uma curadoria matemática de produtos culturais que pode reduzir a diversidade e o contato com títulos que fujam



daquele padrão de consumo. Alguns autores destacam que a tendência de usar os dados a favor da lógica comercial da Netflix encontra respaldo nas reflexões sobre um novo tipo de capitalismo. Zuboff (2020) lembra que o atual estágio faz da informação humana uma matéria-prima, sujeita a ser comercializada para manutenção do lucro. O comportamento online gera rastros virtuais, por sua vez capazes de gerar produtos de predição, que “[...] passaram a ser caçados de modo agressivo e obtidos por meio de vigilância” (*op. cit.*, p. 121), como acontece com a Netflix. E, ainda que colete “poucos dados”, a relação que se estabelece entre eles é um dos mecanismos de poder e de lucro da empresa.

Cabe destacar que a visão sobre como funcionam os algoritmos é uma forma de poder: “vital para a participação no discurso público, essencial para ganhar visibilidade on-line, constituidor de credibilidade e das oportunidades que resultam dela” (GILLESPIE, 2018, p.111). Por isso mesmo, apesar do texto apresentar uma “linguagem para usuário mais leigo” sobre o tema, não se esconde que o algoritmo em si não é exposto e se constitui num segredo comercial do serviço. Assim, reconhece-se que a Netflix combina dois mercados californianos, o da lógica que impera no Vale do Silício (com o extrativismo dos dados) e o de Hollywood (com a produção em larga escala de produtos audiovisuais). Reconhece-se que quanto mais dados em profundidade do usuário, maior a “possibilidade” de corrigir distorções do sistema e ter uma recomendação mais assertiva das produções.

Por fim, embora não seja este o ponto central deste artigo, pusemo-nos a questionar quais as reais necessidades, os interesses e as consequências práticas da empresa criar uma sessão em sua página oficial para falar sobre o funcionamento dos algoritmos em uma “linguagem para leigos”. Muito aquém de fomentar uma alfabetização midiática ou uma discussão crítica, parece-nos que esse tipo de divulgação serve centralmente para criar uma narrativa em torno de supostos benefícios dos algoritmos. De fato, esse conteúdo não é suficiente para que os usuários criem uma crítica sobre o modo de funcionamento do sistema, porque inclusive o texto é produzido pela própria plataforma. Para a maioria dos usuários, o conhecimento sobre os algoritmos ainda é vago, simplista, e, como lembra Gillespie (2018), às vezes, equivocado. Dessa forma, sugerimos que seja necessário promover uma alfabetização midiática pelos próprios usuários para pensar no modo como a Netflix (bem como outras plataformas de conteúdo como Spotify, YouTube, Twitter, etc.) operacionalizam suas recomendações.

Para que seja crítico, tal processo precisa ser conduzido de forma independente das empresas, buscando confrontar as narrativas por elas criadas e as consequências econômicas, políticas e culturais com base na realidade. Entendemos que tal ação pode



e deve ser desenvolvida nas escolas e nas universidades, em conjunto com movimentos políticos e culturais que fazem questionamentos sobre os monopólios da indústria cultural.

Referências

AGGARWAL, C. C. **Recommender Systems**. Cham: Springer International Publishing, 2016.

ALEXANDER, N. Catered to your Future Self: Netflix's "Predictive Personalization" and the Mathematization of Taste. *In*: SMITH-ROWSEY, D.; MCDONALD, K. (Org.). **The Netflix Effect: Technology and Entertainment in the 21st Century**. London: Bloomsbury Academic Publishing, 2016. p. 81-100.

ALVINO, C.; BASILICO, J. Learning a Personalized Homepage. **Netflix TechBlog**, 9 abr. 2015. Disponível em: <https://netflixtechblog.com/learning-a-personalized-homepage-aa8ec670359a>. Acesso em: 10 mai. 2022.

BEER, D. Algorithms: Shaping Tastes and Manipulating the Circulations of Popular Culture. **Popular Culture and New Media**, p. 63-100, 2013.

BRAGHINI, K.; MONTAÑO LA CRUZ, S. E. Software, dado e algoritmo como formas culturais na Netflix. **Intexto**, n. 44, p. 161-183, 1 jan. 2019. Disponível em: <https://seer.ufrgs.br/intexto/article/view/77671>. Acesso em: 10 fev. 2022.

BRASIL. **Lei nº 13.709/2018 (Lei Geral de Proteção de Dados Pessoais)**. Brasília, DF: Presidência da República. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 1 fev. 2022.

BURROUGHS, B. House of Netflix: Streaming media and digital lore. **Popular Communication**, v. 17, n. 1, p. 1-17, 2 jan. 2019. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/15405702.2017.1343948>. Acesso em: 5 abr. 2023.

CARLAN DA SILVEIRA, S.; IUVA, P. Tecnoestética algorítmica e hibridismos audiovisuais nos trailers da Netflix. **Comunicação & Informação**, v. 23, p. 1-18, 30 dez. 2020. Disponível em: <https://www.revistas.ufg.br/ci/article/view/66235>. Acesso em: 5 abr. 2022.

COHN, J. **The Burden of choice: recommendations, subversion, and algorithmic culture**. New Brunswick: Rutgers University Press, 2019.

CURTIN, M.; HOLT, J.; SANSON, K. **Distribution Revolution: Conversations about the Digital Future of Film and Television**. Los Angeles: University of California Press, 2014.

FERNÁNDEZ-MANZANO, E.-P.; NEIRA, E.; CLARES-GAVILÁN, J. Data



management in audiovisual business: Netflix as a case study. **El Profesional de la Información**, v. 25, n. 4, p. 568, 2016. Disponível em: <<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2016.jul.06>>. Acesso em: 10/11/2022

GAW, F. Algorithmic logics and the construction of cultural taste of the Netflix Recommender System. **Media, Culture & Society**, v. 44, n. 4, p. 706-725, 25 mai. 2022. Disponível em: <http://journals.sagepub.com/doi/10.1177/01634437211053767>. Acesso em: 5 mar. 2023.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6 ed. São Paulo: Atlas, 2008.

GILLESPIE, T. The Relevance of Algorithms. In: GILLESPIE, T.; BOCZKOWSKI, P. J.; FOOT, K. A. (Org.). **Media Technologies: Essays on Communication, Materiality, and Society**. Cambridge, MA: MIT Press, 2014. p. 167-194.

GILLESPIE, T. A relevância dos algoritmos (tradução). **Parágrafo**, v. 6, n. 1, p. 95-121, 2018. Disponível em: <http://revistaseletronicas.fiamfaam.br/index.php/recicofi/article/view/722>. Acesso em: 1 fev. 2022.

GOFFEY, A. Algorithm. In: FULLER, M. (Org.). **Software Studies: a lexicon**. Cambridge, MA: MIT Press, 2008. p. 15-20.

GOLDBERG, D. *et al.* Using collaborative filtering to weave an information tapestry. **Communications of the ACM**, v. 35, n. 12, p. 61-70, dez. 1992. Disponível em: <https://dl.acm.org/doi/10.1145/138859.138867>. Acesso em: 1 fev. 2022.

GOMEZ-URIBE, C. A.; HUNT, N. The Netflix Recommender System. **ACM Transactions on Management Information Systems**, v. 6, n. 4, p. 1-19, 14 jan. 2016. Disponível em: <https://dl.acm.org/doi/10.1145/2843948>. Acesso em: 10 fev. 2022.

HALLINAN, B.; STRIPHAS, T. Recommended for you: The Netflix Prize and the production of algorithmic culture. **New Media & Society**, v. 18, n. 1, p. 117-137, 23 jan. 2016. Disponível em: <http://journals.sagepub.com/doi/10.1177/1461444814538646>. Acesso em: 3 mar. 2023.

JENNER, M. Is this TVIV? On Netflix, TVIII and binge-watching. **New Media & Society**, v. 18, n. 2, p. 257-273, 7 fev. 2016. Disponível em: <http://journals.sagepub.com/doi/10.1177/1461444814541523>. Acesso em: 1 fev. 2022.

LATOUR, B. **Ciência em ação: como seguir cientistas e engenheiros sociedade afora**. São Paulo: Editora UNESP, 2000.

LOZIC, J. Financial Analysis of Netflix Platform At the Time of Covid 19 Pandemic. In: HAMMES, K.; VOLAREVIC, H.; KAUROVA, O. (Eds.). **66th International Scientific Conference on Economic and Social Development**. Book of Proceedings. Rabat: Varazdin Development and Entrepreneurship Agency, 2021. p.



78-86. Disponível em: <https://www.bib.irb.hr/1118937>. Acesso em: 5 abr. 2023.

MOROZOV, E. **Big tech**: a ascensão dos dados e a morte da política. São Paulo: Ubu Editora, 2018.

MORRIS, J. W. Curation by code: Infomediaries and the data mining of taste. **European Journal of Cultural Studies**, v. 18, n. 4-5, p. 446-463, 16 ago. 2015. Disponível em: <http://journals.sagepub.com/doi/10.1177/1367549415577387>. Acesso em: 10 mai. 2022.

NETFLIX. Como funciona o sistema de recomendações da Netflix. **Netflix - Central de Ajuda**, [2019?]. Disponível em: <https://help.netflix.com/pt/node/100639>. Acesso em: 10 mai. 2022.

NETFLIX. **About Netflix - Página inicial**. Disponível em: https://about.netflix.com/pt_br. Acesso em: 10 mai. 2022.

O'NEIL, C. **Weapons of math destruction**: how big data increases inequality and threatens democracy. New York: Broadway Books, 2016.

PAJKOVIC, N. Algorithms and taste-making: Exposing the Netflix Recommender System's operational logics. **Convergence: The International Journal of Research into New Media Technologies**, v. 28, n. 1, p. 214-235, 17 fev. 2022. Disponível em: <http://journals.sagepub.com/doi/10.1177/13548565211014464>. Acesso em: 10 mai. 2022.

PARISER, E. **O filtro invisível**: o que a internet está escondendo de você. Rio de Janeiro: Zahar, 2012.

PASQUALE, F. **The Black Box Society**: The Secret Algorithms That Control Money and Information. Cambridge, MA: Harvard University Press, 2016.

PINCH, T. J.; BIJKER, W. E. The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. **Social Studies of Science**, v. 14, n. 3, p. 399-441, 29 ago. 1984. Disponível em: <http://journals.sagepub.com/doi/10.1177/030631284014003004>. Acesso em: 12 nov. 2021.

RITTER, O. I Know What You Saw Last Summer: Uma Análise do Uso de Dados Pessoais pela Plataforma Netflix. In: BELLI, L.; RAMOS, B. (Org.). **Políticas Digitais no Brasil**: Acesso à Internet, Proteção de Dados e Regulação. Rio de Janeiro: FGV Direito Rio, 2021. p. 158-172. Disponível em: <https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/30688/0>. Acesso em: 10 mai. 2022.

SEAVER, N. Algorithms as culture: Some tactics for the ethnography of algorithmic systems. **Big Data & Society**, v. 4, n. 2, p. 1-12, 9 dez. 2017. Disponível em: <http://journals.sagepub.com/doi/10.1177/2053951717738104>. Acesso em: 5 abr. 2023.

SIMÕES, A. L.; CUNHA, M. DO N. Comunicação na Internet e a Violação do Direito

à Privacidade: uma análise avaliativa das políticas e termos de uso na Internet. **Comunicação & Inovação**, v. 19, n. 40, p. 39-55, 25 ago. 2018. Disponível em: http://seer.uscs.edu.br/index.php/revista_comunicacao_inovacao/article/view/5093. Acesso em: 3 mar. 2022.

SRNICEK, N. **Plataform Capitalism**. Cambridge, UK: Policy Press, 2016.

SUMPTER, D. **Dominados pelos números: do Facebook e Google às Fake News - os algoritmos que controlam nossa vida**. Rio de Janeiro: Bertrand Brasil, 2019.

TALIB, S. *et al.* Perception analysis of social networks' privacy policy: Instagram as a case study. *In: The 5th International Conference on Information and Communication Technology for The Muslim World (ICT4M)*, 5., nov. 2014 Kuching, Malaysia. Rabat-Morocco: IEEE, nov. 2014. p. 1–5. Disponível em: <http://ieeexplore.ieee.org/document/7020612/>. Acesso em: 1 fev. 2022.

TIKU, N. Why Netflix Features Black Actors in Promos to Black Users. **Wired**, 24 out. 2018 20:20. Culture. Disponível em: <https://www.wired.com/story/why-netflix-features-black-actors-promos-to-black-users/>. Acesso em: 10 mai. 2022.

WILSON, S. *et al.* The Creation and Analysis of a Website Privacy Policy Corpus. *In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Berlim: Association for Computational Linguistics, 2016. p. 1330-1340. Disponível em: <http://aclweb.org/anthology/P16-1126>. Acesso em: 5 fev. 2022.

WINKLER, S.; ZEADALLY, S. Privacy Policy Analysis of Popular Web Platforms. **IEEE Technology and Society Magazine**, v. 35, n. 2, p. 75-85, jun. 2016. Disponível em: <http://ieeexplore.ieee.org/document/7484849/>. Acesso em: 3 fev. 2022.

ZUBOFF, S. **A era do capitalismo de vigilância**. Rio de Janeiro: Intrínseca, 2020.

Recebido em: 21/12/2022. Rodada 1: Revisora A 20/02/2023. Revisora B 07/03/2023. Rodada 2: Revisora A 17/07/2023. Aprovado em: 22/09/2023.

Informações sobre coautoria

Concepção e desenho do estudo:

Tiago Franklin Rodrigues Lucena e Thiago Fanelli Ferraiol.

Aquisição, análise ou interpretação dos dados:

Eduarda Carretero Garcia e Mariana Maronezzi Brezovsky.

Redação do manuscrito:

Tiago Franklin Rodrigues Lucena, Thiago Fanelli Ferraiol, Eduarda Carretero Garcia e Mariana Maronezzi Brezovsky.

**Informações sobre o artigo**

Resultado de projeto de pesquisa:

(1419/2021 UEM) - Mídias algorítmicas: uma revisão sobre os algoritmos de recomendação em plataformas streaming / (824/2021 UEM) - Elementos matemáticos dos algoritmos de recomendação em plataformas de streaming e suas consequências sociais e culturais.

Fontes de financiamento:

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq.

Considerações éticas:

Não se aplica.

Declaração de conflitos de interesse:

Não se aplica.

Apresentação anterior:

Os resultados foram apresentados de forma oral no XXXI EAIC: Encontro Anual de Iniciação Científica realizado na Universidade Estadual de Maringá (UEM) em 2022.