



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS  
ESCOLA DE COMUNICAÇÃO  
JORNALISMO

**DOS *GATEKEEPERS* HUMANOS AOS COMPUTACIONAIS:  
O SISTEMA DE RECOMENDAÇÃO “PARA VOCÊ”, DO  
GOOGLE NOTÍCIAS, E SEUS IMPACTOS SOBRE A  
FORMAÇÃO DE BOLHAS INFORMACIONAIS**

**CAROLINA FARIA NALIN**

Rio de Janeiro

2021



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS  
ESCOLA DE COMUNICAÇÃO  
JORNALISMO

**DOS *GATEKEEPERS* HUMANOS AOS COMPUTACIONAIS:  
O SISTEMA DE RECOMENDAÇÃO “PARA VOCÊ”, DO  
GOOGLE NOTÍCIAS, E SEUS IMPACTOS SOBRE A  
FORMAÇÃO DE BOLHAS INFORMACIONAIS**

Monografia submetida à Banca de Graduação  
como requisito para obtenção do diploma de  
Comunicação Social – Jornalismo.

**CAROLINA FARIA NALIN**

**Orientador: Prof. Dr. Paulo César Castro de Sousa**

Rio de Janeiro  
2021

# FICHA CATALOGRÁFICA

## CIP - Catalogação na Publicação

NN171g Nalin, Carolina Faria  
Dos gatekeepers humanos aos computacionais:  
o sistema de recomendação "Para Você", do  
Google Notícias, e seus impactos sobre a  
formação de bolhas informacionais / Carolina  
Faria Nalin. --  
Rio de Janeiro, 2021.  
102 f.

Orientador: Paulo César Castro de Sousa.  
Trabalho de conclusão de curso (graduação) -  
Universidade Federal do Rio de Janeiro, Escola  
da Comunicação, Bacharel em Comunicação Social:  
Jornalismo, 2021.

1. Google Notícias. 2. gatekeeper. 3.  
algoritmos. 4. filtro-bolha. 5. diversidade  
noticiosa. I. Sousa, Paulo César Castro de,  
orient. II. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
ESCOLA DE COMUNICAÇÃO

**TERMO DE APROVAÇÃO**

A Comissão Examinadora, abaixo assinada, avalia a Monografia **Dos *gatekeepers* humanos aos computacionais: o sistema de recomendação “Para Você”, do Google Notícias, e seus impactos sobre a formação de bolhas informacionais**, elaborada por Carolina Faria Nalin.

Monografia examinada:

Rio de Janeiro, no dia 16/11/2021

Comissão Examinadora:

Orientador: Prof. Dr. Paulo César Castro de Sousa  
Doutor em Comunicação e Cultura pela Escola de Comunicação – UFRJ  
Departamento de Comunicação – UFRJ

Prof. Dr. Paulo Faltay Filho  
Doutor em Comunicação e Cultura pela Escola de Comunicação – UFRJ  
Departamento de Comunicação – UFRJ

Prof. Dr. Wedencley Alves Santana  
Doutor em Linguística – UNICAMP  
Faculdade de Comunicação – UFJF

Rio de Janeiro

2021

## AGRADECIMENTOS

Escrever o trabalho de conclusão de curso em meio à pandemia, sem dúvidas, não foi uma tarefa fácil. A escrita pode parecer, inclusive, um exercício muito solitário. Mas felizmente aprendi que ela só se desenvolve com o coração, a partir da contribuição de muitas pessoas. Tive a sorte de poder contar com o apoio de muitas delas e é justamente isso que dá sentido à caminhada.

Em primeiro lugar agradeço a Deus, por permitir que cada coisa aconteça no seu devido tempo e ordem. “Por isso não temas, pois estou com você” (Isaías 41:10).

À minha mãe, Simone Vieira, que sempre me amparou e me deu forças para lutar pelo que acredito. Ao meu pai, Francisco Nalin, *in memoriam*, que sempre acreditou que eu seria jornalista antes mesmo de eu ter certeza. Aos meus irmãos, Camilla Nalin e Lucas Nalin, por todo apoio e incentivo. À minha avó, Maria Aparecida, pela acolhida nos momentos de alegria e também nos de incertezas e dificuldades.

Aos meus amigos que viveram comigo esses quase cinco anos de graduação, dividindo as aflições e as angústias, mas também comemorando cada etapa vencida, incluindo a de conclusão deste trabalho. Agradeço também os meus colegas de turma por todo o aprendizado ao longo desses anos.

À Profa. Dra. Zilda Martins, que me fez despertar o interesse pela pesquisa acadêmica durante as excelentes aulas do Laboratório de Comunicação Crítica ainda quando cursava o Ciclo Básico na ECO-UFRJ.

Ao Programa de Educação Tutorial (PET), do Ministério da Educação, pelo benefício da bolsa de auxílio ao desenvolvimento de pesquisas, estudos em grupo e projetos de extensão. A todos os colegas e amigos com quem pude caminhar lado a lado no Programa de Educação Tutorial da ECO-UFRJ, onde aprendi a valorizar ainda mais a pesquisa acadêmica e a universidade pública. Em especial, à Anna Beatriz Lima, que aceitou dividir comigo o processo de criação deste trabalho ainda como projeto de pesquisa no PET-ECO.

Ao Prof. Dr. Paulo César Castro, pelos ensinamentos, primeiro como tutor do PET-ECO e agora como orientador. Obrigada pela generosidade, paciência e atenção no processo de escrita deste trabalho.

Aos Professores Doutores Paulo Faltay Filho e Wedencley Alves Santana, por aceitarem integrar a minha banca e realizarem uma leitura e análise cuidadosa deste trabalho.

À Escola de Comunicação da UFRJ, seus professores e funcionários, por terem contribuído para uma formação cidadã, dentro e fora da sala de aula. Gratidão aos que lutam pelo acesso público, universal e gratuito do ensino superior no Brasil. Desejo que cada vez mais jovens tenham a oportunidade de ingressar nessas e outras universidades de excelência, assim como tive o privilégio de vivenciar essa experiência.

NALIN, Carolina Faria. **Dos *gatekeepers* humanos aos computacionais: o sistema de recomendação “Para Você”, do Google Notícias, e seus impactos sobre a formação de bolhas informacionais.** Orientador: Paulo César Castro de Sousa. Trabalho de conclusão de curso (Graduação em Comunicação Social – Jornalismo). Rio de Janeiro: ECO/UFRJ, 2021.

## RESUMO

Se o consumo de notícias é tradicionalmente marcado pela seleção e categorização de informações realizadas por *gatekeepers* humanos, constituindo-se, estes processos, como marcas fundamentais da produção jornalística, com a internet, tais atividades ganham novos contornos. As chaves dos portões por onde passam (ou não) os conteúdos noticiosos que serão disponibilizados ao público estão nas mãos, agora, de outros tipos de guardiões, mais precisamente dos algoritmos computadorizados. Ao atuarem como agregadores automatizados de notícias, eles passam à condição de novos intermediários entre os leitores e os acontecimentos. Para estudar esse fenômeno de *gatekeeping* algorítmico, foi escolhido o agregador Google Notícias, mais especificamente através da categoria “Para Você”, de modo a avaliar até que ponto a plataforma, ao fazer recomendações personalizadas de notícias aos usuários, é responsável por limitar a pluralidade de conteúdos e, dessa forma, contribuir para submeter seus leitores a verdadeiras bolhas informativas. Diante da dificuldade de estudo de algoritmos cujos códigos são verdadeiras caixas pretas, como é o caso do agregador da Google, a pesquisa usou o recurso da engenharia reversa como proposta por Kitchin (2017), que consiste na análise dos resultados da ação computacional do agregador (o chamado *output*) na busca de entender o processo de seleção da ferramenta. O recurso metodológico foi aplicado aos perfis de quatro usuários do Google Notícias, para, considerando seus hábitos de navegação na internet, observar quais notícias lhes foram recomendadas e, assim, avaliar se seus consumos noticiosos foram impactados pela oferta de uma baixa diversidade de temas.

**Palavras-chave:** Google Notícias; *gatekeeper*; algoritmos; filtro-bolha; diversidade noticiosa.

## LISTA DE FIGURAS

Figura 1 – Incorporação de <i>links</i> de notícias aos resultados das pesquisas dos termos “Obama” e “Gordon Brown” no Google Pesquisa em 2007.....	28
Figura 2 – Captura de tela da aba “Para Você” exibida ao usuário no Google Notícias em versão <i>mobile</i> .....	34
Figura 3 – Captura de tela da aba “Manchetes” exibida ao usuário no Google Notícias em versão <i>desktop</i> .....	34
Figura 4 – Captura de tela da aba “Para Você” exibida ao usuário no Google Notícias em versão <i>desktop</i> .....	35
Figuras 5A e 5B – Representação de dois cenários de algoritmo inacessíveis, com diferentes níveis de observação (total ou apenas a saída)	40
Figura 6 – Editorias da plataforma Google Notícias no <i>desktop</i> .....	46

## LISTA DE GRÁFICOS

Gráfico 1 – Participação de mercado dos mecanismos de pesquisa em <i>desktops</i> em todo o mundo, de janeiro de 2010 a setembro de 2021.....	27
Gráfico 2 – Editorias exibidas ao perfil A entre os dias 19/04/2021 e 23/04/2021...	51
Gráfico 3 – Editorias exibidas ao perfil A entre os dias 17/05/2021 e 21/05/2021...	52
Gráfico 4 – Editorias exibidas ao perfil A entre os dias 07/06/2021 e 10/06/2021...	54
Gráfico 5 – Editorias exibidas ao perfil B entre os dias 19/04/2021 e 23/04/2021....	55
Gráfico 6 – Editorias exibidas ao perfil B entre os dias 17/05/2021 e 21/05/2021....	57
Gráfico 7 – Editorias exibidas ao perfil B entre os dias 07/06/2021 e 10/06/2021....	59
Gráfico 8 – Editorias exibidas ao perfil C entre os dias 19/04/2021 e 23/04/2021....	61
Gráfico 9 – Editorias exibidas ao perfil C entre os dias 17/05/2021 e 21/05/2021....	63
Gráfico 10 – Editorias exibidas ao perfil C entre os dias 07/06/2021 e 10/06/2021..	64
Gráfico 11 – Editorias exibidas ao perfil D entre os dias 17/05/2021 e 21/05/2021..	65
Gráfico 12 – Editorias exibidas ao perfil D entre os dias 07/06/2021 e 11/06/2021 (exceto o dia 10/06).....	67
Gráfico 13 – Veículos exibidos ao perfil A entre os dias 19/04/2021 e 23/04/2021..	73
Gráfico 14 – Veículos exibidos ao perfil A entre os dias 17/05/2021 e 21/05/2021..	74
Gráfico 15 – Veículos exibidos ao perfil A entre os dias 07/06/2021 e 10/06/2021..	76
Gráfico 16 – Veículos exibidos ao perfil B entre os dias 19/04/2021 e 23/04/2021..	77
Gráfico 17 – Veículos exibidos ao perfil B entre os dias 17/05/2021 e 21/05/2021..	78
Gráfico 18 – Veículos exibidos ao perfil B entre os dias 07/06/2021 e 10/06/2021..	79
Gráfico 19 – Veículos exibidos ao perfil C entre os dias 19/04/2021 e 23/04/2021..	81
Gráfico 20 – Veículos exibidos ao perfil C entre os dias 17/05/2021 e 21/05/2021..	82
Gráfico 21 – Veículos exibidos ao perfil C entre os dias 07/06/2021 e 10/06/2021..	83
Gráfico 22 – Veículos exibidos ao perfil D entre os dias 17/05/2021 e 21/05/2021..	84
Gráfico 23 – Veículos exibidos ao perfil D entre os dias 07/06/2021 e 10/06/2021..	85

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>12</b>
<b>2 DOS GATEKEEPERS HUMANOS AOS GATEKEEPERS ALGORÍTMICOS: O TRABALHO DE MEDIAÇÃO ENTRE A NOTÍCIA E O LEITOR</b>	<b>16</b>
2.1 O porteiro da redação: um breve histórico do gatekeeper no jornalismo	16
2.1.1 A influência da estrutura organizacional na seleção de notícias	18
2.2 O algoritmo no processo de seleção jornalística: gatekeeping no século XXI	19
2.2.1 Cultura algorítmica e gatekeeper algorítmico	21
2.2.2 A vez dos agregadores: o case Google Notícias	28
2.2.3 Personalização e diversidade no Google Notícias: estamos reféns das bolhas noticiosas?	31
<b>3 DESAFIOS METODOLÓGICOS AO ESTUDO DOS ALGORITMOS</b>	<b>38</b>
3.1 Engenharia reversa	38
3.2 Procedimento de seleção dos perfis	41
3.3 Coleta de dados	44
3.4 Classificação das notícias	45
<b>4 O ALGORITMO DO GOOGLE NOTÍCIAS EM AÇÃO: ENTRE A DIVERSIDADE EDITORIAL E A BOLHA INFORMACIONAL</b>	<b>50</b>
4.1 A diversidade de editorias resultante da recomendação personalizada	50
4.1.1. Temáticas recomendadas ao perfil A	51
4.1.2. Temáticas recomendadas ao perfil B	55
4.1.3. Temáticas recomendadas ao perfil C	61
4.1.3. Temáticas recomendadas ao perfil D	65
4.1.4 Discussão sobre a diversidade de temáticas recomendadas aos usuários a partir dos resultados	68
4.2 A pluralidade de veículos resultante da recomendação personalizada	72
4.2.1. Fontes recomendadas ao perfil A	73
4.2.2. Fontes recomendadas ao perfil B	77
4.2.3. Fontes recomendadas ao perfil C	81
4.2.4. Fontes recomendadas ao perfil D	84
4.2.5 Discussão sobre os veículos recomendados aos usuários a partir dos resultados	86
<b>5 CONSIDERAÇÕES FINAIS</b>	<b>88</b>
<b>6 REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>93</b>
<b>7 APÊNDICES</b>	<b>97</b>
7.1 Apêndice A - Participação das editorias de acordo com os perfis	97
7.2 Apêndice B - Participação dos veículos de acordo com os perfis	99
7.2 Apêndice C - Planilha com as recomendações do Google Notícias para os perfis	104

## 1 INTRODUÇÃO

O presente trabalho se propõe a investigar as mudanças no processo de seleção das notícias diante do desenvolvimento tecnológico e do surgimento dos meios digitais, considerando que, no campo da comunicação, e, principalmente, do jornalismo, o advento da Internet em meados da década de 1990 passou a promover significativas mudanças no que se refere ao acesso dos leitores às notícias publicadas pelos veículos jornalísticos no ciberespaço.

Conforme veremos no primeiro capítulo, neste recente cenário de hiperinformação, novas formas de controle dos fluxos de informação passam a agir por meio dos algoritmos privados de empresas no ambiente digital. No âmbito das notícias, o vasto conteúdo disponível tornou propícia a criação de agregadores de notícias *on-line* que têm como base sistemas de recomendação automatizados, tornando-se esses, potencialmente, novos intermediários entre a notícia e o leitor. O uso de algoritmos passou a ser, no entanto, questionado por autores como Pariser (2012) e Sunstein (2009), que demonstram preocupação com o quanto o seu uso para personalização de notícias pode se tornar prejudicial para a diversidade de pontos de vista e, conseqüentemente, para a democracia.

Desta forma, partimos para o seguinte problema de pesquisa: neste processo de filtragem das notícias no ambiente digital, estamos sob a atuação de novos *gatekeepers*, agora algorítmicos? Se sim, quais efeitos este novo “portão” pode ter sobre a informação que chega ao leitor? Essa indagação parte, em um primeiro momento, de um interesse pessoal em buscar compreender os impactos de uma parte que compõe a cultura algorítmica (STRIPHAS, 2015) na qual estamos cada vez mais mergulhados nessas duas primeiras décadas do século XXI. Mas há também um interesse profissional, cujo objetivo é buscar identificar o papel dos agregadores de notícias *on-line* enquanto novos atores responsáveis por mediar o acesso entre o que os veículos noticiosos publicam e o que os usuários visualizam. Além disso, há uma vontade de propor mais discussões no ambiente acadêmico sobre o impacto dessas filtragens algorítmicas a partir do que pensaram pesquisadores como Eli Pariser e Cass Sunstein.

Nesse sentido, o trabalho partiu do objetivo de entender o papel do agregador Google Notícias como um possível *gatekeeper* algorítmico, ou seja, sua ação - tomada muitas vezes como não sujeita às subjetividades humanas - no processo de seleção de notícias, analisando as recomendações exibidas pelo agregador a diferentes perfis de usuários por meio da

personalização implícita. (THURMAN; SCHIFFERES, 2012)

Essa intenção se desdobra em objetivos específicos, que são eles: identificar quais são as principais temáticas e quais veículos são recomendadas a cada um dos perfis; de que forma, quantitativamente e qualitativamente, os interesses dos usuários previamente identificados pela plataforma estão sendo ali contemplados; investigar se essa personalização pode levar a uma redução da diversidade de ideias que tem como consequência mais imediata o “aprisionamento” dos usuários a bolhas informativas.

Para compreender esse complexo fenômeno, partimos, no primeiro capítulo, de uma revisão acerca da teoria do *gatekeeping*, utilizada no campo da comunicação primeiramente a partir de um estudo desenvolvido por David White, em 1950, para medir a influência da ação pessoal na seleção das notícias publicadas por um jornal nos Estados Unidos. Em seguida, retomamos as discussões realizadas por Wolf (2003) e Shoemaker e Vos (2011), que propuseram atualizações sobre o conceito de *gatekeeping*.

Em um subcapítulo ainda desta primeira parte, retomamos o papel histórico que empresas de tecnologia, munidas de seus algoritmos privados, assumiram no ciberespaço a partir das observações feitas por Doctor (2011) e Wu (2012). Nesse ambiente digital, em que as mídias tradicionais não se ajustaram completamente a ele, a aceitação à Internet tornou propício o surgimento de novos intermediários, como os agregadores de notícias, conforme explicam Águila-Obra, Padilla-Meléndez e Serarols-Tarrés (2007).

Em um segundo subcapítulo, destacamos a definição do termo algoritmo proposta por Gillespie (2014) e discutimos o papel da empresa Google, a partir das análises de Vaidhynathan (2011) sobre o impacto dos serviços dessa companhia no âmbito da esfera pública global. Considerando o poder da gigante do Vale do Silício, ao qual o autor dá o nome de “Googlelização”, apresentamos um histórico sobre o Google Notícias, agregador da plataforma, e como a empresa detalhou o funcionamento de seus algoritmos ao longo do tempo. Apresentamos ainda, ao final deste capítulo, as previsões de Negroponte, que, já em 1995, defendia que as tecnologias digitais levariam ao surgimento do que chamou de *Daily Me*, consubstanciado na edição individualizada de um jornal cujas notícias seriam apenas as que realmente interessariam a cada leitor. Ou seja, o co-fundador do Media Lab do Massachusetts Institute of Technology (MIT), atento ao potencial da internet, anteviu o que o Google Notícias e outros agregadores fazem, todos os dias, através de suas recomendações.

Considerando o trabalho de seleção e classificação realizado por esses agregadores, dialogamos com os questionamentos feitos por Pariser (2012) e Sunstein (2009), que acreditam que os algoritmos de diferentes plataformas que atuam no campo da comunicação podem contribuir para a redução do contato com uma ampla diversidade de pontos de vista.

O segundo capítulo apresenta a metodologia da engenharia reversa proposta por Kitchin (2017) para justificar o ponto de partida da pesquisa, de modo a observar quais e como as notícias e os respectivos veículos responsáveis por elas acabam sendo recomendados aos usuários através da aba “Para Você” do Google Notícias. Nesta parte, destacamos o procedimento de seleção dos perfis estudados com um breve resumo sobre cada um deles, bem como a forma como se desenvolveu a coleta dos dados e como os resultados obtidos foram organizados.

No terceiro capítulo, apresentamos uma análise do levantamento, dividindo este capítulo em dois subcapítulos. No primeiro, observamos, a partir dos gráficos, a evolução das notícias recomendadas aos perfis no decorrer da pesquisa, buscando identificar quais foram os assuntos mais recorrentes. Ao fim deste subcapítulo, realizamos uma observação geral do que foi possível constatar na pesquisa para verificar se houve redução da diversidade em termos de temáticas apresentadas. Em seguida, no segundo subcapítulo, usamos gráficos para mostrar como se deu o nível de participação de cada um dos veículos recomendados pelo agregador. Por fim, realizamos apontamentos gerais no que diz respeito à representação dos veículos observados durante o período de coleta das informações, identificando se houve baixa diversidade de fontes jornalísticas.

Passados 70 anos do lançamento da teoria que buscou lançar luz sobre os critérios de seleção de notícias pelos jornalistas e, com isso, sobre a diversidade informacional a que os leitores estavam expostos, a internet, com as radicais transformações comunicacionais e informacionais que inaugurou, tem colocado o problema sob novos ângulos. Diante da tsunami de conteúdos que nos atinge cotidianamente, é aos algoritmos computacionais que tem sido confiado o trabalho de triagem do que é útil e importante para nós, individualmente. E mais: com eles, não precisaríamos mais ficar sujeitos às escolhas subjetivas e enviesadas dos *gatekeepers* de carne e osso, pois, regidos pela lógica matemática, agem racional e objetivamente. Confrontar esse tipo de discurso já é, por si só, uma boa razão para a pesquisa que nos propusemos a desenvolver neste trabalho. Mas ao falar do papel dos algoritmos

computacionais sobre a produção e o consumo de notícias, também queremos tratar do futuro do jornalismo, que, impactado pelas novas condições advindas do ciberespaço, certamente não é mais o mesmo de sete décadas atrás. Esperamos que esta monografia seja uma pequena contribuição a este debate.

## **2 DOS GATEKEEPERS HUMANOS AOS GATEKEEPERS ALGORÍTMICOS: O TRABALHO DE MEDIAÇÃO ENTRE A NOTÍCIA E O LEITOR**

O presente capítulo tem como finalidade discutir de que forma a dinâmica de seleção de informações e, conseqüentemente, de notícias, realizadas desde o surgimento da atividade jornalística pelos meios tradicionais de comunicação – rádio, jornalismo impresso e televisão – tem transformado o conceito de *gatekeeper* ao longo das últimas duas décadas com a popularização do uso de algoritmos na Internet e o crescimento dos agregadores de notícias on-line. Em primeiro lugar, será realizada uma revisão da teoria do *gatekeeper*, restaurando as discussões já realizadas pelos pesquisadores. Em seguida, serão introduzidas noções sobre cultura algorítmica e a influência das seleções automatizadas na divulgação de matérias jornalísticas no século XXI, ao que chamamos de *gatekeeper* algorítmico. Ao final do capítulo, apresentaremos o *case* do Google Notícias como um dos agregadores mais utilizados e relevantes, sobretudo no Brasil, nos últimos anos.

### **2.1 O porteiro da redação: um breve histórico do *gatekeeper* no jornalismo**

Desde que se firmou como atividade profissional remunerada nos moldes do capitalismo, a partir do final do século XIX, o jornalismo tem sido alvo de questionamentos por pesquisadores das áreas das ciências humanas e sociais. No contexto de consolidação do pensamento positivista e da sociedade de massa, o jornalismo passou a ser defendido como um ofício cujas práticas e produtos resultavam da objetividade.

Como descreveu Oliveira (2003, p. 2), “Do estilo panfletário dos jornais do final do século XIX, o jornalismo passou a amparar-se em fatos e a recorrer a dispositivos para ‘objetivar’ o discurso como, por exemplo, o uso de aspas e a responsabilização de fontes, pelas citações”. Para Amaral (1996), o desenvolvimento industrial, o surgimento das agências de notícias, as duas guerras mundiais e o advento da publicidade foram os quatro principais fatores que contribuíram para que as empresas de comunicação aderissem à objetividade e às práticas capitalistas de gestão empresarial, algo que experimentamos até hoje. No Brasil, a adesão a esse ideário, que resulta em novas práticas, começa na passagem dos anos 1940 aos anos 1950, depois do intercâmbio cultural nos Estados Unidos de que participaram jornalistas como Pompeu de Souza, Alberto Dines, Danton Jobim, entre outros (SILVA, 1990).

No entanto, ao passo em que a objetividade passou a ser percebida como um valor

inerente ao jornalismo, era preciso discutir desmistificar essa ideologia, considerando que a produção de notícias também é sujeita à subjetividade. E foi a partir da teoria do *gatekeeping*, originária de uma pesquisa da área da psicologia, realizada por Kurt Lewin em 1947, que David Manning White realizou um estudo em 1950 para medir a influência da ação pessoal nas redações. O objetivo era compreender o que estava por trás do processo de seleção das notícias publicadas pelos veículos de comunicação da época.

Ao acompanhar o trabalho realizado pelo nominado editor Mr. Gates – que representava metaforicamente um jornalista com cerca de 25 anos de experiência que atuava como editor no jornal *Peoria Star*, nos Estados Unidos –, o sociólogo americano cunhou o termo *gatekeeper* para expressar o funcionamento do trabalho diário do profissional. Ele tinha a missão de selecionar, entre milhares de conteúdos fornecidos diariamente por três agências de notícias, quais matérias escritas por repórteres seriam publicadas na primeira página e nas páginas internas de um jornal com circulação de 30 mil exemplares.

O editor era tomado, assim, como o último “portão” de um fluxo de produção de notícias do qual participavam vários emissores comunicacionais, ou seja, diferentes jornalistas - tomados como espécies de “cancelas” -, através dos quais as informações passavam à medida que eram pautadas, coletadas, organizadas, editadas e publicadas. Nessa cadeia produtiva, eram esses profissionais que rejeitavam ou permitiam que os conteúdos seguissem seu fluxo e posteriormente fossem tratados como notícias, adquirindo valor suficiente para serem consumidas como tal. Os “portões”, como bem definido pelo autor, são os próprios repórteres que transmitem até o último *gatekeeper* - o editor principal ou o editor-chefe - a notícia, ficando a cargo dele publicá-la ou não. White sugere, ainda, que, ao analisar o comportamento do *gatekeeper*, seria possível compreender os fatores que condicionam as decisões durante a escolha das notícias que chegam até os leitores.

Uma história é transmitida de um *gatekeeper* para outro na cadeia de comunicações. Do repórter ao redator, passando pelo chefe da agência e pelos editores estaduais de arquivos em vários escritórios da associação de imprensa, o processo de escolha e descarte ocorre continuamente. (WHITE, 1950, p. 384, tradução nossa, destaque nosso)<sup>1</sup>

Ao partir desse processo, o autor aponta que as escolhas feitas pelo *gatekeeper* - entendido, portanto, como aquele que seleciona - são compostas por critérios subjetivos e arbitrários. Com base em dados e entrevistas, White concluiu que “muitas das razões pelas

---

<sup>1</sup> No original: “Thus a story is transmitted from one ‘gate keeper’ after another in the chain of communications. From repórter to rewrite man, through bureau chief to ‘state’ file editors at various press association offices, the process of choosing and discarding is continuously taking place.”

quais o Sr. Gates rejeita histórias se enquadram na categoria de julgamentos de valor altamente subjetivos” (WHITE, 1950, p. 386, tradução nossa).<sup>2</sup>

Dessa forma, é possível constatar que os processos de seleção jornalística carecem, por vezes, de clareza e objetividade. Isso nos leva a compreender que reconhecer a subjetividade no fluxo torna-se uma parte importante do processo comunicacional de produção e emissão, e que diretores, editores e repórteres podem atuar como guardiões dos portões por onde passam, ou são barradas, as notícias que serão consumidas a cada nova publicação realizada pelo veículo jornalístico.

### **2.1.1 A influência da estrutura organizacional na seleção de notícias**

Após o termo *gatekeeper* ter sido introduzido por David White para expressar o processo de produção jornalística, novas discussões acerca do seu procedimento metodológico e do poder de ação do conceito surgiram. Como consequência, a teoria do *gatekeeper* passou a ser discutida através de novas pesquisas, que propuseram novas reflexões. Wolf (2003), por exemplo, mostra que certas normas profissionais como “falta de espaço e sobreposição com histórias já selecionadas superavam as distorções subjetivas” (HIRSCH, 1977, p. 22 apud WOLF, 2003, p. 79).

Desta forma, diversas pesquisas na área da comunicação consideraram a influência da estrutura organizacional na seleção e na filtragem das notícias, cujo impacto é superior às preferências pessoais. Wolf (2003, p. 80) aponta, ainda, que as ações dos *gatekeepers* seriam distorções inconscientes que, “dia após dia, realçam uma determinada representação da realidade social”. Entretanto, o autor afirma que mesmo as distorções inconscientes incluem “formas de controle da informação, que podem estabelecer-se nas decisões acerca da codificação das mensagens, da seleção, da formação da mensagem, da difusão, da programação, da exclusão de toda a mensagem ou de suas componentes (DONOHUE, TICHENOR; OLIEN, 1972, p. 43 apud WOLF, 2003, p. 79).

As observações e estudos sobre essa teoria da comunicação continuaram produzindo novas ponderações. Aqui, consideramos a análise de Shoemaker e Vos (2011, p. 11) sobre o conceito, que define *gatekeeping* como “o processo de seleção e transformação de vários pequenos pedaços de informação na quantidade limitada de mensagens que chegam às pessoas diariamente”.

---

<sup>2</sup> No original: “Thus we see that many of the reasons which Mr. Gates gives for the rejection of the stories fall into the category of highly subjective value-judgments.”

## 2.2 O algoritmo no processo de seleção jornalística: *gatekeeping* no século XXI

Ainda que o atual contexto comunicacional seja diferente do que se tinha durante as primeiras e posteriores discussões sobre a teoria do *gatekeeping*, os autores supracitados ressaltam que a ideia de seleção de informações permanece relevante para compreender a comunicação no século XXI e deve, ainda, ser aprimorada. Shoemaker e Vos (2011, p. 183) apontam que alterações no ambiente comunicacional precisam ser consideradas pelos pesquisadores ao realizar análises sobre o processo de *gatekeeping* no tempo presente. O fenômeno da globalização e as próprias características que constituem a mídia online – como o seu dinamismo em tempo real – impõem novas questões metodológicas e modelam o conceito de *gatekeeping* (SHOEMAKER; VOS, 2011, p. 184).

Já Castells (2020) lembra que o início da segunda metade do século XX começou a dar os contornos do que hoje é chamado de Era ou Sociedade da Informação, e, que, em seu progresso, levou a um ambiente virtual dinâmico, característico da “sociedade em rede”. Em termos simples, a sociedade em rede é uma estrutura social baseada em redes operadas por tecnologias de comunicação e informação fundamentadas na microeletrônica e em redes digitais de computadores que geram, processam e distribuem informação a partir de conhecimento acumulado nos nós dessas redes (CASTELLS, 2020, p. 20).

Empresas e instituições, no âmbito das tecnologias de informação e comunicação, sobretudo no território estadunidense, deram início a uma revolução tecnológica, a partir principalmente da década de 1970, que foi capaz de configurar uma nova economia em escala global, além de estabelecer transformações políticas, sociais, econômicas e culturais que reestruturaram o capitalismo.

Este novo paradigma tecnológico, como sugere Castells (2020, p. 123), difere dos anteriores pois agora as tecnologias estão à disposição para “agir sobre a informação, e não apenas a informação para agir sobre a tecnologia” (2020, p. 124). Nesse sentido, o advento de ferramentas digitais e a ampliação do uso de computadores, *smartphones* e dispositivos com capacidade de acesso à Internet em escala global tornaram possível não apenas a conexão entre cidadãos de diferentes regiões do planeta por meio dos sistemas de telecomunicações, mas também impulsionaram a transmissão de informações que visam à exploração das potencialidades tecnológicas, criando novos ambientes informacionais.

Nesse contexto, a ascensão da Internet propiciou novas formas de interação amparadas na esfera virtual (LÉVY, 1999). O ciberespaço, como define o filósofo, emerge como um

novo ambiente comunicacional cuja sociabilidade possibilita e estimula a interação entre diversos atores sociais, como pessoas, empresas, governos e demais organizações. Este novo espaço comunicacional, ancorado na virtualização por meio das telecomunicações, tem como caráter natural o aumento da criação e do consumo exponencial de dados. Desse modo, ao passo em que permitem novas formas de interação, as redes digitais, e outras plataformas baseadas na Internet, têm sido responsáveis pela produção de gigantescas quantidades de dados, às quais foi dado nome de *big data*<sup>3</sup>. Neste ambiente, que promove formas novas de controle e domínio (LÉVY, 1999, p. 30), os fluxos de informação são processados por algoritmos instalados nas próprias plataformas.

Vaidhyathan (2011) ressalta que a Internet começou a se consolidar nos anos 1990 e, com isso, passou a conquistar relevância no cotidiano das pessoas. Entretanto, seus recursos ainda não eram voltados aos usuários como se apresentam hoje, décadas após o seu assentamento.

No começo, a World Wide Web (Rede Mundial) era uma coletânea intimidadora, interligada, mas não indexada. A confusão e a desordem reinavam. Era impossível separar o joio do trigo, o confiável do oportunista, e o verdadeiro do falso. A rede era estimulante e democrática – a ponto de ser anárquica. (VAIDHYANATHAN, 2011, p. 15)

Mesmo assim, as grandes empresas de comunicação, segundo Wu (2012, p. 308), chegaram a considerar a Internet uma ameaça existencial ao longo de seu surgimento e evolução. De todo modo, os veículos tentaram integrar o modelo de negócio da mídia tradicional ao funcionamento da rede mundial de computadores, que, multifacetada, não concebia apenas o transporte do conteúdo.

Doctor (2011) detalha algumas das ações da imprensa diante do surgimento da Internet e, sob a mesma ótica que Wu, compreende que a consolidação do ciberespaço privilegiou o surgimento de outras empresas, que ocuparam espaço significativo ao qual as mídias tradicionais não conseguiram totalmente se ajustar.

O que aconteceu com o velho mundo? A Internet, preferências modificadas dos leitores e anunciantes, e o azar puro e simples conspiraram para causar grandes estragos na imprensa tradicional – tanto nos jornais quanto nas revistas – e na mídia de broadcast. Apesar das suas iniciativas às vezes resolutas, e às vezes cômicas, essa mídia tradicional simplesmente não se adaptou o suficiente ao novo mundo comandado pelo Google, Facebook, YouTube e iPhone. (...) Estavam convencidos de que como o computador pessoal se tornará outra maneira de acessar as notícias “deles”, eles seriam o centro do novo veículo. (DOCTOR, 2011, p. 110)

---

<sup>3</sup> No Dicionário Oxford (BIG DATA, 2021), o termo significa “conjuntos de informações que são grandes ou complexos demais para serem manipulados, analisados ou usados com métodos padrão”.

É neste contexto que empresas de tecnologia se desenvolvem e assumem o controle de fluxos da informação na rede, sendo muito mais ágeis para realizar o gerenciamento - desde a coleta até o processamento e tratamento - de *big data* por meio de sofisticados sistemas computacionais. Ou seja, empresas de tecnologia como Google e Facebook apropriaram-se do ciberespaço e, a partir do volume de dados que elas administram, compreendidos como o novo petróleo (THE WORLD’S..., 2017), refletem na esfera pública global a sua capacidade de influência.

Estes processos de gerenciamento de informações têm influência sobre uma vasta gama de serviços e produtos oferecidos na Internet, inclusive naqueles tradicionalmente associados às empresas jornalísticas. Se estas continuam sendo detentoras do conteúdo que produzem para os seus sites, não são as únicas a controlarem o fluxo dele no vasto espaço virtual. As organizações noticiosas precisam lidar, cada vez mais, com algoritmos de terceiros que indexam seus conteúdos, como os agregadores de notícias, e criam novos espaços de troca de informações, a exemplo das redes sociais online, através das quais se dá em grande parte a circulação contemporânea de informações.

Barsotti (2019) reforça esse aspecto ao lembrar que o conteúdo jornalístico no ambiente digital continua sob influência dos jornalistas em razão da seleção de notícias que são publicadas nos sites e compartilhadas nas plataformas. Contudo, “[...] nas redes sociais, os regimes de visibilidade das notícias são norteados, em última instância, pelos algoritmos, e não pelos critérios jornalísticos” (BARSOTTI, 2019, p. 46).

### **2.2.1 Cultura algorítmica e *gatekeeper* algorítmico**

Ainda que continuemos a ter os *gatekeepers* convencionais – diretores, editores, chefes de reportagem, repórteres – ocupados, nas redações, com a seleção de notícias, o crescimento da Internet sugere que os guardiões humanos têm, cada vez mais, dividido espaço com os *gatekeepers* automatizados, chamados de algoritmos. São estes conjuntos de códigos de programação os responsáveis pela coleta, seleção, ordenação e exibição de conteúdos considerados relevantes aos internautas nas suas interações com os mais diferentes equipamentos digitais conectados à grande rede.

Convém acrescentar que algoritmo, em sua definição mais geral, diz respeito a um conjunto de etapas estruturadas para processar instruções ou dados, considerando um certo problema, para, assim, produzir uma saída (resultado) desejada (GILLESPIE, 2014, p. 167). Além disso, cabe compreender que os algoritmos possuem múltiplas aplicações no mundo

contemporâneo. Sandvig et al. (2014, p. 3) ressaltam que sistemas tecnológicos têm sido transformados em “interações inteligentes”, com grande poder de computação incorporado.

No núcleo desses sistemas estão os algoritmos que proporcionam funções como classificação social, segmentação de mercado, personalização, recomendações e gerenciamento de fluxos de tráfego para carros. Transformar essas estruturas, mais recentemente, em computacionais levou-as a se tornarem muito mais poderosas, mas também muito mais opacas para o escrutínio público e a compreensão. A história da triagem e da discriminação em uma variedade de contextos levaria a acreditar que o escrutínio público dessa transformação é crítico. (SANDVIG *et al.*, 2014, p. 3, tradução nossa)<sup>4</sup>

A capacidade de abrangência e o impacto dos algoritmos têm sido ainda mais percebidos desde a invenção do computador e, principalmente, nas duas últimas décadas, com o *big data* que resulta da produção de dados por pessoas e máquinas conectadas à internet. Os computadores (e outros diferentes dispositivos que vêm compondo a Internet das Coisas) podem ser considerados “máquinas algorítmicas” (GILLESPIE, 2014, p. 167), capazes de executar tarefas extensas e complexas praticamente impossíveis de serem feitas manualmente ou por dispositivos analógicos, considerada a quantidade de dados que são capazes de processar.

Gillespie (2014) aponta ainda que estes processos computacionais podem ser classificados como algoritmos de relevância pública. O autor argumenta que eles se apresentam a partir dos mesmos procedimentos matemáticos, mas seus processos são realizados a partir de uma avaliação da informação, que, por sua vez, representa uma lógica de conhecimento construída com base em presunções específicas sobre o que é conhecimento e como identificar seus componentes mais relevantes. Para o autor, há uma preocupação sobre como estamos utilizando os algoritmos para definir o que precisamos saber, da mesma forma como confiamos em especialistas, no método científico ou no senso comum.

Negroponte (1995) já havia identificado a possibilidade do surgimento de “mordomos digitais”, espécies de “agentes de interface”, frutos do interesse do ser humano em criar funcionalidades e delegar tarefas que, poucos anos depois, realizadas por processos computacionais, poderiam acontecer em larga escala.

Eu não tenho o menor interesse em acessar um sistema e ficar examinando protocolos para descobrir seu endereço na Internet. Eu só quero mandar minha

---

<sup>4</sup> No original: “At the core of these systems sit algorithms that provide functions like social sorting, market segmentation, personalization, recommendations, and the management of traffic flows from bits to cars. Making these infrastructures newly computational has made them much more powerful, but also much more opaque to public scrutiny and understanding. The history of sorting and discrimination across a variety of contexts would lead one to believe that public scrutiny of this transformation is critical.”

mensagem para você. Do mesmo modo, não quero ter de ler milhares de boletins a fim de me certificar de que não estou perdendo alguma coisa. Quero é que meu agente de interface o faça por mim. (NEGROPONTE, 1995, p. 133)

Vinte anos depois, Striphas introduziu o conceito de “cultura algorítmica”, resultado desta situação em que “os seres humanos têm delegado o trabalho da cultura - a ordenação, classificação e hierarquização de pessoas, lugares, objetos e ideias - para processos computacionais com o uso intensivo de dados” (STRIPHAS, 2015, p. 398, tradução nossa).<sup>5</sup>

A utilização destes processos computacionais, por consequência, tem moldado o que nós, enquanto indivíduos e usuários de plataformas e redes sociais, visualizamos nos catálogos, *feeds* de notícias, linhas do tempo e resultados de busca de plataformas como Amazon, Netflix, Facebook, Twitter e Google.

É também sob essa perspectiva de influência das plataformas e seus algoritmos sobre a nossa forma de visualização da informação que Vaidhyanathan (2011) esclarece o impacto do Google no âmbito da esfera pública global. Segundo o pesquisador, os serviços oferecidos pelo Google, controlada pela *holding* Alphabet Inc., quase se confundem com a própria concepção de Internet, exercendo um poder sobre a conduta humana ao que ele dá o nome de “Googelização”.

O Google põe ao nosso alcance recursos até pouco tempo inimagináveis – bibliotecas imensas, arquivos, um enorme arsenal de documentos governamentais, uma imensidão de produtos e mercadorias, a vasta movimentação de boa parte da humanidade. É isso que entendo por googelização de “tudo”. A googelização atinge três grandes áreas de interesse e conduta humanas: “nós” (através dos efeitos do Google sobre nossas informações pessoais, nossos hábitos, opiniões e juízos de valor); “o mundo” (através da globalização de um estranho tipo de vigilância e daquilo que chamo de imperialismo infraestrutural); e o “conhecimento” (através de seus efeitos sobre o uso de um vastíssimo agregado de conhecimentos acumulados em livros, bases de dado on-line e na Internet). [...] Ao catalogar nossos juízos individuais e coletivos, nossas opiniões e (ainda mais importante) nossos desejos, a empresa também vai se transformando numa das mais importantes instituições globais. (VAIDHYANATHAN, 2011, p. 16)

Convém destacar que o Google foi desenvolvido em um cenário de expansão da *World Wide Web* propício à criação de produtos que fossem capazes de organizar as imensas quantidades de informação virtual para os usuários. Com novos sites e páginas publicados na grande rede diariamente, foi transferido a *crawlers* e metabuscadores computacionais o trabalho que antes estava nas mãos de profissionais especializados em coletar, recuperar e indexar informações, como bibliotecários, arquivistas, cientistas da informação, jornalistas, entre outros. A proposta inicial do cientista da computação Lawrence (Larry) Page (1997),

<sup>5</sup> No original: “[...] many ways human beings have been delegating the work of culture – the sorting, classifying and hierarchizing of people, places, objects and ideas – to data-intensive computational processes”.

apresentada na publicação *PageRank: Bringing Order to the Web*, durante o doutorado na Universidade de Stanford, na Califórnia, era, diante da já então grande quantidade de conteúdos disponíveis no ambiente online, facilitar o acesso a eles por parte dos usuários, usando o seu sistema de busca para indexar e classificar as páginas da internet por meio de sua estrutura de links. Para alcançar tal objetivo, Page propôs o uso do algoritmo *PageRank*, que determinava a relevância de cada site a ser exibido nos resultados de uma busca ao contabilizar a quantidade de links externos apontados para ele.

O detalhamento do funcionamento do algoritmo foi publicado dois anos depois, em um relatório de pesquisa assinado, além de Page, por Rajeev Motwani, Terry Winograd e pelo sócio-fundador da Google, Sergey Brin (1998). No texto, os autores apresentam a lógica por trás do *PageRank*, que, através de seus cálculos, eleva um *site* aos lugares mais altos do ranking quanto mais outras páginas estivessem apontadas para ele; ou, também, se algumas das páginas que apontavam para o *site* tivessem uma credibilidade - ou um *PageRank* - alta.

A missão do Google reforça que a empresa iniciou seus serviços – sendo o primeiro deles o Google Search, sua ferramenta de busca, lançada em 1997 – com o objetivo de organizar todas as informações que chegavam à Internet. De acordo com o site da companhia, sua missão é “organizar as informações do mundo para que sejam universalmente acessíveis e úteis para todos” (NOSSA..., [s. d.]).

Para Vaidhyathan (2011), porém, apesar da missão do Google expressar desejo de tornar universal o conhecimento, seus serviços e resultados de busca reforçam a fragmentação do conhecimento. O pesquisador compreende que a empresa pode até estar organizando a informação mundial e tornando-a acessível, mas não realiza o mesmo sobre o “conhecimento do mundo”.

Com o tempo, à medida que usuários de um conjunto tão diversificado de países treinam os algoritmos do Google para responderem a pesquisas especializadas com resultados locais, cada lugar do planeta terá uma lista diferente do que é importante, verdadeiro ou “relevante” na resposta a qualquer busca. [...] O Google conhece a localização geral do buscador e estrutura os resultados de modo a refletirem os hábitos revelados por outras pessoas da mesma localidade. (VAIDHYANATHAN, 2011, p. 154).

Não por acaso a crítica realizada pelo autor tem relação com o modo como o Google classifica sites e informações e os oferta aos usuários em seus resultados de buscas. Pouco depois de o Google realizar sua oferta pública inicial de ações (IPO, na sigla em inglês) em 19 de agosto de 2004, a empresa do Vale do Silício lançou, no dia 28 de junho de 2005, a Pesquisa Personalizada (KAMVAR, 2005) a todos os usuários conectados a um e-mail Google

com histórico da web ativado em suas contas<sup>6</sup>.

Em termos simples, a introdução da personalização trouxe uma mudança significativa ao conteúdo recomendado: quando um usuário realiza uma pesquisa no Google, os resultados deixam de se basear apenas no algoritmo do *PageRank* – que estimava a relevância das páginas para aquele termo de pesquisa – e passam a ser influenciados também pelo histórico de pesquisas anteriores. O intuito, de acordo com a empresa, era fornecer melhores resultados ao usuário numa tentativa de contextualizar sua experiência.

Em 2009, a Pesquisa Personalizada foi expandida para todos os usuários, passando a incluir aqueles que não estavam vinculados a uma conta Google. Assim, tornou-se possível para a empresa personalizar resultados das pesquisas a todos os usuários, segundo ela, com base em 180 dias de atividade de pesquisa através de um *cookie* anônimo no navegador.

Por exemplo, como eu sempre procuro pelo termo [receitas] e frequentemente clico nos resultados de *epicurious.com*, o Google pode classificar *epicurious.com* em posição superior na página de resultados na próxima vez que eu procurar receitas. Outras vezes, quando procuro notícias sobre os times esportivos da Cornell University, procuro [big red]. Como frequentemente clico em *www.cornellbigred.com*, o Google pode me mostrar esse resultado primeiro, em vez da empresa de refrigerantes Big Red ou outros. (PERSONALIZED..., 2009).

Outros produtos e algoritmos foram anunciados pela Google desde que a empresa decidiu vender ações na bolsa de valores Nasdaq, e alguns deles têm sido observados e discutidos por empresas jornalísticas com bastante preocupação, como será exposto a seguir. Dentre as múltiplas capacidades dos processos computacionais em realizar tarefas para os indivíduos no século XXI, pretendemos nos ater neste trabalho aos processos que agem como agregadores de notícias, em razão da sua influência sobre o jornalismo contemporâneo, como é o caso do produto Google Notícias.

### **2.2.2 A vez dos agregadores: o case Google Notícias**

Na medida em que passamos a delegar aos algoritmos a classificação e seleção de informações, um novo modelo de negócio começou a ganhar espaço na Internet a partir dos anos 2000. Conforme explicam Águila-Obra, Padilla-Meléndez e Serarols-Tarrés (2007), a crescente aceitação e uso da Internet fizeram emergir novas atividades econômicas baseadas

---

<sup>6</sup> Cabe ressaltar que a implementação da personalização da busca na Web também foi estratégia de outros motores de busca na época, como o caso do Yahoo, que lançou no mesmo ano uma ferramenta de busca personalizada (YAHOO..., 2005). Contudo, o Google se tornou a empresa com maior valor de mercado no planeta naquela ocasião, o que justifica sua importância no mercado de sistemas de busca e sobre as inovações tecnológicas que surgiram na época (GOOGLE..., 2005).

na produção, distribuição e venda de informação.

De acordo com os autores, a cadeia de valor de diversas indústrias, como é o caso da de notícias, se tornou mais complexa e fragmentada. Isso porque os tradicionais atores de diferentes setores passaram a precisar de novos intermediários, chamados de “cybermediários”, para ampliar as possibilidades de fornecimento e distribuição de seus produtos e serviços. Uma das modalidades de atuação desses intermediários são os agregadores *web*, um fenômeno típico da internet capaz de coletar, organizar, avaliar e personalizar volumosos conjuntos de informações dispersas.

[...] um agregador da web *é uma entidade que pode coletar e analisar de forma transparente informações de várias fontes de dados da web.* [...] Assim, eles combinam conteúdo e aplicativos de várias fontes online cooperantes ou não [...] e geralmente sem acordos prévios. (ÁGUILA-OBRA; PADILLA-MELÉNDEZ; SERAROLS-TARRES, 2007, p. 189, grifo do autor, tradução nossa)<sup>7</sup>

A esses novos agentes automatizados foi dado o nome de infomediários, classificados como agregadores ou distribuidores de conteúdo. O que os diferencia, sobretudo, é o estabelecimento ou não de acordos com as fontes de dados. Os distribuidores estabelecem acordos legais com suas fontes, o que faz com que a integração de novas bases de dados seja menos rápida e mais cara. Já os agregadores fornecem principalmente tecnologia e uma plataforma de serviço (ÁGUILA-OBRA; PADILLA-MELÉNDEZ; SERAROLS-TARRES, 2007, p. 189).

Segundo Doctor (2011), enquanto a Internet foi capaz de eliminar intermediários, a exemplo das agências de publicidade presentes nos distribuidores de jornais e nos fabricantes de prensas, inevitavelmente, a própria web criou novos intermediários: os agregadores de notícias, como o Google Notícias, Apple Notícias, Digg, Feedly, Flipboard, Pocket, News360, Panda, Techmeme, Inoreader, entre outros.

Dados reunidos por Ken Doctor mostram que cerca de 75% dos 23 bilhões de dólares da Internet eram destinados inicialmente a quatro grandes infomediários: Google, Yahoo, MSN e AOL.

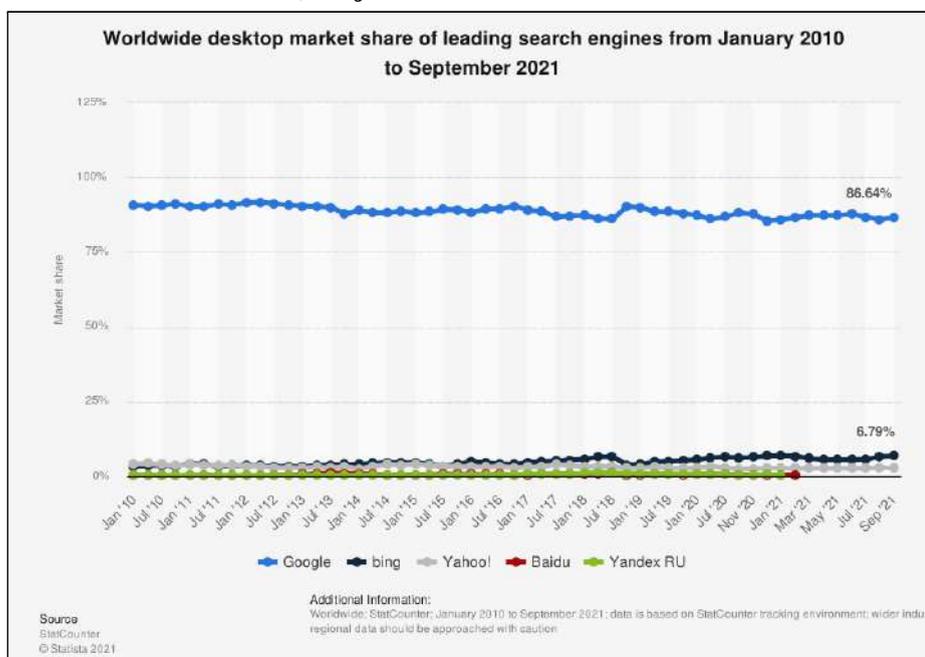
Essas quatro companhias dominam. O know-how delas: reunir em um só lugar uma grande quantidade de conteúdo de outras pessoas, esquadrinhando incessantemente sites de notícias e da rede. Elas produzem muito pouco e talvez nenhum conteúdo original. (DOCTOR, 2011, p. 147)

---

<sup>7</sup> No original: “Accordingly, a web aggregator is an entity that can transparently collect and analyse information from multiple web data sources (Madnick & Siegel, 2002, p. 36). Thus, they combine content and applications from multiple online cooperating or non-cooperating sources (Madnick & Siegel, 2002) and generally without prior agreements.”

Cabe destacar que, segundo informações divulgadas pelo *Statista*, empresa alemã especializada em dados de mercado, a participação do Google Pesquisa no mercado de mecanismos de busca global é predominante desde o seu surgimento, em 1997. De acordo com gráfico disponibilizado pela companhia, o Google detinha 92,47% do mercado de buscas na versão *desktop* em junho de 2021. Apesar do leve recuo, a empresa segue como líder mundial disparada, com 86,64% do mercado dos mecanismos de busca em setembro de 2021. Ela é seguida logo depois por Bing e Yahoo, respectivamente com 6,79% e 2,75%. (GLOBAL..., 2021)

**Gráfico 1 – Participação de mercado dos mecanismos de pesquisa em *desktops* em todo o mundo, de janeiro de 2010 a setembro de 2021**



Fonte: Statista (GLOBAL..., 2021)

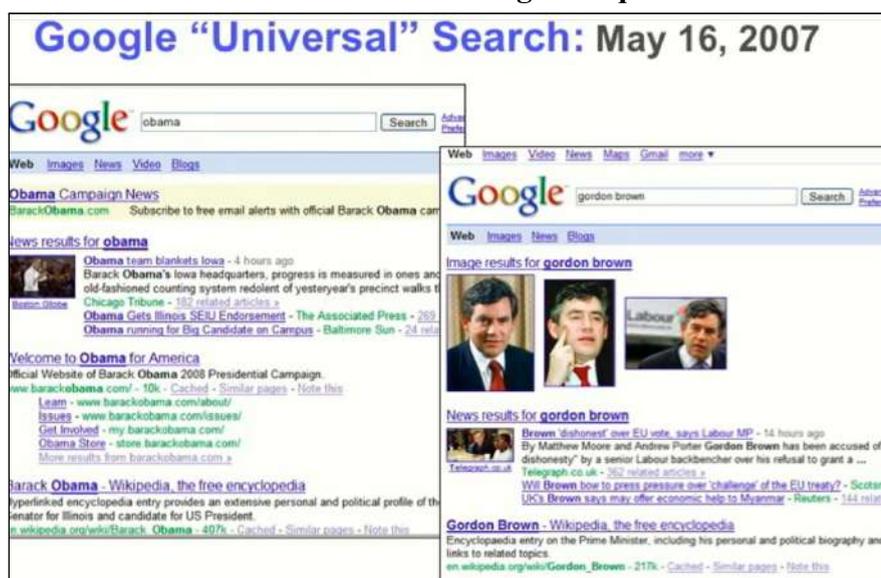
De acordo com informações do blog do Google (SHIELS, 2018), o agregador Google Notícias foi um dos primeiros produtos que a empresa lançou após a Pesquisa. O mecanismo foi disponibilizado aos usuários em 22 de setembro de 2002, dois anos após o Google Pesquisa ter versões em português, alemão, francês, italiano, sueco, finlandês, espanhol, holandês, norueguês e dinamarquês. Sua criação foi motivada pela experiência de Krishna Bharat, cientista indiano e ex-funcionário da Google Inc., ao ver que pessoas procuravam notícias sobre os ataques de 11 de setembro em Nova York, mas encontravam apenas informações antigas.

Na ocasião, o *Google News* foi lançado com 4.000 fontes de notícias, permitindo que usuários encontrassem conteúdos mais recentes sobre qualquer história que pesquisassem. Isso porque o algoritmo até então utilizado no Google Notícias era o mesmo do Google Pesquisa, o que não dava ênfase à lógica do jornalismo, pois priorizava o número e a qualidade dos links dirigidos às páginas de notícias e não levava em consideração o fato de uma notícia ser recente.

A fim de valorizar as *hard news* (notícias mais importantes e recentes), inicialmente Bharat decidiu agrupá-las em tópicos, quando diziam respeito a um mesmo conteúdo, para que este fosse adicionado de forma rápida à Pesquisa. O feito acabou gerando críticas de editores, já que um determinado furo (informação exclusiva divulgada inicialmente por apenas um veículo) poderia ser empurrado para o final dessa lista de notícias à medida que outros sites iam contando suas próprias histórias horas depois. De todo modo, este e outros processos ligados às notícias foram sendo refinados ao longo dos anos.

O Google manteve por 16 anos o Google Notícias como uma categoria localizada próxima ao Google Pesquisa, mesmo que ano a ano fossem acrescentadas atualizações ao seu modo de funcionamento, como a incorporação de notícias em resultados de pesquisa em maio de 2007 (Figura 1), a implementação de uma linha do tempo em abril de 2009, o *Google Fast Flip* em setembro de 2009 e outras.

**Figura 1 – Incorporação de *links* de notícias aos resultados das pesquisas dos termos “Obama” e “Gordon Brown” no Google Pesquisa em 2007**



Fonte: KRAZIT, 2010.

Em maio de 2018, o Google Notícias foi relançado sob a perspectiva de ser um produto “totalmente novo”. Duas mudanças foram cruciais para que fosse considerado diferente do que já havia: passou a ter um aplicativo próprio e começou a contar com um conjunto de técnicas de inteligência artificial para sintetizar informações e agrupá-las a fim de possibilitar ao leitor uma melhor forma de desencadeamento dos fatos. Depois da sexta patente da aplicação, concedida em outubro de 2019 (CURTISS; BHARAT; SCHMITT, 2019), a empresa anunciou que os algoritmos do Google Notícias usavam “[...] centenas de diferentes fatores para identificar e organizar os conteúdos que os jornalistas estão cobrindo, para tornar relevantes informações diversificadas e confiáveis” (GINGRAS, 2019)<sup>8</sup>.

Se pelo menos até 2011 o Google se orgulhava de dizer que o Google Notícias estava “livre da mácula do julgamento humano, com o algoritmo decidindo o que mostrar e em que ordem” (DOCTOR, 2011, p. 210), depois de uma série de críticas por parte de editores de redação e pesquisadores da comunicação, o agregador passou a contar com uma parte de curadoria humana. Ainda assim, foi a implementação da inteligência artificial que ganhou destaque durante a sua divulgação, já que esta é considerada capaz de realizar uma curadoria semelhante à de um editor humano.

Quando criamos o Google Notícias original, há 15 anos, simplesmente organizamos as notícias para facilitar a visualização de uma variedade de fontes sobre o mesmo tópico. O Google Notícias repensado usa um novo conjunto de técnicas de IA [Inteligência Artificial]. [...] Essa abordagem significa que o Google Notícias entende as pessoas, lugares e coisas envolvidas em uma história conforme ela se desenvolve e se conecta como elas se relacionam umas com as outras. (UPSTILL, 2018, tradução nossa)<sup>9</sup>

Tal atualização nos permite inferir que o Google Notícias delega boa parte do seu trabalho de curadoria aos processos computacionais, com uso de IA, enquanto uma parcela menor deste exercício é realizada por editores e equipes do aplicativo. Segundo o site de suporte do Google Notícias, são os “algoritmos de computador” que selecionam o que o usuário vê no Google Notícias, exceto em “alguns casos”, embora não especificados. “Os algoritmos determinam quais matérias, imagens e vídeos são exibidos, bem como a ordem de exposição” (COMO..., 2021). O próprio site do agregador afirma que alguns dos conteúdos são exibidos de modo personalizado, e atribui à personalização uma forma de ajudar o usuário

<sup>8</sup> No original: “The algorithms used for our news experiences analyze hundreds of different factors to identify and organize the stories journalists are covering, in order to elevate diverse, trustworthy information.”

<sup>9</sup> No original: “When we created the original Google News 15 years ago, we simply organized news articles to make it easier to see a range of sources on the same topic. The reimagined Google News uses a new set of AI techniques. [...] This approach means Google News understands the people, places and things involved in a story as it evolves, and connects how they relate to one another”.

a encontrar “conteúdo relevante de maneira rápida e fácil”.

O modelo de funcionamento do Google Notícias, criado em 2002, já havia sido idealizado por estudiosos do campo da ciência da computação em meados dos anos 1990, como já indicado anteriormente a partir das previsões de Negroponte (1995). O co-fundador e primeiro diretor do Media Lab, do Massachusetts Institute of Technology (MIT), projetava que a vida digital mudaria o modelo econômico da seleção de notícias. Os agentes de interface – como o autor classificou os processos computacionais – seriam capazes de coletar notícias e construir, diariamente, uma espécie de jornal personalizado para cada leitor, abrindo espaço para que os interesses de quem o lesse ganhassem mais destaque.

Imagine um futuro no qual seu agente de interface vai poder ler todos os jornais e captar todos os noticiários de TV e rádio do planeta, construindo a partir daí um sumário personalizado para você. Esse tipo de jornal terá uma tiragem de uma única cópia. [...] Tal edição misturaria manchetes de grande repercussão com matérias “menos importantes” sobre pessoas que você conhece, que você vai ver amanhã, e sobre lugares para os quais você está para ir ou dos quais acabou de voltar. Um jornal assim traria reportagens sobre companhias que você conhece. Na realidade, sob tais condições, você poderia até se dispor a pagar mais por dez do que por cem páginas do *Globe* de Boston, contanto que você confie em que o jornal esteja lhe entregando o subconjunto certo de informações. Desse jornal, você consumiria cada bit (ou literalmente, cada pedacinho). Pode chamá-lo de *Minha Gazeta*. (NEGROPONTE, 1995, p. 134)

A previsão de Negroponte pode ser considerada concretizada, já que o próprio Google Notícias passou a ser capaz, em alguma medida, de mesclar notícias de grande apelo com notícias de menor importância. No site de suporte do agregador, a empresa explica que as categorias “Manchetes” e “Para Você” representam seleções diferentes. Na primeira, os algoritmos escolhem os assuntos por idioma e região, o que significa que outros usuários que usam os mesmos dois critérios veem a cobertura dos mesmos assuntos. Já na segunda categoria, de notícias personalizadas, os algoritmos escolhem conteúdos que serão exibidos com base nas configurações do usuário no Google Notícias e de suas atividades em sua conta do usuário Google, como no mecanismo de busca e no YouTube.

A partir desta transformação no modo como as notícias são selecionadas para serem exibidas, Negroponte projetou que o processo de *gatekeeping* passaria por uma revisão. Neste novo modelo econômico da vida digital, como afirma o autor, a nova forma de selecionar as matérias faria com que o leitor não fosse mais obrigado a consumir somente aquilo que outros julgam ser notícia (NEGROPONTE, 1995, p. 153). Apesar de o autor ponderar que as empresas jornalísticas poderiam colocar à disposição o seu capital humano e tecnológico para entregar edições únicas sob a forma de bits a cada um de seus leitores, foram principalmente

as empresas de tecnologia que passaram a dominar este mecanismo para fazer a mediação entre a notícia e o leitor.

Como a experiência e engenhosidade da agregação de busca do Google eram tão superiores a qualquer coisa que a indústria de notícias poderia inventar, as pessoas haviam migrado e passado a utilizar o Google como um dos seus principais métodos de encontrar as notícias. (DOCTOR, 2011, p. 152)

O próprio Google se utiliza dessa condição de ser uma empresa de tecnologia, que possui a capacidade de criação e gerência de uma série de algoritmos, para organizar não somente as “informações do mundo”, como se propõe em sua missão, mas para utilizar tais processos de curadoria algorítmica em prol da organização de todo o conteúdo jornalístico publicado por uma série de meios de comunicação ao redor do globo.

Com tantas notícias disponíveis de tantas editoras, pode ser um desafio entender tudo isso. As experiências de notícias do Google visam resolver isso usando tecnologia para ajudá-lo a descobrir notícias relevantes e confiáveis de uma variedade de fontes em toda a web. (OUR APPROACH, [s. d.]

Segundo o site do projeto Google News Initiative, criado em 2018<sup>10</sup>, a empresa reconhece a atividade jornalística como vital para as sociedades, mas destaca o seu papel de utilizar sua tecnologia para otimizar a série de notícias que chegam aos usuários diariamente.

### **2.2.3 Personalização e diversidade no Google Notícias: estamos reféns das bolhas noticiosas?**

Ao passo em que o uso de algoritmos para organizar e ofertar informações foi se popularizando nas últimas duas décadas, discussões em torno das lógicas destes processos computacionais privados passaram a ser suscitadas. Além dos infomediários (ou *gatekeepers* algorítmicos) trazerem novos dilemas para as empresas jornalísticas, pesquisadores do campo da comunicação começaram a questionar o modo de funcionamento desses sistemas automatizados voltados para a recomendação de conteúdo jornalístico.

O pesquisador e ativista Eli Pariser (2012) acredita que os algoritmos podem contribuir para a redução da diversidade de pontos de vista no campo da internet. Para o autor, o fato de empresas como Google e Facebook utilizarem algoritmos para conhecer as preferências dos usuários e oferecer conteúdo personalizado pode colocar os internautas em “bolhas informacionais”, fazendo com que seja exibido somente aquilo que estes desejam

---

<sup>10</sup> Segundo o blog do Google, a iniciativa Google News Initiative foi apresentada ao público em março de 2018. <https://blog.google/outreach-initiatives/google-news-initiative/announcing-google-news-initiative/>

visualizar, deixando de fora informações que podem ser mais importantes a longo prazo. É o fenômeno que o autor chama de “bolha dos filtros”.

Partindo dessa teoria, a bolha informacional ocorre a partir dos próprios rastros digitais que usuários deixam na rede. Uma simples escolha de clicar ou não em um link já pode ser um dado relevante para compor as bases de dados destas mega empresas, levando-as, a partir do trabalho realizado pelos seus algoritmos, a compreenderem se um usuário tem interesse ou não em acompanhar determinado assunto. E, como os algoritmos são programados para exibir mais daquilo que um perfil clica ou com o qual interage, é possível que isso crie um ciclo em que a diversidade é ainda mais reduzida. Segundo Pariser (2012, p. 14), essas ferramentas “criam e refinam constantemente uma teoria sobre quem somos e sobre o que vamos fazer ou desejar seguir”.

No campo do jornalismo, a consequência de um usuário estar inserido em um filtro-bolha pode ser ainda mais preocupante, já que o autor argumenta que as notícias “moldam a consolidação da nossa visão de mundo, do que é importante, da escala, tipo e caráter dos problemas que enfrentamos” (PARISER, 2012, p. 49). O ativista da internet avalia que essa curadoria algorítmica também pode atuar como uma lente de aumento, sendo útil quando um indivíduo deseja receber mais informação sobre uma área específica do conhecimento. Afinal, há uma imensa quantidade de notícias disponíveis na Internet e a capacidade de um indivíduo acompanhar toda a sua distribuição tornou-se humanamente impossível. No entanto, há uma tendência de os filtros personalizados limitarem o nosso campo de visão.

Eli Pariser também pondera que os algoritmos não são tão diferentes de um editor de jornal, ao qual está associado um poder de escolha, seleção e hierarquização. Entretanto, o autor defende que os novos *gatekeepers* não têm necessariamente um senso ético bem definido como o do editor humano. Na prática, isso pode significar que estas tecnologias priorizem o ganho de receita ao exibir em seus resultados de buscas principalmente matérias que proporcionem mais acessos, sem considerar questões éticas e funções cívicas que as empresas de comunicação, em princípio, tenderiam a seguir. Por consequência, muitos podem ser os desdobramentos para o usuário, como no caso de matérias sobre um conflito acabarem não sendo ofertadas em razão do algoritmo ter “entendido” que aquele assunto não era de seu interesse por não ter sido pré-selecionado pelo seu universo amostral.

A preocupação com a possibilidade de usuários ficarem submetidos a um filtro-bolha se tornou também uma preocupação do Google. A atualização do Google Notícias em maio de

2018, que, pela primeira vez, ofereceu uma versão na forma de aplicativo, disponibilizou um recurso adicional que gera um layout diferente do exibido no *desktop*. No *app*, ou seja, em sua versão *mobile*, o Google Notícias exibe na categoria “Para Você” (Figura 2) um resumo personalizado com cinco matérias que mesclam as manchetes mais importantes, notícias locais e matérias recentes personalizadas com assuntos de interesse pessoal do leitor.

Segundo Vidhya Bhat, funcionária do Google e criadora do recurso de *Full Coverage*<sup>11</sup>, a empresa não queria que as pessoas ficassem presas em um filtro-bolha.

Embora partes do novo aplicativo Google Notícias sejam personalizadas, não queríamos que as pessoas ficassem presas em um filtro-bolha. Queríamos ajudar as pessoas a verem perspectivas diversas e também tornar agradável o contexto de uma história. [...] Para combater os silos de informação, o resumo diário inclui as grandes histórias que todos veem junto com notícias que são pessoalmente importantes para você. A guia Manchetes também oferece uma visão totalmente não personalizada do que está acontecendo no mundo. A cada passo do caminho, tornamos mais fácil tirá-lo da sua zona noticiosa de conforto. (SHIELS, 2018, tradução nossa).<sup>12</sup>

Uma pesquisa no aplicativo do Google Notícias nos permite verificar que esse mecanismo “Resumo de notícias”, criado para combater os filtros-bolha e localizado na aba “Para Você” – contendo cinco notícias classificadas como Manchetes (Figura 2), personalizadas ou notícias locais –, segue em funcionamento. Contudo, apesar da porta-voz do Google demonstrar que a empresa tem se preocupado, publicamente desde 2018, em combater os filtros-bolha, o recurso “Resumo de notícias” – que de acordo com a empresa colabora para a supressão do fenômeno explicado por Eli Pariser – ainda não está disponível na versão *desktop* para os usuários brasileiros, sendo esse o nosso objeto de pesquisa.

---

<sup>11</sup> Segundo o blog do Google, o recurso “Cobertura completa” no Google Notícias fornece uma visão completa de como uma história é relatada a partir de uma variedade de fontes. Ao visualizar a cobertura completa, o usuário verá as principais manchetes de diferentes fontes, vídeos, notícias locais, perguntas frequentes, comentários sociais e uma linha do tempo para histórias que foram reproduzidas ao longo do tempo. O recurso não é personalizado e todos os usuários veem o mesmo conteúdo.

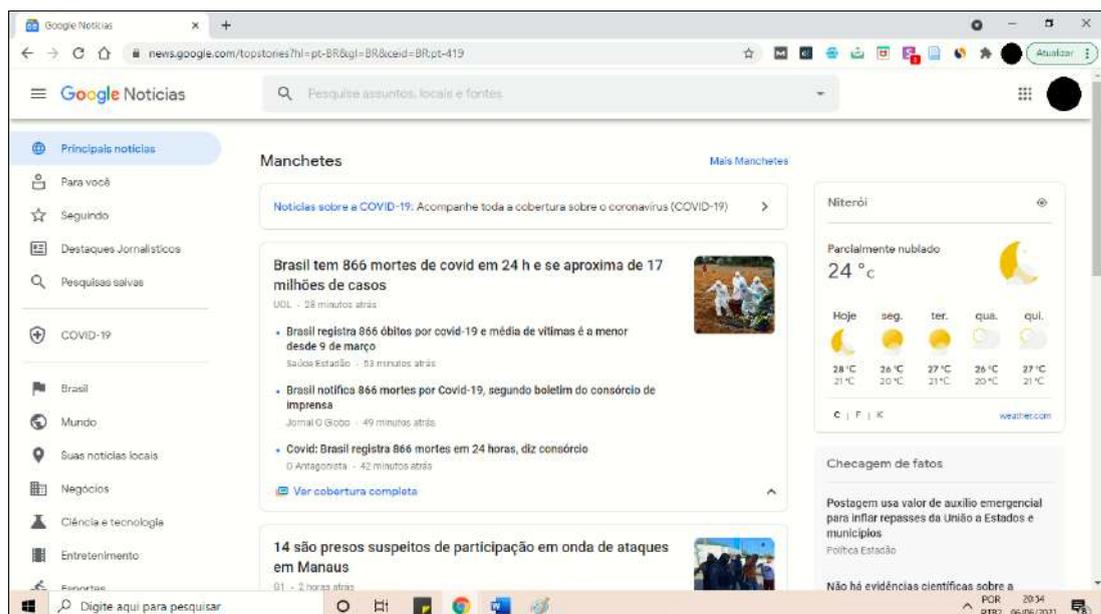
<sup>12</sup> No original: “Although parts of the new Google News app is personalized, we didn’t want people to get stuck in a filter bubble. We wanted to help people see diverse perspectives and also make it enjoyable to get context around a story. [...] To combat information silos, the daily briefing includes the big stories that everyone sees, alongside news that is personally important to you. The Headlines tab also gives you a completely unpersonalized view of what’s going on in the world. Every step of the way we make it easy to get you out of your news comfort zone.”

**Figura 2 – Captura de tela da aba “Para Você” exibida ao usuário no Google Notícias em versão *mobile***



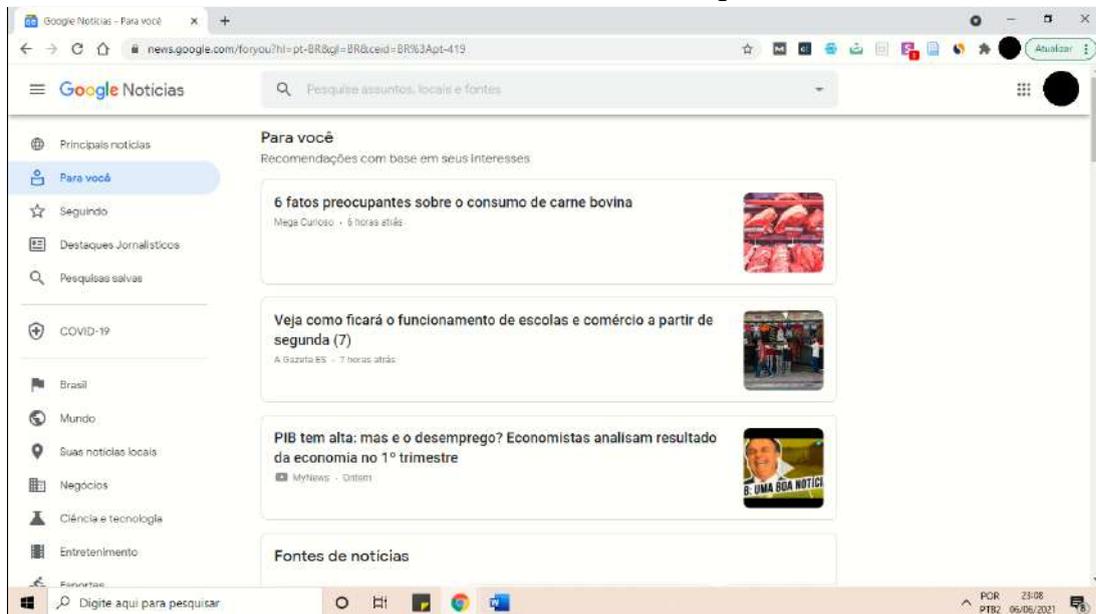
Fonte: Google Notícias. Acesso em: 6 jun. 2021, às 20h54.

**Figura 3 – Captura de tela da aba “Manchetes” exibida ao usuário no Google Notícias em versão *desktop***



Fonte: Google Notícias. Acesso em: 6 jun. 2021, às 20h54.

**Figura 4 – Captura de tela da aba “Para Você” exibida ao usuário no Google Notícias em versão *desktop***



Fonte: Google Notícias. Acesso em: 6 jun. 2021, às 23h09.

As capturas de tela na versão *desktop* apontam que, ao acessar o site do Google Notícias, o usuário é levado a visualizar primeiro a categoria Manchetes (Figura 3), diferentemente do que acontece na versão *mobile*, que prioriza o contato do usuário com a categoria personalizada em um primeiro momento de uso do aplicativo. No *desktop*, porém, apesar da empresa ofertar a categoria Manchetes, a categoria “Para Você” (Figura 4) não possui o recurso “Resumo das notícias” como acontece na versão *mobile*.

Outro aspecto que pesquisadores consideram relevante para a análise de um agregador de notícias é a sua diversidade. Haim, Graefe e Brosius (2018) lembram que os meios de comunicação de massa devem ser guiados pela diversidade de notícias para permitir que os cidadãos estejam bem informados, e isso inclui a diversidade de conteúdo, fontes e pontos de vista.

Sunstein (2009) também argumenta que a exposição a diferentes pontos de vista é vital para a democracia, e sua ausência gera uma falta de diversidade que é prejudicial ao debate político. Além de sinalizar os riscos aos quais as democracias podem estar submetidas quando as pessoas não possuem contato com materiais que não sejam do seu interesse imediato, o autor aponta que os sistemas de comunicação têm caminhado para uma personalização completa e lembra que a profecia de Nicholas Negroponte com o *Daily Me* não foi ambiciosa o suficiente.

Isso porque, em vez do próprio usuário ou jornal terem a capacidade de criar uma

edição jornalística única a partir de um sistema próprio, a personalização tem sido feita por outras empresas e seus respectivos produtos – como o Google Notícias, ao qual nos atemos neste trabalho. Por isso, o que acontece na prática é que essa seleção não é somente personalizada, mas previamente automatizada e realizada por outros, sem que os indivíduos ou empresas jornalísticas tenham amplo controle sobre o processo.

Acontece que você não precisa criar um Daily Me. Outros podem criá-lo para você. Se as pessoas souberem um pouco sobre você, elas podem descobrir e dizer do que “pessoas como você” tendem a gostar - e podem criar um Daily Me, só para você, em questão de segundos. (SUNSTEIN, 2009, p. 4, tradução nossa)<sup>13</sup>

A fim de oferecer maior transparência em meio a questionamentos sobre desinformação e propagação de notícias falsas (IN RACE..., 2017), o Google passou a indicar, em parte, como funcionam os seus sistemas de classificação para conteúdo de notícias no Google Notícias e no YouTube Notícias. De acordo com o blog do Google News Initiative, o agregador Google Notícias usa a mesma tecnologia de rastreamento e indexação da Pesquisa do Google para identificar e organizar artigos de notícias, observando uma série de fatores para classificação (ORGANIZAR..., [s. d.]).

Entre estes estão a relevância, interesses, localização, proeminência, autoridade, frescor e usabilidade (EXIBIR..., [s. d.]). No âmbito da relevância, a empresa aponta que um indicador básico para o aparecimento de uma notícia é quando um artigo contém as mesmas palavras-chave da consulta. No entanto, há outras formas mais avançadas para determinar a relevância do conteúdo pelos algoritmos que a empresa não detalha.

Já quanto aos interesses, o Google diz que seu agregador exhibe notícias de acordo com os gostos indicados ou inferidos com base na “interação com os produtos do Google, dependendo das configurações da conta”. A localização também é um fator explicitamente tido como importante, e a plataforma diz que utiliza o local onde o usuário está para encontrar conteúdo relevante sobre a região, como o recurso “Matérias locais” no Google Notícias.

Outro fator destacado pela empresa é a proeminência, que considera, entre outros aspectos, se os veículos estão publicando muitos artigos sobre uma determinada notícia e exibindo esse conteúdo em destaque nos próprios sites, além do quanto uma matéria ou artigo está em alta. O critério de Legitimidade é adotado pelos algoritmos a fim de que o agregador recomende informações de fontes confiáveis. Para isso, os sistemas identificam por meio de

---

<sup>13</sup> No original: “As it turns out, you don’t need to create a Daily Me. Others can create it for you. If people know a little bit about you, they can discover, and tell you, what “people like you” tend to like—and they can create a Daily Me, just for you, in a matter of seconds.

“sinais”, não detalhados, que apontam para legitimidade e confiabilidade nas páginas sobre um determinado assunto, “com base no feedback dos avaliadores da Pesquisa”, embora não fique claro quais são esses avaliadores. Ainda sob esse fator, o Google diz que valida fontes a partir da opinião de pessoas sobre determinadas fontes para consultas próximas ou a presença de links que são dirigidos ao conteúdo por meio de outros sites reconhecidos - algo que nos remete ao modo de funcionamento do algoritmo do PageRank, embora não fique claro se a lógica desse processo é a mesma.

O critério de Atualização leva em consideração a data de publicação do artigo e a importância da atualização do conteúdo para o assunto em questão. Por fim, a usabilidade leva em conta aspectos como a facilidade de visualização do conteúdo de um site e tempo de carregamento da página. Segundo o Google, cada um dos fatores “é pesado ao lado dos demais e não seria o único motivo para a exibição de um artigo”, o que reforça o grau de opacidade que os algoritmos possuem apesar de suas descrições.

De todo modo, a empresa enfatiza que seus sistemas não classificam o conteúdo com base em aspectos políticos ou ideológicos e que não tentam inferir a opinião dos leitores ou do material classificado. No âmbito das notícias personalizadas, a empresa diz que sua proposta é ajudar o leitor a manter-se informado sobre os assuntos que são importantes para ele, o que inclui “seus interesses e a comunidade local”.

Embora a Google não detalhe nestes termos, Thurman e Schifferes (2012) explicam que as personalizações para recomendação de conteúdo podem ser explícitas – com base em informações que o usuário decidiu fornecer voluntariamente para colaborar com a classificação das notícias que gostaria de receber – e implícitas – com base em informações capturadas a partir do comportamento on-line deste usuário. É comum que os agregadores de notícias combinem a personalização explícita e implícita, como no Google Notícias.

Desse modo, nota-se que o jornalismo digital tem sido permeado por novas formas de seleção e distribuição de conteúdo quando analisamos os agregadores de notícias que distribuem os conteúdos produzidos pelos veículos - e que estes, como as próprias empresas já reconhecem, possuem poder sobre a forma como as matérias jornalísticas são classificadas e recomendadas aos leitores. É sobre essa forma de seleção de conteúdo que iremos tratar no presente trabalho, com foco na personalização implícita do agregador Google Notícias.

### **3 DESAFIOS METODOLÓGICOS AO ESTUDO DOS ALGORITMOS**

A fim de investigar o funcionamento da personalização implícita do Google Notícias e as possíveis formações de filtros-bolha como consequência do uso da ferramenta “Para Você” em seu agregador de notícias, analisamos como o conteúdo personalizado se comporta diante de diferentes perfis de usuários. Dessa forma, é possível observar as diferenças entre os supostos interesses dos perfis reconhecidos pela plataforma, identificar semelhanças e diferenças entre os conteúdos exibidos aos usuários e verificar o que parece determinar tanto o conteúdo quanto a fonte. Sendo assim, veremos quais os efeitos da personalização implícita na diversidade de notícias sob a perspectiva das editorias e veículos exibidos.

Mas como analisar a ação de algoritmos computacionais, ainda mais quando não se tem acesso ao seu código - e, mesmo tendo, o que fazer se sua programação se transformar em um empecilho para pesquisadores não versados em linguagens de programação?

Para tanto, o método escolhido para a realização deste trabalho é a engenharia reversa. Nesse sentido, o presente capítulo busca esclarecer no que consiste a metodologia e por que esse método é importante para pesquisas nas áreas das ciências humanas e sociais. Em seguida, serão detalhados os processos de coleta, classificação e visualização dos dados obtidos para a pesquisa.

#### **3.1 Engenharia reversa**

A escolha da metodologia se deu a partir da tentativa de aplicação da engenharia reversa, proposta pelo pesquisador Rob Kitchin (2017). A investigação de algoritmos em pesquisas científicas é um terreno relativamente recém-explorado, onde as metodologias não estão ainda amplamente consolidadas, sobretudo no que se refere ao campo das ciências humanas e sociais e suas produções de análises sobre o funcionamento de sistemas algorítmicos. Por isso, antes de explicarmos o que é a engenharia reversa, devemos nos ater ao contexto que leva à aplicação dessa metodologia na presente área do conhecimento.

De acordo com Kitchin (2017), existem três desafios principais que dificultam a pesquisa sobre algoritmos. Um deles é o ganho de acesso a sua formulação, uma vez que muitos dos algoritmos que são utilizados no cotidiano das pessoas têm sua codificação realizada em ambientes privados e é preservada em arquivos de difícil acesso por terceiros. Como Pasquale (2015) observa, os algoritmos são mantidos em verdadeiras “black boxes” (caixas-pretas), pois quase nada se sabe sobre seu funcionamento e apenas “podemos observar

entradas e saídas, mas não podemos dizer como um se torna o outro” (PASQUALE, 2015, p. 3).

E, mesmo quando pesquisadores conseguem acesso ao código-fonte de um algoritmo, raramente sua codificação é simples de ser desconstruída. Segundo Kitchin (2017, p. 20, tradução nossa)<sup>14</sup>, este seria o segundo principal desafio, já que os algoritmos “estão inseridos em complexos conjuntos sociotécnicos compostos de um conjunto heterogêneo de relações, incluindo potencialmente milhares de indivíduos, conjuntos de dados, objetos, aparelhos, elementos, protocolos, padrões, leis, etc. que enquadram seu desenvolvimento”.

Além disso, o autor destaca que, ao estudar algoritmos, é preciso compreender que estes processos computacionais requerem atenção prática e epistemológica, o que traz uma complexidade ainda maior ao pesquisar este fenômeno. Isso porque os algoritmos estão em constante desenvolvimento, o que significa que seu código costuma ser “frequentemente refinado, retrabalhado, estendido e corrigido, iterando em várias versões” (MIYAZAKI, 2012) (KITCHIN, 2017, p. 21, tradução nossa)<sup>15</sup>.

Outro fator que deve ser considerado é que um algoritmo pode ser programado para ser reativo e mutável às entradas (*inputs*). Isso significa que ele não age de maneira estática sobre um conjunto de usuários em uma plataforma, por exemplo. Um mesmo algoritmo pode produzir saídas (*outputs*) diferentes a depender de como cada usuário, único em sua forma de utilizar uma plataforma, interage no espaço virtual em determinado contexto.

Os algoritmos, então, estão frequentemente “fora de controle” no sentido de que seus resultados às vezes não são facilmente antecipados, produzindo efeitos inesperados em termos de seu trabalho no mundo [...]. O que isso significa na prática é que engajamentos únicos ou limitados com algoritmos não podem ser simplesmente extrapolados para todos os casos e que um conjunto de estudos de caso comparativos precisam ser empregados, ou uma série de experimentos realizados com o mesmo algoritmo operando sob diferentes condições. (KITCHIN, 2017, p. 4, tradução nossa)<sup>16</sup>

Considerados os desafios acima, Kitchin expõe seis abordagens que podem ser aplicadas para analisar o fenômeno do funcionamento de um sistema algorítmico. Uma delas

<sup>14</sup> No original: “Moreover, they are embedded within complex socio-technical assemblages made up of a heterogeneous set of relations including potentially thousands of individuals, data sets, objects, apparatus, elements, protocols, standards, laws, etc. that frame their development.”

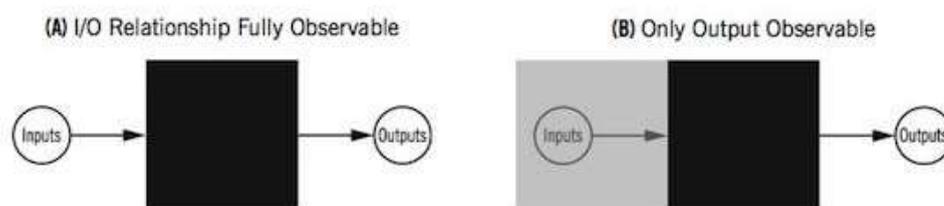
<sup>15</sup> No original: “In other cases, algorithms and their instantiation in code are often being refined, reworked, extended and patched, iterating through various versions”.

<sup>16</sup> No original: “As it turns out, you don’t need to create a Daily Me. Others can create it for you. If people know a little bit about you, they can discover, and tell you, what “people like you” tend to like—and they can create a Daily Me, just for you, in a matter of seconds.”

é a engenharia reversa<sup>17</sup>, que frequentemente é utilizada em casos em que o código-fonte permanece inacessível. Neste método, o pesquisador busca fazer o caminho inverso, analisando quais dados são alimentados e qual saída é gerada, a fim de identificar quais critérios um algoritmo tende a elencar e ponderar para o seu funcionamento.

Este é também o método estudado por Nicholas Diakopoulos (2015), que determina duas possibilidades para o uso da engenharia reversa. Conforme ilustrado na Figura 5A, há casos em que é possível realizar a engenharia reversa quando se tem a capacidade de observação ampla de todas as entradas (*inputs*) e saídas (*outputs*) de um algoritmo. Segundo o autor, isso ocorre quando os algoritmos estão acessíveis por meio de uma API<sup>18</sup> online, por exemplo. Este processo também é comum em casos em que um algoritmo possui recursos como os de autocompletar ou autocomplemento (em inglês, *autocomplete* ou *code completion*), de autocorreção ou discriminação de preços.

**Figuras 5A e 5B – Representação de dois cenários de algoritmo inacessíveis, com diferentes níveis de observação (total ou apenas a saída)**



**Figure 1. Two black box scenarios with varying levels of observability.**

Fonte: DIAKOPOULOS, 2015

Por outro lado, Diakopoulos também avalia que é possível realizar a engenharia reversa mesmo quando apenas as saídas (*outputs*) do algoritmo são visíveis, como ilustrado na Figura 5B. “Este é um caso comum que jornalistas de dados encontram: um grande conjunto de dados de saída está disponível, mas as informações sobre como os dados foram transformados por algoritmos são limitadas” (DIAKOPOULOS, 2015, p. 11, tradução

<sup>17</sup> Segundo Kitchin (2017, p. 23), “Diakopoulos (2013, p. 13) explica que “engenharia reversa é o processo de articular as especificações de um sistema através de um exame rigoroso baseado no conhecimento, observação e dedução do domínio para descobrir um modelo de como esse sistema funciona”.

<sup>18</sup> API é a sigla de Application Programming Interface, o que significa um conjunto de protocolos criados por um aplicativo ou plataforma web e que, assim, permitem que outros aplicativos possam se conectar a ele, seja para acessar os seus dados ou para oferecer novas funcionalidades.

nossa).<sup>19</sup> Neste cenário, o pesquisador considera que a observação parcial de *inputs* e *outputs* pode ser acrescida de realização de entrevistas, raspagem de dados e investigação de documentos para levar a resultados interessantes, incluindo a técnica de *crowdsourcing* como forma de contornar a dificuldade de coletar os dados de entrada, em que estes podem ser usados como supostas entradas para o algoritmo.

Há ainda situações em que não é possível ter controle sobre os *inputs* porque estes são dados originados a partir do uso feito pelos usuários das plataformas, o que não é possível de simular fielmente. De todo modo, o autor pondera que estes dados podem ser aproveitados, mas sua utilização pode ser mais eficaz se a amostra obtida puder ser delimitada de modo que se proponha a responder diretamente a questão colocada pelo investigador.

### 3.2 Procedimento de seleção dos perfis

No presente trabalho, aplicamos a metodologia da engenharia reversa considerando os *outputs* (saídas), ou seja, os resultados apresentados pelo Google Notícias na categoria personalizada “Pra Você”, na versão *desktop*, referentes às cinco primeiras notícias e seus respectivos veículos de comunicação de origem.

Apesar de o agregador permitir que o usuário selecione quais conteúdos está mais ou menos interessados em ler - ao indicar se deseja receber mais notícias sobre futebol e menos sobre política, por exemplo -, o que concretiza a chamada personalização explícita, buscamos analisar o funcionamento da personalização implícita (THURMAN; SCHIFFERES, 2012), ou seja, quando o usuário não sinaliza nenhuma preferência de forma objetiva.

Desta forma, o propósito é averiguar como o algoritmo do Google usa o histórico de navegação e os hábitos de acesso da internet pelo usuário - seja através dos registros gerados com o uso do próprio e-mail e outros serviços da empresa ou a partir da interação em sites de redes sociais, por exemplo - e realiza as recomendações das notícias ao leitor.

A escolha pela análise da versão *desktop* se deu em razão desta modalidade não ter disponível até o momento o recurso “Resumo de notícias”, como ocorre na versão *mobile* do aplicativo no Brasil citado no capítulo anterior, o que pode significar que o usuário talvez esteja em uma condição mais propensa ou vulnerável para lidar com um filtro-bolha na categoria personalizada.

A primeira etapa consistiu na seleção de quatro perfis de usuários com conta Gmail a

---

<sup>19</sup> No original: “This is a common case that data journalists encounter: a large output dataset is available but there is limited information about how that data was transformed algorithmically.”

fim de tornar possível a utilização do recurso de personalização das notícias. Dada a necessidade de acesso às notícias e veículos ofertados em cada conta Google, a pesquisa contou com duas formas de coleta dos dados: o acesso direto a três das quatro contas de usuários, mediante concessão de *login* e senha pelos administradores; e o repasse das notícias exibidas por um dos próprios usuários.

Convém destacar que a necessidade de repasse das notícias ou acesso direto às contas mediante *login* para coleta foi identificado como um fator limitante para o tamanho da amostra da pesquisa e posterior análise do material. Entretanto, tal método para obtenção das notícias foi mantido em razão da dificuldade técnica para criação de *bots*, o que levou a pesquisa a lidar com resultados da personalização implícita sobre as notícias com base no comportamento humano e genuíno dos usuários a partir de suas respectivas contas. Em seguida foram documentados aspectos socioeconômicos dos usuários, sendo levados em consideração o sexo, a faixa etária, grau de instrução, localização e breve resumo sobre os interesses de cada perfil em relação ao consumo de notícias e a outros temas gerais nas suas navegações pela internet (Tabela 1).

**Tabela 1 – Perfis dos participantes da amostra segundo o sexo, faixa etária, grau de instrução, localização e interesses para avaliação da personalização implícita**

<b>Usuários</b>	<b>Sexo</b>	<b>Faixa etária</b>	<b>Grau de Instrução</b>	<b>Localização</b>	<b>Resumo sobre os interesses do perfil</b>
<b>Perfil A</b>	Feminino	23	Ensino superior incompleto na Universidade Federal do Rio de Janeiro (UFRJ)	Niterói, RJ	O usuário acompanha diariamente notícias sobre economia (com foco em macroeconomia) e tecnologia, em razão da cobertura jornalística como repórter que exige pesquisas sobre os temas do momento.
<b>Perfil B</b>	Feminino	28	Pós-graduação ( <i>stricto sensu</i> ) incompleta na Universidade Federal do Estado do Rio de Janeiro (UNIRIO)	Niterói, RJ	Assiste vídeos sobre astrologia no YouTube e se interessa por conteúdos sobre o uso da <i>cannabis</i> para fins medicinais e de negócios. Também se interessa por tecnologia, viagens, alimentação saudável e cuidados com animais de estimação, por ter um cachorro.
<b>Perfil C</b>	Feminino	48	Pós-graduação ( <i>lato sensu</i> ) completa	Niterói, RJ	Tem interesse em conteúdos sobre decoração, jardinagem, culinária, entretenimento e o andamento da vacinação contra

					a Covid-19 em Niterói. Por dividir o uso do celular com o filho, sua conta Google também serve de base para pesquisas sobre cultura geek, que com maior frequência são feitas no YouTube.
<b>Perfil D</b>	Masculino	24	Ensino superior incompleto na Universidade Federal Fluminense (UFF)	Niterói, RJ	Tem interesse em conteúdos sobre cultura geek (principalmente filmes e séries de super heróis), além de música pop e conteúdos em plataformas de streaming. Também pesquisa sobre equipamentos de tecnologia e cuidados com animais de estimação, por ter cinco gatos.

Fonte: Elaborado pela autora.

Os quatro usuários são moradores de Niterói, cidade da região metropolitana do Rio de Janeiro. É preciso destacar que o perfil B é utilizado por pessoa do sexo feminino que mora com os filhos e que um deles, menor de idade e do sexo masculino, utiliza a conta do Google de sua responsável para acessar a Internet. Portanto, trata-se de um perfil com acesso misto, o que tende a refletir os interesses e as preferências de duas pessoas e não apenas uma, como ocorre nos demais perfis participantes da pesquisa.

A seleção de usuários da mesma localidade se deu não só por conveniência - já que seria mais desafiador lidar com perfis inseridos em outros espaços geográficos e, conseqüentemente, com possíveis aspectos socioculturais distintos -, como também pela possibilidade de a localização do usuário influenciar no conteúdo exibido, assim como acontece na seção “Manchetes”, “Cobertura Completa” e “Resultados da Pesquisa” descrito pela própria empresa (COMO..., 201-?).

Além disso, considerando uma eventual possibilidade de o algoritmo do Google Notícias utilizar o IP (*Internet Protocol Address*, em inglês) para sugerir notícias locais dentro da seção “Para Você”, analisada na pesquisa, consideramos apenas os resultados vistos sempre no mesmo dispositivo eletrônico, garantindo que, sob o mesmo IP, o algoritmo não tivesse que lidar com grandes mudanças nos dados dos usuários.

Também optou-se por utilizar o próprio navegador Google Chrome para realizar a pesquisa, com a obrigatoriedade de acessar as contas e coletar as informações sob a aba “Visitante” a fim de que o perfil e as informações de outros perfis não pudessem influenciar

de algum modo as notícias vistas em cada conta Google, já que a própria empresa sinaliza na Central de Ajuda do Google Chrome que o modo visitante “limita a visualização e alteração das informações de outros perfis do Chrome” (NAVEGAR..., 20-?).

### 3.3 Coleta de dados

Para se obter os dados das notícias recomendadas pelo Google Notícias foram feitas coletas a partir do site do agregador, sob a URL da categoria “Para Você”<sup>20</sup>. Não há uma API da própria plataforma que permita que os dados sejam extraídos, portanto, eles foram coletados manualmente. Evidentemente, tal método de coleta não é tão rápido e eficiente se comparado com outras possibilidades. Entretanto, foi a melhor maneira encontrada para atender as necessidades da pesquisa e encontrar respostas adequadas aos objetivos propostos. Convém ponderar que tal abordagem pode ser considerada, por outro lado, uma forma mais fidedigna de coleta dos dados, uma vez que a forma de acesso é a mais próxima de um comportamento efetivamente genuíno e humano, que poderia ser realizado pelo próprio usuário.

Sendo assim, foi adotado o critério de documentar as cinco primeiras notícias exibidas aos quatro usuários sempre entre às 22h00 e 23h00 (horário de Brasília), com anotações sobre cada perfil observado em questão, os títulos das cinco primeiras notícias exibidas a ele e o veículo aos quais estas pertenciam. Faz-se a ressalva de que as notícias exibidas ao perfil D foram obtidas por meio de capturas de tela feitas pelo participante da pesquisa que, sempre sob o mesmo IP, realizou o envio à pesquisadora para posterior agrupamento à pesquisa.

O horário escolhido para a coleta da pesquisa, foi determinado a partir da tentativa de minimizar eventuais aparições de notícias cujos assuntos fossem considerados “urgentes”. Ou seja, o horário fixo tenderia a diminuir uma eventual possibilidade de o agregador ofertar notícias mais “quentes” em um momento e depois mais “frias” em outro, a depender da quantidade de notícias que os veículos estão produzindo em cada um dos horários do dia, por exemplo. Além disso, foi escolhido um horário fixo ao final do dia para que pudesse facilitar, em alguma medida, a realização de uma análise sobre as notícias recomendadas para cada perfil que fosse capaz de englobar quaisquer matérias publicadas pelos veículos ao longo de todo o dia.

A coleta dos dados foi realizada a partir das notícias exibidas aos perfis em três períodos. Primeiro, foi realizada uma coleta entre os dias 19/04/2021 e 23/04/2021 (segunda a

<sup>20</sup> Cf. <https://news.google.com/foryou?hl=pt-BR>

sexta-feira) com os perfis A, B e C. O perfil D não esteve presente na primeira coleta por razões externas, já que a coleta das notícias exibidas a este perfil dependia do repasse das informações à pesquisadora e não foi possível que o agente realizasse o repasse neste período. Sendo assim, foi encontrada uma dificuldade pela pesquisa com relação à forma de coleta indireta do material. Conseqüentemente, o material analisado do perfil D possui uma amostra ligeiramente menor que as demais.

Em seguida, foi realizada a segunda coleta das cinco primeiras notícias aos perfis A, B, C e D entre os dias 17/05/2021 e 21/05/2021, de segunda a sexta-feira. Por fim, foi realizada uma terceira coleta entre os dias 07/06/2021 a 10/06/2021 (segunda a quinta-feira). É preciso ressaltar que o perfil D, por dificuldade de acesso ao Google Notícias via *desktop* durante o período das 22h no dia 10/06, não pôde enviar a captura de tela com as notícias exibidas.

### **3.4 Classificação das notícias**

A partir dos critérios supracitados foi realizada a coleta do material, composta pelas notícias que foram exibidas aos usuários da plataforma durante o período. Em seguida, essas informações foram organizadas em uma planilha no *Google Sheets*. A esse conjunto de notícias foi realizada uma classificação do conteúdo, de modo a categorizar cada um dos assuntos exibidos, segundo as seguintes rubricas: Astrologia, Animais, Ciência, Coronavírus, Cultura, Cultura Geek, Economia, Entretenimento, Estilo de Vida, Internacional, Música, Notícias Locais, Política, Polícia, Saúde e Tecnologia.

As editorias foram escolhidas a partir da compreensão sobre as temáticas das notícias encontradas durante a coleta da pesquisa. Optou-se por classificar as notícias conforme a rubrica que melhor as representavam, diferentemente, portanto, das categorias elencadas pelo Google Notícias e exibidas na plataforma, como ilustra a Figura 6 (o que sugere que a pesquisadora, aqui, também atua como *gatekeeper*).

**Figura 6 – Editorias da plataforma Google Notícias no *desktop***



Fonte: Google Notícias. Acesso em: 16 ago. 2021, às 20h50.

Tal escolha se deu em razão da constatação de que estas oito editorias, além da rubrica “Covid-19”, não dão conta de todo o conteúdo encontrado nas matérias durante o período analisado. Cabe destacar ainda que a plataforma não deixa claro se utiliza ou não estas mesmas categorias para recomendar as notícias aos usuários na categoria “Para Você”. Portanto, há a possibilidade de a classificação do conteúdo nesta aba do Google Notícias ser ainda mais complexa dada a sua intenção de oferecer conteúdo personalizado.

Também houve a tentativa de classificar as notícias conforme as editorias de dois grandes veículos tradicionais de alcance nacional, *O Globo* e *Folha de S. Paulo*. Entretanto, as divisões estabelecidas por ambos os jornais são não apenas distintas entre si, como insuficientes para dar conta do universo da pesquisa. Desse modo, após a coleta do material, foi constatado que era preciso eleger rubricas novas que fossem mais claras e específicas.

Foram consideradas como matérias de Astrologia todas aquelas que tinham como foco conteúdos que tratassem sobre a crença da influência dos astros na vida dos indivíduos, como notícias sobre horóscopo, tarologia, previsões astrológicas, personalidades dos indivíduos de diferentes signos, entre outros. Já as matérias que tratavam sobre animais domésticos e silvestres, cuidados com bichos e fatos curiosos ficaram sob a editoria Animais.

As matérias relacionadas a estudos e descobertas científicas, além de notícias que tratassem sobre o espaço e o clima, foram incluídas na editoria Ciência. Diante do impacto da pandemia de Covid-19 sobre a população global e, conseqüentemente, sobre o noticiário

jornalístico, as notícias que tratavam sobre número de casos, mortes, vacinados - além de calendários de vacinação, estudos científicos sobre o coronavírus e suas variantes, entre outros - foram agrupadas na editoria Coronavírus.

A editoria Cultura incluiu matérias que tinham como foco reflexões sobre os movimentos culturais e históricos. Optamos por separar, porém, as notícias que tratavam sobre um movimento cultural em específico e, por isso, foi criada também a editoria “Cultura Geek”. O objetivo foi concentrar notícias que pertenciam a uma mesma identidade cultural, considerada por pesquisadores como uma subcultura<sup>21</sup> que inclui conteúdos sobre videogames, ficção científica, jogos, histórias em quadrinhos, mangás e outros.

Já a editoria Economia concentrou notícias sobre indicadores macroeconômicos (inflação, taxa de desemprego, taxa de juros, PIB etc.), negócios, finanças, emprego, programas de transferência de renda (a exemplo das notícias sobre calendário de pagamento do auxílio emergencial, Bolsa Família etc.) e matérias que tratavam sobre descontos em produtos e leilões da Receita Federal.

Optou-se por diferenciar Cultura do Entretenimento por entender que, baseado na compreensão sobre Cultura descrita acima, há conteúdos que mais podem ser associados ao Entretenimento. Enquanto a editoria Cultura contemplou reflexões sobre movimentos culturais e históricos, a rubrica Entretenimento concentrou matérias sobre vida pessoal e profissional das celebridades, *realities shows*, conteúdos em plataformas de streaming.

Foram categorizadas como Estilo de Vida conteúdos que tratavam sobre dicas de alimentos para melhorar o sono, receitas saudáveis, decoração do lar e truques para limpeza da casa com produtos naturais. Reportagens em que o objeto central da notícia foi algum assunto ou fato externo ao país ficaram sob a editoria Internacional. Já a rubrica Música contemplou as matérias que tratavam sobre álbuns, clipes e músicas de artistas, dentro e fora do universo das plataformas digitais de streaming de áudio.

As notícias que exploravam a questão da localização em seus títulos ficaram sob a editoria Notícias Locais. A escolha pela rubrica se deu a partir da compreensão de que o Google Notícias tende a valorizar, em alguma medida, o aparecimento de notícias que tenham a ver com localidades visitadas pelo usuário. Isso pode ser percebido de duas formas. No

---

<sup>21</sup> De acordo com McCain et al. (2015, p. 1), “A cultura geek é uma subcultura de entusiastas que é tradicionalmente associada a mídias obscuras (animação japonesa, ficção científica, videogames, etc.). No entanto, a cultura geek está se tornando uma influência cada vez mais dominante no mundo contemporâneo cultural. A cultura geek inclui uma série de atividades, como RPG jogos (por exemplo, *masmorras e dragões*), ficção científica (por exemplo, *Star Trek*), histórias em quadrinhos e fantasias (por exemplo, *cosplay*)”.

próprio *layout* do Google Notícias via *desktop* a empresa elenca “Notícias Locais” como uma das editorias da plataforma conforme exibido na Figura 6. Além disso, sabe-se que toda conta Google possui uma modalidade de configuração intitulada “Histórico de localização”. Ela não é ativada por padrão, porém, quando o usuário permite que sua conta Google rastreie dados de localização, sua conta salva todos os dados dos locais visitados. De acordo com a empresa, essa configuração registra os locais visitados por meio de dispositivos móveis conectados à Conta Google com histórico ativado e, mesmo que esse recurso seja desativado pelo usuário em algum momento, “alguns dados de local poderão continuar sendo salvos em outras configurações, como na ‘Atividade na Web e de apps’, quando você usar outros serviços, como a Pesquisa Google e o Maps” (GERENCIAR..., 20-?).

Na sequência da classificação das notícias, as matérias com foco em debates, discussões, atos políticos e notícias relacionadas aos governos e seus representantes foram classificadas como Política. Já as notícias que abordaram acidentes, casos de violência e investigações policiais ficaram sob a rubrica Polícia. É importante destacar, neste caso, que a escolha do termo para classificar esse conteúdo foi feita a partir do teor das notícias exibidas. As matérias que focavam na divulgação de casos de segurança pública contavam com títulos que proviam somente o mínimo para deixar o leitor curioso e não contemplavam nos títulos a trama social, política e econômica no qual estavam inseridos os fatos. De modo geral, tais conteúdos podem ser também analisados à luz da lógica do *clickbait*<sup>22</sup>.

Por fim, matérias que tratavam sobre saúde do corpo humano e descobertas científicas sobre o mesmo foram categorizadas como Saúde. Já as notícias que incluíram conteúdos sobre inovações tecnológicas, internet, *smartphones*, aplicativos e equipamentos eletrônicos ficaram sob a editoria Tecnologia.

Feita a apresentação da classificação, a planilha com o conteúdo coletado foi preenchida e organizada por meio das seguintes colunas: Título da Notícia, Link, Navegador, Data, Hora, Editoria e Veículo (como exemplificado na Tabela 2). Já nas linhas foram contempladas as cinco primeiras notícias exibidas a cada perfil.

---

<sup>22</sup> De acordo com o Dicionário Oxford (CLICKBAIT, 2021), *clickbait* significa “material colocado na internet para atrair a atenção e encorajar os visitantes a clicarem em um link para uma página da web específica”.

**Tabela 2 – Exemplo de classificação da amostra analisada a partir das informações dos perfis e notícias personalizadas exibidas**

Perfil A							
Ordem	Título da Notícia	Link	Navegador	Data	Hora	Editoria	Veículo
1	Mercado prevê inflação maior para 2021, aponta Boletim Focus	<a href="https://trademap.com.br/blog-noticia-mercado-preve-inflacao-maior-para-2021-aponta-boletim-focus/">https://trademap.com.br/blog-noticia-mercado-preve-inflacao-maior-para-2021-aponta-boletim-focus/</a>	Chrome (Visitante)	19/04/21	22h	Economia	TradeMap
2	Facebook anuncia clone do Clubhouse, suporte a podcasts e Soundbites	<a href="https://tecnoblog.net/433850/facebook-anuncia-clone-do-clubhouse-suporte-a-podcasts-e-soundbites/">https://tecnoblog.net/433850/facebook-anuncia-clone-do-clubhouse-suporte-a-podcasts-e-soundbites/</a>	Chrome (Visitante)	19/04/21	22h	Tecnologia	Tecnoblog
3	Magazine Luiza lança cartão de crédito sem anuidade e com cashback	<a href="https://exame.com/marketing/magazine-luiza-lanca-cartao-de-credito-sem-anuidade-e-com-cashback/">https://exame.com/marketing/magazine-luiza-lanca-cartao-de-credito-sem-anuidade-e-com-cashback/</a>	Chrome (Visitante)	19/04/21	22h	Economia	Exame

Fonte: Elaborado pela autora.

## **4 O ALGORITMO DO GOOGLE NOTÍCIAS EM AÇÃO: ENTRE A DIVERSIDADE EDITORIAL E A BOLHA INFORMACIONAL**

O terceiro e último capítulo deste trabalho será dedicado à análise quantitativa e qualitativa dos dados obtidos por meio da coleta das notícias exibidas na versão *desktop* do Google Notícias para os quatro perfis selecionados. Após ter sido detalhado o processo de visualização dos dados a partir da planilha descrita no capítulo anterior, foram gerados gráficos que pudessem ser apresentados e analisados à luz dos questionamentos sobre *gatekeeper*, filtro-bolha e diversidade, definidos no capítulo de referencial teórico.

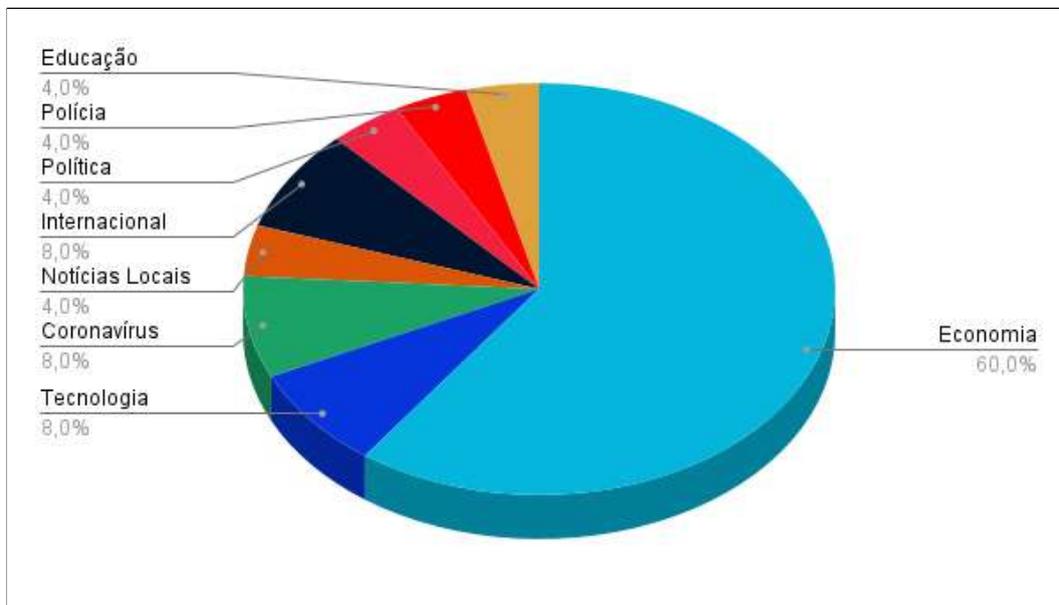
Primeiro, será analisada cada uma das amostras dos perfis individualmente. Posteriormente, iremos comparar os resultados dos perfis uns com os outros. Assim, iremos observar como as recomendações exibidas aos usuários se comportam em termos de diversidade de editorias e diversidade de veículos. Desse modo, olhar para estes aspectos das notícias recomendadas nos ajuda a entender o papel do algoritmo do Google Notícias, seu funcionamento e o impacto dessa curadoria nos resultados de acordo com as noções sobre filtro-bolha e diversidade de notícias.

### **4.1 A diversidade de editorias resultante da recomendação personalizada**

A subseção deste capítulo detalha os resultados obtidos a partir da análise da diversidade de editorias. Sendo assim, as seguintes figuras e explicações indicam a pluralidade de temáticas apresentadas para cada uma das quatro contas Google, tendo como medida a porcentagem de editorias por notícias recomendadas.

#### 4.1.1. Temáticas recomendadas ao perfil A

**Gráfico 2 – Editorias exibidas ao perfil A entre os dias 19/04/2021 e 23/04/2021**



Fonte: Elaborado pela autora.

O período de cinco dias de coleta das cinco primeiras notícias, realizada entre os dias 19/04/2021 e 23/04/2021, resultou em um total de 25 notícias para o perfil A (Gráfico 2). Dentro desse espaço temporal, foi possível observar que 60% das notícias eram vinculadas à Economia. Sob essa editoria foram concentrados artigos sobre macroeconomia (com projeções para inflação, crescimento econômico e reajustes de contratos a partir de indicadores); finanças (ao citar o Pix, meio para realização de transferências bancárias); negócios (vinculados ao comércio eletrônico e suas estratégias e riscos); artigos sobre empregos em *startups* e vagas a partir de concursos do IBGE; além de posicionamento do Ministério da Economia sobre a realização do Censo Demográfico do IBGE e calendário de pagamento do INSS.

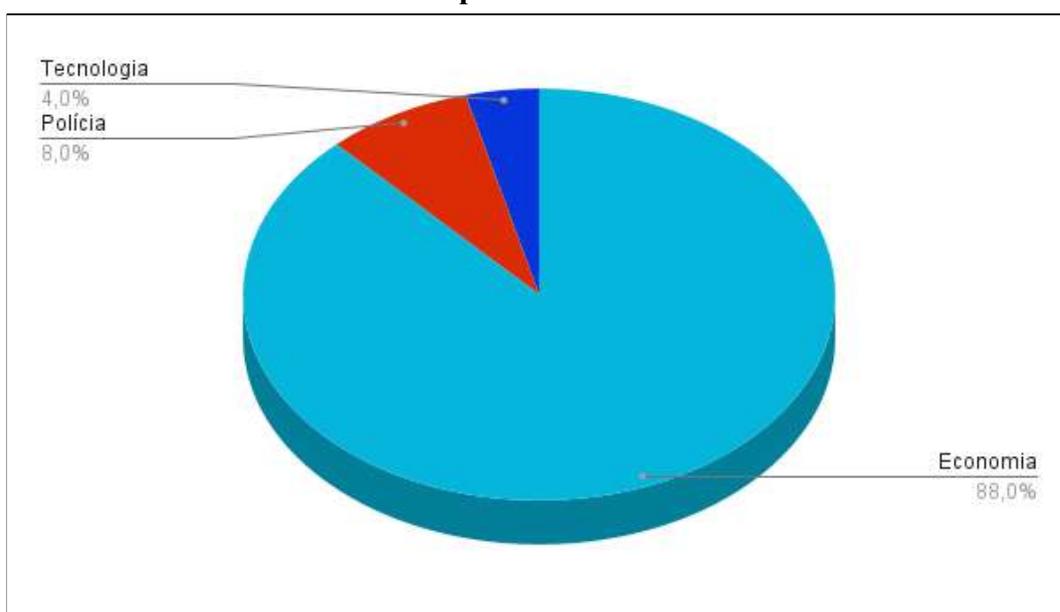
Já a editoria Tecnologia teve participação de 8% na amostra, com duas matérias, sendo uma sobre redes sociais e outra sobre marcas de equipamentos eletrônicos. Com igual participação (8%) apareceu a editoria Coronavírus, com artigos que tratavam sobre decretos com medidas restritivas no Rio de Janeiro e curva de mortos pela doença em estados do país. Um artigo classificado como pertencente a “Notícias Locais” tratava sobre a Universidade Federal do Rio de Janeiro (UFRJ), o que nos leva a considerar que o Google Notícias utiliza

os locais supostamente visitados ou reconhecidos como de interesse do usuário para oferecer notícias que envolvam tal localidade, conforme previsto no capítulo anterior.

Também foram ofertadas para este perfil, nesta coleta, notícias sobre Educação (4%), Polícia (4%), Política (4%) e Internacional (8%). A matéria sobre educação tratava sobre *homeschooling*, assunto que não está no rol de interesses do perfil como princípio. Já as matérias de Internacional tratavam sobre futuros previstos para a sociedade pela Comunidade de Inteligência dos Estados Unidos (CI) e um marinheiro em um navio abandonado. A notícia sobre polícia trouxe um caso sobre uma influenciadora digital encontrada morta em seu apartamento, e a notícia sobre política falava sobre a Lei Geral de Proteção de Dados (LGPD) a partir de uma fala do vereador Carlos Bolsonaro (Republicanos-RJ). O Gráfico 1 mostra a proporção de artigos por editoria.

Os resultados indicam que a personalização implícita do Google Notícias, de fato, ocorre. Isso porque estão presentes os dois grandes interesses do usuário na amostra (Economia e Tecnologia) - inclusive, com domínio significativo de 68% destes. Cabe destacar, contudo, que há um percentual de 32% (resultado da soma das outras editorias) de notícias que não estão diretamente relacionadas às preferências do perfil e também se fazem presentes. Sendo assim, há uma quantidade razoável de editorias fora das preferências do usuário A que são apresentadas.

**Gráfico 3 – Editorias exibidas ao perfil A entre os dias 17/05/2021 e 21/05/2021**



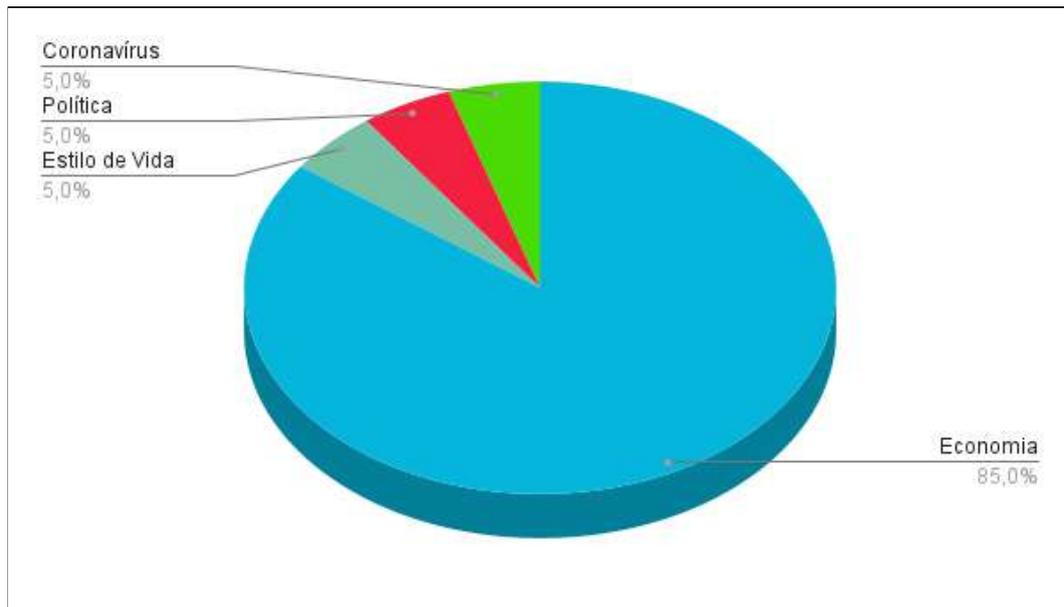
Fonte: Elaborado pela autora.

A segunda pesquisa, realizada entre os dias 17/05/2021 e 21/05/2021, também teve duração de cinco dias e resultou em um total de 25 notícias oferecidas para o perfil A (Gráfico 3). Desta vez, porém, a concentração de artigos sobre Economia foi maior: passou de 60% da primeira coleta para para 88%. Ou seja, quase todo o conteúdo exibido pertencia a uma mesma editoria, o que compreende uma diversidade de notícias ainda menor do que a apresentada na primeira coleta.

As temáticas apresentadas na editoria econômica não mudaram, apenas foram acrescentados mais assuntos. Estiveram presentes matérias que traziam projeções para indicadores macroeconômicos (PIB e IGP), além de estratégias do comércio eletrônico, Pix e desdobramentos da realização do Censo Demográfico pelo IBGE – assuntos que também apareceram na primeira coleta. Já os novos assuntos econômicos tratavam sobre a situação dos frigoríficos e do consumo de carne bovina pelos brasileiros; valores de criptomoedas; taxa de investimentos; fornecimento de energia no Brasil e riscos de apagão; regras do Banco Central para aquisições de cartões de crédito e negócios envolvendo compra de empresas.

Em seguida, apareceu a editoria Polícia com participação de 8%, o equivalente a duas notícias, e posteriormente Tecnologia, com 4%, o que representa apenas uma notícia sobre o tema. A matéria de Tecnologia, novamente, tratava sobre rede social, ainda que sobre outra plataforma nessa amostra. Já uma matéria sobre Polícia era referente a uma investigação policial da morte de MC Kevin e a outra sobre a prisão de uma mulher acusada de chefiar o tráfico no Jacarezinho, comunidade no Rio de Janeiro. Tais resultados nos levam a considerar que o agregador de notícias parece buscar que o usuário tenha contato com algumas notícias de editorias externas aos seus interesses ao longo do tempo.

**Gráfico 4 – Editorias exibidas ao perfil A entre os dias 07/06/2021 e 10/06/2021**



Fonte: Elaborado pela autora.

A terceira coleta de artigos noticiosos do perfil A foi realizada entre os dias 07/06/2021 e 10/06/2021 e resultou em 20 notícias ofertadas (Gráfico 4). A participação da editoria Economia ocupou a maior parte da amostra, o que indica a manutenção da dominância dos assuntos econômicos ao usuário, apesar de a participação ter diminuído (de 88% na segunda coleta para 85% na terceira coleta). Sendo assim, a diversidade de editorias para o perfil é considerada pequena.

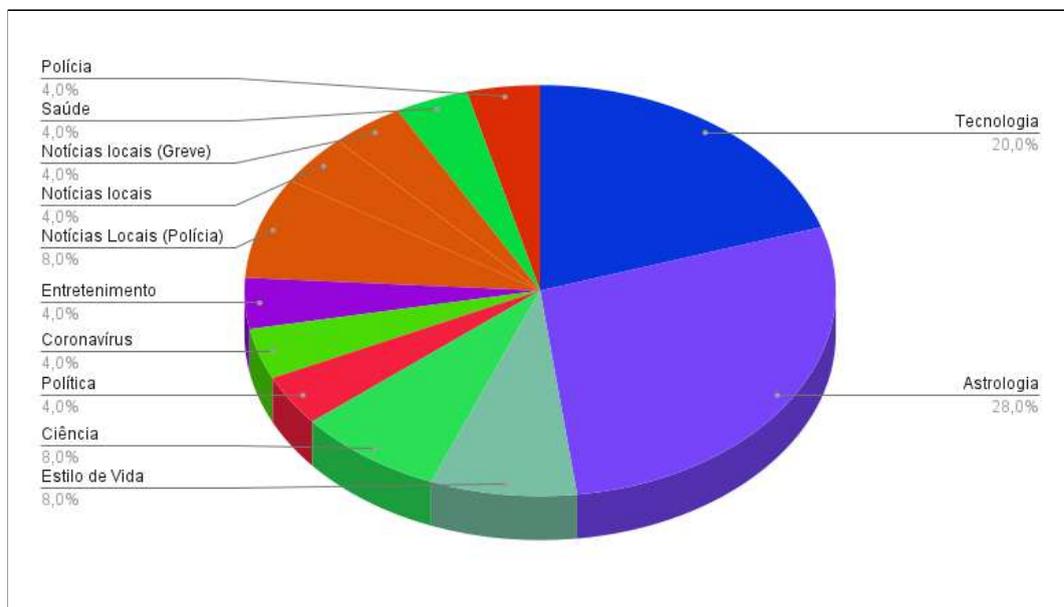
No entanto, cabe destacar que a matéria considerada como de “Estilo de Vida” trata de um assunto não muito distante da temática econômica, pois o objeto central se refere ao consumo de carne bovina – assunto identificado como de interesse do usuário na coleta anterior, em que uma matéria tratava sobre a queda do consumo desse insumo pela população. Desta vez, porém, o conteúdo ofertado tratava das implicações que o consumo de carne bovina pode trazer ao ser humano, como o risco de doenças.

Tal resultado observado no perfil A sugere que o agregador Google Notícias reconhece os termos de interesse possivelmente pesquisados pelo usuário e busca oferecer a ele um leque mais amplo de recortes noticiosos sobre uma mesma temática. Assim, se um usuário acompanha como a questão econômica impacta o consumo de carne bovina, talvez se interesse pelos seus eventuais impactos sobre o corpo humano. Não deixa de ser também uma forma de ofertar uma visão diferente para não restringi-lo a uma bolha, ainda que trate sobre um mesmo objeto relacionado ao interesse de um tema previamente identificado.

Por fim, as notícias ligadas à Política e Coronavírus também corresponderam a 5%, cada, do total da amostra, com um conteúdo sugerindo que Gleisi Hoffmann deseja proibir viagens espaciais de Jeff Bezos e outro sobre a discussão da vacinação de adolescentes pela prefeitura do Rio. Não foi possível identificar uma relação clara entre os interesses do usuário e o que pode ter levado à aparição da notícia sobre política, exceto o fato de a notícia ter como um dos atores o empresário Jeff Bezos, dono da Amazon, que notadamente é uma personalidade alvo de notícias no âmbito de seus negócios. A mesma ausência de relação entre um claro interesse do usuário e o aparecimento da notícia ocorreu sobre a editoria Coronavírus, já que o perfil A não costuma pesquisar sobre o assunto, porém, o andamento da vacinação no país pode ser considerado um assunto de interesse público da população. Assim, fica em aberto a suposição de que o Google Notícias tenha sugerido matérias que o perfil A poderia vir a se interessar para evitar a totalidade de notícias vindas de Economia.

#### 4.1.2. Temáticas recomendadas ao perfil B

**Gráfico 5 – Editorias exibidas ao perfil B entre os dias 19/04/2021 e 23/04/2021**



Fonte: Elaborado pela autora.

O período de cinco dias de coleta das cinco primeiras notícias, realizada entre os dias 19/04/2021 e 23/04/2021, resultou em um total de 25 notícias para o perfil B (Gráfico 5). A editoria Astrologia apareceu com participação de 28%, seguido de Tecnologia (20%), Notícias Locais (16%), Ciência (8%), Estilo de Vida (8%), Coronavírus (4%), Saúde (4%), Política (4%), Entretenimento (4%), Notícias Locais (Polícia) (8%), Notícias Locais (4%), Notícias Locais (Greve) (4%), Saúde (4%), Polícia (4%).

(4%) e Polícia (4%). Nota-se a presença de dez editorias na presente amostra, o que tende a fazer com que a diversidade de notícias seja maior neste período.

Os artigos noticiosos vinculados à Astrologia tratavam de previsões de horóscopo para os signos, bem como as características de cada um deles e os seus reflexos na vida pessoal. Já as notícias sobre Tecnologia trouxeram comparações entre equipamentos eletrônicos e suas respectivas marcas.

Na sequência, as Notícias Locais foram subdivididas na amostra do gráfico apenas para facilitar a visualização do que se tratava. Esta categoria abarcou conteúdos sobre assassinato em São Gonçalo, explosão de um carro em Niterói, inauguração de um espaço pela prefeitura e a paralisação de rodoviários na cidade niteroiense. Nota-se que o Google Notícias identificou localidades possivelmente registradas pelo usuário e próximas às identificadas (como é o caso da cidade vizinha a Niterói, São Gonçalo) com o intuito de ofertar notícias sobre essas regiões. Vale destacar que, das quatro notícias locais, duas trataram sobre casos de polícia com títulos que podem trazer um certo apelo ao leitor.

Também chama a atenção a presença da editoria Polícia, com a seguinte notícia: “Vídeos mostram momento em que piscina desaba em edifício de luxo no ES”, assunto esse que não é de interesse direto do perfil B e cujo título pode ser entendido como *clickbait*, já que induz o usuário a clicar na matéria para assistir o vídeo de um acidente.

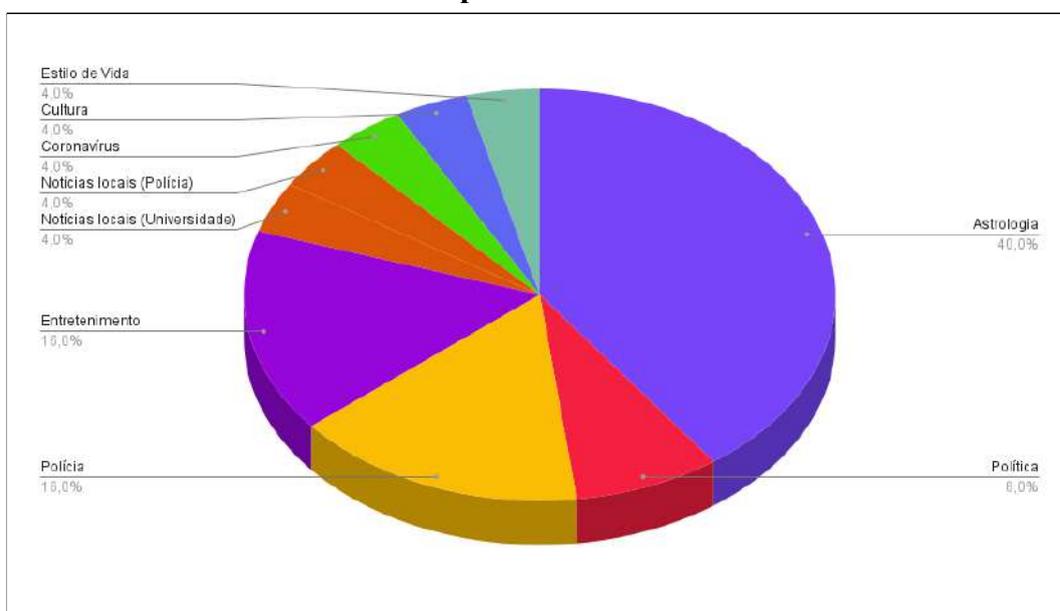
Já a editoria Política trata da discussão no âmbito político sobre cultivo da maconha, conteúdo que se relaciona com o interesse do usuário em acompanhar os avanços na legislação a respeito do tema. A editoria Coronavírus, por sua vez, apresentou uma matéria sobre a distribuição de *cannabis* aos vacinados em Nova York, o que traz um certo apelo no título, ainda que o objeto da matéria se relacione com os interesses previamente descritos do usuário com a *cannabis* e suas aplicações medicinais.

A editoria Saúde apresentou uma matéria que trata do comportamento humano no contexto das emoções, possivelmente um conteúdo sugerido pelo agregador para o usuário que acompanha matérias sobre astrologia que tratam do amor e do comportamento dos indivíduos nessa esfera pessoal.

Duas matérias que tratavam sobre alimentação saudável foram categorizadas como Estilo de Vida, cujo conteúdo parece estar diretamente relacionado com os gostos do perfil B. Já a editoria Entretenimento trouxe uma matéria sobre um participante de um *reality show*, assunto este que não faz parte das preferências do usuário. Por fim, a editoria Ciência contemplou duas notícias, sendo uma sobre um drone da Nasa (agência espacial

estadunidense) e outra sobre um fenômeno avistado em um planeta. Chama a atenção a oferta de conteúdos científicos sobre o espaço. Imagina-se que a plataforma tenha identificado um interesse do usuário pela astrologia, pseudociência que relaciona o céu com o cotidiano da vida humana. Por isso, a aparição da editoria Ciência a partir dos recortes supracitados pode ter a ver com a tentativa de o agregador oferecer um conteúdo que o usuário possa vir a se interessar, assim como foi visto nos resultados exibidos do perfil A.

**Gráfico 6 – Editorias exibidas ao perfil B entre os dias 17/05/2021 e 21/05/2021**



Fonte: Elaborado pela autora.

A segunda coleta das cinco primeiras notícias exibidas ao perfil B, realizada entre os dias 17/05/2021 e 21/05/2021, também resultou em um total de 25 notícias (Gráfico 6). Desse total, há o destaque para as matérias sobre Astrologia, que tiveram participação de 40% na amostra - 12% a mais do que na primeira coleta. O conteúdo das matérias sob essa editoria se mantiveram os mesmos. Cabe destacar, contudo, que, na medida em que a concentração de notícias sobre Astrologia aumentou mesmo com a repetição de abordagens, também foram ofertados mais conteúdos sobre Política (que passou de 4% para 8%), Polícia (de 4% para 16%) e Entretenimento (de 4% para 16%). Sendo assim, foram oferecidas oito editorias nesta amostra ante dez na pesquisa anterior.

Nota-se também no Gráfico 6 que a segunda amostra não apresenta mais notícias sobre Tecnologia, o que pode sugerir que o interesse do perfil tenha sido alterado. Se na primeira coleta apareceram matérias que comparavam equipamentos eletrônicos, nesta não houve

nenhum conteúdo sobre. Assim, é possível especular que determinado interesse do usuário pelo assunto possa ter sido pontualmente sanado - eventualmente com a compra de um produto, já que o Google considera o histórico de buscas do usuário e também pode rastrear o comportamento do perfil em diversos sites.

Ao destrinchar o conteúdo vinculado à Política, percebe-se que uma mesma matéria foi exibida para o usuário em datas diferentes. Embora não seja possível identificar a razão disso, podemos especular que os algoritmos da plataforma podem ter considerado que o assunto era de interesse do perfil ou que, dado a sua quantidade de visualizações ou viralização nas redes sociais, por exemplo, o perfil B não deveria deixar de vê-lo ou pelo menos ter contato com o material. Evidentemente que também não é possível descartar, por exemplo, que possa ser um *bug* da plataforma, muito embora sejam apenas hipóteses.

A mesma duplicação de conteúdo ocorre quando se observam dois dos quatro artigos exibidos vinculados à Polícia. Desta vez, as matérias recomendadas não era exatamente iguais pois não pertenciam aos mesmos veículos, mas nota-se que o agregador de notícias ofereceu duas vezes o mesmo assunto para o perfil B, ao indicar as notícias “Antes de queda, MC Kevin brigou com amigos por sexo e preservativo” e “MC Kevin brigou com amigos por sexo e preservativo antes de queda, relata modelo”.

Outras matérias relacionadas à editoria foram: “Polícia prende acusados de roubo na Ilha Grande, em Angra dos Reis” e “Polícia encontra droga ilegal em hortas de cultivo autorizado de maconha”. Assim como em outros momentos da pesquisa, chama a atenção uma quantidade significativa de conteúdo policial, sendo um deles vinculado à apreensão de droga ilegal relacionado a um cultivo da *cannabis* - mais uma vez, objeto de interesse do usuário, mas não com essa abordagem previamente descrita em seu perfil.

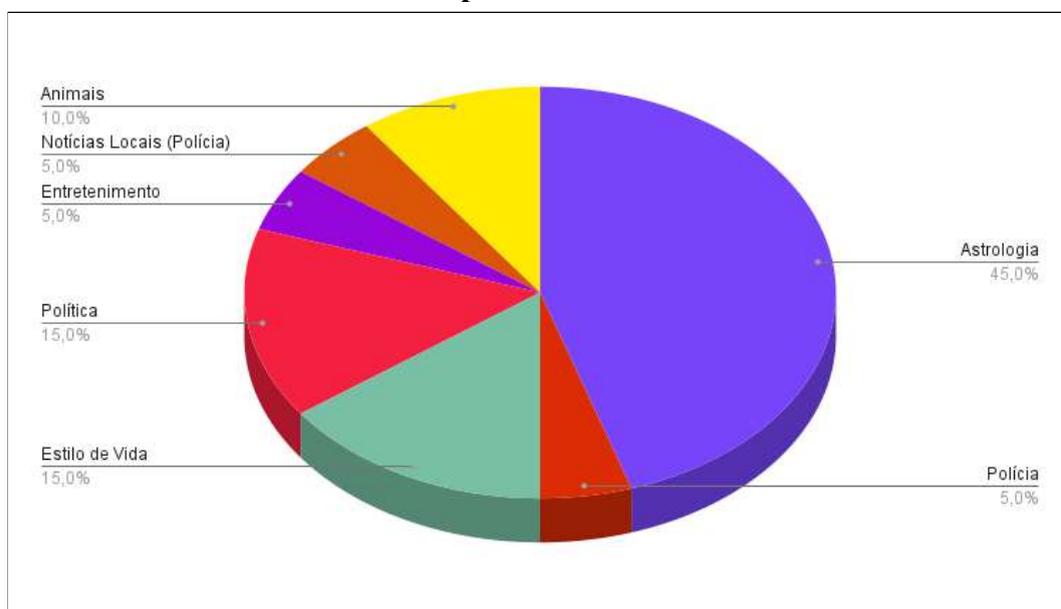
Outro ponto evidenciado na segunda amostra foi a presença de uma matéria da editoria Cultura (“Maconha e a vanguarda brasileira”) - cujo link de origem é um canal do YouTube e não um *site* -, que narra processos históricos sem influências das abordagens policiais, como visto acima, ou com fim medicinal como evidenciado na primeira coleta.

No âmbito das Notícias Locais, foram exibidas duas matérias. Enquanto uma trata sobre segurança pública, ao noticiar uma notícia sobre trocas de tiros constantes em Niterói, a outra notícia trata da possibilidade de greve na UniRio, universidade de interesse do usuário dado a sua relação com a instituição, cuja localização possa ter sido visitada e rastreada pelo Google ou cujo termo tenha sido pesquisado e identificado pelos algoritmos. Sendo assim,

fica evidente a tentativa da plataforma de sugerir matérias que tenham a ver com os lugares que o usuário frequenta e a região em que mora.

Por fim, dois artigos de notícia apresentaram conteúdos relacionados aos gostos do perfil B. É o caso da editoria Estilo de Vida, que exibiu uma matéria sobre culinária e da rubrica Coronavírus, que ofertou uma notícia sobre uma vacina contra a Covid-19 feita a partir de tabaco, que se relaciona com o interesse do usuário de saber sobre avanços científicos no que tange ao uso medicinal da *cannabis*.

**Gráfico 7 – Editorias exibidas ao perfil B entre os dias 07/06/2021 e 10/06/2021**



Fonte: Elaborado pela autora.

A terceira coleta de matérias exibidas ao perfil B foi realizada entre os dias 07/06/2021 e 10/06/2021 e resultou em 25 notícias recomendadas (Gráfico 7). Se comparada com as duas primeiras coletas, a terceira amostra apresentou uma quantidade menor de editorias - foram sete ante oito da segunda coleta. Houve um predomínio ainda maior de matérias sobre Astrologia, com participação de 45% do total de artigos, frente 40% e 28% da segunda e primeira pesquisas, respectivamente. O teor das notícias sobre astrologia se repetem em sua maioria: são apresentadas previsões diárias e semanais de horóscopo, além de comportamentos dos signos em determinadas esferas da vida, com exceção de uma matéria que dessa vez estaria mais relacionada à superstição ou crença, chamada “Amor e casamento: confira 12 simpatias de amor para Santo Antônio”.

Nesta pesquisa, Estilo de Vida e Política tiveram, cada, participação de 15%. Em seguida aparece a editoria Animais (10%), seguido de Notícias Locais (5%), Entretenimento (5%) e Polícia (5%). Em Estilo de Vida, foram recomendadas duas matérias, sendo uma sobre culinária e outra com dicas para lavagem de roupas - tais temáticas podem ser consideradas de interesse específico do usuário, de acordo com o resumo sobre o perfil.

Sob a rubrica Animais, duas matérias recomendadas ao perfil trataram sobre raças de cães. O perfil B tende a se interessar por assuntos que envolvem cães, e estes conteúdos podem ter sido ofertados em razão de alguma pesquisa do usuário, assim como ocorreu com a editoria Tecnologia, na primeira coleta. Isso porque ambas as editorias estiveram pontualmente presentes nas amostras do perfil e com percentual ligeiramente significativo. No entanto, a amostra não é suficiente para comprovar tal hipótese.

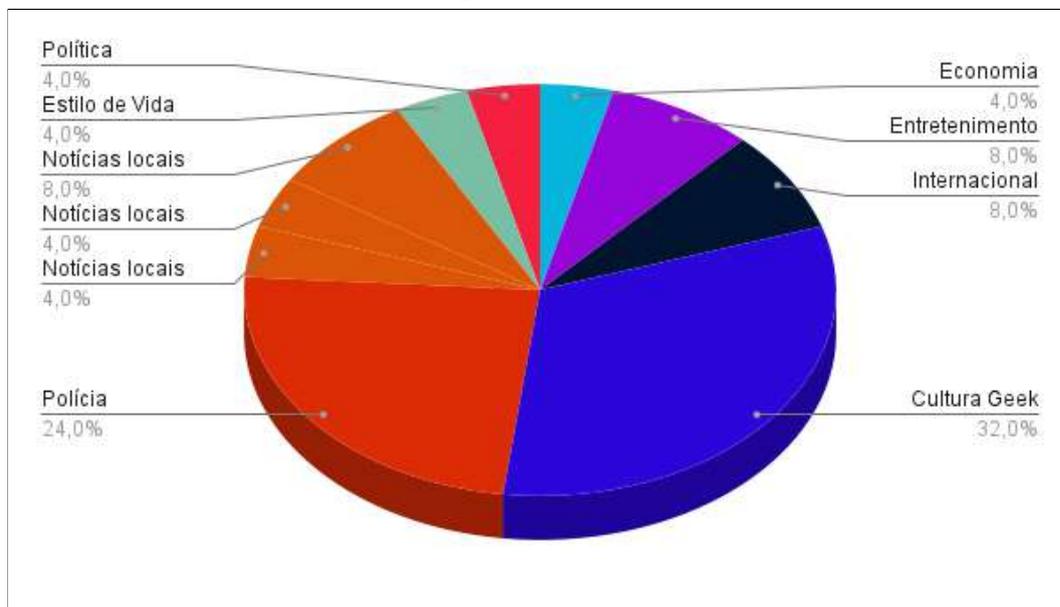
Já em Política, as matérias recomendadas trataram de projetos sobre cultivo da maconha para fins medicinais e industriais, que tratam de temas diretamente ligados ao interesse do perfil B. Apesar do aumento da participação de editorias como Astrologia e Estilo de Vida, e também do aparecimento de uma nova editoria (Animais), conteúdos sobre Entretenimento, Polícia e Notícias Locais não deixaram de aparecer.

Em Polícia, foi recomendado um conteúdo que contava o caso de um homem preso por possuir plantação ilegal de maconha. Assim como observado nas outras duas coletas, mais uma vez a plataforma sugeriu conteúdo que tratava de um assunto de interesse do usuário, ainda que por meio de outro recorte ou abordagem.

Já em Entretenimento, uma matéria sobre uma celebridade foi recomendada ao usuário, embora esta não haja uma descrição clara de interesse do perfil pelo tema. Por fim, em Notícias Locais, foi sugerida uma matéria sobre assassinato em um bairro de Niterói, o que torna clara a identificação pelo agregador dessa cidade como relevante para o usuário. Assim, o Google Notícias procura recomendar acontecimentos que ocorreram no município em que ele reside.

### 4.1.3. Temáticas recomendadas ao perfil C

**Gráfico 8 – Editorias exibidas ao perfil C entre os dias 19/04/2021 e 23/04/2021**



Fonte: Elaborado pela autora.

A coleta das cinco primeiras notícias recomendadas ao perfil C ocorreu entre os dias 19/04/2021 e 23/04/2021 e resultou em um total de 25 notícias (Gráfico 8). A editoria Cultura Geek representou 32% do total de editorias, seguido de Polícia (24%), Notícias Locais (16%), Internacional (8%), Entretenimento (8%), Estilo de Vida (4%), Economia (4%) e Política (4%). Observa-se, portanto, a presença de oito editorias no período.

As matérias classificadas como Cultura Geek abordavam jogos disponíveis para PlayStation 4, filmes de super heróis e acontecimentos envolvendo seus atores. Assim, fica evidente a presença de matérias que contemplam parte dos interesses do perfil C, conforme evidenciado no resumo sobre o perfil no capítulo anterior.

Já as notícias sobre Polícia abordaram diferentes casos policiais. Os títulos das matérias, desta vez, foram diversos, sendo os seguintes: “Incêndio em Fazenda Rio Grande deixa prejuízo superior a R\$ 1 milhão”; “Carro é esmagado por duas carretas no norte do País”; “Mulher desaparecida é encontrada dentro de cobra píton”; “Garota de 14 anos esconde câmera e filma abusos de pastor em Goiânia”; “Após impedir assalto, jovem é morto a tiros no Maranhão; vídeo” e “Vídeos mostram momento em que piscina desaba em edifício de luxo no ES”. A quantidade de matérias vinculadas a esta categoria chamam a atenção, sobretudo pelo fato de que o usuário não costuma realizar pesquisas ativas sobre casos como os citados

acima - muito embora seja do interesse do usuário C jogos que possuem algum tipo de nível de violência - como é o caso do jogo Grand Theft Auto (GTA), recomendado, inclusive, em uma das matérias de Cultura Geek.

Apesar da possível relação entre o interesse por jogos dessa modalidade e, conseqüentemente, uma eventual predisposição do usuário em consumir fatos que carregam algum nível de violência, não é possível aferir que estes aspectos estão propriamente interligados. Apenas é possível afirmar que a plataforma recomendou uma quantidade expressiva desse tipo de conteúdo e que seus títulos podem vir a ser considerados *clickbait*, de acordo com as definições descritas no capítulo anterior.

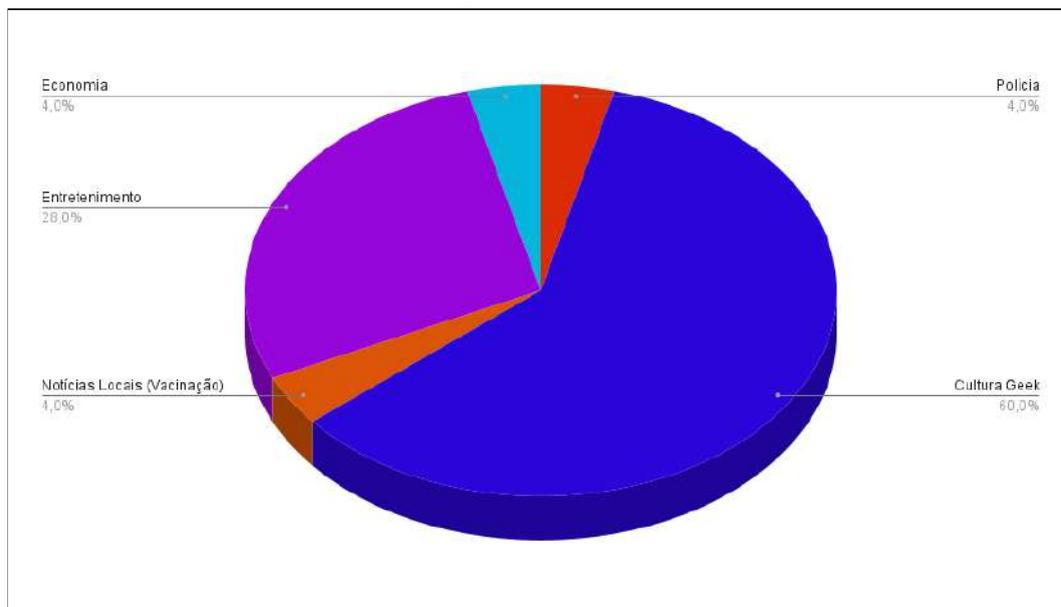
As Notícias Locais, que representaram 16% da amostra, abarcavam conteúdo sobre vacinação contra a Covid-19, paralisação do transporte público e explosão de um automóvel. Assim como visto no perfil B, o perfil C também apresentou matérias relacionadas à cidade onde reside o usuário. Fica evidente, portanto, que o Google Notícias atribuiu certa relevância à localidade e busca sugerir notícias relacionadas à tal região.

Em seguida, os artigos de notícias recomendados em Internacional foram “Susto no ar e no mar” e “Funcionário salta nos trilhos e salva criança de trem na Índia”, reforçando a quantidade de notícias na amostra que são voltadas para acidentes cujos enunciados também podem ser associados à lógica *clickbait* assim como os conteúdos de Polícia.

Já em Entretenimento foram ofertadas duas notícias, uma sobre a participação de uma atriz em um seriado e outra sobre o desfecho do enredo desta mesma série citada na primeira matéria. Em Economia, foi recomendada uma notícia sobre leilão de um console de *videogame* com destaque para o termo “PS4” no título, frequentemente encontrado nas matérias sobre Cultura Geek, de interesse do perfil C. Assim, é possível deduzir que o Google Notícias sugeriu uma editoria que, embora não estivesse normalmente sendo acompanhada pelo usuário, ele poderia vir a gostar já que o objeto da notícia tem relação direta com uma de suas preferências.

Já em Política, foi sugerida uma matéria sobre uma discussão entre um deputado federal e um economista, cuja editoria e seu respectivo conteúdo não costumam ser de interesse do usuário C. Por fim, a editoria Estilo de Vida apresentou uma matéria sobre decoração de varandas com plantas, temática descrita como relevante no resumo sobre o usuário no capítulo anterior.

**Gráfico 9 – Editorias exibidas ao perfil C entre os dias 17/05/2021 e 21/05/2021**



Fonte: Elaborado pela autora.

A segunda coleta das cinco primeiras notícias recomendadas ao perfil C ocorreu entre os dias 17/05/2021 e 21/05/2021 e resultou em um total de 25 notícias (Gráfico 9). Desta vez, notícias vinculadas à Cultura Geek apresentaram quase o dobro de participação na pesquisa (60% ante 32%, na segunda coleta). Em seguida apareceram matérias sobre Entretenimento (28%), seguido de Economia (4%), Polícia (4%) e Notícias Locais (4%). Dessa forma, nota-se que apenas cinco editorias estiveram presentes nesta amostra, uma queda com relação à segunda coleta, que ofereceu oito editorias, e à primeira, que sugeriu sete assuntos.

Sob a editoria Cultura Geek foram recomendadas matérias sobre filmes de zumbis, jogos de *videogame* envolvendo super heróis ou que incluem ação e violência, como Battlefield, Mortal Kombat, Fortnite, Resident Evil, Red Dead e Grand Theft Auto (GTA), citados nas matérias.

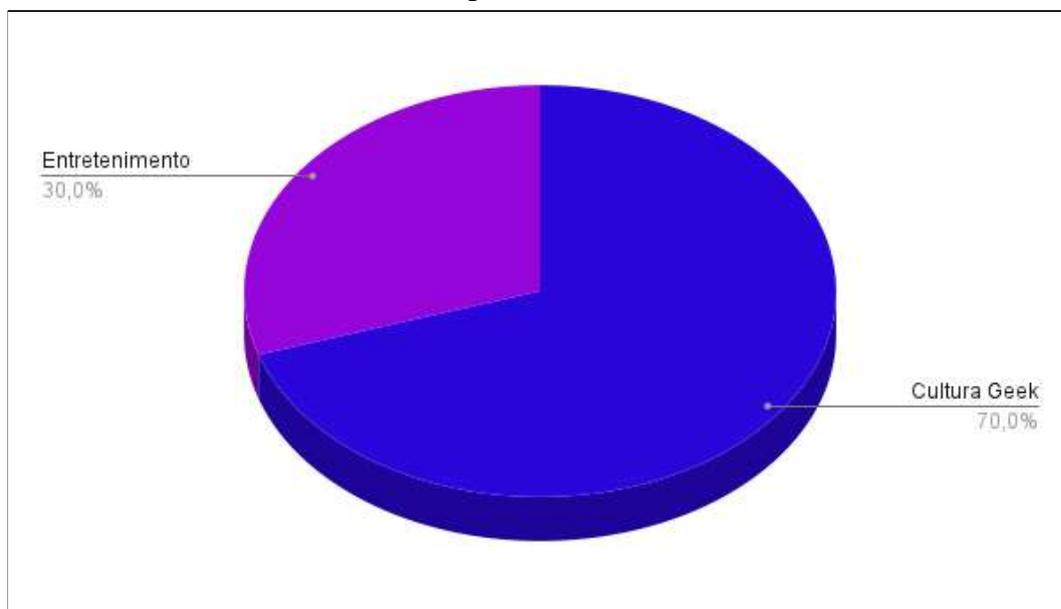
De acordo com o resumo sobre o perfil e os resultados da coleta, é possível identificar na presente pesquisa uma relação entre o que o usuário tem interesse, considerado como Cultura Geek, frente ao conteúdo que ele recebeu na aba personalizada do agregador de notícias. O mesmo pode ser observado na participação da rubrica Entretenimento, que representou 28% do total da pesquisa. Há matérias que tratam de filmes sobre zumbis em plataforma de streaming e o cantor Travis Scott - mais ligados aos interesses do menor de idade que utiliza a conta do responsável pelo perfil C, assim como há a presença de

conteúdos que são do interesse do próprio usuário C. É o caso das matérias que tratam do reality show “No Limite”, que apareceram quatro vezes na presente coleta.

Na sequência, em menor participação, apareceu uma matéria vinculada à Polícia, que não pode ser relacionada ao interesse prévio do usuário. Já a matéria pertencente à Notícia Local, que trata do andamento da vacinação em Niterói está diretamente ligada ao interesse do presente perfil analisado.

Por último, foi sugerida uma matéria sobre Economia que trata da disputa judicial entre a Apple e a Fortnite. Cabe destacar que entre os atores principais da notícia está o Fortnite, jogo que com frequência apareceu em matérias ligadas à Cultura Geek. Sendo assim, nota-se em mais um perfil a tentativa por parte do Google Notícias de sugerir notícias cujas abordagens são diferentes, mas que tratam de objetos/assuntos que já tendem a terem sido identificados anteriormente como de interesse do perfil, o que pode fazer com que aquele usuário se interesse por uma editoria diferente, a fim de minimizar a possibilidade de deixá-lo em um filtro-bolha.

**Gráfico 10 – Editorias exibidas ao perfil C entre os dias 07/06/2021 e 10/06/2021**



Fonte: Elaborado pela autora.

A terceira coleta das cinco primeiras notícias recomendadas ao perfil C ocorreu entre os dias 07/06/2021 e 10/06/2021. Chama a atenção no Gráfico 10, que traz a proporção de notícias por editoria, a concentração de matérias sob apenas duas rubricas, o que coloca o perfil C em uma condição de diversidade de notícias menor do que o observado nas coletas

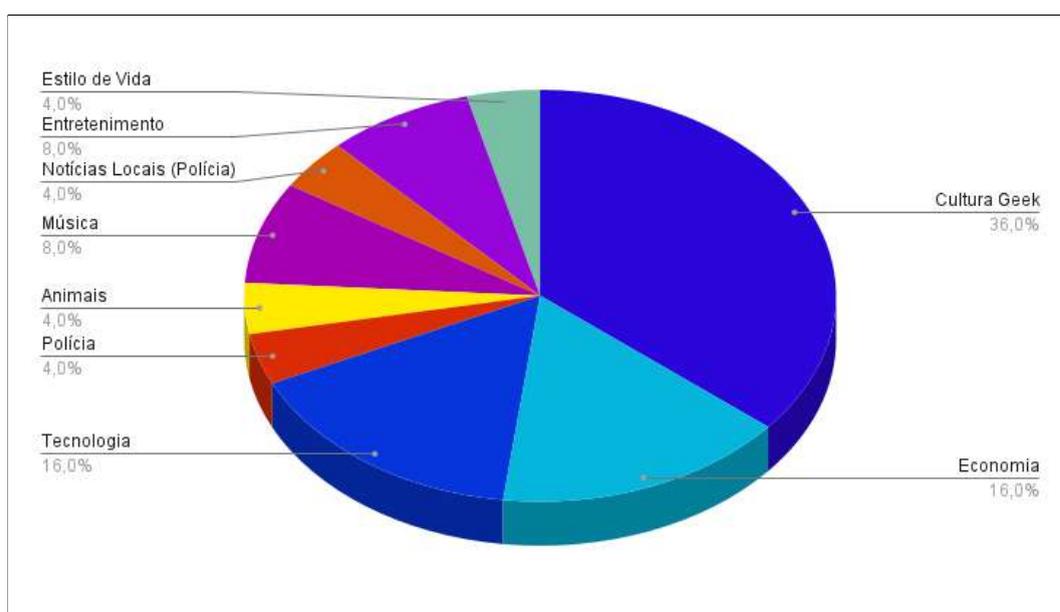
anteriores. Enquanto na primeira e segunda pesquisas foram recomendadas notícias categorizadas como de oito e cinco editorias, respectivamente, nesta coleta foram sugeridos apenas conteúdos de duas categorias: Cultura Geek (70%) e Entretenimento (30%).

Sob a editoria Cultura Geek, estiveram presentes notícias sobre jogos de videogame e filmes de super heróis. Já sob a editoria Entretenimento foram recomendadas notícias sobre celebridades, suas participações em conteúdos de *streaming* e reality shows.

Chama a atenção que, neste período analisado, o Google Notícias tenha indicado apenas duas categorias e que representam os principais interesses do perfil C. Tal resultado nos leva a deduzir que o agregador tende a identificar quais são os assuntos de interesse do perfil e quais deles são mais perenes, como é o caso das notícias vinculadas à Cultura Geek e Entretenimento, que agradam tanto o usuário menor de idade que utiliza o perfil quanto a sua responsável, que administra esse perfil. Apesar da busca por contemplar significativamente o usuário, a pequena diversidade de editorias sugere que o usuário esteja limitado, ainda que por um circunscrito espaço de tempo, a receber, naquele momento, apenas notícias de seu interesse, sem quaisquer pontos de fricção com assuntos fora do seu agrado.

#### 4.1.3. Temáticas recomendadas ao perfil D

**Gráfico 11 – Editorias exibidas ao perfil D entre os dias 17/05/2021 e 21/05/2021**



Fonte: Elaborado pela autora.

A primeira coleta das cinco primeiras notícias recomendadas ao perfil D ocorreu entre os dias 17/05/2021 e 21/05/2021 e resultou em um total de 25 notícias (Gráfico 11). A presença majoritária foi de matérias relacionadas à Cultura Geek, que representaram 36% do total de notícias da pesquisa. Em seguida, as notícias classificadas como parte das editorias Tecnologia e Economia apareceram, ambas, com 16%. Entretenimento e Música tiveram, cada, 8% de participação, seguidas de Estilo de Vida (4%), Notícias Locais (4%), Animais e Polícia (4%). Desse modo, foram contabilizadas nove editorias na amostra, quantidade significativa de assuntos diversos, ainda que três deles tenham aparecido com maior concentração.

Em Cultura Geek foram exibidos conteúdos sobre séries de super heróis e os respectivos atores que atuam nesses programas - temáticas recomendadas pelo agregador e que se relacionam com os interesses previamente descritos do perfil D. O mesmo ocorreu com a notícia categorizada como da editoria Animais, onde foi sugerido conteúdo com dicas de como deixar seu gato mais calmo, conteúdo de gosto do leitor conforme evidenciado no capítulo anterior.

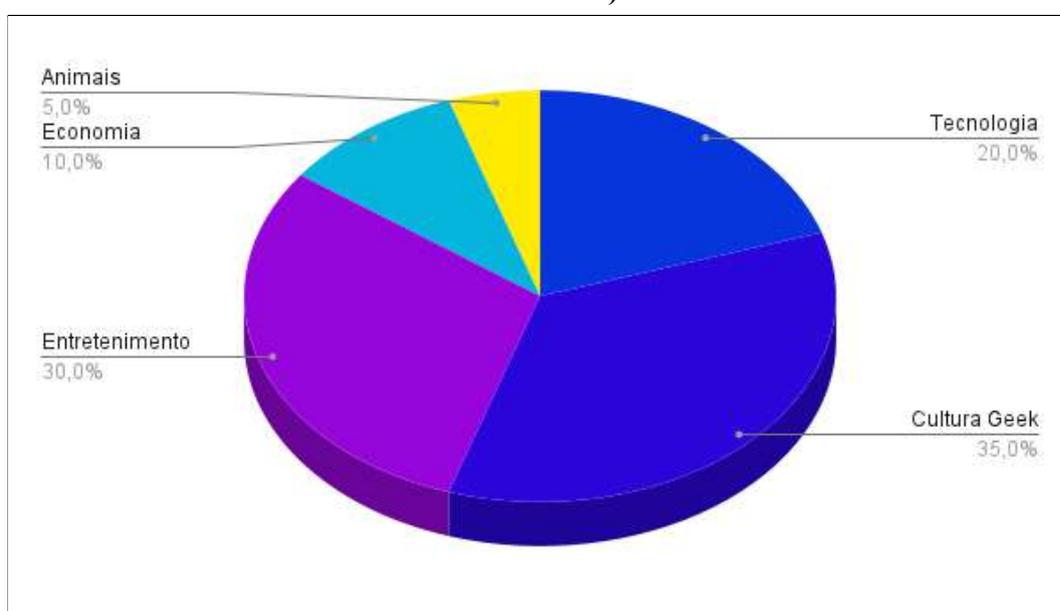
Sob a editoria Economia, foram recomendadas matérias que tratavam de descontos em equipamentos eletrônicos e programas de transferência de renda. Não se sabe se o usuário tem interesse em conteúdo a respeito do pagamento do auxílio emergencial, contudo, é de se esperar que ele se interesse em promoções que envolvam *smartphones* e *notebooks*, dado o seu gosto por notícias sobre equipamentos eletrônicos em Tecnologia. Na editoria Tecnologia, notícias sobre como realizar ajustes nas configurações do iPhone; comparações entre softwares de smartwatch de Google, Samsung e Apple; análise de chip processador dos Macs e personalização de iPhone estão entre as recomendações realizadas ao usuário.

Já em Entretenimento apareceu uma matéria sobre cuidados de beleza feitos por uma celebridade e outra sobre o lançamento de uma série em uma plataforma de *streaming* - sendo este segundo conteúdo previamente identificado como de interesse do perfil D. Em Música, duas matérias sobre a cantora Olivia Rodrigo reforçam o interesse do perfil em música pop (“Olivia Rodrigo: "good 4 u" sobe ao #1 do Spotify Global” e “Olivia Rodrigo estreia com 'Sour', álbum angustiante, irresistível e com Disney rock alternativo”). Em Estilo de Vida, foi recomendada a matéria “11 penteados práticos para cabelos cacheados usando apenas 5 itens”, embora este não tenha sido considerado um interesse prévio do presente usuário.

Por fim, foram ofertados dois conteúdos, sendo um de Polícia (Vídeo mostra momentos após queda de MC Kevin de varanda de hotel na Barra da Tijuca) e outro de

Notícias Locais (“Guerra sem fim em Niterói: mais uma madrugada de tiroteio no Santo Cristo”). Tais resultados ressaltam o que foi observado em outros perfis: o Google Notícias tende a sugerir um conteúdo cujo interesse não está diretamente ligado ao histórico de atividade do usuário, mas dado algum critério como relevância ou grande quantidade de visualizações por outros usuários, o agregador decidiu recomendá-lo ao perfil. Além disso, nota-se o aparecimento de uma notícia que tem relação com a cidade onde o usuário reside.

**Gráfico 12 – Editorias exibidas ao perfil D entre os dias 07/06/2021 e 11/06/2021 (exceto o dia 10/06)**



Fonte: Elaborado pela autora.

A segunda coleta das cinco primeiras notícias recomendadas ao perfil D ocorreu entre os dias 07/06/2021 e 11/06/2021, com exceção do dia 10/06, e resultou em um total de 20 notícias (Gráfico 12). Foi observado que matérias jornalísticas vinculadas à Cultura Geek representaram 35% do total de notícias coletadas no período. Em seguida, aparece a categoria Entretenimento (30%), Tecnologia (20%), Economia (10%) e Animais (5%).

Em Cultura Geek, foram recomendados conteúdos sobre séries de super heróis, ligadas ao interesse do perfil D. Já em Entretenimento foram exibidas matérias que tratavam sobre séries de ação, comédia ou terror e seus respectivos enredos, bem como duas matérias sobre cantoras *pop*. Em Tecnologia, o agregador recomendou quatro notícias que tratavam sobre a marca: três sobre o lançamento de um novo sistema operacional e outra sobre a conferência realizada pela companhia.

Duas matérias sobre Economia foram exibidas ao usuário, com enfoque no auxílio emergencial. Entretanto, não é possível afirmar se tal conteúdo é de interesse direto do usuário. Por fim, uma notícia sobre Animais foi recomendada ao perfil D, cujo interesse está diretamente ligado às preferências descritas no resumo do usuário no capítulo anterior.

#### **4.1.4 Discussão sobre a diversidade de temáticas recomendadas aos usuários a partir dos resultados**

Os resultados obtidos a partir da análise indicam, em um primeiro momento, que os quatro perfis obtiveram recomendações de notícias com assuntos diferentes entre si. Evidentemente, como a categoria analisada “Para Você”, do Google Notícias, trata da personalização implícita, era de se esperar que a curadoria algorítmica se expressasse, a partir das editorias, de modo distinto em cada perfil.

O que significa, também, que estabelecer padrões de influência dos algoritmos não é simples e pode inclusive depender de uma série de outros aspectos, como os termos exatos pesquisados pelos usuários em outros produtos Google; os locais que cada usuário frequenta e sua conta Google é capaz de registrar; os cliques que o usuário executa em determinadas notícias, que podem reforçar seu interesse em um certo assunto, por exemplo; o seu tempo de permanência nas notícias; quais fontes ele costuma acessar, entre outros fatores. Sendo assim, não é possível tecer generalizações sobre o funcionamento dos algoritmos do Google Notícias apenas a partir das notícias recomendadas, mas podemos destacar nuances observadas a partir das notícias recomendadas aos perfis.

Quando analisamos a diversidade de editorias à luz do filtro-bolha, o primeiro aspecto que se constata a partir dos resultados é que a classificação da editoria não presume que os usuários irão ver notícias parecidas entre si. Isso significa que, mesmo que alguns perfis tenham visualizado uma mesma editoria, os conteúdos das notícias podem ser amplamente diferentes. É o caso do perfil A, que recebeu recomendações de conteúdo sobre Economia assim como o perfil D, mas os assuntos tratados nas matérias são completamente distintos. Enquanto o perfil A, na sua primeira coleta, recebeu notícias sobre indicadores macroeconômicos, por exemplo, o perfil D recebeu duas notícias sobre promoções de equipamentos eletrônicos. Ou seja, há poucas recomendações comuns entre os usuários, o que reforça a ideia de que os perfis estão sob visões de mundo diferentes, algo que a plataforma parece reforçar.

Isso parece ser explicado pela forma individualizada como o algoritmo trata cada perfil. Trata-se de uma experiência de consumo de conteúdos diferente do que acontece quando duas pessoas lêem um mesmo jornal, acessam a primeira página de um portal de notícias, ou mesmo assistem a um telejornal, por exemplo. Ainda que tenham interesses diferentes, a modalidade de consumo estabelece uma grande probabilidade de que os leitores ou telespectadores esbarrem nos mesmos conteúdos.

Quando consideramos o critério da diversidade de editorias que nos propusemos a analisar, nota-se que os perfis possuem níveis de “isolamento” diferentes. O perfil A apresentou um baixo nível de diversidade de editorias se comparado com os demais. Como o perfil A pesquisa muito sobre um nicho específico e pouco em termos de diversidade de editoria, o seu perfil apresentou um conteúdo menos diverso. Já os perfis B, C e D, que possuem uma gama mais ampla de interesses, tiveram uma diversidade de editorias maior, em geral. Ou seja, a quantidade de assuntos pelos quais o usuário tem interesse ativamente também parece importar para chegarmos aos resultados apresentados. Isso pode nos apontar que, quanto mais específico for o interesse do usuário, mais ele tende a ficar refém de determinadas coberturas editoriais. O que pode significar, por exemplo, que aquele usuário se torna um leitor assíduo de um ou dois recortes (como ocorre no perfil A, que consome principalmente Economia e Tecnologia) e deixa de cruzar com uma série de outros assuntos na aba “Para Você”.

Tal constatação nos faz refletir sobre outro ponto que deve ser levado em consideração quando analisamos as recomendações de artigos: os termos centrais que constam nos títulos das notícias. A terceira coleta realizada no perfil A mostra a recomendação de uma matéria intitulada “6 fatos preocupantes sobre o consumo de carne bovina”, classificada como Estilo de Vida. Entretanto, não há estabelecido um interesse prévio do usuário por essa editoria, nem mesmo por aspectos que tratam sobre alimentação saudável, por exemplo.

Mas é de se pensar que, se o perfil A pode se interessar por informações sobre exportação de carne bovina ou consumo e sua relação com questões econômicas - já que foram oferecidas na coleta anterior as notícias “Confirmado baixo de poder de compra de carne bovina, frigoríficos recuam”; “Argentina suspende exportações de carne bovina por 30 dias”; “Com baixo consumo de carne bovina, atacadistas adotam medidas para sustentar os preços” e “Consumo de carne no Brasil cai ao menor nível em 25 anos com disparada de preços”, e que todas que tratam de aspectos econômicos relacionados à carne bovina -, os

algoritmos podem ter sugerido um outro recorte sobre esse mesmo termo de interesse, que é “carne bovina”.

Isso não é possível de ser afirmado, mas talvez esteja relacionado com o esforço da plataforma de recomendar não apenas notícias de interesse do usuário, mas de sugerir matérias com outros recortes que possam fazer o usuário consumir um conteúdo ligeiramente mais diverso - assim, evita-se a possibilidade dos usuários se manterem em contato apenas com o que lhes interessa, e o fazem ter contato com outras editorias, distanciando-se de um eventual filtro-bolha.

Essa hipótese nos remete ao que descreveu Pariser (2012) sobre a serendipidade. O autor afirma que uma personalização em excesso poderia fazer com que indivíduos só consumissem um determinado conteúdo, em que não haveria serendipidade, que seria o encontro com algo útil que não se estava esperando. Ofertar um conteúdo cujo objeto principal seja conhecido do usuário, mas esteja aplicado a outro contexto, talvez se configure como um mecanismo para combater a falta de serendipidade e fazer o leitor entrar em contato com algo que ele possa vir a gostar, ainda que de forma sutil.

Apesar desse resultado sugerir que a plataforma se preocupa com o nível de concentração das temáticas em sua aba personalizada e busca recomendar algumas matérias sobre conteúdos diversos, essa preocupação não foi materializada ao longo de toda a pesquisa. A terceira coleta do perfil C mostra que, em determinado horário ao longo de uma semana, o agregador exibiu apenas notícias vinculadas ao interesse do usuário, com conteúdos sugeridos sobre Cultura Geek (70%) e Entretenimento (30%). Não foi possível concluir se esse resultado se manteve por mais tempo, dado que a pesquisa contempla apenas um determinado universo de notícias que foram exibidas durante um certo período do dia, ao longo de alguns dias. Ainda assim, é importante sinalizar que a preocupação que o pesquisador Eli Pariser (2012) quanto aos algoritmos tornarem usuários reféns de certos pontos de vista não é equivocada.

Apesar da baixa diversidade de editorias - ou até casos de completa ausência, como visto acima -, os diferentes períodos de coleta das notícias recomendadas aos perfis mostram que a quantidade de editorias tende a variar ao longo do tempo, ainda que com alguma concentração das temáticas mais acompanhadas por aquele usuário. Isso demonstra que estamos diante de um objeto de difícil capacidade de compreensão uma vez que os algoritmos, como descreveu Kitchin (2017), não são estáticos e podem ser programados para serem reativos e mutáveis para as entradas.

Isso se explica pelo fato de que lidamos ainda com os interesses de usuários reais, que podem mudar ao longo do tempo - considerando ainda que, em poucos segundos, uma pessoa pode expressar um novo interesse rastreável pela sua conta Google (por meio de um vídeo assistido no YouTube ou uma pesquisa feita no Google Pesquisa) e isso depois se refletir, por exemplo, em recomendações novas no Google Notícias. Uma possível mudança no interesse de um usuário pode ser observada nos Gráficos 4, 5 e 6 do perfil B. Se na primeira coleta foram sugeridos conteúdos sobre Tecnologia, com destaque para matérias sobre comparações de preços de fones de ouvido, as duas coletas seguintes deixaram de recomendar notícias sobre Tecnologia, incluindo essa temática supracitada.

Outro aspecto observado a partir das recomendações de notícias aos usuários foi a presença de matérias que apareceram para mais de um perfil. A matéria “Vídeos mostram momento em que piscina desaba em edifício de luxo no ES”, do G1, foi recomendada no dia 23/04 para os perfis B e C. Já a matéria “Vídeo mostra momentos após queda de MC Kevin da varanda de hotel na Barra da Tijuca” foi recomendada no dia 17/05 tanto para o perfil C quanto para o perfil D. Nota-se que as notícias sugeridas pertencem à Polícia - editoria esta que foi recomendada, inclusive, a todos os usuários em alguns períodos das amostras, mesmo que nenhum dos perfis tenha interesse prévio estabelecido na temática.

Em primeiro lugar, a aparição da mesma notícia a perfis diferentes, no mesmo dia e horário, nos leva a questionar se o agregador tende a sugerir matérias que estão em alta, por alguma razão não explicitada, a usuários, propensos ou não, a se interessar pelo conteúdo. Cabe lembrar que o blog do Google News Initiative elenca como um dos critérios para classificação de conteúdo a Proeminência - cujos algoritmos levam em conta, entre outras motivações, o quanto uma matéria ou artigo está em alta - ainda que não seja bem descrito como essa “viralização” é mensurada, se a partir da quantidade de acessos que uma página obteve, o tempo de permanência dos usuários ou outras razões.

Em segundo lugar, a presença da editoria Polícia a todos os perfis chama a atenção, pois nos leva a considerar como hipótese que a plataforma recomenda notícias que vão além do interesse do usuário e não necessariamente também têm a ver com o seu interesse prévio. Isso tenderia a ser algo positivo, uma vez que o usuário teria contato com editorias que não abordem suas preferências pessoais de imediato. No entanto, o fato de ter sido ofertada mais a editoria de Polícia e não editorias outras que variassem ao longo do tempo pode sinalizar uma eventual inclinação do algoritmo em recomendar conteúdos a partir apenas do que está sendo mais clicado, por exemplo. E, como o algoritmo não faz juízo de valor sobre a informação,

leitores podem acabar tendo contato com matérias que privilegiam algumas perspectivas que abordam mais violência em detrimento de outras que podem estimular outros conhecimentos.

Cabe destacar que as notícias classificadas como sendo da editoria Polícia possuem títulos mais instigantes e que por vezes noticiam fatos mais alarmantes, a exemplo desses exibidos ao perfil C: “Mulher desaparecida é encontrada dentro de cobra píton”; “Garota de 14 anos esconde câmera e filma abusos de pastor em Goiânia”; “Após impedir assalto, jovem é morto a tiros no Maranhão; vídeo”. O conteúdo da última notícia citada acima diz apresentar uma cena de violência explícita, por exemplo. O mesmo ocorreu sob a editoria Internacional. Dos quatro conteúdos recomendados de Internacional ao longo de toda a pesquisa, dois tratavam sobre acidentes surpreendentes fora do país, sendo “Susto no ar e no mar | AFP” e “Funcionário salta nos trilhos e salva criança de trem na Índia”.

Embora não seja possível identificar o que leva a plataforma a recomendar tais conteúdos, cabe refletir que é preciso uma investigação mais a fundo acerca dos conteúdos recomendados aos usuários e como eles são classificados - se os algoritmos são capazes de identificar títulos no modelo *clickbait*; se eles aparecem por terem relação com as métricas das matérias, sendo essas as mais lidas ou acessadas por internautas ao longo do dia; entre outros possíveis fatores.

Por fim, um último aspecto observado por meio dos resultados foi a presença de notícias locais nos quatro perfis. Tal presença confirma a premissa descrita pelo blog do Google News Initiative de que a categoria de notícias personalizadas “Para Você” recomenda conteúdos relacionados à localização do usuário. Mas, para além disso, foi observado que uma notícia sobre Polícia, intitulada “Carro explode e assusta moradores de Santa Rosa, em Niterói”, apareceu nos perfis B e C na primeira coleta de ambos.

#### **4.2 A pluralidade de veículos resultante da recomendação personalizada**

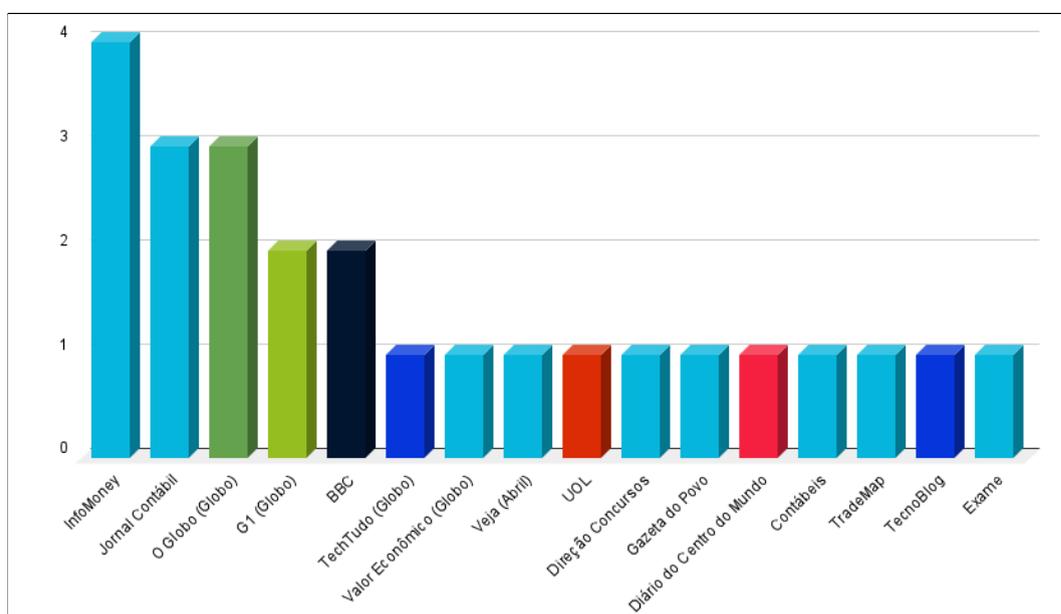
A subseção deste capítulo detalha os resultados encontrados a partir da coleta e análise da pluralidade de veículos. A escolha desse recorte se deu ao considerarmos a importância da diversidade não apenas por meio das notícias propriamente recomendadas, mas também da diversidade de abordagens dos conteúdos que cada meio de comunicação apresenta. Tal aspecto pode representar, para os leitores, uma visão mais ampla sobre a forma como os acontecimentos são abordados, considerando que cada veículo jornalístico, através de sua linha editorial, representa um enquadramento dos temas que elege em suas pautas. Ou seja,

além do conteúdo em si (ou o enunciado), o consumo noticioso é impactado também pela forma (ou a enunciação) como o conteúdo é elaborado pelo periódico.

A seguir, são apresentados gráficos e explicações que indicam a pluralidade de veículos para cada uma das quatro contas Google analisadas pela pesquisa, tendo como medida a participação dos meios por volume de notícias sugeridas.

#### 4.2.1. Fontes recomendadas ao perfil A

**Gráfico 13 – Veículos exibidos ao perfil A entre os dias 19/04/2021 e 23/04/2021**



Fonte: Elaborado pela autora.

A primeira coleta das cinco primeiras notícias recomendadas ao perfil A, entre os dias 19/04/2021 e 23/04/2021, apresentou 16 veículos diferentes, de um total de 25 notícias (Gráfico 13). Nota-se que o veículo *InfoMoney* foi recomendado quatro vezes por meio de notícias sobre Economia, sendo essa fonte a com maior participação no período. Em seguida, o *Jornal Contábil* aparece em segundo lugar, com uma representação a partir de três notícias sobre economia. O jornal *O Globo*, do grupo Globo<sup>23</sup>, também aparece por meio de três notícias, a partir das editorias Notícias Locais, Coronavírus e Economia. Com duas notícias, o *G1*, também do grupo Globo, obteve a terceira maior participação na amostra, junto com a *BBC*. O *G1* ofereceu uma notícia sobre coronavírus e outra sobre educação. Já as matérias

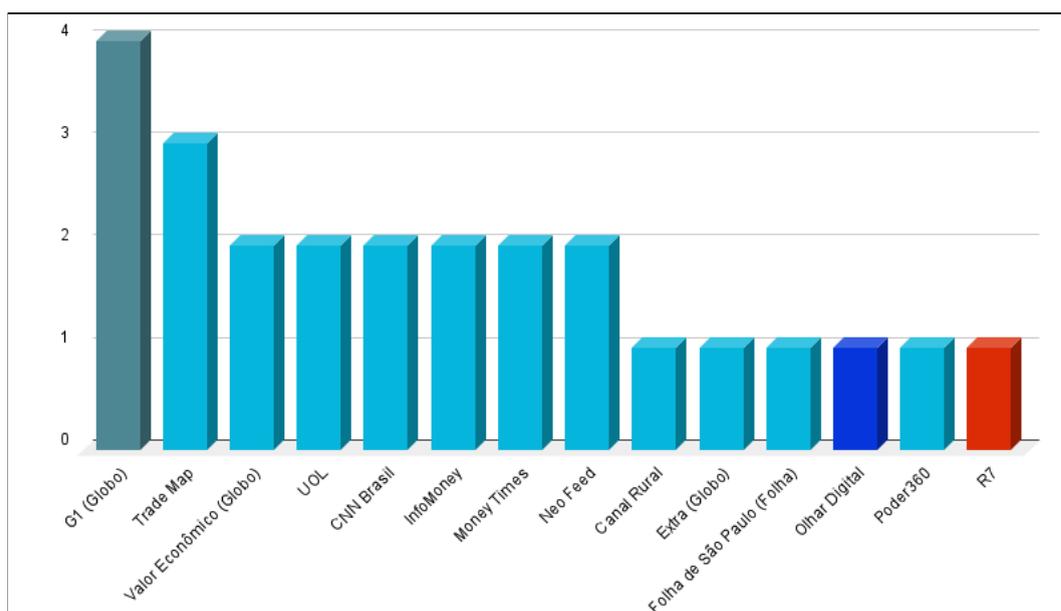
<sup>23</sup> De acordo com o site do Grupo Globo, o Grupo Globo (Organizações Globo Participações S.A) é composto pelas empresas Globo (Globo Comunicações e Participações S.A.), Editora Globo, Sistema Globo de Rádio e Globo Ventures. (GRUPO GLOBO, [s. d.]

exibidas pela *BBC* são de Internacional.

Os demais veículos da amostra apareceram, cada um, com apenas uma notícia, o que configura uma menor participação de cada um desses na pesquisa. Nesse conjunto, percebe-se a presença de outros dois veículos do grupo Globo (*TechTudo*, com matéria sobre tecnologia; *Valor Econômico*, com matéria sobre economia). O grupo Abril apareceu por meio da presença da revista *Veja*, com matéria sobre economia. O grupo UOL<sup>24</sup> (Universo Online) apareceu por meio de uma matéria sobre Polícia, enquanto os veículos *Direção Concursos*, *Gazeta do Povo*, *Contábeis*, *TradeMap* e *Exame* apareceram com uma matéria cada sobre Economia e, por fim, o *Diário do Centro do Mundo* aparece com conteúdo sobre Política e o TecnoBlog, tecnologia.

Como conclusão desta primeira coleta, é possível constatar que o veículo *InfoMoney*, que pode ser entendido como um veículo não-tradicional de comunicação no país, obteve a maior participação na amostra. Entretanto, cabe destacar certo aspecto relevante ao investigar a presença de veículos pertencentes a um grande grupo de comunicação no país. A amostra contém a presença de fontes que pertencem a um mesmo grupo de comunicação, o que deve ser levado em consideração, já que estes veículos podem estar sob uma mesma linha editorial. Na presente coleta, os sites *O Globo*, *G1*, *TechTudo* e *Valor Econômico* obtiveram, juntos, sete notícias - número maior do que o contabilizado pelo *InfoMoney*.

**Gráfico 14 – Veículos exibidos ao perfil A entre os dias 17/05/2021 e 21/05/2021**



<sup>24</sup> De acordo com site do UOL, o grupo é a maior empresa brasileira de conteúdo, tecnologia, serviços e meios de pagamentos e seu site é pioneiro na produção de conteúdo noticioso na internet brasileira. (SOBRE UOL, [s. d.])

Fonte: Elaborado pela autora.

A segunda coleta das cinco primeiras notícias recomendadas ao perfil A, entre os dias 17/05/2021 e 21/05/2021, apresentou 14 veículos diferentes, de um total de 25 notícias (Gráfico 14). Desta vez, o veículo *GI*, do grupo Globo, foi recomendado quatro vezes por meio de três notícias da Economia e uma de Polícia. O *Trade Map* aparece em segundo lugar, com três notícias sobre economia. Na sequência, as fontes *Valor Econômico*, do grupo Globo; *UOL*; *CNN Brasil*; *InfoMoney*; *Money Times* e *Neo Feed* aparecem com duas notícias sobre Economia cada.

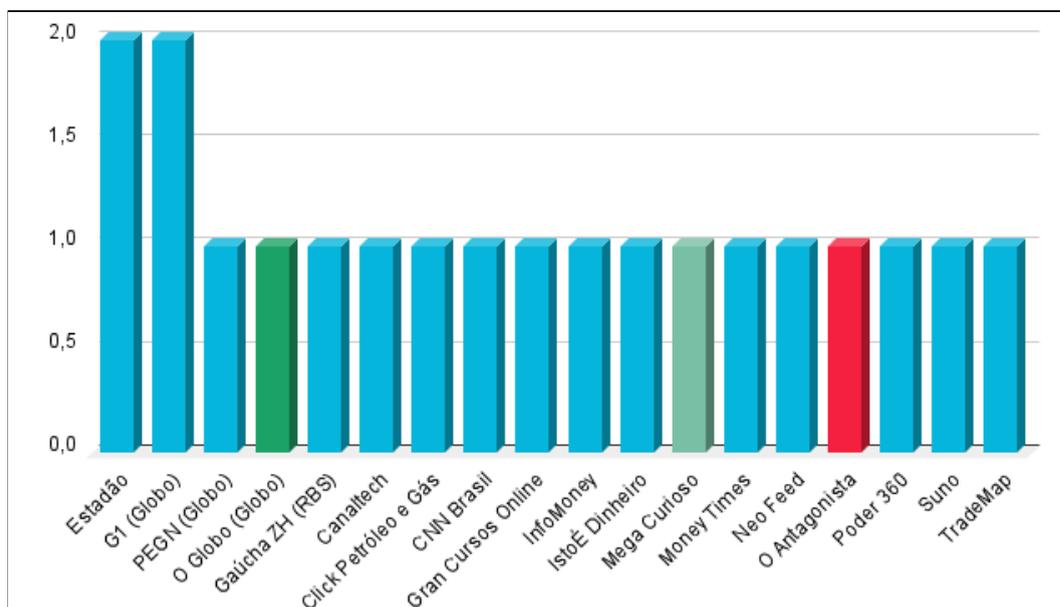
Os demais veículos da amostra apareceram cada um com apenas uma notícia, o que configura uma menor participação de cada um desses na pesquisa. Nesse conjunto, percebe-se a presença dos veículos *Canal Rural*; *Extra*, do Grupo Globo; *Folha de S. Paulo*, do Grupo Folha<sup>25</sup>; e *Poder 360*, todos com matérias sobre Economia. Nota-se que o site *Olhar Digital* aparece na amostra a partir de uma matéria sobre tecnologia, enquanto o *R7* surge com uma matéria sobre Polícia.

Quando contabilizamos a participação relativa das fontes, constatamos que os veículos *GI*, *Valor Econômico* e *Extra* somam, juntos, um total de sete notícias do grupo Globo. Isso daria à organização a maior participação na pesquisa não só de forma absoluta, como vimos acima, mas também de forma relativa. O Grupo UOL, por sua vez, aparece com duas matérias, o que lhe mantém no posto de terceiro lugar após o *Trade Map*.

---

<sup>25</sup> De acordo com site da Folha, o Grupo Folha é um dos principais conglomerados de mídia do país. Controla, entre outras empresas, o jornal Folha de S.Paulo, seu site, o DataFolha e a agência de notícias Folhapress. O Grupo Folha tem participação minoritária, indireta e em ações sem direito a voto no UOL. (CONHEÇA..., [s. d.])

**Gráfico 15 – Veículos exibidos ao perfil A entre os dias 07/06/2021 e 10/06/2021**



Fonte: Elaborado pela autora.

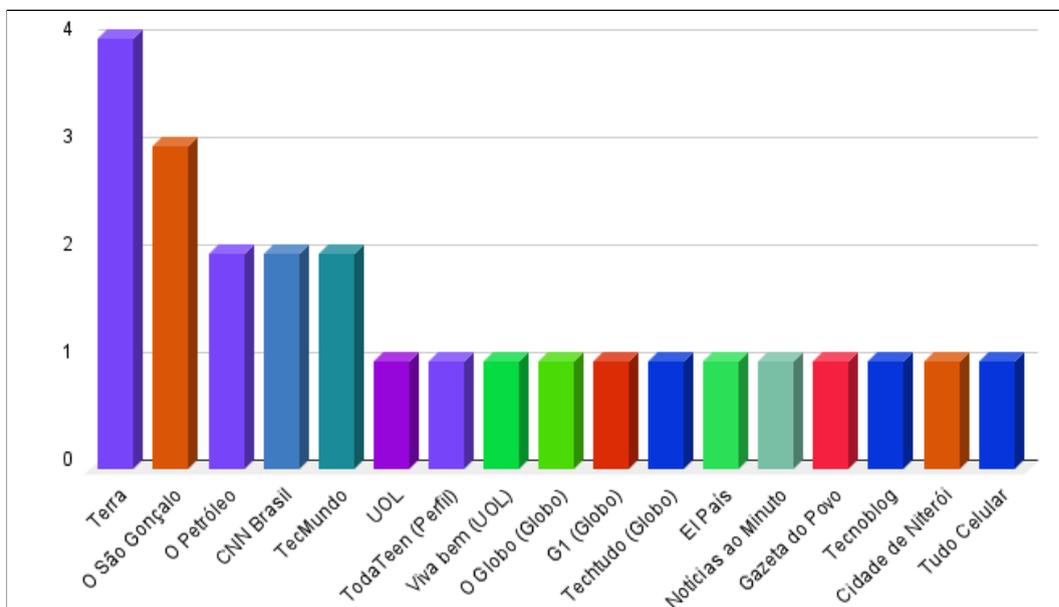
A terceira coleta das cinco primeiras notícias recomendadas ao perfil A, entre os dias 07/06/2021 e 10/06/2021, apresentou 18 veículos diferentes, de um total de 25 notícias (Gráfico 15). Em comparação com as duas amostras anteriores, essa foi a coleta com o maior número de fontes para esse perfil. O *Estadão* e o *G1*, do grupo Globo, apresentaram a maior participação absoluta na amostra ao serem recomendados por meio de duas notícias da Economia cada. Em seguida, todos os demais veículos aparecem apenas com uma notícia cada. Os veículos *PEGN* (sigla para Pequenas Empresas, Grandes Negócios), *Gaúcha ZH*, *Canaltech*, *Click Petróleo e Gás*, *CNN Brasil*, *Gran Cursos Online*, *InfoMoney*, *IstoÉ Dinheiro*, *Money Times*, *NEo Feed*, *Poder 360*, *Suno* e *Trade Map* apareceram por meio de notícias sobre economia. Já *O Globo* surgiu na amostra por meio de uma notícia sobre Coronavírus; o *Mega Curioso*, sobre Estilo de Vida, e *O Antagonista* sobre Política.

Quando investigamos a participação relativa das fontes, constatamos que os veículos *G1*, *PEGN*, *O Globo*, somam, juntos, quatro notícias para o grupo Globo. Isso configura que a organização obteve a maior participação relativa na pesquisa, ultrapassando a presença do *Estadão* que, de forma absoluta, apareceu com duas notícias na amostra. O Grupo RBS<sup>26</sup>, apesar de aparecer por meio do site *Gaúcha ZH*, foi recomendado apenas uma vez.

<sup>26</sup> Segundo site do RBS, o Grupo RBS é um dos maiores grupos empresariais multimídia do país. Entre as suas marcas está o *ZeroHora* e o *GaúchaZH*. (NOSSAS MARCAS | GRUPO RBS, [s. d.])

#### 4.2.2. Fontes recomendadas ao perfil B

**Gráfico 16 – Veículos exibidos ao perfil B entre os dias 19/04/2021 e 23/04/2021**



Fonte: Elaborado pela autora.

A primeira coleta das cinco primeiras notícias recomendadas ao perfil B, entre os dias 19/06/2021 e 23/04/2021, apresentou 17 veículos diferentes, de um total de 25 notícias (Gráfico 16). O site que obteve maior participação absoluta foi o *Terra*, um dos portais mais antigos da internet<sup>27</sup>, com quatro notícias sobre Astrologia. Em seguida apareceu o jornal *O São Gonçalo*, que, por meio da oferta de conteúdo local, obteve a segunda maior participação. Já com a terceira maior participação, cada, apareceram os sites *O Petróleo*, *CNN Brasil* e *TecMundo*. O site *O Petróleo* ofertou duas notícias sobre astrologia, enquanto a *CNN Brasil* sugeriu dois conteúdos, um sobre Tecnologia e outro sobre Estilo de Vida; e o *TecMundo*, Tecnologia e Ciência.

Os demais veículos sugeridos possuem apenas uma notícia cada, consolidando uma menor participação de cada um desses na amostra. Três veículos do grupo Globo marcaram presença com as fontes *O Globo*, *G1* e *TechTudo*, com conteúdos sobre Coronavírus, Polícia e Tecnologia, respectivamente. O grupo UOL apareceu por meio dos sites *UOL* e *VivaBem*, com

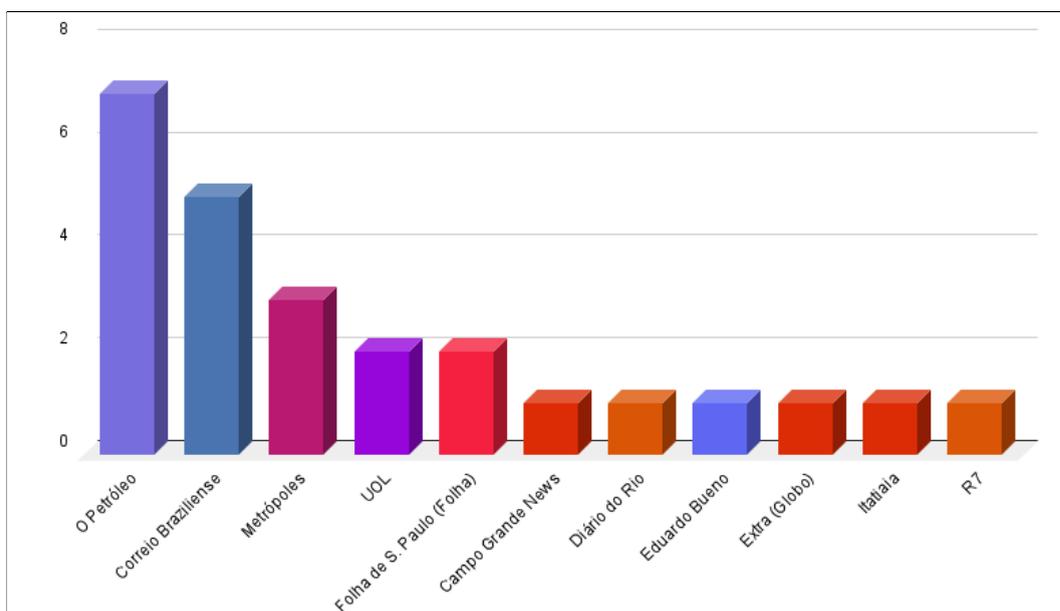
<sup>27</sup> Segundo site do Jornal do Commercio, o portal Terra foi um dos primeiros grandes portais que surgiram no início da internet. No Brasil, foi lançado em 1996, quando o provedor NutecNet foi comprado pela RBS, o que gerou o lançamento do portal gaúcho ZAZ. Em 2000, o Zaz foi comprado pela empresa espanhola Telefónica e ganhou a marca Terra. A partir da operação brasileira, a Terra deu início a expansão pela América Latina, num processo com quase duas décadas de duração. (PORTAL TERRA..., 2017)

conteúdos sobre entretenimento e saúde. O site *TodaTeen*, do grupo Perfil, apareceu por meio de uma notícia astrologia. Já o jornal *El País* apareceu com uma matéria sobre Ciência, enquanto o portal *Notícias ao Minuto* sugeriu um conteúdo de Estilo de Vida; na sequência, o *Gazeta do Povo* recomendou um artigo sobre política, seguido do *Tecnoblog*, que sugeriu uma notícia sobre tecnologia; o site *Cidade de Niterói* ofertou conteúdo local e, por fim, o *Tudo Celular* recomendou uma matéria sobre Tecnologia.

Ainda que a maior participação, tanto absoluta quanto relativa (se classificados os conteúdos por grupo de comunicação), tenha sido do *Terra*, quando contabilizamos a participação geral do grupo Globo por meio das fontes *O Globo*, *G1* e *TechTudo*, o grupo de mídia possui três notícias dentro da amostra, o que lhe daria o posto de segundo lugar.

Já o grupo UOL, por meio das fontes *UOL* e *Viva Bem*, teria a terceira maior participação na pesquisa, ao lado de outros veículos *O Petróleo*, *CNN Brasil* e *Tecmundo*.

**Gráfico 17 – Veículos exibidos ao perfil B entre os dias 17/05/2021 e 21/05/2021**



Fonte: Elaborado pela autora.

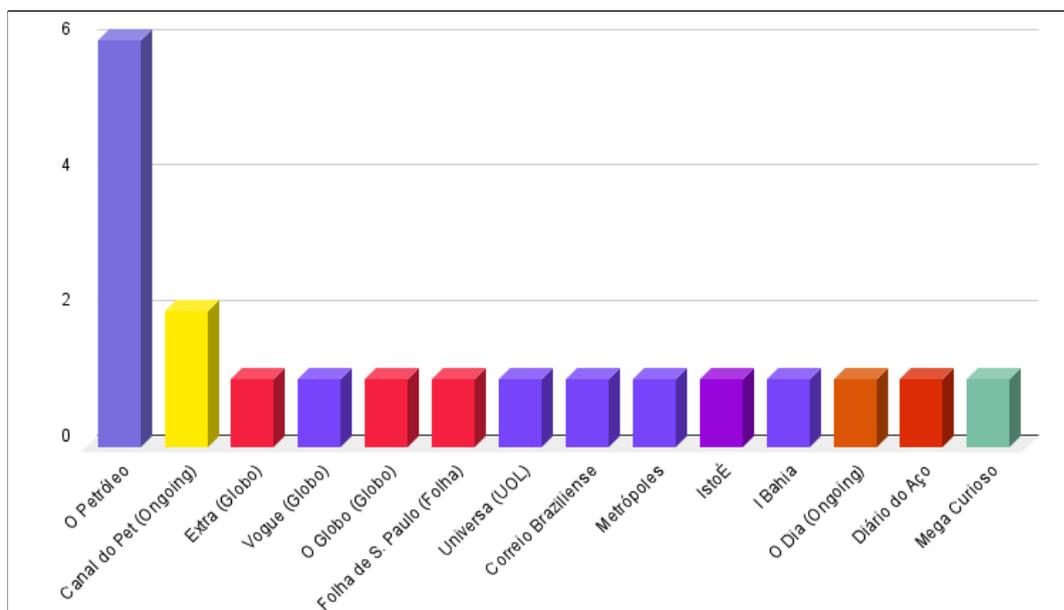
A segunda coleta das cinco primeiras notícias recomendadas ao perfil B, entre os dias 17/05/2021 e 21/05/2021, apresentou 11 veículos diferentes, de um total de 25 notícias (Gráfico 17). Em comparação com a amostra anterior, essa foi a coleta com a menor pluralidade de fontes para esse perfil. O site *O Petróleo* apresentou a maior participação absoluta na amostra ao ser recomendado por meio de sete notícias, sendo seis sobre

Astrologia e uma sobre Estilo de Vida. Em seguida, o *Correio Braziliense* aparece em segundo lugar, por meio de cinco notícias, sendo quatro sobre Astrologia e uma sobre Coronavírus.

O site *Metrópoles* aparece com três notícias, duas sendo sobre Entretenimento e uma sobre Polícia. Com duas notícias, o *UOL* apareceu na amostra com conteúdos sobre Entretenimento, enquanto o jornal *Folha de S. Paulo*, do grupo Folha, recomendou duas matérias sobre Política. Os demais veículos apareceram apenas com uma notícia cada, com destaque para as fontes *Campo Grande News*; *Extra*, do grupo Globo; *Itatiaia* e *Metrópoles*, que apareceram por meio de matérias de Polícia, enquanto os sites *Diário do Rio* e *R7* ofertaram notícias locais e o canal *Eduardo Bueno*, do YouTube, apareceu com conteúdo sobre Cultura.

Chama a atenção nessa amostra que, quando investigamos a participação relativa das fontes, observamos que o site *O Petróleo* permanece com a maior representação na amostra, mesmo que esse seja um veículo considerado não tradicional. Uma hipótese que podemos levantar é que o site pode ter se tornado um veículo de grande autoridade - conforme descreve o Google News Initiative -, quando o assunto é Astrologia.

**Gráfico 18 – Veículos exibidos ao perfil B entre os dias 07/06/2021 e 10/06/2021**



Fonte: Elaborado pela autora.

A terceira coleta das cinco primeiras notícias recomendadas ao perfil B, entre os dias

07/06/2021 e 10/06/2021, apresentou 14 veículos diferentes, de um total de 25 notícias (Gráfico 18). Em comparação com as amostras anteriores, essa coleta apresentou uma quantidade de fontes ligeiramente maior do que a segunda, porém menor que a primeira. O site *O Petróleo*, mais uma vez, apresentou a maior participação absoluta na amostra ao ser recomendado por meio de seis notícias, sendo quatro a respeito de Astrologia e duas sobre Estilo de Vida. Em seguida, o *Canal do Pet*, site do Internet Group (sob a sigla iG), que pertence ao grupo Ongoing<sup>28</sup>, aparece em segundo lugar, com duas matérias sobre Animais.

Os demais veículos aparecem por meio de uma notícia cada, o que indica que a presente amostra apresentou uma concentração maior de notícias do veículo *O Petróleo* de forma geral. O Diário do Aço apareceu por meio de uma notícia sobre Polícia; o *Mega Curioso*, com uma sobre Estilo de Vida; *O Globo*, *Extra* e *Folha de S. Paulo*, com uma notícia sobre Política cada; o *IstoÉ*, com Entretenimento; *O Dia* - também do grupo Ongoing<sup>29</sup> -, por meio de Notícias Locais e os sites *Universa*, *I Bahia* e *Vogue*, com uma notícia sobre Astrologia cada.

Quando investigamos a participação relativa das fontes, observamos que o site *O Petróleo*, assim como na coleta anterior, permanece com a maior representação, mesmo que esse seja um veículo considerado não tradicional. Ainda assim, a participação total do grupo Globo soma três notícias, por meio dos sites *Extra*, *Vogue* e *O Globo*, o que lhe daria a segunda maior participação na pesquisa. O mesmo ocorre com o grupo Ongoing, que soma três notícias, sendo duas por meio do Canal do Pet e uma de O Dia.

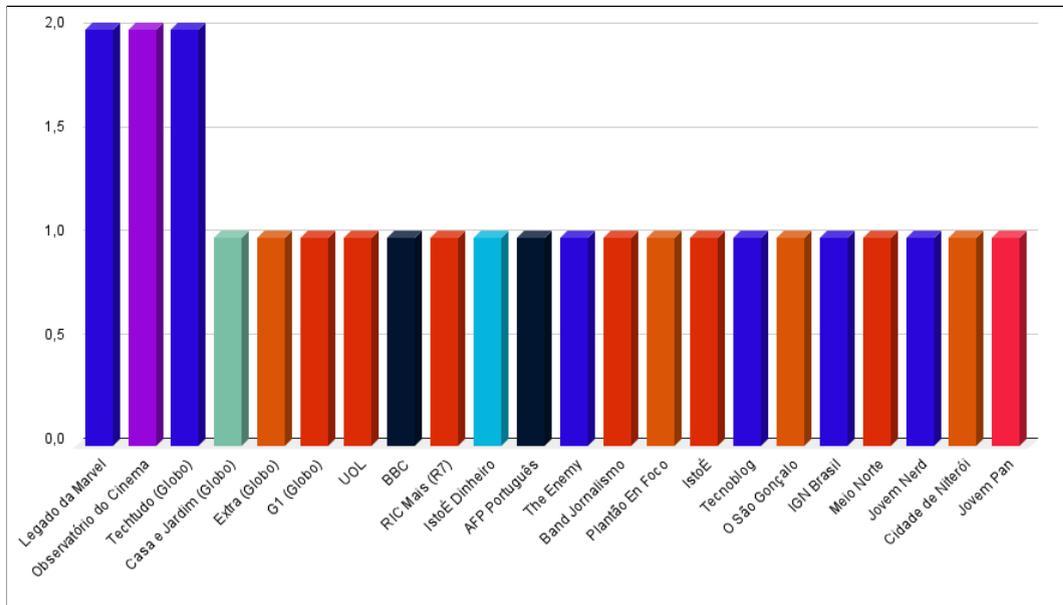
---

<sup>28</sup> Segundo notícia da *Época Negócios*, o grupo português Ongoing comprou o portal IG em 2012 com objetivo de utilizar as notícias de seus jornais Brasil Econômico, Marca, O Dia e Meia Hora no portal. (GRUPO ONGOING..., 2012)

<sup>29</sup> Em 2010, a Empresa Jornal Econômico S.A., participada do grupo Ongoing acordou a aquisição do grupo O Dia. (EMPRESA..., 2010)

### 4.2.3. Fontes recomendadas ao perfil C

**Gráfico 19 – Veículos exibidos ao perfil C entre os dias 19/04/2021 e 23/04/2021**



Fonte: Elaborado pela autora.

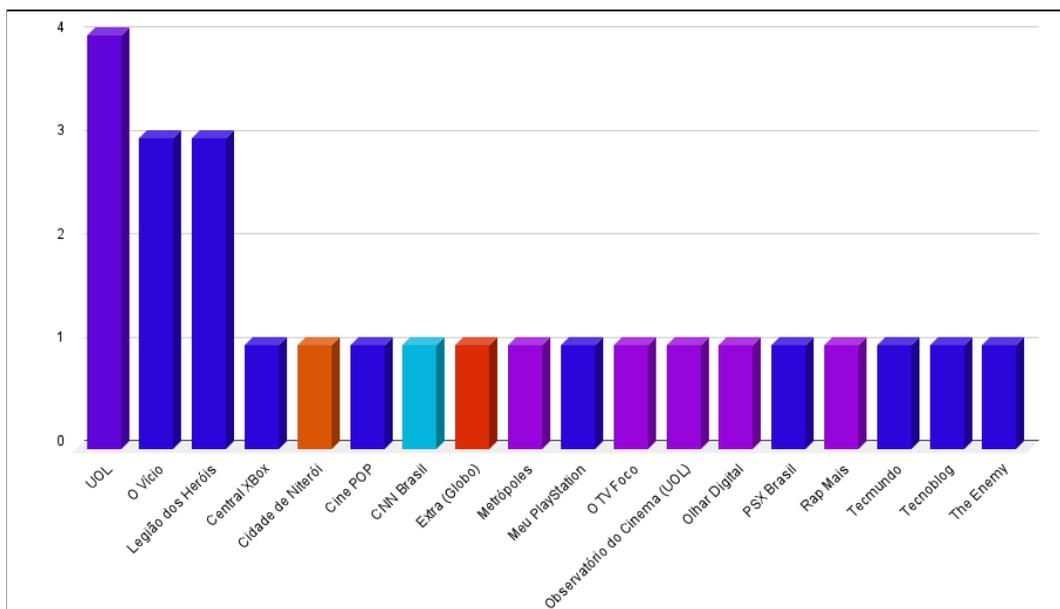
A primeira coleta das cinco primeiras notícias recomendadas ao perfil C, entre os dias 19/04/2021 e 23/04/2021, apresentou 22 veículos diferentes, de um total de 25 notícias (Gráfico 19). Os veículos *Legado da Marvel*, *Observatório do Cinema* (grupo UOL) e *TechTudo* (grupo Globo) foram exibidos a partir de duas notícias cada, o que nos faz constatar que esses três canais tiveram participação ligeiramente maior na amostra em comparação com os demais.

Convém destacar que o site *Legado da Marvel* é um site alternativo, ou seja, não vinculado à mídia tradicional. Essa fonte ofertou conteúdo sobre Cultura Geek, assim como o *TechTudo*, embora o segundo possa ser considerado sob o eixo de uma organização de mídia tradicional e o primeiro não. Já o *Observatório do Cinema* consolidou sua participação na pesquisa a partir de duas notícias sobre Entretenimento. Os demais veículos que aparecem na amostra possuem menor participação, com apenas uma notícia cada.

Quando investigamos a diversidade de veículos por grupos de comunicação, a presente coleta aponta que três matérias pertencem ao grupo UOL e cinco estão sob domínio do grupo Globo. Dessa forma, o resultado da participação relativa dos veículos aponta nessa coleta para uma diversidade ainda menor de fontes - uma vez que dois grupos (UOL e Globo) são detentores, juntos, de 32% do total das notícias recomendadas.

Caso a participação do grupo Globo seja considerada a partir da soma dos seus veículos, ele possui a maior representação dentro da pesquisa, o que lhe daria a hegemonia de primeiro lugar. Em seguida apareceria o grupo UOL, com três notícias; o *Legado da Marvel*, com duas notícias; e assim por diante.

**Gráfico 20 – Veículos exibidos ao perfil C entre os dias 17/05/2021 e 21/05/2021**



Fonte: Elaborado pela autora.

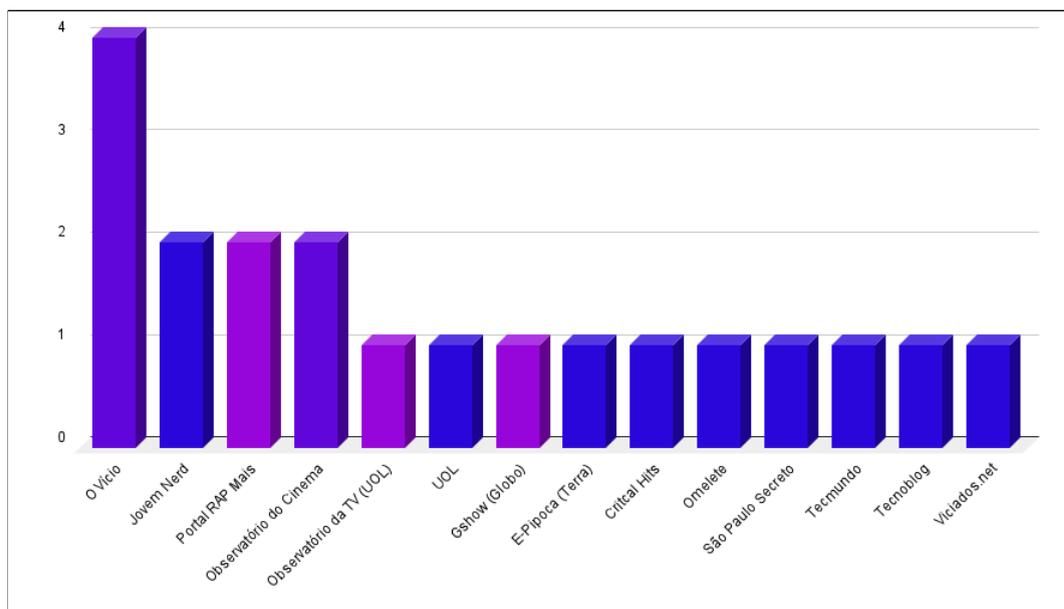
A segunda coleta das cinco primeiras notícias recomendadas ao perfil C, entre os dias 17/05/2021 e 21/05/2021, apresentou 18 veículos diferentes, de um total de 25 notícias (Gráfico 20). Em comparação com a amostra anterior, essa coleta apresentou uma quantidade de fontes ligeiramente menor. Desta vez, o *UOL* apresentou a maior participação absoluta na amostra ao ser recomendado por meio de quatro notícias, sendo duas sobre Entretenimento e duas sobre Cultura Geek. Em seguida, os sites *O Vício* e *Legião dos Heróis* apareceram cada um com duas matérias sobre Cultura Geek.

Os demais veículos aparecem por meio de uma notícia cada. As fontes *Central XBox*, *PSX Brasil*, *Cine POP*, *Meu PlayStation*, *The Enemy*, *Tecmundo* e *Tecnoblog* apareceram por meio da editoria Cultura Geek, com destaque para a participação desses veículos não tradicionais sobre o universo *geek*. Já sob a editoria Entretenimento, apareceram os sites *Observatório do Cinema*, do grupo UOL; *Olhar Digital*; *O TV Foco*; *Rap Mais*; *CNN Brasil* e *Metrópolis*. O *Extra*, do grupo Globo, apareceu por meio de Polícia e o site *Cidade de*

*Niterói*, por meio de Notícias Locais.

Quando investigamos a participação relativa das fontes, nota-se que apenas o grupo UOL sofreria alteração na quantidade de notícias sob seu escopo, uma vez que o site *Observatório do Cinema* também pertence ao grupo, totalizando cinco notícias para a organização. Chama a atenção a quantidade de veículos não-tradicionais nessa amostra, o que nos faz levantarmos a hipótese de que, sendo um perfil misto que consome bastante conteúdo no YouTube de fontes não-tradicionais como os criadores de conteúdo, a plataforma pode ter certa tendência de recomendar veículos alternativos. Outra possibilidade é que, como veículos tradicionais não possuem longas coberturas sobre o universo dos games e dos super heróis, é possível que a plataforma recomende sites especializados que possuam maior Autoridade sobre essas temáticas.

**Gráfico 21 – Veículos exibidos ao perfil C entre os dias 07/06/2021 e 10/06/2021**



Fonte: Elaborado pela autora.

A terceira coleta das cinco primeiras notícias recomendadas ao perfil C, entre os dias 07/06/2021 e 10/06/2021, apresentou 14 veículos diferentes, de um total de 25 notícias, a menor quantidade de fontes se comparado com as amostras anteriores (Gráfico 21). Desta vez, o site *O Vício* apresentou a maior participação absoluta na amostra ao ser recomendado por meio de quatro notícias, sendo três sobre Cultura Geek e uma sobre Entretenimento. Em seguida, os sites *Jovem Nerd*, *Rap Mais* e *Observatório do Cinema* apareceram cada um com

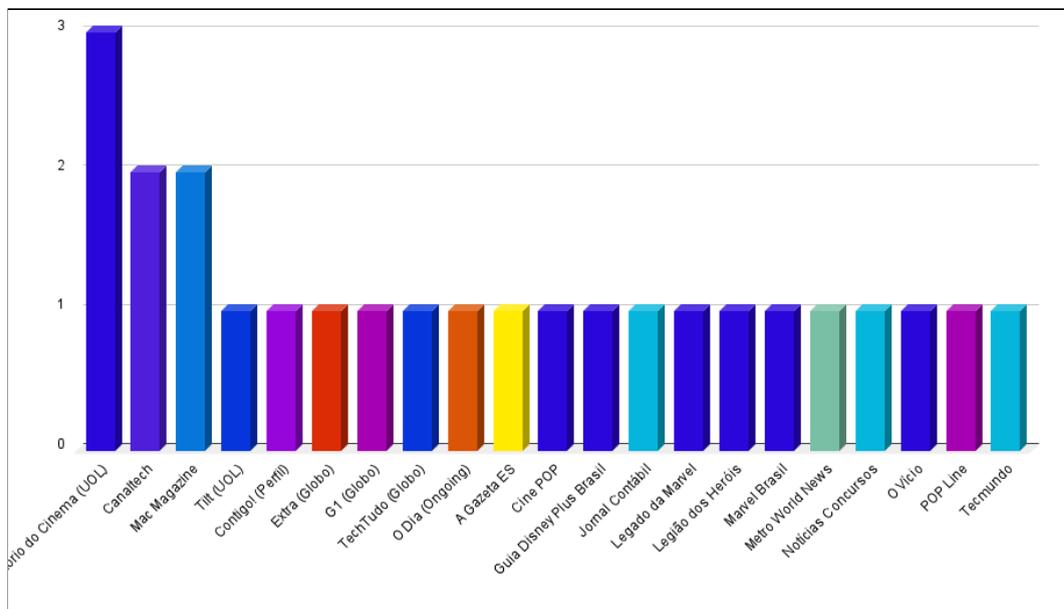
duas matérias. Enquanto o *Jovem Nerd* exibiu duas matérias sobre Cultura Geek, o *Rap Mais* apresentou duas matérias sobre Entretenimento. Já o *Observatório do Cinema*, do grupo UOL, ofereceu um conteúdo sobre Cultura Geek e outro sobre Entretenimento.

Os demais veículos aparecem por meio de uma notícia cada. As fontes *Viciados.net*, *Critical Hits*, *São Paulo Secreto*, *Tecnoblog*; *E-Pipoca*, do portal Terra, *UOL*, *Tecmundo* e *Omelete*, apareceram sob a editoria Cultura Geek. Já sob a editoria Entretenimento foram sugeridas matérias dos sites *Observatório da TV*, do grupo UOL, e *Gshow*, do grupo Globo.

Quando investigamos a participação relativa das fontes, nota-se que o grupo UOL somaria quatro notícias para a organização, sendo duas do site *Observatório do Cinema*, uma do *Observatório da TV* e uma do próprio site *UOL*. O *Gshow* e o *e-Pipoca*, ainda que pertençam aos grupos Globo e Terra, respectivamente, aparecem com apenas uma notícia cada.

#### 4.2.4. Fontes recomendadas ao perfil D

**Gráfico 22 – Veículos exibidos ao perfil D entre os dias 17/05/2021 e 21/05/2021**



Fonte: Elaborado pela autora.

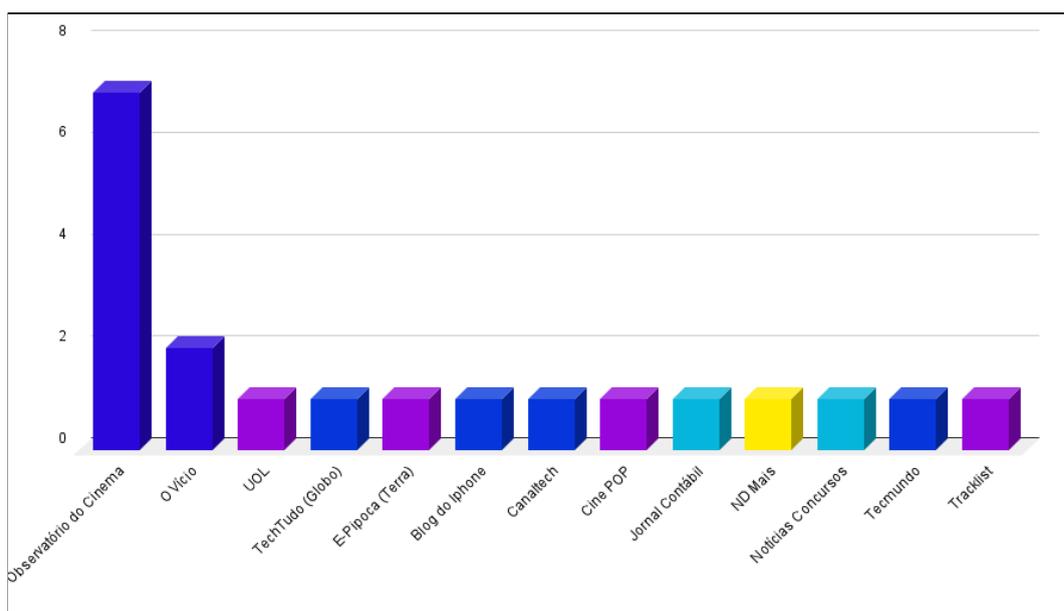
A primeira coleta das cinco primeiras notícias recomendadas ao perfil D, entre os dias 17/05/2021 e 21/05/2021, apresentou 21 veículos diferentes, de um total de 25 notícias (Gráfico 22). O site *Observatório do Cinema*, do grupo UOL, apresentou a maior participação absoluta na pesquisa. O *Canaltech* e o *Mac Magazine* apareceram cada um com duas

matérias, sendo o primeiro com notícias sobre Tecnologia e Entretenimento e o segundo com conteúdos sobre Tecnologia e Economia.

Os demais veículos aparecem por meio de uma notícia cada. As fontes *Legado da Marvel*; *O Vício*; *Guia Disney Plus Brasil*; *Marvel Brasil* (canal do YouTube), *Cine POP* e *Legião dos Heróis* apareceram sob a editoria Cultura Geek. Já sob a editoria Entretenimento foi sugerida uma matéria do site *Contigo!*, do grupo Perfil, enquanto em Economia foram recomendadas matérias do *Tecmundo*, *Jornal Contábil* e *Notícias Concursos*. Em Tecnologia, nota-se a presença dos sites *Tilt*, do grupo UOL, e *TechTudo*, do grupo Globo. Na editoria Música, foi constatada a presença dos sites *POP Line* e *G1*, sendo o segundo do grupo Globo. A editoria Polícia apareceu por meio do site *Extra*, do grupo Globo. Já sob a editoria Animais, foi recomendada uma matéria pertencente ao site *A Gazeta ES*. Por fim, em Notícias Locais foi sugerida uma matéria do jornal *O Dia*, do grupo Ongoing, e em Estilo de Vida foi recomendado um conteúdo do site *Metro World News*.

Quando investigamos a participação relativa das fontes, nota-se que o grupo UOL soma quatro notícias para a organização, sendo três do site Observatório do Cinema e uma do site Tilt. O grupo Globo contabiliza duas notícias, sendo uma do *Extra* e outra do *G1*, o que lhe daria a segunda maior participação na pesquisa ao lado das fontes *Canaltech* e *Mac Magazine*. Os demais não sofrem alteração nessa análise.

**Gráfico 23 – Veículos exibidos ao perfil D entre os dias 07/06/2021 e 10/06/2021**



Fonte: Elaborado pela autora.

A segunda coleta das cinco primeiras notícias recomendadas ao perfil D, entre os dias 07/06/2021 e 10/06/2021, apresentou 13 veículos diferentes, de um total de 25 notícias (Gráfico 23), o que aponta para uma diversidade menor de fontes em comparação com a coleta anterior. O site *Observatório do Cinema*, do grupo UOL, mais uma vez apresentou a maior participação absoluta na pesquisa, desta vez, com sete notícias sobre Cultura Geek. Em seguida, o site *O Vício* aparece com duas matérias sobre a mesma editoria.

Os demais veículos aparecem por meio de uma notícia cada. As fontes *TechTudo*, do grupo Globo; *Blog do Iphone*; *Canaltech* e *Tecmundo* apareceram sob a editoria Tecnologia. Já sob a editoria Entretenimento foi sugerida uma matéria do *UOL*; uma do *E-Pipoca*, do portal Terra; uma do *Cine POP* e uma do *Tracklist*. Também foram sugeridas duas matérias sobre Economia, sendo uma delas pertencente ao site *Jornal Contábil* e outra do *Notícias Concursos*. Por fim, na categoria Animais, foi recomendado um conteúdo do site *ND Mais*.

Quando investigamos a participação relativa das fontes, nota-se que o grupo UOL soma oito notícias para a organização, sendo sete do site *Observatório do Cinema* e uma do próprio site *UOL*. O grupo Globo contabiliza apenas uma notícia, assim como o portal Terra. Nessa avaliação, fica evidente a concentração de notícias a partir das fontes do grupo UOL para o perfil D.

#### **4.2.5 Discussão sobre os veículos recomendados aos usuários a partir dos resultados**

Os resultados obtidos a partir das amostras apontam que, para os quatro perfis, alguns veículos foram mais representados do que outros. Não é possível afirmar com clareza o que leva isso a acontecer e se isso depende de fatores como, por exemplo, o que cada um dos perfis tem o costume de ler em um determinado portal, pois tais informações não fizeram parte da coleta de dados. De todo modo, percebe-se que há um predomínio de certos veículos ao longo do tempo.

Tal constatação nos remete ao princípio de Autoridade, descrito pelo Google News Initiative, que parece ser evidentemente um fator levado em consideração, uma vez que a plataforma indica sites que possuem páginas confiáveis e parece construir determinada autoridade sobre o tema. Chama a atenção que nos perfis C e D a incidência de uma representação ligeiramente maior de veículos alternativos, o que nos leva a questionar se, por acaso, assuntos não cobertos pela mídia tradicional abrem janelas de possibilidade para que fontes não tradicionais apareçam. Isso porque nos perfis B e C, sites como *O Petróleo* e *O Vício* apareceram como os mais acessados e tratam de assuntos não convencionais.

Ainda assim, esse não parece ser o único fator preponderante para a sobre-representação de alguns veículos na participação absoluta, uma vez que o perfil A consome notícias de economia - editoria esta que costuma ter grande cobertura jornalística por parte dos veículos tradicionais - e, mesmo assim, os sites InfoMoney e Jornal Contábil, por exemplo, apareceram com certo destaque. Tal constatação nos leva a questionar também se o fato desses perfis serem especializados em um assunto, como é o caso dos supracitados, o coloca à frente dos veículos tradicionais em um certo momento, já que os algoritmos do Google podem entender que sua cobertura é mais completa a respeito de uma determinada temática em comparação com outros veículos.

Quando analisamos os resultados a partir da ótica da participação relativa dos veículos, percebemos que os veículos tradicionais aparecem de forma mais enfática nas pesquisas somadas. Isso significa que, quando analisamos a participação relativa das fontes, estes grupos de comunicação estão melhor representados do que outros que não são pertencentes a grandes grupos, já que eles podem aparecer nas pesquisas com uma distribuição de temáticas maior. Ou seja, dado o seu caráter extensivo na cobertura jornalística, o G1 pode noticiar tanto matérias sobre economia, tecnologia e saúde. Ou ainda, enquanto em uma mesma coleta o site Extra, que também pertence ao grupo Globo, recomenda uma matéria de Polícia, o site G1 recomenda uma matéria sobre Música, por exemplo. Juntos, esses sites somam uma participação maior para o domínio principal *globo.com*, o que pode talvez conferir maior autoridade para tal endereço central comum aos veículos do grupo. Isso nos leva a pensar que talvez os grandes conglomerados consigam, de forma geral, alcançar mais participações nos resultados do que os sites focalizados e específicos sobre um determinado nicho.

Por fim, cabe ainda constatar que, assim como visto na discussão sobre a diversidade de editorias, a representação de cada um dos veículos oscila ao longo do tempo. Isso confirma o que Kitchin (2017) apontou sobre os algoritmos não serem estáticos. No entanto, o motivo que leva a isso não fica claro a partir das informações que temos de *output*. Seria preciso um monitoramento mais profundo do comportamento do usuário ao longo do tempo para tentar detectar, por exemplo, se essas recomendações tiveram relação com os cliques do usuário ou não, se parte do intuito da plataforma de buscar um equilíbrio dinâmico para não deixar que o usuário consuma apenas alguns veículos, e outros possíveis fatores.

## 5 CONSIDERAÇÕES FINAIS

Diante da tempestade informacional a que estamos expostos principalmente desde a web 2.0, a promessa de entrega personalizada de notícias úteis aos leitores é, em princípio, um verdadeiro presente. Com o nosso tempo cada vez mais escasso, é compreensível que queiramos apenas o trigo (e nada de joio). Mas qual é exatamente o trigo que está sendo entregue? Se os pesquisadores em comunicação buscaram entender os processos de seleção de notícias quando estas eram exclusividade dos profissionais das redações, da mesma forma, hoje, é fundamental investigarmos o trabalho dos *gatekeepers* computacionais, ainda mais quando, mesmo estando a serviço de grandes empresas globais, se anunciam como objetivos e livres de vieses.

Mas investigar o funcionamento de algoritmos em um sistema de recomendação de notícias e suas implicações éticas, com consequências das mais variadas naturezas, esbarra em algumas limitações. Primeiro porque tal fenômeno é baseado em processos computacionais cuja programação não somente é de acesso privado da plataforma, como também, ainda que fossem acessíveis, exigiria conhecimentos específicos que vão, no geral, além da formação de quem pesquisa no campo das Ciências Humanas e Sociais e que são tão necessários para analisá-los de forma integral. Além disso, o uso de inteligência artificial por parte da plataforma neste processo de recomendação influencia os resultados da personalização do agregador de notícias, o que nos faz considerar que a análise carece de uma informação mais clara sobre como essa inteligência artificial age e o que ela se propõe a detectar. Sendo assim, revela-se um grande desafio estudar este fenômeno a partir das saídas (*outputs*) geradas pela forma como os algoritmos do Google Notícias “compreendem” o comportamento de cada um dos usuários.

Tal dificuldade leva a uma segunda limitação com relação à pesquisa, que envolve o método de coleta dos dados. A captação das notícias recomendadas se deu de forma manual, o que faz com que material de análise, que são as notícias sugeridas, represente um recorte de um espectro mais amplo. Nesse sentido, a gama de informações analisadas constitui apenas o retrato de um momento da plataforma, que não esgota a complexidade do sistema de recomendação de notícias. Por isso, pesquisas futuras devem contemplar o uso de um *script* (um programa) para ampliação do *corpus*, permitindo uma investigação contínua da evolução dos perfis e contribuindo para uma análise mais profunda sobre o que os resultados podem de fato configurar sobre a ação dos algoritmos do agregador - ainda que os dados das pesquisas

precisem lidar com a ausência de garantia de que a versão com a qual o usuário acaba de interagir a cada momento seja a mesma segundos depois, conforme destaca Kitchin (2017).

Apesar dessas restrições, analisar as notícias e os respectivos veículos recomendados a partir do método da engenharia reversa trouxe *insights* acerca da atuação dos algoritmos do Google Notícias baseada na atividade dos usuários na internet - o que se relaciona com o objetivo geral deste trabalho de refletir sobre a influência do *gatekeeping* durante o processo de mediação entre a notícia e o usuário leitor, proposto no primeiro capítulo. Ao retomarmos a discussão sobre a teoria do *gatekeeper* a partir do conceito descrito por White (1950), passando pela análise de Wolf (2003) - que considera a influência da estrutura organizacional na filtragem das notícias -, até chegar a Shoemaker e Vos (2011) - que aponta que as mudanças no ambiente comunicacional impõem uma revisão do conceito -, observamos que essa atualização enseja ser revisada considerando os novos atores inseridos neste contexto.

A partir da pesquisa, foi possível constatar que a atuação do agregador Google Notícias tem semelhanças com o trabalho exercido pelos editores e demais jornalistas cujo papel é selecionar o que vai ou não ser publicado. Mas se, no geral, as escolhas realizadas por estes profissionais baseiam-se em um *feeling* a respeito dos interesses dos leitores ou no julgamento acerca do que pode ser mais importante para a melhoria da sociedade, seja baseados na deontologia da profissão ou nas suas éticas pessoais, o que exatamente está por trás das filtragens realizadas pelos *gatekeepers* algorítmicos? Evidentemente, não foi possível responder a tudo, mas podemos dizer que há diferenças para os *gatekeepers* humanos.

É esse mecanismo que, sendo capaz de rastrear o histórico de atividade do usuário, atua sobre esse comportamento previamente identificado e decide quais serão as próximas notícias a serem oferecidas ao leitor na categoria “Para Você”, sobretudo no que se refere à personalização implícita, foco deste trabalho. Nesse sentido, é possível concluir que a ideia trazida por Negroponte (1995) de que no futuro teríamos um *Daily Me* - edição de notícias personalizada para cada leitor - foi aprimorada por Sunstein (2009), que compreendeu que as empresas poderiam aprender sobre os gostos do leitor e entregá-lo, em questão de segundos, notícias que atendam a interesses mais específicos. Dessa forma, parte do processo de controle desses filtros algorítmicos se concentrou mais nas mãos das companhias do que nos indivíduos, como imaginava Negroponte.

Ainda assim, ponderamos que o *gatekeeper* humano, presente nas redações, continua sendo relevante no processo de filtragem das notícias. Devemos considerar ainda que os sites dos jornais também podem capturar informações dos leitores (como ver onde há maior

incidência de cliques, tempo de permanência e conteúdos mais lidos, por exemplo), a partir dos mecanismos de rastreamento como o uso de *cookies e beacons*. Mas a capacidade de lidar com essas informações pelo Google Notícias tende a ser superior, uma vez que ele leva em conta os outros serviços da empresa vinculados à conta Google do usuário para aprender sobre o usuário e seus interesses, conforme descrevemos no primeiro capítulo. Sendo assim, chegamos à conclusão de que o Google Notícias atua como um “meta gatekeeper”, ou seja, ele se torna mais um portão pelo qual passa as notícias fluem até chegarem aos leitores.

Essa atuação do *gatekeeper* algorítmico enquanto mecanismo que filtra as notícias em uma determinada plataforma no ambiente digital está permeada de contestações, conforme ressaltam os autores Pariser (2012) e Sunstein (2009) ao detalharem que os algoritmos voltados para a personalização podem ser prejudiciais à diversidade de ideias e, conseqüentemente, para a democracia. No exposto por Eli Pariser, esses filtros poderiam ainda levar à formação de bolhas, pois, a partir da atuação dos algoritmos como verdadeiros garçons, o usuário só visualizaria aquilo que lhe agrada. Nosso objetivo específico foi identificar se estes apontamentos poderiam ser visualizados por meio dos resultados obtidos através da coleta de dados dos perfis.

Quando analisadas as editorias recomendadas, os resultados apontam que a personalização indica que houve uma redução de diversidade nos quatro perfis analisados diante do predomínio de alguns assuntos em detrimento de outros. Isso pôde ser visto tanto no Apêndice A como nos Gráficos, do 2 ao 12. O mesmo foi observado com relação aos veículos sugeridos, em que alguns apareceram com maior frequência, conforme ilustram os Gráficos, do 13 ao 23. No entanto, essas editorias e veículos que obtiveram maior relevância não tiveram a participação quantitativa constante ao longo das três coletas da pesquisa. Ou seja, os conteúdos recomendados e suas respectivas editorias não foram quantitativamente estáticos à medida que a coleta ocorreu.

Não fica claro, contudo, até que ponto essa oscilação tem relação com uma possível ação do usuário - que pode ter mudado os seus interesses ao longo do tempo -, ou se isso também parte, em alguma medida, de uma tentativa da plataforma de sugerir conteúdos não vinculados aos interesses do usuário para combater a falta de serendipidade e não deixá-lo preso em um filtro-bolha, conforme sugerido no capítulo anterior e evidenciado pela plataforma, como descrito no primeiro capítulo. Não foram analisados os históricos de atividade dos usuários porque esbarraríamos em uma questão de privacidade dos usuários que já contribuíram gentilmente para a pesquisa. Portanto, não foi possível realizar essa

diferenciação.

Embora tenha sido observada uma redução significativa na diversidade de notícias e veículos, não é possível apontar um padrão de filtro-bolha pelos perfis que acessaram o Google Notícias. Conforme evidencia o Apêndice A, o perfil A teve 77% de notícias voltadas para Economia somando todos os períodos analisados, enquanto o perfil B teve 66% voltado para Astrologia neste tempo; o perfil C teve 53% voltado para Cultura Geek e o perfil D, 36% voltado para esta editoria. Isso nos leva a concluir que o processo de *gatekeeping* pelo agregador Google Notícias possui resultados diferentes em cada perfil, o que exige que cada usuário deva ser analisado individualmente, já que não há por padrão um filtro-bolha.

Nesse sentido, chegamos à mesma conclusão apresentada por Loiola (2018), cuja pesquisa aponta para a falta de um padrão claro sobre a formação de filtro-bolhas, embora tenha sido visualizada uma redução da diversidade de conteúdos nos perfis analisados. Apesar de o autor ter utilizado objeto e metodologia diferentes, sua investigação sobre filtro-bolha no YouTube sugere que os filtros algorítmicos utilizados entre esses produtos do Google possuem mecanismos semelhantes.

De toda forma, a presente pesquisa evidencia a importância de investigar o funcionamento de algoritmos e não esgota os questionamentos acerca da redução da diversidade e das discussões sobre o fenômeno do filtro-bolha. Um futuro trabalho poderia investigar mais a fundo o aparecimento de notícias tendo como base as recomendações feitas a um determinado perfil na categoria personalizada “Para Você” a partir da identificação de possíveis critérios para a presença desses conteúdos, como por exemplo: métricas que tratem da quantidade de *pageviews* desses links, o tempo de permanência dos visitantes nessas páginas e o número de compartilhamentos dessas páginas em redes sociais. A partir disso poderia ser estabelecido, por exemplo, um indicador de relevância para as notícias recomendadas. Isso poderia trazer uma noção mais clara dos critérios que o blog do Google News Initiative detalhou como relevantes, buscando entender como a plataforma efetivamente pondera esses fatores descritos em seu site.

Também aconselha-se para futuras pesquisas a análise dos títulos das notícias recomendadas aos usuários sob a perspectiva do conceito do *clickbait*, buscando verificar se títulos de notícias que possuam maior apelo e sejam considerados “caça-cliques” possam, de fato, virem a ser mais clicados pelos usuários e, portanto, serem mais recomendados pelas plataformas. Nesse sentido, estimula-se também uma investigação sobre os possíveis efeitos do agregador Google Notícias no que se refere à produção jornalística, uma vez que o

jornalismo passou a lidar com os novos regimes de visibilidade propostos pelos algoritmos, conforme evidenciou Barsotti (2019). Seria preciso investigar se os meios de comunicação podem estar operando sob a lógica desses algoritmos para construir suas reputações e como isso impacta a construção dos materiais posteriormente publicados e direcionados aos leitores.

É sugerido ainda que uma próxima pesquisa investigue se a diversidade de veículos jornalísticos recomendados aos usuários é capaz de, efetivamente, tornar o consumo de notícias dos leitores mais eclético e diverso em termos de conteúdo - ainda que essa diversidade não se dê apenas pelos veículos. Mas, dado que o ambiente on-line possibilita que um cidadão tenha mais facilidade de acessar diferentes jornais, investigar se o conteúdo dos veículos apresentados seria, de fato, mais diverso e plural, pode apontar para novos horizontes à luz dos questionamentos sobre filtros-bolha e democracia.

Por fim, espera-se que o trabalho tenha clarificado algumas perspectivas acerca do funcionamento do algoritmo do Google Notícias e sua implicação para personalização implícita sobre as notícias. Concluimos que a pesquisa enseja discussões a respeito da atuação dos *gatekeepers* algorítmicos neste ambiente comunicacional digital, onde há a presença de infomediários como o Google Notícias, e, entre as mais diversas perguntas, nos contentamos, por enquanto, em fazer estas: Até que ponto esses *gatekeepers* nos ajudam a encontrar diferentes pontos de vista? Há algum nível de serendipidade nessas plataformas? E, se existe, ela é suficiente?

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

ÁGUILA-OBRA, A. R. del; PADILLA-MELÉNDEZ, A.; SERAROLS-TARRES, C. Value creation and new intermediaries on Internet. An exploratory analysis of the online news industry and the web content aggregators. **International Journal of Information Management**, [S. l.], v. 27, p. 187–199, jul. 2007. Acesso em: 23 maio 2021.

AMARAL, L. **A objetividade jornalística**. 1. ed. Porto Alegre: Sagra - D.C. Luzzato, 1996 (Coleção Comunicação Viva).

BARSOTTI, A. Quando a visibilidade da notícia depende dos algoritmos: os riscos para a sociedade. **Trama: indústria criativa em revista**, [S. l.], v. 8, n. 1, p. 44–63, 2019. Acesso em: 17 abr. 2021.

CASTELLS, M. **A sociedade em rede**. trad. Roneide Venâncio Majer. 21. ed. São Paulo: Paz e Terra, 2020. v. 1, (A era da informação: economia, sociedade e cultura).

COMO AS MATÉRIAS DO GOOGLE NOTÍCIAS SÃO SELECIONADAS - AJUDA DO GOOGLE NEWS. 2021. **Google**. Disponível em: <<https://support.google.com/googlenews/answer/9005749?hl=pt-BR>>. Acesso em: 30 maio 2021.

CONHEÇA O GRUPO FOLHA. [s. d.]. **Folha de S.Paulo - Grupo Folha**. Disponível em: <<http://www1.folha.uol.com.br/institucional/>>. Acesso em: 9 nov. 2021.

CURTISS, M.; BHARAT, K. A.; SCHMITT, M. **Systems and methods for improving the ranking of news articles**. Mountain View: [s. n.], 29 out. 2019. Disponível em: <<https://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=10,459,926.PN.&OS=PN/10,459,926&RS=PN/10,459,926>>. Acesso em: 24 jul. 2021.

DIAKOPOULOS, N. Algorithmic Accountability: Journalistic investigation of computational power structures. **Digital Journalism**, [S. l.], v. 3, n. 3, p. 398–415, 4 maio 2015. Disponível em: <<http://www.tandfonline.com/doi/full/10.1080/21670811.2014.976411>>. Acesso em: 18 jul. 2021.

DOCTOR, K. **Newsonomics: Doze Novas Tendências Que Moldarão As Notícias E O Seu Impacto Na Economia Mundial**. São Paulo: Cultrix, 2011.

EMPRESA DO BRASIL ECONÔMICO REFORÇA INVESTIMENTO COM AQUISIÇÃO DO GRUPO O DIA. [S. l.], 1 abr. 2010. Disponível em: <[https://web.archive.org/web/20150924060426/http://www.ongoing.com/pdf/Aquisicao\\_do\\_Grupo\\_O\\_DIA\\_01ABR10\\_PT.pdf](https://web.archive.org/web/20150924060426/http://www.ongoing.com/pdf/Aquisicao_do_Grupo_O_DIA_01ABR10_PT.pdf)>. Acesso em: 9 nov. 2021.

EXIBIR CONTEÚDO ÚTIL E RELEVANTE. [s. d.]. **How News Works**. Disponível em: <[https://newsinitiative.withgoogle.com/hownewsworks/pt\\_br/approach/surfacing-useful-and-relevant-content/](https://newsinitiative.withgoogle.com/hownewsworks/pt_br/approach/surfacing-useful-and-relevant-content/)>. Acesso em: 12 out. 2021.

GERENCIAR O HISTÓRICO DE LOCALIZAÇÃO - AJUDA DO CONTA DO GOOGLE.

20-? Disponível em: <<https://support.google.com/accounts/answer/3118687?hl=pt>>. Acesso em: 29 ago. 2021.

GILLESPIE, T. The Relevance of Algorithms. *In*: GILLESPIE, T.; BOCZKOWSKI, P. J.; FOOT, K. A. (org.). **Media technologies: essays on communication, materiality, and society**. Inside technology. Cambridge, Massachusetts: MIT Press, 2014. p. 167–194. Disponível em:

<[https://www.microsoft.com/en-us/research/wp-content/uploads/2014/01/Gillespie\\_2014\\_The-Relevance-of-Algorithms.pdf](https://www.microsoft.com/en-us/research/wp-content/uploads/2014/01/Gillespie_2014_The-Relevance-of-Algorithms.pdf)>. Acesso em: 30 maio 2021.

GINGRAS, R. **A look at how news at Google works**. 6 maio 2019. **Google**. Disponível em: <<https://blog.google/products/news/look-how-news-google-works/>>. Acesso em: 24 jun. 2021.

GLOBAL MARKET SHARE OF SEARCH ENGINES 2010-2021. 8 out. 2021. **Statista**. Disponível em: <<https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/>>. Acesso em: 26 out. 2021.

GRUPO GLOBO. [s. d.]. Disponível em: <<https://grupoglobo.globo.com/>>. Acesso em: 9 nov. 2021.

GRUPO ONGOING CONFIRMA A COMPRA DO PORTAL IG. 2012. **Época Negócios**. Disponível em: <<https://epocanegocios.globo.com/Informacao/Acao/noticia/2012/04/grupo-ongoing-confirma-compra-do-portal-ig.html>>. Acesso em: 9 nov. 2021.

HAIM, M.; GRAEFE, A.; BROSIUS, H.-B. Burst of the Filter Bubble? Effects of personalization on the diversity of Google News. **Digital Journalism**, [S. l.], v. 6, n. 3, p. 330–343, 16 mar. 2018. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/21670811.2017.1338145>>.

IN RACE AGAINST FAKE NEWS, GOOGLE AND FACEBOOK STROLL TO THE STARTING LINE (PUBLISHED 2017). 25 jan. 2017. Disponível em: <<https://www.nytimes.com/2017/01/25/technology/google-facebook-fake-news.html>>. Acesso em: 12 out. 2021.

KAMVAR, S. **Search gets personal**. 28 jun. 2005. **Official Google Blog**. Disponível em: <<https://googleblog.blogspot.com/2005/06/search-gets-personal.html>>. Acesso em: 30 maio 2021.

KITCHIN, R. Thinking critically about and researching algorithms. **Information, Communication & Society**, [S. l.], v. 20, n. 1, p. 14–29, 2 jan. 2017. Disponível em: <<http://futuredata.stanford.edu/classes/cs345s/handouts/kitchin.pdf>>.

KRAZIT, T. Google News turns 8 (images). 22 set. 2010. **CNET**. Disponível em: <<https://www.cnet.com/pictures/google-news-turns-8-images/>>. Acesso em: 5 jun. 2021.

LÉVY, P. **Cibercultura**. trad. Carlos Irineu Da Costa. São Paulo: Ed. 34, 1999. . Acesso em:

15 jan. 2021.

LOIOLA, D. F. E. **Recomendado para você: O impacto do algoritmo do YouTube na formação de bolhas**. 2018. 165 f. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Minas Gerais, 2018. Disponível em:

<[https://repositorio.ufmg.br/bitstream/1843/BUOS-B6GEZC/1/disserta\\_o\\_daniel\\_loiola\\_final\\_.pdf](https://repositorio.ufmg.br/bitstream/1843/BUOS-B6GEZC/1/disserta_o_daniel_loiola_final_.pdf)>. Acesso em: 15 set. 2021.

MCCAIN, J.; GENTILE, B.; CAMPBELL, W. K. A Psychological Exploration of Engagement in Geek Culture. **PloS one**, [S. l.], v. 10, p. 38, 18 nov. 2015.

NAVEGAR NO GOOGLE CHROME COMO VISITANTE - COMPUTADOR - AJUDA DO GOOGLE CHROME. [s. d.]. Disponível em:

<<https://support.google.com/chrome/answer/6130773?hl=pt-BR&co=GENIE.Platform%3DDesktop>>. Acesso em: 9 ago. 2021.

NEGROPONTE, N. **A vida digital**. trad. Sérgio Tellaroli. São Paulo: Companhia das Letras, 1995.

NOSSA MISSÃO É ORGANIZAR AS INFORMAÇÕES DO MUNDO PARA QUE SEJAM UNIVERSALMENTE ACESSÍVEIS E ÚTEIS PARA TODOS. [s. d.]. Disponível em:

<[https://about.google/intl/ALL\\_br/](https://about.google/intl/ALL_br/)>. Acesso em: 17 abr. 2021.

NOSSAS MARCAS | GRUPO RBS. [s. d.]. Disponível em:

<<http://www.gruporbs.com.br/areas-de-atuacao/>>. Acesso em: 9 nov. 2021.

OFFICIAL GOOGLE BLOG: PERSONALIZED SEARCH FOR EVERYONE. 4 dez. 2009. Disponível em:

<<https://webcache.googleusercontent.com/search?q=cache:io1trgl7rwYJ:https://googleblog.blogspot.com/2009/12/personalized-search-for-everyone.html+&cd=1&hl=pt-BR&ct=clnk&gl=br>>. Acesso em: 1 ago. 2021.

OLIVEIRA, S. M. V. de. O jornal Zero Hora e as eleições de 2002: A construção do discurso anti-petista no Jornal Zero Hora, durante as eleições de 2002, para o governo do Estado do Rio Grande do Sul. In: XXVI CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, a 6/09 2003. Belo Horizonte: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, a 6/09 2003. p. 1–14. Disponível em:

<<http://www.portcom.intercom.org.br/pdfs/8557120044766284664620594929986934554.pdf>>.

ORGANIZAR AS NOTÍCIAS DE TODA A WEB. [s. d.]. **How News Works**. Disponível em:

<[https://newsinitiative.withgoogle.com/hownewsworks/pt\\_br/approach/organizing-news-from-around-the-web/](https://newsinitiative.withgoogle.com/hownewsworks/pt_br/approach/organizing-news-from-around-the-web/)>. Acesso em: 12 out. 2021.

OUR APPROACH. [s. d.]. **How News Works**. Disponível em:

<<https://newsinitiative.withgoogle.com/hownewsworks/approach/>>. Acesso em: 12 out. 2021.

PAGE, L. **The PageRank Citation Ranking: Bringing Order to the Web**. California:

Stanford Digital Library Project, 18 ago. 1997. Disponível em:  
<<https://web.archive.org/web/20020506051802/http://www-diglib.stanford.edu/cgi-bin/WP/get/SIDL-WP-1997-0072?1>>. Acesso em: 30 maio 2021.

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

PASQUALE, F. **The black box society: the secret algorithms that control money and information.** Cambridge: Harvard University Press, 2015.

PORTAL TERRA IRÁ ENCERRAR OPERAÇÕES PELO MUNDO. 28 jun. 2017. **JC Online.** Disponível em:  
<<https://jc.ne10.uol.com.br/canal/tecnologia/noticia/2017/06/28/portal-terra-ira-encerrar-operacoes-pelo-mundo-292322.php>>. Acesso em: 9 nov. 2021.

SANDVIG, C. *et al.* Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. *In*: 64TH ANNUAL MEETING OF THE INTERNATIONAL COMMUNICATION ASSOCIATION, 22 maio 2014., pré-conferência. **Data and Discrimination: Converting Critical Concerns into Productive Inquiry** [...]. Seattle: [s. n.], 22 maio 2014. p. 23. Disponível em:  
<<http://www-personal.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf>>. Acesso em: 30 maio 2021.

SHIELDS, M. **Building Google News for everyone — a retrospective.** 18 set. 2018. **Google.** Disponível em:  
<<https://blog.google/products/news/building-google-news-everyone-retrospective/>>. Acesso em: 4 jun. 2021.

SHOEMAKER, P. J.; VOS, T. P. **Teoria do gatekeeping: seleção e construção da notícia.** trad. Vivian Nickel. Porto Alegre: Penso, 2011.

SILVA, C. E. L. da. **O adiantado da hora: a influência americana sobre o jornalismo brasileiro.** São Paulo, SP: Summus Editorial, 1990(Novas buscas em comunicação, vol. 39).

SOBRE UOL: CONHEÇA A MAIOR EMPRESA BRASILEIRA DE CONTEÚDO, TECNOLOGIA, SERVIÇOS E MEIOS DE PAGAMENTOS DIGITAIS. [s. d.]. Disponível em: <<https://sobreuol.noticias.uol.com.br/>>. Acesso em: 9 nov. 2021.

STRIPHAS, T. Algorithmic culture. **European Journal of Cultural Studies**, [S. l.], v. 18, n. 4–5, p. 395–412, 1 ago. 2015. Disponível em: <<https://doi.org/10.1177/1367549415577392>>. Acesso em: 7 jun. 2020.

SUNSTEIN, C. R. **Republic.com 2.0.** 2. print., and 1. paperback print. Princeton, N.J: Princeton Univ. Press, 2009.

THE WORLD'S MOST VALUABLE RESOURCE IS NO LONGER OIL, BUT DATA | THE ECONOMIST. **The Economist**, [S. l.], London, 6 maio 2017. Disponível em:  
<<https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-lo>>

[nger-oil-but-data](#)>. Acesso em: 17 abr. 2021.

THURMAN, N.; SCHIFFERES, S. The future of personalization at news websites: lessons from a longitudinal study. **Journalism Studies**, [S. l.], v. 13, n. 5–6, p. 775–790, 1 out. 2012. Disponível em: <<https://doi.org/10.1080/1461670X.2012.664341>>. Acesso em: 8 jan. 2021.

UPSTILL, T. The new Google News: AI meets human intelligence. 8 maio 2018. **Google**. Disponível em: <<https://blog.google/products/news/new-google-news-ai-meets-human-intelligence/>>. Acesso em: 4 jun. 2021.

VAIDHYANATHAN, S. **A Googlelização de Tudo: (e por que devemos nos preocupar): a ameaça do controle total da informação por meio da maior e mais bem-sucedida empresa do mundo virtual**. trad. Jeferson Luiz Camargo. São Paulo: Cultrix, 2011.

WHITE, D. M. The “Gate Keeper”: A Case Study in the Selection of News. **Journalism Quarterly**, [S. l.], v. 27, n. 4, p. 383–390, set. 1950. Disponível em: <<http://journals.sagepub.com/doi/10.1177/107769905002700403> . Acesso em: 4 ago. 2021.

WOLF, M. **Teorias da comunicação**. Lisboa: Presença, 2003. Disponível em: <<https://www.inovaconsulting.com.br/wp-content/uploads/2016/09/teorias-da-comunicacao-by-mauro-wolf.pdf>>. Acesso em: 15 jan. 2021.

WU, T. **Impérios da comunicação: do telefone à internet, da AT&AT ao Google**. trad. Claudio Carina. Rio de Janeiro: Zahar, 2012.

## 7 APÊNDICES

### 7.1 Apêndice A - Participação das editorias de acordo com os perfis

Editorias	PERFIL A							
	19 a 23/04	%	17 a 21/05	%	07 a 10/06	%	Todos os períodos	%
Economia	15	60%	22	88%	17	85%	54	77%
Tecnologia	2	8%	1	4%	0	0%	3	4%
Internacional	2	8%	0	0%	0	0%	2	3%
Coronavírus	2	8%	0	0%	1	5%	3	4%
Notícias Locais	1	4%	0	0%	0	0%	1	1%
Política	1	4%	0	0%	1	5%	2	3%
Polícia	1	4%	2	8%	0	0%	3	4%
Educação	1	4%	0	0%	0	0%	1	1%
Estilo de Vida	0	0%	0	0%	1	5%	1	1%

Total	25	100%	25	100%	20	100%	70	100%
-------	----	------	----	------	----	------	----	------

**PERFIL B**

Editorias	19 a 23/04	%	17 a 21/05	%	07 a 10/06	%	Todos os períodos	%
Astrologia	7	28%	22	88%	17	85%	46	66%
Tecnologia	5	20%	1	4%	0	0%	6	9%
Estilo de Vida	2	8%	0	0%	0	0%	2	3%
Ciência	2	8%	0	0%	1	5%	3	4%
Política	1	4%	0	0%	0	0%	1	1%
Coronavírus	1	4%	0	0%	1	5%	2	3%
Entretenimento	1	4%	2	8%	0	0%	3	4%
Notícias Locais	4	16%	0	0%	0	0%	4	6%
Saúde	1	4%	0	0%	1	5%	2	3%
Polícia	1	4%	0	0%	0	0%	1	1%
Total	25	100%	25	100%	20	100%	70	100%

**PERFIL C**

Editorias	19 a 23/04	%	17 a 21/05	%	07 a 10/06	%	Todos os períodos	%
Cultura Geek	8	32%	15	60%	14	70%	37	53%
Polícia	6	24%	1	4%	0	0%	7	10%
Internacional	2	8%	0	0%	0	0%	2	3%
Entretenimento	2	8%	7	28%	6	30%	15	21%
Economia	1	4%	1	4%	0	0%	2	3%
Notícias Locais	4	16%	1	4%	0	0%	5	7%
Estilo de Vida	1	4%	0	0%	0	0%	1	1%
Política	1	4%	0	0%	0	0%	1	1%
Total	25	100%	25	100%	20	100%	70	100%

**PERFIL D**

Editorias	19 a 23/04	%	17 a 21/05	%	07 a 11/06	%	Todos os períodos	%
-----------	---------------	---	---------------	---	---------------	---	----------------------	---

Cultura Geek	0		9	36%	7	35%	16	36%
Economia	0		4	16%	2	10%	6	13%
Tecnologia	0		4	16%	4	20%	8	18%
Polícia	0		1	4%	0	0%	1	2%
Entretenimento	0		2	8%	6	30%	8	18%
Animais	0		1	4%	1	5%	2	4%
Música	0		2	8%	0	0%	2	4%
Notícias Locais	0		1	4%	0	0%	1	2%
Estilo de Vida	0		1	4%	0	0%	1	2%
Total	0		25	100%	20	100%	45	100%

## 7.2 Apêndice B - Participação dos veículos de acordo com os perfis

PERFIL A								
Fontes	19 a 23/04	%	17 a 21/05	%	07 a 10/06	%	Todos os períodos	%
InfoMoney	4	16%	2	9%	1	5%	7	10%
Jornal Contábil	3	12%	0		0		3	4%
O Globo (Globo)	3	12%	0		1	5%	4	6%
G1 (Globo)	2	8%	4	18%	2	10%	8	12%
BBC	2	8%	0		0		2	3%
TechTudo (Globo)	1	4%	0		0		1	1%
Valor Econômico (Globo)	1	4%	2	9%	0		3	4%
Veja (Abril)	1	4%	0		0		1	1%
UOL	1	4%	2	9%	0		3	4%
Direção Concursos	1	4%	0		0		1	1%
Gazeta do Povo	1	4%	0		0		1	1%
Diário do Centro do Mundo	1	4%	0		0		1	1%
Contábeis	1	4%	0		0		1	1%
TradeMap	1	4%	0		1	5%	2	3%
TecnoBlog	1	4%	0		0		1	1%
Exame	1	4%	0		0		1	1%
CNN Brasil	0		2	9%	1	5%	3	4%

Money Times	0		2	9%	1	5%	3	4%
Neo Feed	0		2	9%	1	5%	3	4%
Canal Rural	0		1	5%	0		1	1%
Extra (Globo)	0		1	5%	0		1	1%
Folha de São Paulo (Folha)	0		1	5%	0		1	1%
Olhar Digital	0		1	5%	0		1	1%
Poder360	0		1	5%	1	5%	2	3%
R7	0		1	5%	0		1	1%
Estadão	0		0		2	10%	2	3%
PEGN (Globo)	0		0		1	5%	1	1%
Gaúcha ZH (RBS)	0		0		1	5%	1	1%
Canaltech	0		0		1	5%	1	1%
Click Petróleo e Gás	0		0		1	5%	1	1%
Gran Cursos Online	0		0		1	5%	1	1%
IstoÉ Dinheiro	0		0		1	5%	1	1%
Mega Curioso	0		0		1	5%	1	1%
O Antagonista	0		0		1	5%	1	1%
Suno	0		0		1	5%	1	1%
TradeMap	0		0		1	5%	1	1%
Total	25	100%	22	100%	21	100%	68	100%

PERFIL B								
Fontes	19 a 23/04	%	17 a 21/05	%	07 a 10/06	%	Todos os períodos	%
Terra	4	16%	0		0		4	6%
O São Gonçalo	3	12%	0		0		3	4%
O Petróleo	2	8%	7	28%	6	30%	15	21%
CNN Brasil	2	8%	0		0		2	3%
TecMundo	2	8%	0		0		2	3%
UOL	1	4%	0		0		1	1%

Viva bem (UOL)	1	4%	0		0		1	1%
TodaTeen (Perfil)	1	4%	0		0		1	1%
O Globo (Globo)	1	4%	0		1	5%	2	3%
G1 (Globo)	1	4%	0		0		1	1%
TechTudo (Globo)	1	4%	0		0		1	1%
El País	1	4%	0		0		1	1%
Notícias ao Minuto	1	4%	0		0		1	1%
Gazeta do Povo	1	4%	0		0		1	1%
Tecnoblog	1	4%	0		0		1	1%
Cidade de Niterói	1	4%	0		0		1	1%
Tudo Celular	1	4%	0		0		1	1%
Correio Braziliense	0		5	20%	1	5%	6	9%
Metrópoles	0		3	12%	1	5%	4	6%
UOL	0		2	8%	0		2	3%
Folha de S. Paulo (Folha)	0		2	8%	1	5%	3	4%
Campo Grande News	0		1	4%	0		1	1%
Diário do Rio	0		1	4%	0		1	1%
Eduardo Bueno	0		1	4%	0		1	1%
Extra (Globo)	0		1	4%	1	5%	2	3%
Itatiaia	0		1	4%	0		1	1%
R7	0		1	4%	0		1	1%
Canal do Pet (Ongoing)	0		0		2	10%	2	3%
Vogue (Globo)	0		0		1	5%	1	1%
Universa (UOL)	0		0		1	5%	1	1%
IstoÉ	0		0		1	5%	1	1%
I Bahia	0		0		1	5%	1	1%
O Dia (Ongoing)	0		0		1	5%	1	1%
Diário do Aço	0		0		1	5%	1	1%
Mega Curioso	0		0		1	5%	1	1%
Total	25	100%	25	100%	20	100%	70	100%

PERFIL C								
Fontes	19 a 23/04	%	17 a 21/05	%	07 a 10/06	%	Todos os períodos	%
Legado da Marvel	2	8%	0		1	5%	3	8%
Observatório do Cinema (UOL)	2	8%	1	4%	2	10%	5	14%
Techtudo (Globo)	2	8%	0		0		2	5%
Casa e Jardim (Globo)	1	4%	0		0		1	3%
Extra (Globo)	1	4%	1	4%	0		2	5%
G1 (Globo)	1	4%	0		0		1	3%
UOL	1	4%	4	17%	1	5%	6	16%
BBC	1	4%	0		0		1	3%
RIC Mais (R7)	1	4%	0		0		1	3%
IstoÉ Dinheiro	1	4%	0		0		1	3%
AFP Português	1	4%	0		0		1	3%
The Enemy	1	4%	0		0		1	3%
Band Jornalismo	1	4%	0		0		1	3%
Plantão En Foco	1	4%	0		0		1	3%
IstoÉ	1	4%	0		0		1	3%
Tecnoblog	1	4%	1	4%	1	5%	3	8%
O São Gonçalo	1	4%	0		0		1	3%
IGN Brasil	1	4%	0		0		1	3%
Meio Norte	1	4%	0		0		1	3%
Jovem Nerd	1	4%	0		2	10%	3	8%
Jovem Pan	1	4%	0		0		1	3%
Cidade de Niterói	1	4%	1	4%	0		2	5%
O Vício	0		3	13%	4	19%	7	19%
Legião dos Heróis	0		3	13%	0		3	8%
Central XBox	0		1	4%	0		1	3%
Cine POP	0		1	4%	0		1	3%
CNN Brasil	0		1	4%	0		1	3%
Metrópolis	0		1	4%	0		1	3%
Meu PlayStation	0		1	4%	0		1	3%
O TV Foco	0		1	4%	0		1	3%

Olhar Digital	0		1	4%	0		1	3%
PSX Brasil	0		1	4%	0		1	3%
Rap Mais	0		1	4%	2	10%	3	8%
Tecmundo	0		1	4%	1	5%	2	5%
Observatório da TV (UOL)	0		0		1	5%	1	3%
Gshow (Globo)	0		0		1	5%	1	3%
E-Pipoca (Terra)	0		0		1	5%	1	3%
Critical Hits	0		0		1	5%	1	3%
Omelete	0		0		1	5%	1	3%
São Paulo Secreto	0		0		1	5%	1	3%
Viciados.net	0		0		1	5%	1	3%
Total	25	100%	24	100%	21	100%	37	100%

PERFIL D								
Fontes	19 a 23/04	%	17 a 21/05	%	07 a 11/06	%	Todos os períodos	%
Observatório do Cinema (UOL)	0		3	12%	7	33%	10	22%
Canaltech	0		2	8%	1	5%	3	7%
Mac Magazine	0		2	8%	0		2	4%
Tilt (UOL)	0		1	4%	0		1	2%
Contigo! (Perfil)	0		1	4%	0		1	2%
Extra (Globo)	0		1	4%	1	5%	2	4%
G1 (Globo)	0		1	4%	0		1	2%
TechTudo (Globo)	0		1	4%	1	5%	2	4%
O Dia (Ongoing)	0		1	4%	0		1	2%
A Gazeta ES	0		1	4%	0		1	2%
Cine POP	0		1	4%	1	5%	2	4%
Guia Disney Plus Brasil	0		1	4%	0		1	2%
Jornal Contábil	0		1	4%	1	5%	2	4%
Legado da Marvel	0		1	4%	0		1	2%
Legião dos Heróis	0		1	4%	0		1	2%

Marvel Brasil	0		1	4%	0		1	2%
Metro World News	0		1	4%	0		1	2%
Notícias Concursos	0		1	4%	1	5%	2	4%
O Vício	0		1	4%	2	10%	3	7%
POP Line	0		1	4%	0		1	2%
Tecmundo	0		1	4%	1	5%	2	4%
UOL	0		0		1	5%	1	2%
E-Pipoca (Terra)	0		0		1	5%	1	2%
Blog do Iphone	0		0		1	5%	1	2%
ND Mais	0		0		1	5%	1	2%
Tracklist	0		0		1	5%	1	2%
Total	0		25	100%	21	100%	46	100%

## 7.2 Apêndice C - Planilha com as recomendações do Google Notícias para os perfis

A planilha com os dados das recomendações da categoria “Para Você”, do Google Notícias, aos perfis estão disponíveis em:

<[https://docs.google.com/spreadsheets/d/18xjc90qH3F4SuG-2fqEOoT\\_dRK1hXUgdhYn-mz8pyJg/edit?usp=sharing](https://docs.google.com/spreadsheets/d/18xjc90qH3F4SuG-2fqEOoT_dRK1hXUgdhYn-mz8pyJg/edit?usp=sharing)>.