



Filtros-bolha, personalização e transparência nos portais Globo.com e UOL

Ivo Henrique Dantas¹

Resumo: Diante da crescente quantidade de informações disponíveis no ambiente digital, plataformas de redes sociais e portais de notícias têm recorrido cada vez mais à utilização de algoritmos para selecionar o que mostrar para cada usuário. Com o funcionamento baseado na coleta de dados, os algoritmos atuam como filtros-bolha, mostrando conteúdos cada vez mais personalizados de acordo com os interesses de cada pessoa. O presente estudo procura analisar o nível de transparência oferecido pelos dois maiores portais brasileiros, Globo.com e UOL, acerca das formas de coleta e utilização desses dados para oferecer conteúdos individualizados. Apesar dos constantes questionamentos às empresas donas das plataformas de redes sociais e de motores de busca, este estudo demonstra que os portais também são marcados por baixos índices de transparência nessa área.

Palavras-chave: Webjornalismo; Portais; Filtros-bolha; Transparência; Personalização;

1. Introdução

O impacto da adoção de tecnologias algorítmicas no jornalismo é inegável. Ao longo da última década, a ascensão do jornalismo em base de dados e das técnicas de individualização de conteúdos ajudaram a transformar as práticas de produção, distribuição e consumo de notícias em todo o mundo (BARSOTTI, 2019).

Juntamente com a ascensão dos algoritmos como meio de entregar conteúdos personalizados para os usuários vem se acumulando uma série de questionamentos

¹ Doutor em Comunicação Social pela Universidade Federal de Pernambuco. Jornalista formado pela Universidade Católica de Pernambuco com especialização em marketing pela Fundação Getúlio Vargas. E-mail: ivohenrique.dantas@gmail.com.

acerca da necessidade de uma maior transparência sobre as formas de coleta e utilização dos dados que deixamos a cada interação que realizamos no ambiente digital.

Se, por um lado, as plataformas de redes sociais e os mecanismos de busca têm sido o foco desses questionamentos, principalmente diante da proliferação de informações falsas ou deturpadas, comumente denominadas de fake news, por outro, essa transparência por parte dos veículos de comunicação ainda carece de maior visibilidade no debate público.

Apesar da alta difusão de informações através dessas plataformas de redes sociais, os portais mantêm relevante posição no ecossistema digital brasileiro. Segundo dados da ferramenta Alexa Top Sites², de propriedade da Amazon, os portais Globo.com e Uol são os 5º e 6º sites no ranking nacional, que leva em consideração a média diária de visitantes e pageviews.

No mesmo sentido, a crescente utilização de dados dos usuários a partir desses sites para customizar seus conteúdos levanta um questionamento acerca dos impactos sobre a visibilidade da notícia. Em tempos em que plataformas de redes sociais e motores de busca são constantemente acusados de auxiliar na formação de espaços artificiais de sentido em que os usuários estariam condenados a serem expostos apenas a conteúdos afinados com seus interesses, constituindo filtros-bolha a partir de seus algoritmos (SUMPTER, 2019), podemos afirmar que os portais escapam dessa lógica?

Nesse sentido, o presente artigo procura analisar o cenário atual de circulação de informações através do ecossistema digital, com vistas a avaliar a transparência sobre a coleta e utilização dos dados dos usuários pelos dois maiores portais brasileiros, Globo.com e UOL.

2. Avalanche informacional e personalização

Vivemos em um mundo em que a quantidade de informações disponíveis para os usuários supera e muito o que seria possível consumirmos ao longo de uma vida inteiri-

² Disponível em: <https://www.alexa.com/topsites/countries/BR>. Acesso em: 20 de jul. 2020;

ra³. Não à toa, termos como ‘sobrecarga informacional’ e ‘avalanche informacional’ têm se tornado cada vez mais populares nos estudos acadêmicos nos últimos anos. Durante a pandemia da Covid-19, a Organização Mundial da Saúde (OMS) chegou a utilizar o termo ‘infodemia’ para dar conta da quantidade exorbitante de informação que circula nas redes, muitas vezes sem sequer serem verdadeiras.

Se estamos diante de tantas informações, como conseguimos selecionar aquelas que merecem nossa atenção? Historicamente, o jornalismo desempenhou a função de ser uma instituição cuja definição estava atrelada diretamente ao papel de mediar o nosso acesso aos acontecimentos do mundo, bem como o debate público (HABERMAS, 2014). Hoje, contudo, a democratização do acesso à rede possibilita que outros atores assumam o papel de selecionar as informações que merecem nossa atenção.

Para muitos, a evolução da rede com a ascensão dos mecanismos de busca e das plataformas de redes sociais representa a vitória da desintermediação de conteúdos. Contudo, ao contrário do que uma primeira observação possa indicar, a internet acabou por tornar ainda mais obscuros os mecanismos que atuam entre a produção e consumo das informações. Nesse sentido, Pariser (2012) argumenta que a maioria das pessoas navega pela internet sem ter ideia da quantidade de conteúdos aos quais nunca terão acesso devido aos processos de filtragem aos quais somos expostos de forma invisível.

A esse processo de filtragem dos conteúdos disponíveis, Thorson e Wells (2015) denominam de curadoria. Para eles, curadores são seletores ativos e formadores de conteúdo trabalhando sob a abundância de conteúdo. Nesse contexto, existem cinco processos que compõem os fluxos de seleção: *curadoria jornalística* (baseada em critérios e práticas profissionais estabelecidas), *curadoria individual* (baseada em escolhas dos indivíduos), *curadoria estratégica* (realizada por atores políticos para embasar seus argumentos diante do público), *curadoria social* (estabelecida a partir da rede de contatos a que temos acesso), por fim a curadoria algorítmica (automatizada a partir da análise de dados segundo parâmetros pré-estabelecidos).

Ao mesmo tempo, ao contrário do que uma primeira impressão possa demonstrar, a avalanche informacional que marca a rede após a entrada na web 2.0 não se res-

³ Segundo dados compilados pela Nodegraph, em 2025 devemos atingir a marca de 175zettabytes de informação na internet. Disponível em: < <https://www.nodegraph.se/how-much-data-is-on-the-internet/>>. Acesso em: 05 de ago. 2020;

tringe aos conteúdos disponibilizados por usuários e a grande quantidade de sites segmentados espalhados pela rede. O jornalismo empresarial constituído sobre bases tradicionais, como jornais e portais, também é causador dessa grande oferta de informações. A lógica da redução dos custos de distribuição e do espaço virtualmente infinito também impulsiona o crescimento do número de materiais produzidos por esses veículos.

É bem verdade que no caso dos sites de jornais e portais, o acesso via portas laterais conta com processos de filtragem que fogem ao controle dessas empresas. Buscadores e plataformas de redes sociais, por mais que possam ser influenciadas através de técnicas como o Search Engine Optimization (SEO) e anúncios, se interpõem como mais uma camada entre o jornalismo e o público (MONTENEGRO; DAMASCENO, 2019). Através dos sistemas de análise Alexa e Site Checker, é possível comprovar que ambos os portais utilizam práticas de SEO na organização dos seus conteúdos, sendo bem ranqueados na busca orgânica (resultados não pagos) do Google.

Já quando a filtragem acontece via plataformas de redes sociais, podem ocorrer dois fenômenos. O primeiro está relacionado à publicação na plataforma pelos portais, que publicam links para conteúdos em suas páginas. Já o segundo acaba dando maior agência aos usuários, quando estes compartilham matérias em suas redes. Isso significa uma expansão do alcance do portal para pessoas que talvez não tenham o costume de acessar os endereços. Ao compartilhar uma notícia, o usuário está ativando sua rede de amigos e seguidores e recomendando aquele link. Nesses casos, contudo, precisamos salientar que, para além da recomendação, há também a ingerência dos algoritmos dessas plataformas.

De todo modo, esse compartilhamento abre possibilidade para que ocorram representações do potencial da auto-comunicação em massa defendida por Castells (2016). Ou seja, uma notícia que talvez não tenha recebido destaque pelo portal pode alcançar um público considerável através da ativação do compartilhamento pelos usuários.

Ao mesmo tempo, esse processo de seleção realizado por usuários ao compartilharem em suas redes também demonstra a ideia do gatewatcher, ou seja, um novo nível de filtragem das notícias realizado não por quem produziu a informação, mas por nós que ‘recomendam’ a leitura dessas notícias. Os usuários acabam, desse modo, realizan-

do um processo de curadoria dos conteúdos que eles acreditam ser de interesse de seus amigos e seguidores.

A experiência das redes sociais é completamente determinada pela seleção de pessoas que os usuários escolhem seguir, e quais atualizações irão receber. Ademais, o que seria considerado notícia no sentido tradicional é comumente filtrado através de um grupo escolhido de amigos, associados e estranhos: escolhidos por cada usuário para predeterminar que tipos de atualizações eles irão receber (CRAWFORD, 2011, p. 116).

3. Personalização e transparência nos portais Globo.com e UOL⁴

Segundo dados extraídos através da ferramenta de análise de sites Alexa, grande parte dos acessos dos portais UOL e Globo.com se dá de forma direta⁵. Ou seja, sem intermediários como plataformas de redes sociais, links ou motores de busca. Assim, pode-se afirmar que, mesmo com o constante crescimento dos acessos através de portas laterais (BARSOTTI; AGUIAR, 2017), as opções de personalização oferecidas por estes sites através de suas *homepages* possuem impacto direto na seleção de notícias oferecidas aos usuários.

Observamos ainda que esses dados dizem respeito aos acessos aos referidos portais, produtos jornalísticos oriundos do meio digital, o que pode ajudar a explicar uma relevância maior das *homepages* nesses casos do que nos sites de jornais impressos. Desse modo, os dados de acesso ao Globo.com e ao UOL não se encontram em conflito com observações de outros estudos (BARSOTTI; AGUIAR, 2017) que identificam uma crescente invisibilidade das *homepages* dos jornais.

A seleção dos conteúdos nas páginas de entrada pode ocorrer de forma ativa ou passiva, de acordo com as possibilidades abertas para que o usuário determine os critérios que serão utilizados neste processo. No caso dos portais em questão, a filtragem dos conteúdos das *homepages* é eminentemente definida pelo veículo, relegando poucas opções de customização por parte do usuário.

⁴ O presente estudo faz parte da tese de defesa de doutorado intitulada “Portais, Esfera Pública e Fragmentação: a cobertura do UOL e Globo.com em tempos de filtros-bolha e câmaras de eco” defendida na Universidade Federal de Pernambuco em maio/2020;

⁵ Informações coletadas através da ferramenta em março de 2020 davam conta de que 68,3% dos acessos ao Globo.com e 47,6% ao UOL ocorrem de forma direta;

Dentre as opções limitadas de customização que existem nos portais está a possibilidade de acompanhar os destaques através da assinatura de *newsletters* e *feeds* de RSS (Rich Site Summary). Assim, ambas as tecnologias pressupõem a intencionalidade do utilizador “porque em dado momento ocorreu uma tomada de iniciativa para ‘puxar’ a informação a partir de um certo website, podendo, a qualquer momento, esse canal ser eliminado e a ligação findar” (SÁ, 2005, p. 8).

As opções de RSS em ambos os portais obedecem a lógica da divisão por editoriais, possibilitando ao usuário acompanhar apenas aquelas de seu interesse. O Globo.com ainda acrescenta seções RSS para diversos estados e regiões, claramente possibilitado pela existência de afiliadas da Rede Globo de Televisão nesses locais, gerando conteúdo específicos sobre essas regiões.

No caso das *newsletters*, o UOL oferece apenas uma opção baseada em um resumo diário com alguns dos destaques do dia, mas que não observa preferências do usuário. Já o Globo.com, além do resumo do dia, abre uma paleta de opções baseadas nas principais editoriais, como política, entretenimento, história, universo feminino e gastronomia.

Figura 1: Opções de configuração da newsletter do Globo.com / O Globo



Fonte: Globo.com

A customização de conteúdos diretamente na *homepage* se restringe a editoria de esportes, onde é possível informar ao site qual o time para o qual o usuário torce e passar a receber prioritariamente informações baseadas nesta escolha consciente.

No UOL, ainda é disponibilizada ainda uma opção de marcar conteúdos como favoritos, que ficam salvos em uma lista customizada para serem lidos posteriormente. Como será visto a seguir, essas interações e gravação de preferências podem ser utilizadas para posteriormente alimentar os algoritmos dos portais e aumentar a assertividade da seleção de notícias de acordo com os interesses e comportamentos de cada visitante.

Já a seção das Mais Lidas, que teoricamente traria uma possibilidade de apresentar destaques escolhidos diretamente pela quantidade de acessos a cada matéria, é relegada ao final das páginas. Por um lado, esse movimento pode ser visto como positivo por sobrepor critérios jornalísticos à lógica da reafirmação, que pode criar um círculo vicioso em torno das matérias mais acessadas. Por outro, isso reduz a interatividade com o público, que possibilitaria uma redução do *gap* existente entre o que os jornalistas consideram como digno de destaque e o que o público tem interesse (BOCZKOWSKI; MITCHELSTEIN, 2013).

Tal constatação não significa que os acessos não sejam levados em conta na escolha dos destaques, mas que não há transparência sobre tais critérios. Mesmo em caso de utilização dos números de acesso como um dos elementos observados pelos jornalistas na definição das chamadas, trata-se da preferência por uma possibilidade de filtragem profissional que se complementa com outros mecanismos de personalização, como o uso de algoritmos utilizados também em buscadores e plataformas de rede sociais.

A transparência acerca do uso de elementos de filtragem baseados nos acessos, como a seção Mais Lidas, poderia servir de termômetro para o público sobre o que outros usuários têm acessado, possibilitando o contato com informações que talvez não sejam acessíveis à medida em que uma personalização mais fechada for sendo implementada.

Ao mesmo tempo, sendo portais, eles funcionam como agregadores de conteúdo de diversos sites e editorias. No caso do Globo.com, a proposta se concentra nos materiais gerados pelas redações e iniciativas do Grupo Globo, como o G1, O Globo, Revista Época, TV Globo, entre outros. Não existem assim sites verticais de conteúdo, ou seja,

voltados especificamente para certos públicos. Apenas a clássica divisão por editorias e veículos.

No espectro oposto, o UOL se posiciona como um portal que abarca conteúdos mais amplos, motivado principalmente pelo fato de ter extrapolado as barreiras de ser o portal do Grupo Folha para ser um portal de conteúdo, serviços e tecnologia, incluindo um braço financeiro, com o PagSeguro. Assim, o portal tem criado sistematicamente canais que apostam na segmentação do público em interesses específicos, como o Univera (público feminino), VivaBem (saúde e bem-estar), Start (games), Tilt (tecnologia), MOV (produtora de vídeo) e ECOA (sustentabilidade). Trata-se da aplicação da teoria da Cauda Longa, com a exploração de canais que possuem públicos menores do que a cobertura tradicional noticiosa, mas que produzem alto engajamento por serem direcionados àquelas pessoas (ANDERSON, 2006).

Fundamental para a transparência em uma época de filtros-bolha e algoritmos, ambos os portais analisados disponibilizam para seus visitantes uma declaração de normas de segurança e privacidade explicando quais os dados que são coletados durante o tempo em que o usuário está navegando pelos domínios do site e quais os usos que podem ser feitos a partir deles. Assim, esses documentos acabam sendo “descrições compreensivas sobre as práticas relacionadas a coleta de informações de visitantes” (HONG et al., 2005, p.19).

No UOL, o link fica ao final da *homepage* em meio ao mapa do site, que contém uma gama de outras opções (*Ver figura 21*). Já no caso do Globo.com, o link fica na barra final de todas as páginas internas de seus canais, não apenas na *homepage*. Assim, o usuário que entra nos domínios do portal através de entradas laterais, como buscadores e links diretos para matérias, consegue ter acesso à política de privacidade.

SBPJor – Associação Brasileira de Pesquisadores em Jornalismo
18º Encontro Nacional de Pesquisadores em Jornalismo
3 a 6 de Novembro de 2020

.....
Figura 2: Política de Privacidade e Segurança do Globo.com

Política de privacidade da Globo

A Globo sabe o quanto é importante para você conhecer e estar seguro sobre a utilização dos seus dados pessoais. Por isso, nos preocupamos em esclarecer e divulgar nossa política de utilização dessas informações. Assim, você poderá entender melhor quais informações obtemos e como as utilizamos. Dedique alguns minutos do seu tempo para ler essa política e aproveitar de forma consciente e segura tudo o que a Globo tem para oferecer. Para assegurar regras claras e precisas, a Globo pode alterar essa política de tempos em tempos, e sendo assim, recomenda sua leitura periodicamente.

Esta política permitirá que você saiba:

1. Quais informações obtemos;
2. O que são cookies e qual a sua utilidade;
3. Como serão utilizadas estas informações;
4. Quem coleta as informações;
5. Com quem suas informações podem ser compartilhadas;
6. Quais são as suas opções a respeito da coleta, uso e distribuição de suas informações;
7. Como a Globo trata suas informações relativas a finanças pessoais e saúde;
8. Precauções de segurança estabelecidas para sua proteção em caso de perda, má utilização ou alteração de suas informações e;
9. O que mais você precisa saber sobre sua privacidade na Internet.



Fonte: Globo.com

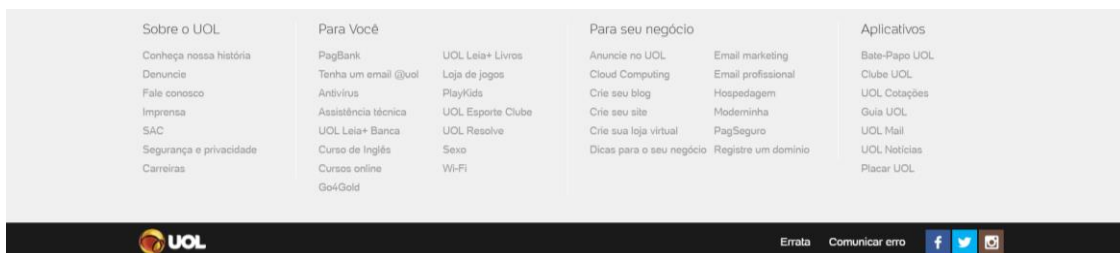
Figura 3: Política de Privacidade e Segurança do UOL

Normas de segurança e privacidade

POLÍTICA DE PRIVACIDADE GLOBAL - GRUPO UOL

Este documento tem por finalidade estabelecer as regras para tratamento de dados que inclui, exemplificativamente, as operações de coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração dos dados coletados dos USUÁRIOS, além do registro de suas atividades, de acordo com as leis aplicáveis.

Quando o USUÁRIO aceita essa Política de Privacidade confere sua livre e expressa concordância com os termos aqui estipulados.



Fonte: Uol.com.br

Nos dois documentos o usuário pode encontrar informações acerca dos principais aspectos da utilização e coleta de seus dados, como quais os dados que podem ser coletados, qual a finalidade e com quem eles podem ser compartilhados. Em nenhum

dos dois sites, contudo, é possível acessar os dados que foram coletados, demonstrando uma falha de transparência. Se por um lado existem informações sobre que dados podem ser guardados, falta acesso a esses dados.

No documento do UOL, a utilização das informações para o oferecimento de conteúdos e serviços personalizados aparece apenas uma vez e é utilizado o termo “customização” (sem realizar uma distinção entre este e “personalização”⁶). Já no documento do Globo.com, as palavras “personalização” e “personalizar” aparecem em cinco ocasiões, deixando claro para o visitante a intencionalidade de oferecer conteúdos cada vez mais voltados para os seus interesses através da utilização das informações coletadas automaticamente pelo sistema.

Também aumentando a troca de informações entre produtores de notícia e audiência está o uso de agentes inteligentes ou *bots* que podem facilitar o processo de procura, filtragem, organização e compartilhamento das informações entre os produtores de notícia e suas audiências (HONG et al. 2005, p. 16).

Não existe em nenhum dos dois casos clareza sobre quais conteúdos estão sendo indicados para todos os usuários e quais fazem parte desse mecanismo de filtragem personalizada. Assim como nas plataformas de redes sociais, os filtros dos portais acabam adquirindo um status de invisibilidade, como denuncia Pariser (2012). Ao mesmo tempo, não é dada ao usuário a opção de não obter conteúdos por meio da personalização de algoritmos que podem ser utilizados pelos portais.

Para explicar a importância dessa transparência sobre as formas como os filtros funcionam dentro dos sites, Pariser (2012) faz uma comparação entre as regras dos algoritmos do Twitter e do Facebook. Enquanto o primeiro possui poucas regras e costuma ser bastante transparente – inclusive com a possibilidade de acessarmos qualquer conta não bloqueada e ver sua timeline sem permissões específicas – o segundo possui regras “incrivelmente turvas. Se postarmos uma atualização, nossos amigos poderão vê-la ou não, e nós talvez vejamos as deles, ou não” (PARISER, 2012, p. 199). Em resumo, o Twitter nos possibilita gerenciar alguns filtros ao escolher quem seguimos e abrir a pos-

⁶ O presente trabalho entende que o termo ‘customização’ está associado a situações de escolhas ativas dos usuários, enquanto ‘personalização’ se refere às escolhas automatizadas ou feitas a partir da coleta de dados dos usuários de forma passiva.



sibilidade de acessarmos os feeds de outras pessoas. Já no Facebook, isso não é possível. Os feeds são individualizados e o acesso exclusivo.

Este é exatamente o comportamento que podemos identificar nos portais estudados. Faltam opções para que usuários selecionem o que querem ver – e mais, entendam o que estão vendo, e por que estão vendo. Esse processo poderia ser facilmente resolvido com uma área de customização, onde usuários pudessem selecionar preferências e editar seus dados coletados.

A única opção que existe para reduzir esse impacto é cancelar a permissão de *cookies*⁷ através do navegador. Opção esta desencorajada pelos portais ao alertarem em suas políticas de privacidade que a exclusão desses dados pode significar o não funcionamento adequado do site e de serviços oferecidos. Ao mesmo tempo, ambos os documentos explicam o que são cookies e procuram deixar claro que há opção do usuário de desativá-los através das configurações do navegador.

Segundo demonstram Hong et al. (2005, p.17, *tradução nossa*), a “coleta de informações pessoas pode ocorrer em diversos pontos dentro de um site, com o registro do consumidor e o preenchimento de formulários sendo os métodos mais óbvios”. Tanto no UOL como no Globo.com, a coleta de dados ocorre basicamente de três modos: 1 – Automatizado, de acordo com padrões de navegação, preferências, cookies, dentre outros. 2 – Intencionalmente, quando o usuário preenche formulários, como o cadastro no site ou assina os conteúdos do portal. 3 – Através de parceiros, desde que haja autorização por parte da pessoa para disponibilização desses dados.

Enquanto o documento do Globo.com é genérico em relação a quais dados podem ser coletados, o do UOL discrimina essas informações de forma mais detalhada. No geral, são armazenadas informações como RG, CPF, endereço, IP (endereço virtual do dispositivo utilizado para o acesso), endereço de e-mail, formas de pagamento, data de aniversário, links e botões clicados, matérias lidas, tempo de permanência, agenda de contatos e páginas visitadas, geolocalização, além de qual o dispositivo e software foi utilizado para o acesso e horários em que ocorrem as visitas e interações. O UOL ainda

⁷ *Cookies* são pequenos arquivos ou pacotes de dados enviados pelo site ao dispositivo do usuário para identificá-lo e coletar informações que auxiliarão o site a personalizar serviços e conteúdos.

acrescenta a essa lista a possibilidade de armazenar informações biométricas, como foto do rosto.

Através dessa coleta de informações pessoais sobre os usuários de notícias, sites podem construir seus perfis, permitindo a oferta de conteúdos personalizados para usuários individuais ou vendendo essas informações para outras empresas (SCHACHTMAN, 2001).

Quanto ao objetivo dessa coleta, o UOL informa que esses dados podem proporcionar “experiências mais pessoais, além de possibilitar eventuais customizações” (UOL, 2020). Já o Globo.com deixa ainda mais clara a intencionalidade de aprimorar os serviços de personalização oferecidos pelo portal.

Serão utilizadas para personalizar o conteúdo e/ou serviços disponibilizados. Isso inclui serviços de personalização, comunicações interativas, compras online e outros tipos de serviços. Sabendo mais sobre você, a Globo pode fornecer conteúdo e serviços mais relevantes e proporcionar uma melhor experiência on-line. Medir certos padrões de navegação, mapeando quais áreas dos Serviços da Globo você visitou e seus hábitos de visita como um todo. Usamos essa informação para verificar a rotina de navegação dos nossos usuários, e assim oferecer conteúdo e/ou serviços cada vez mais personalizados (GLOBO, 2020).

Sobre o uso de algoritmos, um pouco mais de informação encontra-se disponível no blog⁸ do Globo.com que comenta decisões e avanços tecnológicos em seus sistemas. Em post publicado no dia 15 de janeiro de 2019 a utilização de ferramentas que possibilitam a personalização do conteúdo da *homepage* de acordo com as preferências de cada usuário foi um pouco mais detalhada, mesmo sem resolver o problema da transparência sobre quais conteúdos são personalizados para cada usuário. Assim como no caso do UOL, a justificativa mercadológica ganha destaque na motivação do uso dos algoritmos. “Atualmente temos aproximadamente 100 milhões de usuários acessando nossos portais e precisamos dar uma experiência única para cada um deles” (GLOBO.COM, 2019).

Segundo o post, as possibilidades abertas pela tecnologia que passou a ser implementada estão listadas a seguir.

⁸ <https://blog.globo.com/2019/01/15/usando-os-dados-para-transformar-a-globo-com/>. Acesso em: 01 de mar. 2019.



- Classificar usuários em segmentos dado sua navegação. Alguns exemplos de segmentos são: fãs do BBB, amantes de carros.
- Recomendar matérias e vídeos com base nas preferências de cada usuário. Nosso *feed* das páginas iniciais é personalizado e customizado por produto.
- Melhorar o título das matérias dependendo do que os usuários mais querem ler. Ajudamos o editor a encontrar títulos mais pertinentes.
- Identificar automaticamente assunto das matérias e classificar em entidades semânticas. Uma matéria pode falar do Luciano Huck ou do Bolsonaro.
- Identificação de famosos em vídeos. Podemos saber o trecho em que cada artista apareceu em uma série ou novela.

Todavia, apesar dos documentos apresentarem um passo importante em direção ao aumento da transparência sobre a coleta e uso dos dados dos usuários, ainda falta uma maior disponibilidade de informações específicas sobre os modos como esses dados impactam a seleção de conteúdos a que o usuário é exposto.

Enquanto as discussões acerca da coleta e uso de dados por empresas de tecnologia, como Google e Facebook, têm ganho a atenção do mundo acadêmico (CORREIA, 2019), os portais acabam escapando praticamente ilesos desta discussão. Como visto, não são apenas as plataformas de redes sociais e buscadores que coletam dados de seus usuários e utilizam para vender anúncios e oferecer conteúdos cada vez mais personalizados.

Se na customização esse processo fica claro por demandar escolhas ativas da audiência, na personalização em massa tudo acontece de forma automatizada e invisível. Enquanto o próprio diretor executivo de publicidade do UOL Bebeto Pirró reconhece em artigo veiculado no portal a importância da transparência da gestão desses dados nos dias atuais, a prática ainda está longe da teoria. “Num cenário que a partir deste ano será mais regulado em relação à política de dados, o trabalho da publicidade digital dentro de plataformas especializadas de conteúdo tem ainda mais um ganho: o da transparência” (PIRRÓ, 2020).

Ora, essa transparência precisa acontecer tanto na publicidade quanto nos conteúdos jornalísticos. Em tempos de discussão sobre filtros-bolha, o jornalismo profissio-

nal acaba cometendo os mesmos pecados das criticadas plataformas de redes sociais. O Grupo Folha, por exemplo, em seus princípios editoriais reconhece os problemas gerados pela filtragem algorítmica que toma conta dessas plataformas. “As redes sociais, que poderiam ser um ambiente sobretudo de convívio e intercâmbio, são programadas de tal modo que estimulam a reiteração estéril de hábitos e opiniões preexistentes” (FOLHA, 2019)⁹.

Na contramão dessa percepção contudo, ambos os portais afirmam em suas Políticas de Privacidade e Segurança utilizarem os dados para personalizar o conteúdo oferecido, mas não explicam como ocorre essa personalização. Se, por um lado, os portais reconhecem a importância desse quesito, por outro, não disponibilizam informações mais concretas para seus leitores acerca do alcance e critérios utilizados por seus algoritmos.

4. Considerações Finais

A análise dos dois portais demonstra um viés de crescente atenção com os mecanismos de personalização e customização disponíveis para atingir uma audiência cada vez mais reconhecível através da coleta dos dados a partir das interações com os sites. Existe, por outro lado, um claro foco em processos que mantêm o controle da filtragem nas mãos dos media, com limitadas opções de customização baseadas em escolhas efetivas dos usuários.

Nesse sentido, apesar das promessas trazidas pela web 2.0, as empresas continuam sendo as grandes responsáveis pelo controle do conteúdo, abrindo poucas possibilidades nos portais para a ação direta dos usuários no que é defendido como filtragem colaborativa (THURMAN, 2011). Mesmo que a audiência tenha aumentado o seu poder de barganha dentro do universo digital, com a ascensão da figura do *prosumer* (JENKINS, 2009) e a realização de impactos importantes da audiência potente sobre a produção da notícia (BORGES, 2014), quando o assunto são os filtros estabelecidos pelos portais para decidir as notícias que merecem destaque nas suas *homepages*, o controle

⁹ Disponível em: <https://temas.folha.uol.com.br/folha-projeto-editorial/projeto-editorial-folha-de-s-paulo/sua-excelencia-o-consumidor-de-noticias.shtml>. Acesso em: 10 de jan. 2020;

continua nas mãos dos veículos, mesmo que esse processo leve em conta informações coletadas da audiência.

Ao mesmo tempo, a transparência sobre a coleta e utilização dos dados por parte dos portais requer avanços de modo a deixar evidente quais informações são armazenadas, com que finalidades, bem como de que modo os conteúdos acessados são individualizados. Nos moldes atuais, os portais acabam contribuindo para a invisibilidade de processos de personalização (PARISER, 2012), com o crescente risco de suas *homepages* se constituírem como filtros-bolha de forma similar ao que tem sido observado nas plataformas de redes sociais e motores de buscas.

Referências

BARSOTTI, Adriana. **Quando a visibilidade da notícia depende dos algoritmos:** os riscos para a sociedade. Trama: indústria criativa em revista, Vol. 8, Nº1, 2019. Disponível em: <<http://revistaadmmade.estacio.br/index.php/trama/article/viewArticle/7584>>. Acesso em: 10 de jun. 2020.

_____; AGUIAR, L. A invisibilidade da home page e as mudanças nos modos de leitura das notícias. In: Encontro Da Associação Nacional Dos Programas De Pós-Graduação Em Comunicação (Compós), 26.; GT Estudos De Jornalismo, 2017, São Paulo. Anais...São Paulo: Compós, 2017.

BOCZKOWSKI, Pablo; MITCHELSTEIN, Eugenia. **Las brechas de las noticias.** Buenos Aires: Manatíal, 2013.

CASTELLS, Manuel. **O Poder da Comunicação.** São Paulo: Paz e Terra, 2016.

CORREIA, João Carlos. **My Data is Mine:** what is the meaning of participation in data capitalism? In: Glimpse, vol 20, p. 79-87, 2019. Disponível em: <https://www.pdcnet.org/glimpse/content/glimpse_2019_0020_0079_0087>. Acesso em: 02 dez. 2019.

CRAWFORD, Kate. **News to me – Twitter and personal networking of news.** In: REDDEN, Guy; MEIKLE, Graham. News Online. Nova Iorque: Palgrave Macmillan, 2011.

JENKINS, Henry. **Cultura da Convergência.** São Paulo: Aleph, 2009.

MESQUITA, Giovana Borges. **Intervenho, logo existo:** a audiência potente e as novas relações no jornalismo. (Tese de Doutorado), PPGCOM/UFPE, 2014.

MONTENEGRO, Claudia; DAMASCENO, Natanael. **Imprensa, Internet e Redes Sociais:** Uma Análise Sobre o Posicionamento dos Jornais no Twitter. In: 42º Congresso Brasileiro de

Ciências da Comunicação, 2019. Belém. Anais. Disponível em: <<http://portalintercom.org.br/anais/nacional2019/resumos/R14-2236-1.pdf>> Acesso em: 09 dez. 2019.

PARISER, Eli. **O Filtro Invisível: o que a internet está escondendo de você.** Rio de Janeiro: Zahar, 2012.

PIRRÓ, Beбето. **Mercado Digital deve dar sentido a personalização de conteúdo em escala,** jan. 2020. Disponível em: <<https://blog.publicidade.uol.com.br/2020/01/24/mercado-digital-deve-dar-sentido-a-personalizacao-de-conteudo-em-escala/>>. Acesso em: 02 fev. 2020.

SÁ, Alberto. **O código oculto das notícias eletrônicas.** In: 4º Congresso da Associação Portuguesa de Ciências da Comunicação, 2005. Aveiro. Anais. Disponível em: <<https://repositorium.sdum.uminho.pt/handle/1822/4511>>. Acesso em: 10 de out. 2019.

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

THURMAN, N. Making ‘The Daily Me’: **Technology, economics and habit in the mainstream assimilation of personalized news.** In: Journalism, 12(4), 395–415, Sage, 2011. Disponível em: <<https://journals.sagepub.com/doi/10.1177/1464884910388228#articleCitationDownloadContainer>>. Acesso em: 02 dez. 2019.

THORSON, K., & WELLS, C. **Curated Flows: A Framework for Mapping Media Exposure in the Digital Age.** Communication Theory, 26(3), 2015.

TRACI, Hong et al. **Internet privacy practices of news media and implications for online journalism,** Journalism Studies, 6:1, 15-28, 2015.