

DOSSIÊ

QUEM CONTROLA A VOZ?

O uso jornalístico e o domínio informacional em transatores vocais



ANDRÉ FAGUNDES PASE

Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre – Rio Grande do Sul – Brasil

ORCID: 0000-0002-4873-1051

GISELE NOLL

Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Ijuí – Rio Grande do Sul – Brasil

ORCID: 0000-0001-7659-0156

MARIANA GOMES DA FONTOURA

Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre – Rio Grande do Sul – Brasil

ORCID: 0000-0003-0826-7028

LETICIA DALLEGRAVE

Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre – Rio Grande do Sul – Brasil

ORCID: 0000-0001-6438-8727

DOI: 10.25200/BJR.v16n3.2020.1316

Recebido em: 01/05/2020

Desk Review: 21/05/2020

Editores de Desk Review: David Domingo e Laura Storch

Revisado em: 11/10/2020

Approved em: 03/11/2020

RESUMO – O presente artigo tem como objetivo compreender as transformações causadas por novos ecossistemas informacionais no jornalismo contemporâneo. A análise é realizada a partir de notícias acessadas através de assistentes pessoais digitais embarcados em alto-falantes inteligentes. Como procedimento metodológico, adota o estudo de caso múltiplo, definindo como objeto os transatores vocais da Google (Nest Home/Google Assistant) e da Amazon (Echo/Alexa). Observa, portanto, que a inclusão de rotinas algorítmicas e a extensão de conteúdo noticioso para interfaces de voz inteligentes demanda adaptação para a personalização das informações, ecossistema retroalimentado por veículos tradicionais, jornalistas e pessoas que interagem com os artefatos.

Palavras-chave: Jornalismo. GAFAM. Transatores vocais. Alexa. Google Assistant.

WHO CONTROLS THE VOICE?

The journalistic use and the informational domain in vocal transactors

ABSTRACT – This article aims to understand the transformations caused by new informational ecosystems in contemporary journalism. This analysis is performed based on news accessed through personal digital assistants embedded in smart speakers. As a methodological procedure, it adopts a multiple case study, defining the vocal transactors of Google (Nest Home/Google Assistant) and Amazon (Echo/Alexa) as its object. This paper notes, therefore, that the inclusion of algorithmic routines and the extension of news content to intelligent voice interfaces requires adaptation for the personalization of information, an ecosystem that is feedback by traditional vehicles, journalists and people who interact with the artifacts.

Key words: Journalism. GAFAM. Vocal transactors. Alexa. Google Assistant.

¿QUIÉN CONTROLA LA VOZ? EL USO PERIODÍSTICO Y EL DOMINIO INFORMACIONAL EN TRANSACTORES VOCALES

RESUMEN – Este artículo tiene como objetivo comprender las transformaciones causadas por los nuevos ecosistemas informativos en el periodismo contemporáneo. Este análisis se realiza en función de las noticias a las que se accede a través de asistentes digitales personales integrados en altavoces inteligentes. Como procedimiento metodológico, adopta un estudio de caso múltiple, definiendo los transactores vocales de Google (Nest Home/Google Assistant) y Amazon (Echo/Alexa) como su objeto. Señala, por lo tanto, que la inclusión de rutinas algorítmicas y la extensión del contenido de noticias a interfaces de voz inteligentes requiere adaptación para la personalización de la información, un ecosistema que es retroalimentado por vehículos tradicionales, periodistas y personas que interactúan con los artefactos.

Palabras clave: Periodismo. GAFAM. Transactores vocales. Alexa. Google Assistant.

1 Considerações iniciais

A relação entre jornalismo e algoritmos é frequentemente observada através do prisma do papel dos agentes na seleção de quais informações serão apresentadas em *timelines* e espaços análogos em sites de redes sociais (Recuero, 2011; Araújo, 2017). Posteriormente, outro movimento foi realizado abordando a presença de rotinas automáticas nos processos de seleção e construção de conteúdo noticioso (Diakopoulos, 2019; Marconi, 2020). Esses processos culminam com a construção de outro espaço para o emprego de rotinas automáticas jornalísticas, a entrega da informação em assistentes pessoais digitais embarcados em assistentes de voz e demais dispositivos que comportam suas funcionalidades.

Esses artefatos “hospedeiros” possuem uma natureza múltipla, pois estão sempre prontos para apresentar conteúdo de

maneira sonora a partir de um ato evocativo no ambiente em que estão instalados, mas sua pervasividade cobre a vida dos usuários por meio de sua presença permanente no cotidiano através da embarcação do mesmo *software* de reconhecimento de fala em telefones celulares, *smart speakers*, *wearables* e demais artefatos inteligentes. Assistentes pessoais digitais como Alexa e Google Assistant, entre outros, acompanham seus usuários, servindo-os com notícias e informações ao passo que quantificam e analisam a pessoa que constitui um login em nome da personalização dos serviços oferecidos.

Para tanto, esses produtos conectados utilizam redes e estruturas de grandes empresas de tecnologia, como Google, Apple, *Facebook*, Amazon e Microsoft (conhecidas pelo acrônimo GAFAM). São extensões físicas das competências oferecidas on-line ou em outros artefatos conectados e concentram a vida digital dos usuários interligando diversos serviços, como agenda, envio de mensagens, entre outros. Assim, constituem um objeto que enfatiza a formação de um ecossistema de consumo de informação. No caso da Alexa, também é um vetor para compra de bens físicos através da Amazon.

Tais produtos/*software* são, portanto, uma espécie de “Cavalo de Troia” digital, pois adentram os espaços domésticos como facilitadores de processos, porém permitem aos verdadeiros controladores compreender e oferecer serviços para reter e expandir o consumo dentro das suas estruturas. Nesse sentido, são chamados de transatores vocais (Cayley, 2017), pois permitem realizar operações que vão do *streaming* de músicas à compra de produtos cotidianos, como creme dental.

Esse movimento contemporâneo é fundamental para compreender uma mudança na forma como utilizamos a internet sobretudo por dois pontos. O primeiro é a passagem de uma rede baseada na busca por informações para outra que passa a oferecer esse conteúdo de maneira automática. Cabe ressaltar que o processo decorre da cultura contemporânea dos aplicativos (Gardner & Davis, 2013; Pellanda, 2010), que transformaram o acesso generalista em uso a partir de aplicações dedicadas para as mais diversas finalidades. Esse movimento tornou a rede mais tátil, originalmente baseada no uso de teclados, passando para gestos com aparelhos e então modificando essa ação principal para um ato puramente vocal, possibilitando, portanto, a construção de um diálogo sonoro com o universo digital.

A compreensão do panorama não é fácil e demanda novas estratégias de pesquisa. Apesar das empresas divulgarem tutoriais e apoiarem cursos (Amazon Alexa, s.d.; Google Assistant, s.d.) para a construção e implementação dessas rotinas, não há clareza no funcionamento do processo. A passagem para o universo algorítmico elimina a compreensão entre o público de um espaço como o das redações jornalísticas, de um local para produção e apuração de conteúdo. A interface fala, mas nem sempre apresenta seu local de fala. Se as empresas jornalísticas tradicionais compreendem esse universo aos poucos, também as pesquisas observando os dilemas estão em estágio inicial de construção (Pase et al., 2019; Santos, 2019; Noll, 2020). É preciso traçar estratégias que entendam o contexto e explorem as informações possíveis para elucidar como as rotinas operam.

O presente artigo utiliza a metodologia do estudo de caso múltiplo (Yin, 2001, 2018) para definir seus objetos e compor uma compreensão das suas propriedades. Para entender a dimensão das empresas do GAFAM e como essas forças moldam o cenário, em um primeiro momento, como mapeamento inicial, serão analisados dois exemplos pertencentes ao núcleo: a interface Alexa presente no Echo Dot da Amazon e o Google Assistant através do Nest Home. Através desses transatores vocais, especificamente, podemos observar algumas das possibilidades que oferecem para o consumo de conteúdo jornalístico. Os dois produtos são oferecidos para compra no Brasil e, apesar de o último não contar com uma loja oficial, a empresa está presente com seu assistente nos milhares de telefones com sistema operacional Android em funcionamento no país.

A escolha do objeto de estudo se deve ao fato de que, no Brasil, por enquanto, esses são os dois transatores vocais mais popularizados, seja em função dos artefatos ou do *software* de reconhecimento de fala neles embarcados. Da mesma forma, levamos em consideração que Siri e Cortana, embarcadas respectivamente em artefatos Apple e Microsoft, ainda não atingiram o mesmo patamar de popularidade que os dois assistentes pessoais digitais supracitados, enquanto que o assistente de voz Portal, do *Facebook*, leva Alexa, da Amazon, enquanto *software* de reconhecimento de fala. Assim, nessa proposta, faremos uma série de ações nas interfaces da Google e da Amazon para compreender como operam na apresentação de informações jornalísticas, levando em consideração o contexto cultural, econômico, tecnológico e social que também engloba outras empresas do núcleo: Apple, Microsoft e *Facebook*.

2 O ecossistema informacional contemporâneo

Nesse contexto de transformação do jornalismo, tanto em seu consumo como em sua distribuição, é fundamental entender como a mudança para um novo ecossistema digital afeta e interfere nesse processo. Essa nova economia digital foi impulsionada por grandes empresas de tecnologia da Internet como Google, Amazon, Facebook, Apple e Microsoft (GAFAM). Segundo o ranking anual da Forbes, as empresas do núcleo GAFAM são as cinco marcas mais valiosas do mundo em 2019 (Badenhausen, 2019) e formam entre si um estável monopólio econômico.

Essas corporações substituíram grandes conglomerados de mídia e de indústrias culturais com uma nova forma de relacionamento entre tecnologia e pessoas. Elas “competem pela atenção dos usuários, desenvolvem produtos e serviços de mídia e atraem a maioria dos investimentos em publicidade online”¹ (Bustos & Castillo, 2019, p.804). É importante entender, contudo, como o GAFAM funciona para manter essa hegemonia, que deixa de ser somente econômica, uma vez que faz parte do dia a dia da vida dos consumidores.

A economia digital, formada por produtos intangíveis (*software*) e serviços, opera em uma escala globalizada e o GAFAM ocupa posições importantes no contexto. Bustos e Castillo (2019, p.805) separam as atividades em três áreas, sendo: “Infraestrutura (redes, dispositivos, *softwares*, telecomunicações, Internet das Coisas (IoT) e instalações), transações comerciais, e o conteúdo da mídia digital (mídia gratuita e paga, Big Data, etc.)”². Analisando exemplos dados pelos autores, podemos perceber que há uma grande presença de produtos e soluções inteiramente digitais, que dispensam custos com distribuição e produção física tradicional.

Ao contemplar praticamente todas as atividades da economia digital, as empresas formam uma espécie de ecossistema próprio, uma vez que não estão presentes apenas como único produto ou serviço, e que podem ou não ter compatibilidade com produtos e serviços dos concorrentes. Em um contexto geral, oferecem prioridade para que o usuário utilize as ferramentas complementares da própria empresa, alimentando o seu ecossistema em particular.

A competição através dessa formação de ecossistemas, e não somente produtos, é uma das principais características do GAFAM. Nesse cenário, é possível observar a dificuldade da entrada de novas empresas de tecnologia no ambiente digital formado,

visto que já existe hegemonia no mercado. Quando uma empresa nova com tecnologia inovadora surge, muitas vezes é comprada por uma companhia do núcleo que integra a recente expertise ao seu ecossistema, impedindo a ascensão de uma possível concorrência. Além da compra de empresas menores, há vasto investimento em pesquisas internas, uma vez que existe dependência de inovações, seja de produtos ou serviços. Isso torna as empresas do núcleo uma fonte para invenções e patentes, inclusive sendo as principais investidoras em soluções que envolvam a Internet das Coisas (Bustos & Castillo, 2019).

Nesta nova era, os dispositivos portáteis se tornaram vestíveis e estão passando do vestível ao incorporado. Estamos nessa era agora. De um dispositivo que você pode largar quando terminar, a um dispositivo que você tira como joias a um dispositivo que exigiria uma cirurgia ambulatorial para desconectar, os dispositivos inteligentes vão além do pessoal e do intrapessoal. Ao mesmo tempo, esses dispositivos estão profundamente entrelaçados com a Internet, conversando com outros dispositivos em uma nova “Internet das coisas”, ignorando seus lentos e carnosos hosts. As duas tendências de dispositivos menores e mais íntimos e de uma rede cada vez mais onipresente estão se unindo para produzir algo que transcende a tecnologia individual. (Swaine & Freiberger, 2014, p.369, tradução nossa³).

Ao usar *software* e dispositivos, sejam aplicativos, um smartphone ou um transator vocal, as empresas coletam informações sobre seus interatores e sobre o uso de ferramentas digitais. O êxito na disputa pela atenção resulta em uma quantidade imensa de dados que, quando compilados e analisados, culminam no que conhecemos como *Big Data*, que se torna o principal motor das empresas de tecnologia.

A tecnologia digital rapidamente se infiltrou no nosso dia a dia, não somos apenas consumidores, criamos dados, aprofundamos nossas pegadas digitais, deixamos que uma parte de nossas vidas seja gravada e se torne dados. Esse material pode ser transformado em um produto ou um serviço e é nessa transformação que a aprendizagem de máquina está ganhando espaço, como ocorre com assistentes de voz com assistentes pessoais digitais embarcados como o Google Home, da Google, a linha Echo, da Amazon e o HomePod, da Apple. (Noll, 2020, p.50).

Essa presença invisível e constante do GAFAM na vida das pessoas não ocorre apenas através de dispositivos e *software*. Conforme apontado, as empresas de tecnologia conquistaram e substituíram posições que antes pertenciam a grupos de mídia tradicionais, onde o jornalismo – originalmente – está inserido. Essa

conquista de mercado, contudo, transformou o formato e a veiculação de informações jornalísticas, tornando convergente a relação entre jornalismo e plataformas digitais (Bell & Owen, 2017).

3 Jornalismo e algoritmo no contexto desenvolvido pelo GAFAM

O jornalismo, de modo geral, sempre esteve estreitamente conectado às tecnologias (Pavlik, 2000; Deuze, 2006). Ao traçar uma breve linha do tempo dessa relação, Deuze destaca que “[...] o século XIX, está ligado à (aparição, crescimento e solidificação da) imprensa escrita, o século XX está ligado à radiotelevisão e o século XXI muito provavelmente irá ligar-se às plataformas multimídia digitais (sem fios)” (Deuze, 2006, p.17). As plataformas, contudo, fazem parte de um mercado baseado em grandes acúmulos de informação, onde o jornalismo tem uma pequena participação. Essa presença funciona conforme a demanda da plataforma, que é baseada em escala, receita, velocidade – e a necessidade de adaptação ao motor da plataforma transforma a produção jornalística (Bell & Owen, 2017).

Passada mais de uma década, a cronologia de Deuze (2006) continua pertinente. No entanto, é a partir do final do século XX, com a popularização da internet e seu uso comercial, que é possível observar o início das grandes mudanças que nos trazem até a contemporaneidade. Mais do que plataformas multimídia digitais sem fios, como previa o autor, chegamos à automatização e à integração ecossistêmica dessas plataformas em variadas materializações.

Nesse contexto, a automatização pode ser observada tanto pelo eixo da coleta de informações, quanto da produção e da distribuição/ acesso de notícias (Dörr, 2016; Lewis, 2015), todos perpassados por um elemento comum: os algoritmos. O jornalismo computacional, ou jornalismo de Big Data, é um exemplo, quando o trabalho é executado não por uma pessoa, mas sim por inteligências artificiais que coletam dados e informações, e os transformam, rapidamente, através do uso da linguagem natural, em um texto jornalístico (Dörr, 2016).

A emergência desse gênero evoca reflexões que deslocam a centralidade da atividade jornalística do componente humano, onde a tecnologia é vista como puramente instrumental. Primo e Zago (2015) questionam quem faz jornalismo de fato e sugerem a troca de perspectiva, através da teoria ator-rede (TAR) de Bruno

Latour, para uma visão mais horizontal dos aparatos necessários à atividade jornalística em relação aos próprios jornalistas. Enquanto as plataformas funcionam através de algoritmos, evitando o investimento em força humana, o jornalismo precisa do julgamento humano (Bell & Owen, 2017).

Do início da utilização da internet para fins comerciais, nos anos 1990, até a década de 2000, “[...] a internet presenciou mudanças na distribuição de informações” e “os anos 2010 parecem ter dado uma guinada na lógica de organização e curadoria de tais dados” (Dornelas, 2018, p.89). Nesse contexto, em relação ao jornalismo, especificamente, Dornelas (2018, p.85) destaca que “[...] a fragmentação da audiência; a atuação incisiva dos algoritmos; e a possibilidade de customização da informação recebida” também são reflexos dessas mudanças. A segmentação de informações não é algo novo (Dornelas, 2018; Gillespie, 2018): as notícias organizadas em editoriais, suplementos em formato de caderno impresso e *hotsites* em portais são alguns dos exemplos. Mas o que se tem hoje é a hiperpersonalização (Dornelas, 2018) na entrega dos conteúdos.

Comumente associados aos sites de redes sociais, os algoritmos também impactam o jornalismo, sobretudo quando os veículos passam a ocupar espaço nessas plataformas que priorizam e se alimentam de dados e conteúdo dos seus usuários. Os *feeds* de notícias colocam em ação os códigos que trabalham em uma dupla via: a de aprender as preferências do usuário e a de entregar conteúdo e anúncios de acordo com o seu perfil. De posse de tantos dados, as consequências dessas lógicas extrapolaram o âmbito dos sites e reverberaram em esferas maiores, com impactos sociais, econômicos e políticos por todo o mundo. A parcela de responsabilidade dessas plataformas, que insistem em posicionar sua atividade apenas no ramo da tecnologia e não de mídia (Napoli & Caplan, 2018), conduziram às bolhas de informação (Pariser, 2012; Recuero et al., 2017) e à crescente propagação de notícias com desinformação nesses ambientes.

Ao integrar esses espaços, normalmente pertencentes a empresas do núcleo GAFAM, os veículos de comunicação acabam por aceitar os termos e diretrizes impostos pelas plataformas para ter seu conteúdo distribuído nesses locais. Nesse ponto, há um controle da plataforma no conteúdo distribuído, uma vez que ele passa justamente por algoritmos e regras dos aplicativos, por exemplo.

Se o jornalismo é uma atividade que objetiva, em essência, uma transparência nos seus modos de produção, com os algoritmos temos

o extremo oposto: estruturas opacas, propriedades das plataformas de mídia e tecnologia, as quais, muitas vezes, não apenas pagamos para ter acesso, mas alimentamos constantemente com nossos dados.

Sob outra perspectiva, Gillespie (2018) propõe uma análise sociológica dos algoritmos, que busca desviar da compreensão dos códigos para observar o fator humano envolvido não apenas na construção, mas na manutenção dessas estruturas. A visão do teórico engloba a perspectiva dos efeitos dos algoritmos, porém considera que a relação é de um “entrelaçamento”, de troca: os algoritmos modificam as dinâmicas de consumo de informação, mas os usuários também os modificam.

Isso se dá principalmente quando o algoritmo é o instrumento de um negócio que tem como produto a informação que entrega (ou as propagandas a que se associa). Se os usuários falharem ou se recusarem a se encaixar nas suas práticas e a conceder significado a elas, o algoritmo irá falhar. Isso significa que não devemos considerar seus “efeitos” nas pessoas, mas um “entrelaçamento” multidimensional entre algoritmos postos em prática e as táticas dos usuários que fazem face a eles. [...] os algoritmos mudam, assim como as populações de usuários e as atividades com as quais se deparam. [...] À medida que esses algoritmos se abrigam na vida das pessoas e nas suas rotinas informacionais, os usuários moldam e rearticulam os algoritmos com os quais se deparam. Os algoritmos também afetam a maneira que as pessoas procuram informações, como elas percebem e pensam sobre os horizontes de conhecimento, e como elas se compreendem no e pelo discurso público. (Gillespie, 2018, p.110).

Nesse sentido, o conceito de “technogenesis” (Hayles, 2012) poderia ser aplicado, ou seja, haveria um processo mútuo de afeto entre humanos e tecnologia. Como destacado por Noll (2020), essa coevolução é potencializada quando tratamos de assistentes de voz programados para a aprendizagem. Associando essas visões ao pensamento de Gillespie (2018), esse processo entre humanos e estruturas computacionais estabelecem outras dinâmicas de circulação e produção de conhecimento. Mas o fato é que, por mais que as pessoas e suas ações ainda sejam a principal fonte para os algoritmos e suas aplicações, não devemos esquecer que as escolhas feitas por essas estruturas são, em última instância, feitas também pelas empresas que as gerenciam.

Se, por um lado, as plataformas permitem ao jornalismo atingir massivamente pessoas, as informações sobre acesso e perfil do usuário que o conteúdo produzido impacta ficam retidas na plataforma, que compartilha pouco ou nenhum dos dados com os veículos jornalísticos (Bell & Owen, 2017), uma vez que essas informações são fundamentais para seu próprio funcionamento (Bustos & Castillo, 2019). Dessa forma,

mais uma vez, as empresas jornalísticas distribuem seu conteúdo nesse ambiente controlado por empresas de terceiros.

Conforme visto, o GAFAM mantém sua hegemonia no mercado de informação, impedindo, sempre que possível, a entrada de novos concorrentes. Assim, torna inviável aos profissionais das redações a competição com empresas do núcleo para a criação de uma plataforma própria onde teriam autonomia total das informações (principalmente *Big Data*) e das receitas publicitárias. Essa adaptação forçada a uma plataforma externa transforma a maneira como a informação jornalística é produzida, o que altera o jornalismo como um todo (Bell & Owen, 2017).

Essa remodelação é notável quando interfaces de voz contemporâneas com sistemas embarcados que possuem inteligência artificial começam a fazer parte da rotina das pessoas. Ao resgatar, dentro do GAFAM, os produtos e dispositivos que se relacionam com o jornalismo e *software* de assistentes pessoais digitais (Quadro 1), é possível constatar que todas as empresas do núcleo possuem produtos e serviços específicos pensando nesse cenário: notícias que podem ser acessadas por voz através de assistentes pessoais e *chatbots* embarcados em alto-falantes inteligentes, além de informações sobre qualquer tema que seja do interesse dos usuários, oferecidas a partir de uma personalização algorítmica embasada nas buscas dos próprios consumidores nas plataformas do GAFAM.

Quadro 1 – Produtos do GAFAM

			Google	Apple	Facebook	Amazon	Microsoft
Infraestruturas	IoT	Assistente virtual ativado por voz, <i>chatbot</i> , alto-falantes inteligentes	Google Home, Google Assistant	Siri, HomePod	M (2015), Messenger, Bots (2016)	Echo Alexa, Amazon Lex	Cortana, Bot Framework, Tay, Zo, Ruuh etc.
Mídia Digital	Big Data	Notícias	Google News, Google AMP	Apple News	NewsFee, Instant Articles	–	Linked-In

Fonte: quadro feito a partir do desenvolvido por Bustos e Castillo (2019).

Por trás do consumo de notícias ou informações específicas em geral através das plataformas dessas empresas, está a construção de uma rede sociotécnica complexa que reúne algoritmo e um ponto fundamental da inteligência artificial: o *machine learning*, ou aprendizagem de máquina. Esses aspectos causaram

uma revolução silenciosa no campo da comunicação social como um todo e no jornalismo, sendo que, quando máquinas são dotadas dessas características, além da possibilidade de ouvir e falar, as alterações no campo passam a ser observadas com maior frequência e imprevisibilidade. Dessa forma, o que buscamos, portanto, é compreender o que é acessado, escolhido e entregue ao usuário e de que modo essas ações são executadas por assistentes de voz com assistentes pessoais digitais embarcados. Conforme limitações no trabalho, em um primeiro momento, temos como objetivo estudar os transatores vocais produzidos por duas empresas do núcleo GAFAM: Google (Google Nest Home e Google Assistant) e Amazon (Echo Dot e Alexa).

É importante ressaltar que esses artefatos e *software* de reconhecimento de fala não foram produzidos exclusivamente para criar ou reproduzir informações jornalísticas. Esse aspecto/funcionalidade é apenas mais uma das *Skills/Actions* acessadas através dos assistentes de voz. Embora sejam metainterfaces que trabalham com comunicação e informação, nem sempre essas características privilegiam o tratamento da informação enquanto conteúdo jornalístico ou noticioso. Assim, é a pessoa que interage com os transatores vocais quem determina que tipo de conteúdo ou que usos e apropriações fará do artefato e seu assistente pessoal embarcado, embora, no momento que a pessoa solicita informação sobre clima, trânsito ou economia global, por exemplo, são as plataformas que direcionam as fontes desses conteúdos, não deixando a critério do usuário a opção de seleção.

Nesse ponto, observamos que há um jogo entre pessoas, interfaces, empresas de tecnologia, desenvolvedores e empresas jornalísticas que varia de acordo com interesses individuais (o que a pessoa quer consumir) e coletivos (o que as empresas/desenvolvedores querem ofertar). Conforme observado por Noll (2020), em pesquisa qualitativa sobre o tema, é possível perceber que a função notícia poderia ser mais explorada entre aqueles que possuem transatores vocais. O fator que geralmente determina a compra desse tipo de artefato, levando em conta os *smart speakers*, geralmente é o interesse em deixar a casa mais “inteligente” e conectada. Com o tempo, outras funções são mais recorrentes como temporizadores, alarmes, lembretes, listas, clima e reprodução de música. Esses indícios apresentados pela autora apontam para o atual estágio do segmento jornalístico associado à interfaces de voz que

funciona mais como uma espécie de entrada rápida de informações curtas do que de conteúdo noticioso elaborado, crítico e complexo.

4 Transatores vocais na indústria de metainterface

Smart speakers com assistentes pessoais digitais embarcados são chamados por Cayley (2017) de transatores vocais, por fazerem uma série de operações em nome das pessoas e para as pessoas que possuem um ou mais modelos em casa. O surgimento desse tipo de tecnologia faz parte de um conjunto de proposições organizadas e consolidadas por empresas de tecnologia como as do núcleo GAFAM, que operam em regime da computação, chamado por Cayley (2017) de “*Big Software*”.

Nesse sentido, essas empresas atuam como arquitetas do Big Software, constroem sistemas, redes, artefatos, *software* e afins seguindo uma lógica que datafifica a vida humana, desde hábitos de consumo a padrões de fala. Os assistentes de voz são mais um dos integrantes dessa teia computacional contemporânea constituída por fluxos de dados e informações. Esses transatores vocais estão inseridos em uma cultura de interface que tem a possibilidade de desaparecer em meio às práticas humanas cotidianas. Essas interfaces contemporâneas são intituladas por Andersen e Pold (2018) de metainterface, por possuírem um caráter pervasivo, além de presença cada vez mais “invisível”.

A metainterface possui um modo diferente de produção, pois “incorpora uma interface sinal-computador que quantifica e datafifica e, finalmente, transforma o mundo inteiro em uma interface” (Andersen & Pold, 2018, tradução nossa⁴). Dessa forma, transatores vocais podem ser considerados parte dessa indústria de metainterface, cuja forma digital implica em *hardware* e *software*, além de uma capacidade de ação que pode causar transformações nas formas de expressão humanas e, no caso do presente estudo, no modo como o jornalismo passa a ser criado, distribuído e consumido.

Nessa indústria cultural gerenciada por grandes empresas de tecnologia, camadas de *software* passam a ser colocadas em todos os tipos de objetos, que funcionam a partir de uma série de ações programáveis, além daquelas aprendidas pela máquina através do convívio com humanos e da coleta de dados disponíveis na internet. Transatores vocais como Google Home/Google Assistant e Echo Dot/Alexa possuem, portanto, uma:

[...] gramática e forma de inscrição próprias, uma máquina aural que transforma áudio em texto e texto em código-máquina (binário). Na interação de pessoas com o transator vocal e dele com a rede, há o processo de inscrição textual, aural e de código que perpassa um sistema invisível e, muitas vezes, incompreensível aos seus interagentes, como sensores, processadores, chips, placas, dispositivos de localização, entre outros. (Noll, 2020, p.110).

Assim, combinam serviços ofertados pelas empresas Google e Amazon em uma forma física tangível composta por caixa de som com alto-falantes integrados, microfones, luzes de *led* e botões para ligar, desligar e deixar no mudo (os modelos trabalhados neste artigo não possuem tela⁵). Internamente, os artefatos são constituídos por uma placa lógica e circuitos que controlam todas as suas operações.

Ambos os transatores vocais são evocados por palavras-passe como “*Ok Google*”, “*Hey Google*” (Google Home), “*Alexa*”, “*Echo*” e “*Amazon*” (Echo Dot). Para funcionar, os assistentes pessoais digitais precisam ser configurados através de dispositivos com sistema operacional Android ou iOS como smartphones e tablets. A partir disso, é possível vincular e controlar outros objetos com IoT e artefatos digitais, assim como, conteúdo de mídia e *streaming* de vídeo e música, entre outras opções divididas em categorias como: perguntas e respostas, automação, entretenimento e gerenciamento de tarefas.

Outra opção é personalizar a voz do assistente, configurar os locais mais frequentados pelo usuário, fazer com que o sistema reconheça cada interlocutor que interage com ele e escolher o idioma com o qual é possível se comunicar com o assistente. No entanto, para esse estudo, a personalização mais importante é a referente às notícias que cada interator irá ouvir. Nessa configuração, são listadas as fontes de notícias disponíveis e o interator escolhe as que lhe interessam.

Todos esses conteúdos solicitados pelos interatores são acionados, em cada conversa, através da busca por parte dos transatores vocais por *Actions* (Google) e *Skills* (Amazon) pré-programadas para realizar a operação solicitada. Essas transações ocorrem com auxílio de sistemas de reconhecimento automatizado de fala (*Automated Speech Recognition* – ASR) e de nuvem⁶ para nuvem, estando vinculadas a outros serviços oferecidos pelas próprias empresas, parceiros e desenvolvedores.

5 Compreendendo o fluxo noticioso na nuvem

Esse processo de transposição da narrativa noticiosa tradicional para outro que envolve o diálogo entre humanos e máquinas demanda a adaptação do conteúdo para as estruturas das empresas envolvidas, bem como suas rotinas de programação. Ao passo que tanto Google (Google Assistant, s.d.) e Amazon (Amazon Alexa, s.d.) contam com websites voltados para a disseminação do conhecimento, existem poucos relatos sobre o desenvolvimento de produtos jornalísticos para os transatores vocais.

Pase et al. (2019) relatam a experiência de pesquisar e desenvolver uma *Action* para a empresa Globo apresentar seu conteúdo no ambiente do Google Assistant. Além da experimentação com uma nova forma de distribuir notícias, a ação foi classificada como medida estratégica para posicionar a empresa no mercado digital. Essa iniciativa foi o embrião da primeira ação da empresa nesse cenário, a *Action G1* nas Eleições (G1, 2018).

No texto, os autores indicam como maior preocupação não só a transformação do texto produzido para a web em base para o transator vocal, mas também para evitar que o conteúdo eventualmente apresentasse algum viés noticioso. Assim, o projeto desenvolvido em 2017, concentrava a sua abordagem em serviços e informações pontuais, como resultados de pesquisa e apuração de votos. Foi necessário entender como a interface é evocada e qual é a pergunta realizada, chamada no texto de engate. A combinação, no caso “Ok, Google + pergunta” precisava ser compreensível para a máquina e também entender o comportamento de quem aciona. Assim, em um primeiro momento, expressões regionais não foram inseridas na rotina.

Este não é o caso de produzir uma notícia para que outra pessoa leia ou um boletim redigido para que uma interface de voz transforme um texto em som. Há um processamento de instruções de entrada pelo homem, análise pela máquina e consequente construção de um bloco de informações que será lido pela interface de voz e deve ser compreendido por um agente humano. (Pase et al., 2019, p.44).

Outra dificuldade foi a possibilidade de um aplicativo externo utilizar informações pessoais e de localização que a Google possui, mas não permite que serviços externos utilizem. Essa limitação simples, por exemplo, dificulta para que uma empresa de abrangência nacional combine suas informações para apresentar uma notícia

local. Como a empresa proprietária do assistente possui e utiliza tal informação, há aqui uma vantagem para o conteúdo produzido pelos seus aplicativos e não por terceiros. É compreensível que isso ocorra como medida de proteção para a privacidade dos usuários, porém, é necessário compreender esse fato como diferencial competitivo.

Nesse sentido, os autores observaram duas saídas para contornar o problema. A primeira seria desenvolver *Actions* específicas, como uma para notícias sobre o Fluminense e outra para o Botafogo, por exemplo, no caso de times de futebol. Além disso, orientar o público para informar eventuais particularidades no rito evocativo.

Para tanto, foi utilizada a estrutura ENGATE + VERBO + NOTÍCIAS + CONTEXTO/ESPECÍFICO, que permite acionar a ferramenta e granular instruções. Também pode estruturar comandos formais, como voltar para um ponto específico em um bloco de informação lido ou mesmo suspender o serviço. (Pase et al., 2019, p.45).

Para implantar a *Action*, a equipe de desenvolvimento utilizou o serviço *Dialogflow*⁷, também utilizado para *chatbots* (Dialogflow, 2020). Assim, o conteúdo definido seria retirado de informações publicadas em páginas indexadas no protocolo RSS⁸. Dessa forma, um pedido para ouvir as últimas notícias demandaria que o transator vocal ouvisse a pergunta, buscasse em uma lista de palavras-chave cadastradas e, a partir delas, pegasse informações em *feeds* específicos.

Apesar de ser um relato importante para compreender a transformação dos produtores de conteúdo e das estratégias utilizadas, o artigo não explica como a experiência inicial foi implementada posteriormente pela empresa no seu fluxo de produção noticiosa para a integração com seus sistemas de publicação e também o transator vocal. Provavelmente, o processo foi replicado com novas variações a partir do bloco “notícias” na estrutura de requisição de uma informação. A construção de uma árvore de possíveis perguntas dentro de um universo pré-determinado como o das eleições é um desafio, mas o correto mapeamento permite construir o diálogo combinando informações nas bases de dados da empresa. Além disso, a *Action G1* nas Eleições era compatível com o Google Home (então nome do modelo atualmente conhecido como Nest Home) e com o aplicativo do Google para smartphones e tablets.

O desenvolvimento voltado para a máquina e com capacidade de ser traduzido de bits para texto e também áudio, conforme as normas (Google Assistant, s.d.), permite à produtora de conteúdo

explorar o artefato novo e alcançar o público que já conta com seu serviço de assistente pessoal digital embarcado em outros artefatos. Dessa forma, iteradores que atualmente não contam com um artefato da linha Nest Home já estarão familiarizados se utilizarem o objeto no futuro.

6 Procedimentos metodológicos para análise

Após contextualização e observação dos processos internos dos transatores vocais, a análise de um fenômeno contemporâneo como este utiliza uma combinação de técnicas. Diante da impossibilidade de acessar os códigos e rotinas de programação existentes nos servidores da Google e da Amazon, utilizamos a combinação da técnica de estudo de caso múltiplo (Yin, 2001, 2018) com uma exploração do funcionamento das interfaces de voz a partir de evocações pré-definidas e da observação das possibilidades de configuração presentes nos seus aplicativos. Conforme o autor, esta estratégia é válida quando

comportamentos relevantes não podem ser manipulados e quando há uma intenção de estudar um único ou uma série de eventos contemporâneos (por contemporâneo entendemos uma interpretação fluida do passado recente ou do presente, não apenas o presente). (Yin, 2018, p.43).

O trabalho realizado neste artigo adota a exploração e o uso dos artefatos digitais como força para triangular evidências e colocar em prática as observações advindas da pesquisa teórica. Isso permitirá observar a participação das fontes jornalísticas nas práticas cotidianas analisadas, validando ou não percepções derivadas do estudo da influência do grupo de empresas conhecido como GAFAM nesse novo horizonte de veiculação de informações jornalísticas.

Assim, dividimos a exploração prática em duas etapas. A primeira observa os aplicativos para dispositivos móveis que permitem instalar e controlar transatores vocais, enquanto a outra parte das evocações de informações. A observação dos apps busca constatar se é possível visualizar uma lista de *Skills/Actions* de notícias, bem como a identificação das fontes produtoras de notícias. Será observada a presença de produtores de conteúdo não tradicionais junto de empresas de comunicação de renome. Por fim, possíveis separações de categorias entre tempo, trânsito e esportes serão registradas.

Posteriormente, e sem instalar eventuais *Skills/Actions*, Google Assistant e Alexa serão evocadas a partir de uma pergunta simples, “Quais as últimas notícias?”. Isso é propositalmente genérico para entender o que o serviço apresenta como notícia. Como consequência, será observada uma eventual identificação do produtor do conteúdo.

7 As notícias no Google Home e no Echo Dot

Notícias de cunho jornalístico podem ser ouvidas no Google Home e no Echo Dot através de duas ações feitas em conjunto ou isoladamente. A primeira possibilidade, como destacado, é uma pré-seleção feita nos aplicativos Google Home e Amazon Alexa dos veículos de notícia que o interator gostaria de ter como fonte de informação. A segunda ação é perguntar ao assistente sobre notícias de seu interesse (questionar, por exemplo, sobre uma doença em determinado país), ou pedir por determinado campo (solicitando notícias de esporte, cultura, tecnologia etc.).

No aplicativo Google Home, as notícias são consideradas serviços e a lista de fontes de notícias é acessada no item homônimo dentro do app, junto com as *Actions* “Chamada de voz e vídeo”, “Lista de compras”, “Música”, “Vídeos e fotos”, “Agenda”, “Lembretes”, “Ações”, e a ação “Explorar” outros serviços. Em abril de 2020, havia 23 fontes de notícias para o idioma português⁹, tendo 10 “assuntos em destaque”, sendo eles: Boas *Uol* (todo dia uma boa notícia diferente), *Folha* (principais notícias do dia da *Folha de São Paulo*), Repórter *CBN* (áudios do Repórter *CBN*), *Jovem Pan* – Direto da Redação (notícias do dia), *O Globo* (principais notícias do dia), *Uol* Notícias (o melhor do conteúdo *Uol*), *Giro Veja* (podcast de *Giro Veja*), *Panorama CBN* (principais notícias em menos de meia hora), *Resumo R7* (comentários de Camila Busnello e Heródoto Barbeiro sobre as principais notícias do dia) e *Estadão* Notícias (principais fatos do dia no Brasil e no Mundo).

A segunda categoria de notícia é “tecnologia” com informações do *TecMundo* e *Canaltech News*. O primeiro disponibiliza áudios do programa *Hoje TecMundo*, realizado de segunda-feira a sexta-feira e transmitido originalmente no YouTube. Já o *Canaltech* traz notícias do tema em geral, entrevistas, análise de smartphones e tablets, dicas, tutoriais e informações de apps. Em seguida, aparece a categoria

“Esporte”, com notícias do dia do Mercado da Bola *Uol* e *SportBuzz* Resumo. “Cultura e entretenimento” é mais um tipo de informação presente, com notícias da *Caras Brasil*, que publica diariamente um podcast com destaques do mundo do entretenimento e das celebridades.

Na categoria “Saúde” há duas opções de fontes de notícias com especiais sobre o Coronavírus, *Jovem Pan* Especial: Coronavírus e Boletim *Uol* Coronavírus. Para encerrar, há opções de “acontecimentos locais” com *Gaúcha ZH* (Rio Grande do Sul), *Litoral FM* (Espírito Santo), *Gazeta Online* (Espírito Santo), *CBN Vitória* (Espírito Santo), *Rádio Jornal* (Pernambuco) e Boletim *O Povo* (Ceará), onde são transmitidas as principais notícias do dia com enfoque no Brasil e no Estado correspondente.

Em relação aos aplicativos, ambos não vêm com notícias pré-selecionadas, sendo papel do interator escolher as opções de fonte de notícias desejadas, assim como, é possível editar a ordem do conteúdo que vem primeiro quando o interator falar: “*Ok Google, ouvir notícias*”, ou “*Alexa, leia o resumo de notícias*”. No Amazon Alexa é possível encontrar a *Skill* “Resumos de Notícias” no item “configurações”, subtópico “Preferências da Alexa”, mesmo local onde estão listadas as opções de *Skills*: “respostas de voz”, “Música”, “TV e vídeo”, “Trânsito”, “Esportes”, “Calendário” e “Fotos”.

A Amazon não divide, no app do assistente, as notícias por categorias como é feito pela Google. Dessa forma, é o interator quem seleciona as opções a partir de uma lista com 133 resultados de fontes listadas. Essa lista aparece diferente para cada interator, obedecendo critérios como “Relevância”, “Média de avaliação dos clientes”, “Data de lançamento” e “Em destaque”. Uma *Skill* produzida para notícias do clube de futebol Grêmio chamada de “Imortal Tricolor” está indexada ao lado de outras como *G1* e *Repórter CBN*. Além de uma descrição, cada *Skill* apresenta uma lista de comandos de voz sugeridos, “*Comece dizendo*”.

É interessante observar que a caixa do modelo Echo Dot apresenta seis sugestões de *Skills* – Amazon Music, Philips Hue, Spotify, *UOL*, *CBN* e *G1* –, três delas voltadas para notícias. É possível ativar diversas *Skills* de notícias, mas boa parte delas apenas operam a partir do comando “*Alexa, qual o resumo de notícias?*”, tradução da expressão “*flash briefing*” utilizada em inglês. É possível escolher variadas *Skills*, mas o assistente irá apresentar conforme ordem de preferência definida no aplicativo.

No Google Home, inicialmente, parece não ser possível incluir mais itens na lista, contudo, os interatores podem trocar de idioma e

selecionar mais possibilidades de fontes de notícias, sem que itens da lista anterior sejam excluídos. Assim, observamos que, dentro dos aplicativos, no Google Home há maior descrição sobre quais os tipos de informações que cada fonte de notícia trará, seja por um miniresumo logo abaixo ao nome da fonte, ou pela divisão em categorias. Já o app da Amazon Alexa oferece mais opções de fontes de notícias em português, além de possibilitar que o interator veja a avaliação de cada uma, através da média de avaliação dos clientes.

A Amazon realiza a separação entre “Resumos de Notícias” em geral, e notícias sobre “Esporte” e “Trânsito”. Apesar de contar com a primeira categoria, as informações são genéricas, salvo quando há uma *Skill* específica instalada. Isso aparenta ser um pouco diferente do cenário norte-americano, por exemplo, pois a empresa possui convênios com as principais ligas esportivas para calcular estatísticas e processar dados. Quanto a “Trânsito”, é possível definir se as informações serão coletadas a partir do serviço Apple Mapas, Google Maps ou Waze. A controladora da Alexa não possui um serviço próprio para tal e mostra abertura para utilizar informações de outros, inclusive concorrentes.

No aplicativo Google Home, não há categoria isolada para esporte (as informações da categoria são encontradas junto às demais fontes de notícias ofertadas pelo assistente) ou trânsito, como apresentado no Amazon Alexa. No entanto, informações sobre o trânsito podem ser acessadas a partir de ferramentas do Google como o Google Maps, pois existe a possibilidade do interator cadastrar endereços de casa e do trabalho no item “Seus endereços” e utilizar essa configuração para saber como está o trânsito no trajeto ou para que o assistente inicie a navegação até o local (essa ação deve ser executada no smartphone).

É possível perguntar ao assistente, no Google Home, como chegar a um determinado lugar a partir de outro endereço. Nesse caso, ele descreve a distância e quanto tempo, em média, será preciso para realizar o percurso, dependendo do meio de transporte informado pelo usuário. Informações adicionais sobre paradas de ônibus e passos ou metros até o local também são acrescentados durante as instruções.

Devido a sua natureza sonora, os transatores vocais investigados permitem que perguntas sejam feitas diretamente ao assistente. No caso de perguntarmos apenas sobre notícias, se não há nenhuma fonte da preferência do interator cadastrada nos aplicativos,

ambos avisam que existe essa opção de configuração, sendo que, a Alexa, sugere três opções de fontes e a partir da escolha do interator começa a reproduzir o áudio. Já o Google Assistant apresenta o conteúdo a partir da sua lista de fontes que possuem informações do dia, em ordem cronológica de divulgação. Durante a escuta das notícias, é possível trocar a fonte solicitando: “*Ok Google, próximo*”.

Na experiência do uso do Echo Dot e do Google Home com Alexa e Google Assistant embarcados, respectivamente, verificamos que as *Skills/Actions* para notícias apresentam alguns problemas. No caso do Echo Dot, a orientação para o uso da evocativa “*Alexa, qual o meu resumo de notícias?*” orienta muitos desenvolvedores para a produção de pequenos resumos em áudio, dando a impressão de um reaproveitamento do conteúdo criado para outras plataformas, sendo que, o mesmo ocorre com o Google Home, embora haja indicações de podcasts específicos criados pelas fontes, o que demonstra a preocupação em adequar o seu formato para outras plataformas.

Os dados esportivos são muitas vezes desatualizados e muito pontuais. É compreensível que no momento da realização deste estudo muitos eventos foram suspensos em virtude da Covid-19, mas a falta de profundidade no manejo de fontes de informação, bem como a falta de clareza sobre as origens dos dados, parecem ser uma constante, salvo se forem utilizadas *Skills* e *Actions* específicas.

A navegação por voz provoca mensagens de desculpas de ambos os assistentes quando o interator pergunta qual a fonte das notícias depois que ouve uma informação. Com a Alexa, isso também ocorre nas solicitações sobre trânsito e esportes. É compreensível que no processo lógico do desenvolvimento já haja uma escolha do público por um produtor de conteúdo e isso é informado no início da resposta, porém, cabe lembrar que uma notícia pode ser composta por diversas fontes, como reportagens próprias, distribuição por agências ou publicação de material de órgãos institucionais, algo importante no processo jornalístico e que é descartado na operação do assistente (embora apareça dentro dos boletins).

As perguntas de trânsito recebem respostas padrão, com a descrição do trajeto recomendado e do tempo previsto. É uma apresentação apenas elencada sobre rotas, sem aspectos visuais, mas o componente da personalização, com o conhecimento sobre os pontos “casa” e “trabalho” oferecem uma experiência personalizada diferente dos meios de comunicação tradicionais, de natureza massiva. Em um cenário de empresas de comunicação que substituem

seus repórteres presentes nas ruas pelas informações automatizadas dos serviços, os transatores vocais estudados reforçam o ato do seu interator de contar consigo para planejar seu deslocamento e, pouco a pouco, abandonar as fontes de informação do passado.

Em relação ao GAFAM, algo saliente no uso e interação com transatores vocais da Google e da Amazon, é a percepção de que mais importante do que a veiculação de conteúdo noticioso de qualidade, é a disseminação de pequenas informações de qualquer natureza, escolhidas de acordo com o interesse dos interatores. O que preocupa nesse cenário é justamente a fonte desses conteúdos. Ao pedir para a Alexa sobre a população do Brasil, por exemplo, a resposta será: “segundo dados mais recentes de 2017, a população do Brasil é de 207 milhões de pessoas”, já o Google Assistant responde: “a população de Brasil era de 209.469.333 milhões em 2018”. Nenhum dos assistentes divulga a fonte das informações ou apresenta dados de 2020, por voz. No entanto, ao acessar o conteúdo gravado pelo assistente, buscando nas configurações, é possível identificar de onde veio a informação.

Contudo, se o interator tem conhecimento que é o Instituto Brasileiro de Geografia e Estatística (IBGE) que faz essa contagem, é possível solicitar os últimos dados divulgados pelo IBGE, informação que é fornecida de forma mais completa e com alto grau de confiabilidade. Nesse sentido é preciso que os interatores aprendam como fazer perguntas aos transatores vocais, para que as informações solicitadas cheguem conforme o desejado, o que remete ao conceito de *technogenesis*, abordado anteriormente.

Inferimos, portanto, que empresas do núcleo GAFAM, através de seus transatores vocais, atuam como uma espécie de entregadores rápidos de informação, competindo diretamente com empresas jornalísticas. São dados ofertados como serviços, mas que devem ser encarados enquanto componentes essenciais para a construção de conteúdo noticioso que, inclusive, precisaria de um contexto maior para ser compreendido em sua totalidade.

O GAFAM utiliza o rastreamento e personalização, entrando aqui a função algorítmica para otimizar os dados e oferecer “pitadas informativas”, no entanto, falha em algo essencial e caro ao jornalismo: critério e valores notícia. Além disso, pelos transatores vocais não terem sido construídos especificamente para conteúdo noticioso, achar nos aplicativos as *Skills/Actions* específicas pode ser algo deixado em segundo plano para alguns interatores, interessados mais em encontrar resultados rápidos do que em informações mais complexas.

Transatores vocais, devido a sua recente (e em processo) incorporação à rotina cotidiana de parte ainda pequena da população podem ser considerados como novos componentes para veiculação de conteúdo, no entanto, é importante frisar que, devido a sua natureza, arrebatada várias indústrias dentro de uma, reunindo comunicação, música, comércio, notícias, entre outras. Isso não é só uma transposição de serviços, é mais um nó em uma grande rede de serviços e que pode ser usada para os mesmos fins do que o jornalismo.

No atual estágio de desenvolvimento, os transatores vocais analisados evidenciam o agenciamento das empresas Google e Amazon, no caso aqui analisado (mas que pode ser expandido a outras empresas do núcleo), em relação à aplicação da inteligência artificial no que trata de informações curtas oferecidas por *Skills/Actions* nativas. Entretanto, enquanto conteúdo noticioso, atua mais em uma espécie de regulação de formatos e modos de produzir os conteúdos que serão veiculados nos transatores vocais.

Aplicativos e *Skills/Actions* podem usar esse processamento algorítmico, algo que o Google e a Amazon aproveitam para falar sobre o trânsito ou o tempo, por exemplo, contudo, enquanto desenvolvedores externos, não temos como acessar os dados extraídos pela inteligência artificial dessas empresas, algo que já coloca as empresas jornalísticas em desvantagem técnica. Nesse sentido, a partir da análise e uso dos transatores vocais, inferimos que as empresas incentivam a criação de conteúdo para uso de suas estruturas, mas esse acesso e uso é limitado ao que elas estão propondo, o que acaba restringindo a produção de conteúdo à reprodução de podcasts ou a resumos de notícia, sem que haja personalização de conteúdo noticioso de fato.

8 Considerações finais

A transformação digital coloca cenários em construção, sobretudo para o jornalismo. Ao passo que a internet e outras tecnologias permitem experiências diferentes das observadas algumas décadas antes, como a visualização de ambientes em realidade virtual, o desenvolvimento disso depende de novas iniciativas. O investimento do passado em rotativas e estruturas de captação e transmissão de vídeo atualmente possui outro ritmo. Como observado no estudo dos transatores vocais, é necessário

compreender as possibilidades de utilizar processamento de dados em larga escala, uso de rotinas de inteligência artificial, desenvolvimento de algoritmos para gerenciamento de conteúdo, possibilidade de produzir *hardware* e margem para experimentação. Mesmo que seja peça fundamental, o conteúdo aparentemente é colocado como elemento que será inserido por outras empresas interessadas em participar desses *marketplaces*. Assim como nas lojas de aplicativos de smartphones e tablets, as empresas do núcleo GAFAM compreendem que possuem um poder contemporâneo que desafia outras instituições conhecidas do público por mais tempo.

Google e Amazon, representantes do grupo estudados nesta pesquisa, ampliam seus domínios ao controlar *hardware* e *software* e condicionar as formas utilizadas para apresentar conteúdo jornalístico nos seus ambientes e artefatos. Esses aparelhos fascinam pelo seu tom de objeto de ficção científica, pela materialização da internet em uma espécie de mordomo inteligente e por apresentar uma forma fácil de obter conteúdo personalizado apenas com o uso da voz. Para melhor servir, rastreiam comportamento enquanto fornecem dados para suas empresas na busca da otimização da oferta de produtos. O jornalismo é colocado como mais um item nessa galeria de possibilidades de acesso e, em nome da presença nesses espaços, acaba sujeito às normas de empresas que não relevam contextos e cenários locais. É um processo que opera de forma não tão conhecida pelo público. Como observado nos momentos de diálogo com os assistentes, há uma escolha inicial das fontes de informação, porém, em outros níveis não é possível compreender dados sobre a construção das notícias. A adoção dos aplicativos e seu ritmo intenso de produção poucas vezes dá espaço para esse local vital da construção das notícias. A movimentação é intensificada ainda mais com as metainterfaces de voz. Mesmo apresentando os boletins gravados por profissionais, não os observamos, pois é uma voz maior, a nativa do artefato, que apresenta e nos orienta sobre a execução destes conteúdos.

Mesmo os veículos presentes no Google Assistant e Alexa não explicam as suas rotinas, somente apresentam o contexto para o seu público, e mostram que “também estão ali”. Assim, além de informar, o jornalista contribui para formar a compreensão dessas transformações pelo público, através da forma como produz e distribui notícias em transatores vocais. No entanto, o que percebemos com o experimento é que esse formato de distribuição ainda está em fase

de elaboração, o que é evidenciado por não haver um padrão, o que faz com que os veículos de notícias produzam boletins de rádio para os transatores vocais ou podcasts.

Ademais, interfaces passam a oferecer informações personalizadas e capturadas diretamente por sensores e programas, como no caso do trânsito. Câmeras e repórteres em pontos específicos das cidades são massivos e informam sobre os principais espaços da cidade, algo importante para a compreensão do ritmo dos espaços urbanos, mas a conexão direta com aplicativos e serviços de trânsito como o Waze demandam uma transformação desse serviço tradicional. O fenômeno pode ser observado futuramente no esporte, quando a Amazon, por exemplo, processar informações capturadas durante jogos para oferecer diretamente através da Alexa.

Há um ponto importante na aproximação do jornalismo nesses espaços. A informação de uma empresa disponível em uma *Action* ou *Skill* deve ser escolhida e acionada, enquanto serviços nativos oferecidos pelo GAFAM não demandam esse processo. É possível observar aqui uma alta vantagem competitiva das empresas do Vale do Silício perante os veículos que desejam participar desse cenário.

Observando a presença de conteúdo produzido em português, outro ponto constatado corrobora esta questão. Durante as sessões de teste, não foram encontrados conteúdos para compra através do diálogo, como uma notícia aprofundada após a sua leitura em formato resumido. Se por um lado é preciso compreender a estratégia de estar presente nos ambientes e fidelizar o público para posteriormente lucrar com esta operação, fica uma lacuna interessante que pode ser preenchida pela própria GAFAM antes das redações. Algo compreensível em um cenário de dificuldades financeiras e recursos escassos direcionados muitas vezes para a manutenção das operações tradicionais, porém é necessário experimentar, até mesmo antes que outros o façam.

Os algoritmos são postos em prática em uma espécie de “modo de invisibilidade”, aprendendo preferências e entregando conteúdos de acordo com o perfil do usuário. Para a máquina, fica a tarefa de personalizar a escolha humana, ou sugerir essa configuração. Assim, o público consumidor também ajuda a lógica do sistema, mantendo a estrutura apresentada através dos algoritmos. Os transatores vocais aqui trabalhados indicam um cenário que confunde habilidades comuns ao rádio, em função da voz, e da computação. Nesse caminho, meios de comunicação e profissionais se adaptam, pois precisam estar

presentes nas plataformas, contudo, pecam no quesito inovação, que mais uma vez volta às mãos do GAFAM.

Os assistentes de voz são associados com Amazon e Google, porém não são produtos exclusivos. Alternativas com o projeto *Mycroft* (Mycroft, s.d.) permitem desenvolver sistemas parecidos utilizando produtos baratos, como o *Raspberry Pi*. Diferente dos sistemas aqui estudados, seu uso demanda uma construção de rotinas e outros processos digitais, porém é um desafio que pode ser interessante para veículos de médio porte. Mesmo com questões de logística (produção, venda e suporte) envolvidas, há espaço para testar novas formas de entregar o conteúdo.

O desafio que os transatores vocais apresentam para o jornalismo é compreendido aos poucos. O conteúdo surge como uma mistura do trabalho algorítmico com o esforço de pessoas, tanto na apuração como no desenvolvimento de diálogos entre humanos e máquinas. Mesmo que a resistência aparente ser um processo oneroso demais para os orçamentos das instituições, esse processo é necessário para justamente compreender as transformações do jornalismo.

Da mesma forma, a relação entre assistentes pessoais digitais voz, artefatos e jornalismo segue em estabelecimento e constante observação, o presente texto também encontra justamente nesse contexto suas limitações – acentuadas pela opacidade e certa escassez de informações acerca do funcionamento dos sistemas e seus equipamentos. No entanto, se faz pertinente, ainda que em caráter inicial, o registo e a investigação sobre a forma como a produção jornalística vem sendo formatada e distribuída nesses ambientes.

NOTAS

- 1 Tradução livre do original: “These companies compete for the attention of users, develop media products and services and attract most of the online advertising investment”.
- 2 Tradução livre do original: “Infrastructure (network, devices, software, telecommunications, IoT and facilities); commercial transactions; and the digital media content (free and pay media, Big Data supported, etc.)”.

- 3 Tradução livre do original: “In this new era, handheld devices have become wearable, and are moving past wearable to embedded. We’re in that era now. From a device that you can put down when you’re done with it to a device that you take off like jewelry to a device that would require outpatient surgery to disconnect, smart devices are going beyond the personal, to the intra-personal. At the same time, these devices are deeply entwined with the Internet, talking to other devices in a new “Internet of things,”, bypassing their slow fleshy hosts. The two trends of smaller and more intimate devices and of an increasingly ubiquitous network are coming together to produce something that transcends either individual technology”.
- 4 Tradução livre do original: “It incorporates a signal-computer interface that quantifies and datafies, and ultimately turns the whole world into an interface”
- 5 Há outros modelos como Echo Show (Amazon) e Nest Hub (Google) que possuem tela.
- 6 Serviço de computação em nuvem é um conjunto de estratégias utilizadas por empresas de computação para armazenar dados em um conjunto de servidores espalhados pelo mundo. Dessa forma, os dados não ficam armazenados no dispositivo do usuário, mas em servidores externos.
- 7 O Dialogflow é um desenvolvedor de tecnologias de interação humano-computador da Google baseado em linguagem natural que utiliza *machine learning* para entender o propósito da solicitação e o contexto do que o usuário fala para responder de forma mais eficaz.
- 8 Forma simplificada de receber conteúdo na internet, em tempo real. A tecnologia permite a utilização, em sites e afins, de feeds RSS para que o usuário receba atualizações constantes.
- 9 Ao todo, em abril de 2020, o assistente possuía 38 idiomas cadastrados com possibilidade de fontes de notícias diferentes.

REFERÊNCIAS

Amazon Alexa. (s.d.). *O que são Alexa Skills?* Recuperado de <https://developer.amazon.com/pt-BR/alexa/alexa-skills-kit>

Andersen, C. U., & Pold, S. B. (2018). *The metainterface: The art of platforms, cities and clouds*. The MIT Press.

Araújo, W. F. (2017). *As narrativas sobre os algoritmos do Facebook: uma análise dos 10 anos do feed de notícias* [tese de doutorado, Universidade Federal do Rio Grande do Sul]. Repositório Digital LUME – UFRGS.

Badenhausen, K. (2019, maio 22). As 100 marcas mais valiosas do mundo em 2019. *Forbes*. Recuperado de forbes.com.br/listas/2019/05/as-100-marcas-mais-valiosas-do-mundo-em-2019

Bell, E., & Owen, T. (2017, março 29). The Platform Press: How Silicon Valley reengineered journalism. *Tow Center*. Recuperado de www.cjr.org/tow_center_reports/platform-press-how-silicon-valley-reengineered-journalism.php

Bustos, J. C. M., & Castillo, J. I. (2019). Who will control the media? The impact of GAFAM on the media industries in the digital economy. *Revista Latina de Comunicación Social*, 74, 803–821. DOI: 10.4185/RLCS-2019-1358en

Cayley, J. (2017). The Advent of Aurature and the End of (Electronic) Literature. In J. Tabbi (Org.), *The Bloomsbury Handbook of Electronic Literature* (pp. 73–92). Londres: Bloomsbury Publishing.

Deuze, M. (2006). O jornalismo e os novos meios de comunicação social. *Comunicação e Sociedade*, 9 (10), 15–37. DOI: 10.17231/comsoc.9(2006).1152

Dörr, K. N. (2016). Mapping the field of Algorithmic Journalism. *Digital Journalism*, 4 (6), 700–722. DOI: 10.1080/21670811.2015.1096748

Dialogflow. (2020) *Build natural and rich conversational experiences*. Google Cloud. Recuperado de <https://dialogflow.com>

Dornelas, R. (2018). Um jornalismo para chamar de meu? Algoritmos e o fenômeno da customização de notícias. *Revista Parágrafo*, 6 (1), 83–93. Recuperado de revistaseletronicas.fiamfaam.br/index.php/recicofi/article/view/713/562

Diakopoulos, N. (2019). *Automating the News: How Algorithms Are Rewriting the Media*. Cambridge, MA: Harvard University Press.

G1. (2018, setembro 17). G1 responde perguntas sobre eleição para presidente pelo assistente de voz do celular. Recuperado de g1.globo.com/economia/midia-e-marketing/noticia/2018/09/17/g1-responde-perguntas-sobre-eleicao-para-presidente-pelo-assistente-de-voz-do-celular.ghtml

Google Assistant. (s.d.). *Actions on Google*. Recuperado de <https://>

developers.google.com/assistant

Gardner, H., & Davis, K. (2013). *The App Generation*. New Haven, CT: Yale University Press.

Gillespie, T. (2018). A relevância dos algoritmos. *Revista Parágrafo*, 6 (1), 83–93. Recuperado de revistaseletronicas.fiamfaam.br/index.php/recicofi/article/view/722/563

Hayles, K. (2012). *How We Think: Digital Media and Contemporary Technogenesis*. Chicago, IL: The University of Chicago Press.

Lewis, S. C. (2015). Journalism in an Era of Big Data: Cases, concepts, and critiques. *Digital Journalism*, 3 (3), 321–330. DOI: 10.1080/21670811.2014.976399

Marconi, F. (2020). *Newsmakers: Artificial Intelligence and the Future of Journalism*. New York, NY: Columbia University Press.

Mycroft. (s.d.). Control technology with your voice. Recuperado de <https://mycroft.ai>

Napoli, P., & Caplan, R. (2018). Por que empresas de mídia insistem que não são empresas de mídia, por que estão erradas e por que isso importa. *Revista Parágrafo*, 6 (1), 143-163. Recuperado de revistaseletronicas.fiamfaam.br/index.php/recicofi/article/view/724/565

Noll, G. (2020). *Materializações Digitais da Cultura: os transatores vocais e a comunicação contemporânea* [tese de doutorado, Pontifícia Universidade Católica do Rio Grande do Sul]. Biblioteca Digital de Teses e Dissertações – PUCRS.

Pariser, E. (2012). *O filtro invisível: O que a internet está escondendo de você*. Rio de Janeiro, RJ: Zahar.

Pavlik, J. (2000). The Impact of Technology on Journalism. *Journalism Studies*, 1 (2), 229-237. DOI: 10.1080/14616700050028226

Pellanda, E. (2010). *Locast Civic Media: Internet Móvel, cidadania e informação hiperlocal*. Porto Alegre, RS: Edipucrs.

Pase, A., Pellanda, E., Cunha, M. R., & Streck, M. (2019). A Voz como Interface no Sistema Ecoevolutivo dos Assistentes Pessoais. In T. Burgos & R. Cunha (Orgs.), *Interfaces Contemporâneas no Ecossistema Midiático* (pp.62–85). Aveiro: Ria Editorial.

Primo, A., & Zago, G. (2015). Who And What Do Journalism? An actor-network perspective. *Digital Journalism*, 3 (1), 38–52. DOI: 10.1080/21670811.2014.927987

Recuero, R. (2011). *Redes sociais na internet*. Porto Alegre, RS: Sulina.

Recuero, R., Zago, G. S., & Soares, F. B. (2017). Mídia social e filtros-bolha nas conversações políticas no Twitter. *Anais do XXVI Encontro Anual da Associação Nacional de Programas de Pós-Graduação em Comunicação (COMPOS)*. Recuperado de www.compos.org.br/data/arquivos_2017/trabalhos_arquivo_XH5ITT DY1PYGE7PDUQJM_26_5374_18_02_2017_12_53_33.pdf

Santos, L. C. (2019). Máquinas falantes: proposta teórico-metodológica para pensar a agência a partir das assistentes pessoais digitais. In B. Polivanov, W. Araújo, C. Oliveira, & T. Silva (Orgs.), *Fluxos em redes sociotécnicas: das micronarrativas ao big data* (pp.206–232). São Paulo: Intercom. Recuperado de www.ibpad.com.br/wp-content/uploads/2019/12/fluxos-em-redes-sociotecnicas.pdf

Swaine, M., & Freiburger, P. (2014). *Fire in the Valley: The birth and death of the personal computer*. The Pragmatic Bookshelf.

Yin, R. K. (2001). *Estudo de caso: Planejamento e métodos*. Porto Alegre, RS: Bookman.

Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods*. Thousand Oaks, CA: Sage.

ANDRÉ FAGUNDES PASE. Doutor em Comunicação Social. Pesquisador e professor da Escola de Comunicação, Artes e Design – Famecos da PUCRS. Professor do Programa de Pós-Graduação em Comunicação da PUCRS. Pesquisador do Ubilab, Laboratório de Pesquisa em Mobilidade e Convergência Midiática, e coordenador do grupo de pesquisa JEDI - Jogos e Entretenimento Digital Interativo. E-mail: afpase@pucrs.br
Colaboração no artigo: concepção do desenho da pesquisa, desenvolvimento da discussão teórica, coleta e análise dos dados e revisão do texto.

GISELE NOLL. Doutora em Comunicação Social pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Professora dos Cursos de Jornalismo e Comunicação Social–Publicidade

e Propaganda da UNIJUÍ. Integrante do Grupo de Pesquisa JEDI - Jogos e Entretenimento Digital Interativo. E-mail: gisele.noll@gmail.com

Colaboração no artigo: concepção do desenho da pesquisa, desenvolvimento da discussão teórica, coleta e análise dos dados e revisão do texto.

MARIANA GOMES DA FONTOURA. Doutoranda em Comunicação Social pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Bolsista Capes. Integrante do Grupo de Pesquisa JEDI – Jogos e Entretenimento Digital Interativo. E-mail: marigfontoura@gmail.com

Colaboração no artigo: concepção do desenho da pesquisa, desenvolvimento da discussão teórica e revisão do texto.

LETÍCIA DALLEGRAVE. Mestra em Comunicação Social pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Integrante do Grupo de Pesquisa JEDI – Jogos e Entretenimento Digital Interativo. E-mail: leticia.dallegrave@gmail.com

Colaboração no artigo: concepção do desenho da pesquisa, desenvolvimento da discussão teórica e revisão do texto.