
A Ciência de Dados na Estrutura Organizacional das Empresas Jornalísticas no Brasil¹

Verônica SCHEIFER²
Universidade Estadual de Ponta Grossa, Ponta Grossa, PR

Resumo

A Ciência de Dados é uma área que vem sendo explorada por empresas de diversos segmentos. A era Big Data surge com a Internet e possibilita um avanço na criação constante de diversos tipos de dados. O presente trabalho procura fazer uma discussão entre o Jornalismo e a Tecnologia, com base na revisão bibliográfica, para evidenciar como essa ciência tem sido utilizada no Jornalismo. O objetivo principal é demonstrar possíveis usos na estrutura organizacional dos veículos jornalísticos: Estadão, Rede Globo, Jota e Nexo. O principal método para análise é a observação direta da estrutura organizacional dessas empresas e a entrevista. Em conclusão, é possível perceber que o Jornalismo tem buscado se aliar às novas tecnologias para se estruturar como empresa na era Big Data.

Palavras-chave: Ciência de Dados; Estrutura Organizacional; Jornalismo; Jornalismo de Dados; Big Data.

Introdução

A proposta do artigo é apresentar como a Ciência de Dados está sendo utilizada na estrutura organizacional das empresas jornalísticas e de que forma essa ciência influencia na produção e construção do trabalho jornalístico.

O objetivo é verificar práticas jornalísticas baseadas em dados em veículos de comunicação do Brasil. Para isso, decidiu-se fazer uma pesquisa, por meio de revisão bibliográfica, entrevistas e observação direta, junto às equipes de profissionais do Estadão, Globo, Jota e Nexo, empresas que têm trabalhado com esse novo conhecimento. As entrevistas e visitas aos veículos foram realizadas entre os dias 22 a 26 de abril de 2019. Em cada visita foram realizadas entrevistas com os responsáveis da área de tecnologia e, também, com os jornalistas da redação que trabalham com a Ciência de Dados.

¹ Trabalho apresentado no GP Conteúdos Digitais e Convergências Tecnológicas, XX Encontro dos Grupos de Pesquisas em Comunicação, evento componente do 43º Congresso Brasileiro de Ciências da Comunicação.

² Jornalista pela Universidade Estadual de Ponta Grossa e estudante de Big Data e Inteligência Analítica pela Pontifícia Universidade Católica do Paraná, e-mail: vescheifer@gmail.com.

Os jornais brasileiros, que investem na Ciência de Dados para a construção da notícia e para utilização em fins comerciais, sabem que no mercado atual isso é importante e útil à saúde da empresa e do Jornalismo.

A Ciência de Dados aumenta a possibilidade de entender aquilo que as empresas recebem de seus usuários e, também, propor possíveis efeitos para aquilo que elas disponibilizam em seus canais de comunicação.

O Jornalismo se utiliza de mecanismos da Ciência de Dados para aprimorar a estrutura produtiva e fazer frente ao mercado cada vez mais competitivo. Na análise realizada, verificou-se a existência de pelo menos cinco tipos de uso dos dados na estrutura organizacional: Área Interna, Conteúdo Jornalístico, Assinaturas, Marketing, Consumidor Final.

A exemplo de empresas de tecnologia, os veículos jornalísticos e os editores dos jornais, perceberam que, além de gerar informação de qualidade, o uso de dados contribui para a proposição de conteúdos de relevância e interesse público.

História do Jornalismo e Avanços Tecnológicos

O jornalismo teve sua expansão junto com a da imprensa no século XIX. No século seguinte, surgem os novos meios de comunicação de massa como o rádio e a TV, meios que ampliaram o espaço do jornalismo na sociedade. (TRAQUINA, 2005). Os avanços tecnológicos também contribuíram para a expansão e propagação de informação jornalística no mundo.

O impacto tecnológico marcou o jornalismo do século XIX como iria marcar toda a história do jornalismo [...] do século XX até o presente, apertando cada vez mais a pressão das horas de fechamento, permitindo a realização de um valor central da cultura jornalística, o imediatismo. De novas edições dos jornais no mesmo dia à quebra da programação televisiva anunciada como boletins, novos avanços tecnológicos nas últimas décadas do século XX tornaram possível, de longa distância, atingir o cúmulo do imediatismo a transmissão direta do acontecimento. (TRAQUINA, 2005, p. 53)

O desenvolvimento tecnológico contribui para diversos setores da sociedade, principalmente para a área da comunicação, por proporcionar um novo meio para a propagação da informação. Conforme Lima (2012, p. 54), “desde os primórdios da produção tecnológica do jornalismo, através da prensa, a atividade tem sido impactada constantemente por novas introduções tecnológicas”.

A criação da Internet com navegadores, programas e sistemas que automatizam as informações e, também, a Web 2.0, que atua hoje como um meio de divulgação e compartilhamento de informação, foram alguns dos marcos para a comunicação de massa, assim como também para o Jornalismo.

A Web 2.0 é a segunda geração de serviços on-line e caracteriza-se por potencializar as formas de publicação, compartilhamento e organização de informações, além de ampliar os espaços para a interação entre os participantes do processo. A Web 2.0 refere-se não apenas a uma combinação de técnicas informáticas (serviços Web, linguagem Ajax, Websyndication, etc.), mas também a um determinado período tecnológico, a um conjunto de novas estratégias mercadológicas e a processos de comunicação mediados pelo computador. (PRIMO, 2010, p.1)

Os avanços tecnológicos viabilizaram novidades, acrescentadas em várias áreas. A Big Data proporcionou uma modificação também das empresas tradicionais, incentivando novas possibilidades no mundo online. McAfee; Brynjolfsson (2012, p. 4, tradução nossa) mostram que “na verdade, o uso de big data tem o potencial de transformar também os negócios tradicionais. Pode lhes oferecer oportunidades ainda maiores de vantagem competitiva (as empresas on-line sempre souberam que estavam competindo em quão bem entendiam seus dados).”

A era Big Data é um conceito que vem junto com a popularização da Internet e possibilita as pessoas a convergência de informação por meio de dados e a criação de novas possibilidades de comunicação para a sociedade.

A Ciência de Dados e a Tecnologia

Atualmente, as pessoas acessam a Internet quase que instantaneamente, seja pelo computador ou por um Smartphone. Esse acesso, gera milhares de dados de várias origens, em bases de dados de diversas plataformas.

O termo Big Data surge como uma abreviação que indica uma nova abordagem tecnológica, para o entendimento da tomada de decisão. Existe uma quantidade muito maior de dados, que crescem 50% ao ano, segundo a IDC, que é uma empresa de pesquisa em tecnologia e apresenta o aumento de dados. Existem hoje novos fluxos de dados como de automóveis, sensores digitais em equipamentos industriais, caixas de transporte, medidores elétricos. (LOHR, 2012, p. 2)

Dados existem mesmo antes da internet, mas com a Big Data há um crescimento uniforme. Novos dados surgem e possibilitam um volume muito maior de

informação. IBM (2014) afirma que “Big Data é o termo utilizado para descrever grandes volumes de dados e que ganha cada vez mais relevância à medida que a sociedade se depara com um aumento sem precedentes no número de informações geradas a cada dia”

Os novos dados possuem um valor importante para as novas aplicações que surgem para compreender e gerar informações. Lohr (2012, p. 2, tradução nossa) observa que “a riqueza de novos dados, por sua vez, acelera os avanços na computação — um círculo virtuoso de Big Data. Os algoritmos de aprendizado de máquina, por exemplo, aprendem sobre os dados, e quanto mais dados, mais as máquinas aprendem.”

A Ciência de Dados ainda é uma área muito nova na academia e busca pelo seu reconhecimento mesmo antes da criação da Internet e do livre acesso a informações.

Mais de 50 anos atrás, John Tukey pediu uma reforma das estatísticas acadêmicas. Em *The Future of Data Analysis*, ele apontou para a existência de uma ciência ainda não reconhecida, cujo tema de interesse era aprender com dados, ou análise de dados. (DONOHO, 2015, p. 1, tradução nossa)

John Tukey, John Chambers, Bill Cleveland e Leo Breiman são os nomes que deram origem às primeiras discussões da nova ciência. Foi apenas vinte anos atrás que uma discussão sobre a expansão da estatística clássica para além do campo teórico foi reiniciada e levou ao uso do termo “Ciência de Dados” pela primeira vez.

Dez a vinte anos atrás, John Chambers, Bill Cleveland e Leo Breiman, mais uma vez, instaram mais uma vez as estatísticas acadêmicas a expandir suas fronteiras para além do domínio clássico das estatísticas teóricas; Chambers pediu mais ênfase na preparação e apresentação de dados do que na modelagem estatística; e Breiman pediu ênfase na previsão e não na inferência. Cleveland até sugeriu o apelido “Data Science” para seu campo imaginado. (DONOHO, 2015, p. 1, tradução nossa)

Os dados são utilizados pelas organizações e empresas, com o propósito de auxiliar o processo de tomada de decisão e, também, pelo meio acadêmico para produzir conhecimento científico de diversas áreas do conhecimento. Além disso, a importância que a área tem hoje em dia é resultado do percurso de busca pela institucionalização como ciência e educação formal, para formar profissionais que possam ocupar um local no mercado de trabalho na área.

As organizações e a ciência de modo geral dependem diretamente de dados. É notório o papel de dados enquanto insumos essenciais para o processo de tomada de decisão, bem como para o avanço científico. Porém, apenas recentemente a ciência de dados tem se institucionalizado como educação formal, e os profissionais com esta

formação têm ocupado posição de destaque no mundo do trabalho.
(CURTY; SERAFIM, 2016, p. 308)

A institucionalização da área deve-se também ao seu processo de importância na sociedade. O uso feito dos dados aplicados a modelos matemáticos e computacionais traz benefícios às pessoas. A Matemática e a Informática podem ser usadas para auxiliar e, também, manipular diversas bases de dados.

Os dados são domados e compreendidos usando modelos matemáticos e computacionais. Esses modelos, como metáforas na literatura, são simplificações explicativas. Eles são úteis para a compreensão, mas eles têm seus limites. Um modelo pode identificar uma correlação e extrair uma inferência estatística injusta ou discriminatória, com base em pesquisas on-line, afetando os produtos, empréstimos bancários e seguro de saúde oferecidos a uma pessoa, alertam os defensores da privacidade. (LOHR, 2012, p. 5, tradução nossa)

Há diversos usos dos dados, com o intuito de integrar áreas do conhecimento que possuem ferramentas para entender a demanda de informações que aumenta cada vez mais no mundo online.

Análise da influência da Ciência de Dados na Estrutura Organizacional do Estadão, Jota, Nexo e Rede Globo

A pesquisa envolveu quatro empresas jornalísticas. Além de visita e acompanhamento da produção, foram realizadas entrevistas com os profissionais que atuam nas equipes de tecnologias e de dados. Esses profissionais possuem graduação em Jornalismo, Publicidade e Propaganda e Estatística e ocupam cargos como Coordenador de Planejamento de Mercado, Jornalista de Dados/Visual, Editor de Dados, Infografista e Cientista de dados.

No Estadão <<https://www.estadao.com.br/>> há uma equipe de tecnologia, que trabalha diretamente com a Ciência de Dados. Durante a entrevista, o Coordenador de Planejamento de Mercado, Fernando Hideaki Rovaron (2019), que faz parte da equipe de tecnologia e atua principalmente no desenvolvimento de projetos de Ciência de Dados, voltados para a área interna e, também, para o consumidor final do Estadão, explica que no veículo tem-se “um foco muito maior em desenvolver a área interna, por exemplo, entender como são os padrões de consumo dos nossos leitores, dos nossos assinantes e isso em sinergia com o BI (Business Intelligence)”.

Rovaron (2019) comenta que a cultura de trabalhar com dados no Estadão ainda é muito recente, diferente de outras indústrias, como bancos, em que isso está inserido

na estrutura do próprio negócio. Ainda que recente, explica que há um crescente avanço de métodos cada vez mais inovadores para a utilização de dados por parte da empresa.

No Estadão, a proposta de trabalhar com inovação, em conjunto com a equipe de tecnologia, é uma meta. O setor de tecnologia presta consultoria e trabalhos técnicos em projetos e produtos com vista na inovação do jornalismo.

Os principais tipos de dados analisados pelo jornal são os dados navegacionais retirados diretamente do Google Analytics e, também, do sistema SAP, que é o sistema de Gestão Empresarial da empresa. Exemplos de dados dos assinantes são nome, cidade, e-mail e dados cadastrais. A empresa realiza um cruzamento de dados navegacionais e cadastrais para entender comportamentos dos usuários e leitores.

Rovaron (2019) salienta que os dados são utilizados para entender o comportamento a partir da “demografia e perfis, quem são assinantes, não assinantes, pessoas que consomem mais o Estadão, pessoas que consomem menos. Para entender como cada pessoa navega nosso site "e o que prefere, tipos de produtos e informações” são úteis para "podermos tornar nosso conteúdo mais atrativo" e buscar novos assinantes.

Os principais usos da Ciência de Dados pela equipe de tecnologia, atualmente, são modelos de recomendação de conteúdo de Newsletter, que é um projeto piloto, recente da empresa. O modelo utiliza a base de navegação para fazer o usuário assinar a Newsletter. Os conteúdos inseridos lá são personalizados para aquele determinado público, com interesse específico e, também, voltado para a assinatura.

Rovaron (2019) explica, ainda, que o trabalho é direcionado às “pessoas que já assinaram o Estadão, para saber o que elas fizeram antes de assinar e, também, para encontrar quem esteja no mesmo caminho, perto de assinar. É importante ter uma comunicação, junto do mercado de leitor”. Quem cuida disso é o pessoal que trabalha com CRM (Software de Gestão de Relacionamento com o cliente).

Em relação à influência editorial da Ciência de Dados no veículo, Rovaron (2019) observa: “A influência não é direta como um algoritmo do Facebook, que eu acredito que seja o futuro sim, porém devemos tomar muito cuidado com a questão que chamam muito de Ciência de Dados de fairness. Enquanto a gente está jogando um conteúdo, como o algoritmo do Facebook, isso pode se tornar muito mais complicado

quando vai para as cadeias sociais e a gente lida com algoritmos que estão decidindo o futuro de pessoas.”

O que a equipe de tecnologia disponibiliza dessa ciência para os jornalistas da redação são apenas as informações navegacionais de quantas pessoas navegaram em cada notícia, matéria ou editoria. Destaca que “cada chefe editorial, chefe de redação, tem acesso às informações dele e das outras áreas e eles são autossuficientes para tomar decisão a partir das informações que a gente passa”.

As análises mais pontuais são realizadas pela equipe de tecnologia, mas durante o dia a dia a redação recebe, a partir de dashboards, dados de navegação, quantas pessoas estão acessando, quanto tempo que os leitores, em média, ficam naquela página, por exemplo.

Segundo o Coordenador de Planejamento de Mercado, a Ciência de Dados dá um grande suporte para entender o comportamento do leitor e fornecer a informação certa, na hora certa e para o leitor certo, como o Marketing faz.

A segunda entrevista no jornal foi com o Jornalista de Dados/Visual, Rodrigo Menegat Schuinski, que fez uma descrição da Ciência de Dados para a criação de reportagens de Jornalismo, diferente dos trabalhos feitos pela equipe de tecnologia.

O trabalho realizado por ele é construir reportagens baseadas em dados, sem a intenção de medir audiência ou criar estratégias digitais como a equipe de tecnologia e Business Intelligence. Schuinski (2019) explica que “o dia de trabalho é basicamente quase 90% do tempo programando em Python geralmente, mas com esse foco muito claro de tentar encontrar uma coisa, para contar algo de interesse público.”

Conforme Schuinski (2019), o Estadão tem duas equipes dentro da redação que trabalham diretamente com dados. “Tem a gente que trabalha aqui na editoria de Infografia e tem o Estadão Dados, que é o pessoal que toca coisas mais ligadas ao diário. Enquanto a gente faz mais coisas de longo prazo. Trabalhar com dados é essa coisa: ‘o prazo é muito apertado para eu fazer uma investigação mais completa e tal’ [...] nós temos mais tempo e acho, que pela organização da minha equipe, conseguimos dedicar mais tempo a um projeto só. Isso é super legal. Então, acho que meus maiores problemas são basicamente a qualidade do dado público, não é nada na estrutura da redação”.

O Estadão Dados está diretamente ligado à editoria de Política do jornal e as pautas são de acordo com o jornal diário; enquanto a editoria de Infografia possui vínculo com o setor de Arte e pensa em Especiais de longo prazo.

A principal diferença entre as duas equipes é o tempo, o Estadão Dados trabalha com o diário e a Infografia realiza especiais. Schuinski (2019) descreve: “passamos mais tempo fazendo as coisas, dá pra dizer mais aprofundada. A gente tem mais tempo para mergulhar em um tópico. Temos mais tempo para tornar o produto mais sofisticado, talvez.”

A exemplo do Estadão, o Jornal Jota também se preocupada com a análise e uso de dados. O Jornal Jota, <<https://www.jota.info/>>, é um veículo que faz coberturas das instituições públicas e governamentais brasileiras e possui uma equipe de dados que investe cada vez mais nesse tipo de avanço tecnológico.

O veículo faz análises variadas e uma das mais recentes foi sobre o “tempo médio de processos da Lava Jato”. O Editor de Dados, Guilherme Jardim Duarte, explica o processo: “analisamos como que a Lava Jato estava se comportando, coisas desse tipo, até análises mais preditivas, modelagens estatísticas, como o repositório de pesquisas eleitorais do ano passado que fizemos. Agora também estamos com um modelo de previsão legislativa, mas ele não vai para o Jota Info, não é para a área pública, é para assinatura.”

Os dados dos usuários também são utilizados e analisados pelo veículo para buscar maior número de assinantes. Duarte (2019) descreve: “tínhamos todo um modelo para tentar prever que tipo de perfil teria maior probabilidade de assinar o Jota e orientamos os esforços nesse sentido. Fazemos uma análise dos assinantes, para ver quem são, para definir personas e tudo mais.” Antigamente essa era uma das funções do Editor de Dados, mas agora há pessoas específicas do Marketing que lidam com as page views para a produção de uma matéria específica, por exemplo.

Na questão das métricas de audiência, são feitas análises que influenciam editorialmente no jornal, com foco no usuário. Duarte (2019) diz que são feitas análises sim “mas isso não está mais comigo hoje, por exemplo, em termos de buscar novos assinantes para o site. Nós tínhamos todo um modelo para tentar prever que tipo de perfil teria maior probabilidade de assinar o Jota e nós orientávamos os esforços nesse sentido [...] o que tem influências na parte editorial.”

Segundo Duarte (2019), as fontes de dados são coletadas pela equipe ou pelos sistemas. O foco do Jota é o Judiciário ou carreiras dentro do direito e nessa área, no Brasil, é complicado coletar dados, diferente do Congresso, em que há API's. Para a coleta de dados, Duarte (2019) afirma que "temos robôs que utilizam web scraping de tribunais, temos um robô que entra todo dia no site do TJ de SP e verifica todas as decisões de primeira instância e segunda instância relacionadas à liberdade de expressão e dispara todo dia de amanhã e mails para que os repórteres saibam o que saiu. Isso aí gera várias matérias."

O editor também destaca que "a Ciência de Dados no Jornalismo é essencial tanto é que a tendência é o Jota aumentar a equipe, pois isso faz a diferença para a empresa." O jornal já usava a Ciência de Dados para fazer o perfil dos assinantes e agora estão utilizando sobretudo para o Jota Info.

O Nexa <<https://www.nexojournal.com.br/>> é um jornal digital, que trabalha com notícias baseadas em dados e estatísticas. O Infografista do jornal, Rodolfo Almeida, explica que o veículo tem o objetivo de publicar três infográficos por semana e, também, há outros projetos que necessitam de análise de dados que vem da redação. Almeida (2019) esclarece que "às vezes o jornalista tem um mapa, uma tabela ou uma base de dados que ele quer explorar e ele pede ajuda para a gente. Então ajudamos a redação, fazemos os nossos projetos e, também, os interativos do site."

Os gráficos propostos têm o intuito de se tornar uma ferramenta para que o usuário possa utilizar. Almeida (2019) explica: "temos os gráficos que saem na sessão Gráfico que são historinhas fechadas, mas tentamos fazer com que eles sejam como uma ferramenta para explorar. Fizemos um de como a cidade gasta o orçamento, isso foi na época das eleições municipais. Pegamos uma base que em tese está disponível para todo mundo e deixamos disponível de uma maneira mais fácil para explorar e navegar."

A ideia é que o leitor possa entrar e ver a porcentagem dos gastos, como um sistema que fornece uma base de dados que guia a leitura de uma maneira que faça sentido para ele. Almeida (2019) afirma que há "um cuidado com todos os nossos gráficos para serem entendidos por qualquer tipo de pessoa."

A Ciência de Dados no Nexa é bastante explorada pelo Jornalismo de Dados, para criar histórias e facilitar a leitura do leitor de pautas que produzem estatística e gráficos, por exemplo. Além das pautas produzidas normalmente, há um apelo por

novidades e curiosidades. Conforme Almeida (2019) “Tem curiosidades que também se tornam pautas, por exemplo, se os números da mega-sena são aleatórios. Nós procuramos em base de dados e raspamos os dados.”

Um dos principais exemplos foi o projeto de dados que demonstra quanto tempo leva para chegar em qualquer lugar do Brasil, partindo de 5 capitais. Almeida (2019) explica que “usamos uma API open street maps, que vê o desenho de ruas e calcula o tempo que leva para chegar em cada ponto.”

A Rede Globo, outra empresa visitada durante a pesquisa, possui equipes especializadas em Ciência de Dados. O Cientista de Dados da empresa, Luis Guilherme Giordani, trabalha na parte de publicidade digital, que é a venda e colocação de anúncios nas páginas da Globo. Giordani (2019) explica que “se você entrar na globo.com agora, vai ver ali na home que tem dois ou três banners. A área que eu trabalho é a que faz comercialização e a otimização interna dessa publicidade.”

O principal objetivo da equipe é trabalhar com a experiência do usuário nas páginas da Globo e observar como é a sua navegação, para gerar demandas segmentadas. Giordani (2019) aponta que “algo muito comum é, vamos supor, uma empresa de carros querer anunciar um carro novo e quer gerar novos leads ou novas potenciais pessoas a comprar. É uma demanda muito comum deles querer, ao anunciar na nossa página, que a gente entregue o anúncio para um segmento muito específico de pessoas”.

Para realizar essa segmentação, a empresa, com base na análise de dados, precisa identificar o perfil dos consumidores e usuários. Giordani (2019) demonstra um exemplo de abordagem dos possíveis anunciantes que procuram a Globo, “eu vou anunciar um carro popular e eu quero pessoas jovens de 18 a 24 anos.”

Para entregar e fechar contratos, são desenvolvidos primeiramente métodos para encontrar esse tipo de usuário na página. Segundo Giordani (2019), “faz parte do nosso trabalho vasculhar os dados que a gente tem, criar métodos para encontrarmos essas pessoas dentro das nossas páginas e, é claro, existem tecnologias e fontes de dados por trás disso tudo.”

Giordani (2019) explica que “utiliza-se a Ciência de Dados para nunca ficarmos com um espaço em branco, para não acontecer de você estar navegando em uma página e ver um quadrado meio cinza. Isso acontece porque, de alguma forma, aquela página

que você está acessando não conseguiu te entregar publicidade, e então o dono da página está perdendo dinheiro e espaço.”

A equipe da Globo é composta por dois Cientistas de Dados, dois Engenheiros de Dados e um Analista de Negócios. Ela, atualmente, trabalha para buscar soluções com base nos dados fornecidos nos servidores da empresa e, também, pelas informações navegacionais de cada usuário.

As análises realizadas servem segundo Giordani (2019) para traçar perfis ou personas dos usuários e até mesmo auxiliar na construção de conteúdos com composição de teor de humor da página. O trabalho da equipe inicia no problema de negócios, onde começam os estudos de viabilidade. As ideias só se concretizam, se houver dados para serem analisados.

Após a mineração desses dados, podem ser feitas análises para gerar conhecimento de negócio. Giordani (2019) explica que “pode-se observar a correlação entre duas variáveis, posso ver se a quantidade de vezes que a pessoa pausa o vídeo tem alguma correlação com ela compartilhar o vídeo no facebook, para quando eu for construir a técnica ou modelo, eu já ter uma ideia do que vai acontecer”.

O último passo “seria usar uma técnica, seja de Machine Learning, Estatística ou de visualização que resolva o problema. Quando tiver resultados legais, depois de ter testado algoritmos, features, informações, eu posso gerar as predições na amostra de teste e aí, no último passo, vou apresentar o resultado para um gestor.”

Na Rede Globo há também equipes que são especializadas nessa ciência com o intuito de melhorar o trabalho jornalístico e apresentar soluções para as redações que trabalham com Jornalismo de Dados. O Cientista de Dados, Rodolfo Viana, trabalhou anteriormente como Jornalista de Dados e foi transferido para executar a função de Cientista de Dados para desenvolver projetos que podem, também, auxiliar as redações.

Ele trabalha com a parte de projetos mais densos e que geram informações, pautas e conteúdos jornalísticos. Viana (2019) destaca que eles dispõem de uma ferramenta agora que “é para a cópia de todos os dados públicos disponíveis. Nunca sabemos quando esses dados estarão indisponíveis [...] quando eu falo dados públicos, por exemplo, beneficiários do Bolsa Família, a gente tem aproximadamente 13 milhões de beneficiários por mês e aí você tem dados como nome dessa pessoa, CPF, o valor

que ela recebe, quantos filhos essa pessoa tem, a data que fez o saque. Tudo isso o governo disponibiliza.”

O trabalho de acesso desses dados é manipulado por Viana (2019), para que seja automatizado para os jornalistas. Segundo ele “para você acessar é preciso entrar mês a mês para verificar tudo isso, em alguns casos tem que estar estado a estado, ou seja, para você trabalhar se quiser saber a média de recebimentos dos beneficiários do Bolsa Família, não tem como saber isso a não ser perguntando para a assessoria de imprensa.”

As principais ferramentas desenvolvidas por Viana são para a raspagem de dados, ou seja, uma maneira rápida para captar inúmeras linhas de bases de dados disponíveis. Viana (2019) destaca que estão “criando ferramentas para a raspagem de dados, são scripts, programas e aplicações que vão coletando esses dados e salvando no servidor, para quando um Jornalista precisar, ele poder acessar isso e fazer seus próprios cálculos, sem ter que entrar, por exemplo, no site do Bolsa Família.”

Um projeto que está sendo realizado para automatizar o trabalho da redação é sobre mobilidade, para saber quantos ônibus estão rodando em tempo real, em que lugar estão, quanto tempo eles levam de um ponto a outro, quantos passageiros têm para realizar um raio X da mobilidade.

Para realizar esse trabalho, Viana (2019) esclarece que é preciso alguns dados, por exemplo, do GPS. Com ele, a gente consegue saber que, por exemplo, nesse momento o ônibus está na posição X e daqui 5 minutos ele está na posição Y, em 5 minutos ele percorreu tantos metros e isso a gente vai aplicando trecho a trecho. Estamos falando de um ônibus, mas, se a gente utiliza vários ônibus em diversos dias, a gente consegue desenvolver um algoritmo e ensinar para que o próprio sistema entenda que aquele trajeto é feito, em média, em tanto tempo.”

O trabalho então do Cientista de Dados, conforme Viana (2019), é traduzir diversos dados em poucas linhas, para tratá-los em dashboards com algumas análises já realizadas para auxiliar na tomada de decisão do Jornalista. Com base nas entrevistas, as aplicações e usos podem ser divididas em tipos, distribuídos em setores de cada veículo, conforme descritas na Tabela 1.

Tabela 1 - Principais aplicações e usos da Ciência de Dados

Quem	Como	Onde	Tipos de aplicações e usos em setores
------	------	------	---------------------------------------

Fernando Hideaki Rovaron	Ciência de Dados	Estadão	Área Interna e Consumidor Final
Rodrigo Menegat Schuinski	Jornalismo de Dados e Ciência de Dados	Estadão	Conteúdo Jornalístico
Guilherme Jardim Duarte	Jornalismo de Dados e Ciência de Dados	Jota	Conteúdo Jornalístico, Assinaturas, Marketing, Consumidor Final
Rodolfo Almeida	Jornalismo de Dados e Ciência de Dados	Nexo	Conteúdo Jornalístico
Rodolfo Viana	Jornalismo de Dados e Ciência de Dados	Rede Globo	Conteúdo Jornalístico
Luiz Guilherme Giordani	Ciência de Dados	Rede Globo	Assinaturas, Marketing, Consumidor Final

Fonte: a autora.

Verifica-se que há uma demanda da Ciência de Dados na estrutura organizacional da redação jornalística, para a produção de conteúdo jornalístico de qualidade em todos os veículos analisados. Contudo, há também uma influência na estrutura do Marketing do Estadão, Jota e Rede Globo para a venda de assinaturas e com a finalidade aprimorar o atendimento ao consumidor final. A ciência proporciona uma segmentação dos usuários com o objetivo de atraí-los com possibilidades que interessem cada perfil que acessa as páginas online dessas empresas.

Uma empresa possui em sua dimensão divisões que possibilitam a concentração de equipes e departamentos para realizar o trabalho. A estrutura organizacional é que define cada uma desses grupos e divisões.

A estrutura organizacional é o produto das decisões de divisão e coordenação do trabalho, e define não apenas as atribuições específicas, mas também o modo como devem estar interligados os diversos grupos especializados, chamados departamentos (MAXIMIANO *apud* PAULA, 2017, p.24).

Devido aos avanços tecnológicos, muitas empresas estão criando novos setores e equipes para trabalhar com a Internet, seja para melhorar e automatizar o trabalho interno, quanto para criar novas possibilidades em relação aos clientes.

Considerações Finais

Evidenciou-se pelo menos cinco tipos de aplicações e usos da Ciência de Dados: Área Interna, Conteúdo Jornalístico, Assinaturas, Marketing, Consumidor Final. Para todos esses usos há investimentos em tecnologia, o que demonstra a importância que as tecnologias vêm assumindo no campo do jornalismo e nas empresas de comunicação.

É possível identificar que a Ciência de Dados é um campo muito novo ainda no Jornalismo, mas que vem se popularizando cada vez mais. A tecnologia e as novas áreas da ciência, como é a Ciência de Dados, tem potencialidades para mudar a prática do jornalismo e a estrutura empresarial dos veículos jornalísticos como demonstra o trabalho. Com certeza trata-se de um campo que será cada vez mais explorado pelas empresas de comunicação que sempre se colocam na ponta do uso das novas tecnologias e das novas descobertas.

Embora o foco maior seja ainda no Jornalismo de Dados, aos poucos, essa ciência passa a ser explorada por outros setores das empresas jornalísticas. Por ser uma área muito nova, acredita-se que há muito a explorar. A quantidade de dados que surgem atualmente e os avanços tecnológicos tendem a possibilitar muitas outras análises, conforme as demandas vão aparecendo.

A área já possui grande aplicabilidade e potencial para se constituir em soluções à crise do Jornalismo. A adaptabilidade do campo jornalístico e profissional a esses avanços e conhecimentos é primordial, não só do ponto de vista da informação, como também em relação à estrutura das empresas.

É possível elencar alguns possíveis estudos que podem ser feitos em relação a Big Data e a crise do Jornalismo como, por exemplo, uma análise dos modelos de negócios hoje adotados pelo setor e a prática do Jornalismo na Internet.

É perceptível que o Jornalismo possui papel social de extrema importância para a sociedade, para a democracia, em favor de direitos e justiça social. Por isso, precisa se adaptar e aproveitar os novos conhecimentos para ampliar sua ação no meio social em defesa, principalmente, daqueles que não têm voz na sociedade.

É preciso encontrar alternativas para driblar a crise e para continuar atualmente em favor da coletividade. A tecnologia possui um papel de extrema importância nesse sentido e pode auxiliar o campo a superar seus desafios e seus próprios limites.

Referências bibliográficas

CURTY, Renata Gonçalves; SERAFIM, Jucenir da Silva. A formação em ciência de dados: uma análise preliminar do panorama estadunidense. **Informação & Informação**, [S.l.], v. 21, n. 2, p. 307–331, 2016.

DONOHO, David. **50 years of Data Science**. Massachusetts: MIT, 2015.

IBM, 2013. **Saiba o que é o Big Data e os desafios que as empresas enfrentam**. Disponível em: <https://www.ibm.com/midmarket/br/pt/infografico_bigdata.html> Acesso em 07/08/2018.

LIMA JUNIOR, Walter Teixeira. **Big Data, Jornalismo Computacional e Data Journalism: estrutura, pensamento e prática profissional na Web de dados**. São Paulo: Universidade Metodista de São Paulo, 2012.

LOHR, Steve. **The Age of Big Data**. Big Data's Impact in the World. New York: The New York Times, 2012.

MCAFEE, Andrew; BRYNJOLFSSON, Erik. Big Data: The Management Revolution. Exploiting vast new flows of information can radically improve your company's performance. But first you'll have to change your decision-making culture. Massachusetts: **Harvard Business Review**, 2012.

PAULA, Geraldo Marcelio de. **Estruturas Organizacionais: O Papel do gestor de nível intermediário**. Universidade FUMEC, Faculdade de Ciências Empresariais: Belo Horizonte, 2007.

PRIMO, Alex Fernando Teixeira. **A busca por fama na web: reputação e narcisismo na grande mídia, em blogs e no twitter**. Brasil: Intercom, 2009.

TRAQUINA, Nelson. **O estudo do jornalismo no século XX**. São Leopoldo: Editora. Unisinos, 2001.