



UNIVERSIDADE DE BRASÍLIA
FACULDADE DE COMUNICAÇÃO
CURSO DE COMUNICAÇÃO ORGANIZACIONAL

LAURA PEREIRA HIJDRA VAN HAGEN

**PROCESSOS DE *BIG DATA* NA INICIATIVA PRIVADA E COMPORTAMENTO DO
CONSUMIDOR: ESTUDO EXPLORATÓRIO SOBRE CARACTERÍSTICAS,
IMPACTOS E TENDÊNCIAS**

Brasília

2019

LAURA PEREIRA HIJDRA VAN HAGEN

**PROCESSOS DE *BIG DATA* NA INICIATIVA PRIVADA E COMPORTAMENTO DO
CONSUMIDOR: ESTUDO EXPLORATÓRIO SOBRE CARACTERÍSTICAS,
IMPACTOS E TENDÊNCIAS**

Trabalho de Conclusão de Curso apresentado ao curso de Comunicação Social – Comunicação Organizacional da Faculdade de Comunicação da Universidade de Brasília, como requisito parcial para obtenção do grau de Bacharel em Comunicação Social – Comunicação Organizacional.

Orientador: Professor Dr. Sivaldo Pereira da Silva

Brasília

2019

FOLHA DE APROVAÇÃO

LAURA PEREIRA HIJDRA VAN HAGEN

PROCESSOS DE *BIG DATA* NA INICIATIVA PRIVADA E COMPORTAMENTO DO CONSUMIDOR: ESTUDO EXPLORATÓRIO SOBRE CARACTERÍSTICAS, IMPACTOS E TENDÊNCIAS

Trabalho de Conclusão de Curso apresentado ao curso de Comunicação Social – Comunicação Organizacional da Faculdade de Comunicação da Universidade de Brasília, como requisito parcial para obtenção do grau de Bacharel em Comunicação Social – Comunicação Organizacional, orientada pelo prof. Dr. Sivaldo Pereira da Silva.

MEMBROS DA BANCA EXAMINADORA

Orientador: Prof. Dr. Sivaldo Pereira da Silva (FAC/UnB)

Examinador: Prof. Ms. Newton Duarte Molon (IP/UnB)

Examinadora: Prof^a. Dr. Délcia Maria de Mattos Vidal (FAC/UnB)

Prof. Dr. Felipe da Silva Polydoro (FAC/UnB) – suplente

Brasília

2019

AGRADECIMENTOS

Gostaria de agradecer, primeiramente, aos meus queridos tios Patrícia Vieira e Alexandre Zart por me permitirem utilizar seu espaço para desenvolver este trabalho com tranquilidade e concentração, e por todo o apoio e amor que me deram durante os últimos meses.

À minha querida mãe, Marta Vieira, obrigada por ter me ajudado com tanto carinho nesse momento tão desafiador, e por ter me ensinado tanto sobre inteligência emocional e coragem.

Também gostaria de reconhecer meu pai, Peter van Hagen, pelo encorajamento e pela força para enfrentar desafios durante minha trajetória acadêmica e profissional, além de tantas dicas acerca do assunto e o compartilhamento de materiais sobre o tema, que foram muitas vezes utilizados neste trabalho.

À minha querida avó, Maria Vieira, por todo o suporte incondicional e carinhos imensuráveis. Te agradeço pela força que me deu em vários momentos. Você é meu combustível.

Para minha grande amiga, Marina Aguilera, que sempre acreditou na qualidade deste trabalho e no meu potencial de concluí-lo com sucesso, além do carinho de sempre: meu sincero obrigada.

Ao Diogo Pessoa, amigo, colega e mentor para diversos assuntos da vida, obrigada por todos os sábios conselhos que sempre me indicam o caminho certo.

Também agradeço a Juliana Americano, minha chefe, por todo o suporte e pela compreensão com as demandas advindas da monografia e pelas palavras amigas e motivadoras dos últimos meses. Elas fizeram toda a diferença.

Por fim, gostaria de agradecer ao meu orientador, Sivaldo Pereira, pelo encantamento que provocou em mim para com o tema de processos de *Big Data*, pelas orientações de qualidade, pelos conhecimentos transferidos e pela disponibilidade para tirar minhas dúvidas e me auxiliar a finalizar esta monografia com êxito.

RESUMO

A tecnologia de dados está crescendo e se tornando substancial para organizações, sendo acompanhada pela onda da transformação digital e dentro de um contexto de Indústria 4.0. Os processos de Big Data são comumente abordados juntamente à conceitos de Inteligência Artificial, na medida em que essas tecnologias são correlacionadas. A utilização de processos de Big Data é crescente em diversos setores da iniciativa privada, na medida em que o dado torna-se uma matéria-prima cada vez mais relevante. Seus benefícios são inúmeros para os negócios, especialmente no que tange à compreensão do comportamento do consumidor e identificação de novas formas de atingi-lo, por meio da análise preditiva e da personalização, por exemplo. A implantação de ferramentas de Big Data pode ser desafiadora nas organizações, provocando novos aspectos da cultura organizacional, que agora precisa ser voltada para dados, e demandando novos tipos de profissionais, especializados em habilidades inéditas. Além disso, há riscos envolvidos na utilização dessas ferramentas, com foco na privacidade do usuário e controle social, especialmente se relacionados aos conceitos de datificação, que consideram os dados on-line quantificados como forma de entendimento do comportamento humano. Considerando este cenário, as empresas (B2B ou B2C) buscam formas de garantir a segurança dos dados de seus clientes, na medida em que a pauta torna-se cada vez mais relevante e polêmica. Por meio de entrevistas com profissionais de empresas e experts no tema, estudos de caso, sondagem de opinião e revisão bibliográfica, este trabalho buscou entender como organizações privadas utilizam processos de Big Data e seus reflexos no comportamento do consumidor contemporâneo.

Palavras-chave: *Big Data*. Indústria 4.0. Datificação. Comportamento do Consumidor. Riscos.

ABSTRACT

Data technology is growing and becoming substantial for organizations, being accompanied by the wave of digital transformation and within an Industry 4.0 context. Big data processes are commonly approached alongside Artificial Intelligence concepts as these technologies are correlated. The use of big data processes is increasing in various private sector sectors as data becomes an increasingly relevant raw material. Its benefits are countless for business, especially in understanding consumer behavior and identifying new ways to achieve it through predictive analytics and personalization, for example. Deploying big data tools can be challenging in organizations, sparking new aspects of the organizational culture that now needs to be data driven, and demanding new types of professionals who specialize in unprecedented skills. In addition, there are risks involved in using these tools, focusing on user privacy and social control, especially if related to the concepts of datification, which consider quantified online data as a way of understanding human behavior. Considering this scenario, companies (B2B or B2C) are looking for ways to ensure the security of their clients' data, as the agenda becomes increasingly relevant and controversial. Through interviews with business professionals and experts, case studies, opinion polls and bibliographic review, this paper sought to understand how private organizations use Big Data processes and their reflections on contemporary consumer behavior.

Keywords: Big Data. Industry 4.0. Datification. Consumer Behavior. Risks.

LISTA DE ILUSTRAÇÕES

Figura 1 - Relação entre as tecnologias	24
Figura 2 – Nove Tecnologias da Indústria 4.0	32
Figura 3 - Big Data: criação de valor para organização.....	35

LISTA DE QUADROS

Quadro 1 - Tipos de informações coletáveis	17
Quadro 2 – Ganhos de empresas a partir do acesso aos dados compartilhados do Facebook	29

LISTA DE GRÁFICOS

Gráfico 1 - Compreensão sobre processos de Big Data (%).....	64
Gráfico 2 - Presença dos processos de Big Data no cotidiano (%).....	65
Gráfico 3 - Utilização de serviços que possuem Big Data (%).....	66
Gráfico 4 - Processos de Big Data nas matérias cursadas na UnB (%).....	67
Gráfico 5 - Percepção sobre extração dos dados (%).....	68
Gráfico 6 - Termos de privacidade dos aplicativos (%).....	68
Gráfico 7 - Fornecimento do CPF (%).....	69
Gráfico 8 - Correlação entre variáveis.....	70

SUMÁRIO

1. INTRODUÇÃO	11
1.1 PROBLEMA DE PESQUISA.....	12
1.2 JUSTIFICATIVA.....	13
1.3 OBJETIVOS GERAIS E ESPECÍFICOS	14
1.4 ESTRUTURA DA MONOGRAFIA.....	15
2. <i>BIG DATA</i> E INTELIGÊNCIA ARTIFICIAL: BASES CONCEITUAIS, AVANÇOS E RISCOS	15
2.1 DEFINIÇÃO DE PROCESSOS DE <i>BIG DATA</i>	16
2.2 PROCESSOS DE <i>BIG DATA</i> E INTELIGÊNCIA ARTIFICIAL	20
2.3 ASPECTOS NEGATIVOS DA DATIFICAÇÃO E DATAÍSMO.....	24
3. PROCESSOS DE <i>BIG DATA</i> NA INICIATIVA PRIVADA	31
3.1 PROCESSOS DE <i>BIG DATA</i> NOS NEGÓCIOS.....	31
3.2 COMPORTAMENTO DO CONSUMIDOR	38
3.2.1 Análise Preditiva	40
3.2.2 <i>Mass Customization</i> e Personalização	41
3.2.3 Mecanismos de Recomendação	43
3.2.4 Precificação Dinâmica	44
4. ESTUDO DE CASO: CARACTERIZANDO USO POR EMPRESAS E RECEPÇÃO DE ESTUDANTES UNIVERSITÁRIOS	46
4.1 PARÂMETROS METODOLÓGICOS	46
4.2 <i>BIG DATA</i> E EXPERIÊNCIAS DE EMPRESAS.....	48
4.2.1 Foxtrot Systems	49
4.2.2 In Loco	50
4.2.3 AIS Digital	52
4.2.4 Twitter	54
4.3 PRINCIPAIS DESAFIOS PARA IMPLEMENTAÇÃO DE PROCESSOS DE <i>BIG DATA</i> NAS EMPRESAS	57
4.4 RISCOS RELACIONADOS À PRIVACIDADE E ÉTICA NA UTILIZAÇÃO DE DADOS	60
4.5 ESTUDO DE PERCEPÇÃO: ALUNOS DE COMUNICAÇÃO DA UNIVERSIDADE DE BRASÍLIA	63
5. CONCLUSÃO	72
REFERÊNCIAS	75

1. INTRODUÇÃO

O cenário atual do mundo é caracterizado por uma palavra: tecnologia. Avanços desenfreados ocorrem por todo o globo e impressionam pessoas, desafiam profissionais, reestruturam organizações e quebram paradigmas dos moldes por meio dos quais vivíamos no passado.

No mundo dos negócios, organizações privadas e públicas buscam se adaptar à transformação digital. Algumas sofrem e, muitas vezes, quebram, ao passo que outras surfam na onda do digital para alcançarem patamares cada vez mais altos. Na esfera corporativa, especialmente nas gigantes empresas de alta tecnologia como as que estão presentes no Vale do Silício, a utilização de Inteligência Artificial, Ferramentas de *Big Data* e *Analytics*, *Machine Learning*, *Deep Learning* e Internet das Coisas (*IoT*) torna-se cada vez mais exponencial e surpreendente.

O momento atual está em consonância com a onda de *startups* e todo o contexto da Era Digital ou Informacional, período consolidado no fim do século XX e no qual estamos inseridos, o qual corresponde, principalmente, à otimização dos fluxos informacionais do mundo.

A conjuntura socioeconômica na qual os processos de *Big Data* passaram a ser utilizados é marcada por aspectos bastante particulares no que diz respeito à noção de tempo e espaço e à nova forma de relacionar-se, que passou a ser intermediada por plataformas digitais. Dessa forma, os processos de *Big Data* alteram completamente a dinâmica em que vivemos, como nos relacionamos uns com os outros e com os diversos entes ao redor, como organizações públicas, privadas, governo e empresas estatais.

O dado tornou-se a matéria-prima mais relevante do mundo. Essa transformação está refletida na ascensão de empresas de tecnologia, que atualmente superaram as grandes petrolíferas do século XX no que se refere a valor de mercado. A datificação, definida por processos que envolvem a quantificação de dados sobre o comportamento humano extraídos por meio de ambientes digitais, permite que empresas criem produtos e serviços de forma extremamente personalizada e sejam embasadas em análise preditiva por meio do *mass customization*, termo que também se refere à produção em massa envolvendo os elementos personalizáveis do cliente final.

Dessa maneira, os processos de *Big Data* são uma valiosa fonte de identificação de padrões, preferências e hábitos dos mais variados públicos. A partir dessa tecnologia, é possível realizar previsões e utilizá-las a favor do negócio, uma vez que os *insights* obtidos podem direcionar as organizações no sentido de enviar a mensagem correta, no momento correto, para o público mais adequado. Isso ocorre especialmente em empresas de varejo que possuem canais digitais, conforme pontuado por Rebecca Barros, PhD em *Economic Analytics* (informação verbal) ¹.

O trabalho em questão, conduzido por viés exploratório, visa compreender como os processos de *Big Data* estão sendo utilizados na iniciativa privada e analisar o impacto dos mesmos no comportamento do consumidor. São destacados, no estudo em questão, conceitos fundamentais acerca de ferramentas de *Big Data*, bem como de tecnologias correlatas.

As ferramentas de *Big Data* são apresentadas como divisores de águas nos negócios, atualmente, e discute-se como a apropriação dessa tecnologia tem se tornado fundamental em vários setores do mercado. Ademais, é construída uma visão geral sobre empresas que já adotaram processos de *Big Data* e quais foram os resultados obtidos, bem como os desafios enfrentados durante esse processo de mudança.

Além dos aspectos positivos em termos de negócio, os processos de *Big Data* geram certa preocupação no que tange à privacidade dos indivíduos. A personalização de ofertas envolve a captação de dados de usuários nas redes, sendo que esses usuários, muitas vezes, não possuem conhecimento suficiente para compreender como a extração de suas informações está sendo feita e o seu nível de exposição. Outra discussão relevante diz respeito ao controle social intrínseco a essa dinâmica, uma vez que as pessoas se tornam mais suscetíveis à manipulação por parte de diversos entes.

1.1 PROBLEMA DE PESQUISA

Como os processos de *Big Data* estão sendo utilizados atualmente na iniciativa privada e impactos no comportamento do consumidor.

¹ Entrevista concedida por BARROS, Rebecca. [16 set de 2019]. Entrevistador: Laura Pereira Hijdra van Hagen. Brasília, 2019. 1 arquivo .docx. A entrevista na íntegra encontra-se transcrita no Apêndice A desta monografia.

1.2 JUSTIFICATIVA

Este estudo constitui, primeiramente, em uma análise da contemporaneidade no que tange à quantidade de informações existentes, advindas de inúmeras fontes. Isso reflete a mudança ocasionada pelas tecnologias emergentes e a sua influência em vários campos, inclusive nas relações entre sociedade e organizações; sociedade e sociedade; e sociedade e tecnologia.

A presente monografia também fomenta um panorama da utilização de ferramentas de *Big Data* em organizações privadas, englobando objetivos do uso, resultados percebidos e principais desafios enfrentados, incluindo autores com renome na temática, cases de grandes empresas sob o olhar de gestores atuantes na iniciativa privada, que a partir de vivência prática e contato com o mercado, puderam contribuir com suas percepções e experiências relativos a processos de *Big Data* na estrutura de suas respectivas empresas. São abordados, por exemplo, os principais desafios para implementação desta tecnologia e as formas de lidar com questões de privacidade, além do que pode ser proveitoso para indivíduos que estejam à procura de embasamento para desenvolver seus próprios projetos.

Os benefícios dos processos de *Big Data* para as organizações são diversos no que diz respeito ao processo de tomada de decisão; ao acompanhamento de concorrência; à criação de estratégias de *marketing*; ao relacionamento com cliente, dentre outros. A utilização dessas ferramentas tornou-se fundamental para os negócios, ao passo que fomenta vantagens competitivas consideráveis a partir de seu poder preditivo. No *E-Commerce* há inúmeros exemplos: o próximo produto que o usuário pretende adquirir; o momento correto para oferecer passagens de avião; e a quais anúncios um determinado indivíduo é mais suscetível. É possível, além disso, prever tendências para vários nichos de mercado, subsidiar processos de tomada de decisão em companhias, gerar *insights* e criar produtos/serviços com base em padrões de comportamento identificados.

Entretanto, o cenário não é completamente positivo. São levantados alguns dos riscos envolvendo a utilização de processos de *Big Data*. Questões relacionadas à invasão de privacidade e ao potencial de manipulação pelo alto nível de segmentação de públicos-alvo, capacidade inerente à tecnologia, estão bastante em voga atualmente e são trabalhadas por diversos autores da área. Não obstante, outro aspecto negativo concerne ao monitoramento e controle social proporcionados pela

tecnologia. Alguns teóricos da área associam o termo *Big Data* a *Big Brother*, programa de *reality show* onde é possível assistir a todos os movimentos e ações dos participantes, que ficam confinados em uma casa 24 horas por dia.

Por fim, expõe-se também uma análise realizada com os estudantes da Faculdade de Comunicação da Universidade de Brasília (FAC/UnB) no intuito de compreender a percepção de uma amostra composta por indivíduos da geração Z, que nasceram na era digital e possuem, por natureza, facilidade notável no contato com a tecnologia e rapidez na compreensão, além de cursarem uma graduação que possui interface significativa com processos de *Big Data*. Ao mesmo tempo, a aplicação de ferramentas de *Big Data* no *marketing* se dá de forma cada vez mais acentuada e inovadora, além de contribuírem para a compreensão do consumidor em seus diversos públicos e segmentos, do cidadão em sociedade ao usuário digital. Não obstante, também é possível citar a necessidade de profissionais de Comunicação em lidarem e trabalharem com a informação.

A conclusão deste estudo compreende alguns dos principais entendimentos adquiridos por meio da realização da monografia, assim como apresenta sugestões para novos temas e linhas de pesquisa que podem ser conduzidas futuramente.

1.3 OBJETIVOS GERAIS E ESPECÍFICOS

Objetivo Geral:

Compreender como organizações privadas utilizam processos de *Big Data* e seus impactos no comportamento do consumidor, bem como mapear os principais benefícios e riscos da aplicação desta ferramenta.

Objetivos Específicos:

- a) Descrever o impacto de processos de *Big Data* nos negócios;
- b) Identificar como as organizações privadas estão utilizando processos de *Big Data* na relação com seus clientes;
- c) Mapear setores do mercado que utilizam processos de *Big Data* de forma mais acentuada na relação com o consumidor;
- d) Elencar os benefícios e riscos decorrentes do uso de processos de *Big Data*;
- e) Compreender de forma exploratória qual a percepção de consumidores sobre o fenômeno.

1.4 ESTRUTURA DA MONOGRAFIA

Esta monografia é dividida em cinco capítulos principais. O primeiro corresponde ao *Big Data* e Inteligência Artificial, no âmbito de seus conceitos, avanços e riscos relacionados. Já o segundo capítulo demonstra uma visão geral do uso de processos de *Big Data* na iniciativa privada e nos negócios.

Após contextualização acerca do tema, no terceiro capítulo explana-se o comportamento do consumidor dentro da esfera de utilização destas tecnologias, percorrendo acerca de alguns dos principais impactos de ferramentas de *Big Data* e suas influências sobre o consumidor.

O capítulo quatro, por sua vez, contempla um estudo de caso caracterizado por descrições de experiências de empresas privadas com a utilização de processos de *Big Data* e análise de percepção para com o assunto de alunos da Faculdade de Comunicação (FAC) da Universidade de Brasília (UnB).

Finaliza-se o trabalho desenvolvido por meio do capítulo cinco, que aborda as principais conclusões obtidas por meio dos estudos e análises realizadas neste trabalho de conclusão de curso.

2. **BIG DATA E INTELIGÊNCIA ARTIFICIAL: BASES CONCEITUAIS, AVANÇOS E RISCOS**

Este capítulo foi estruturado de forma a abordar, primeiramente, os principais conceitos relacionados a processos de *Big Data*, bem como a analisar a essência dessa nova tecnologia aplicada em seu contexto de surgimento.

Expõe-se também a relação das ferramentas de *Big Data* com Inteligência Artificial (IA) e *Machine Learning*, ao passo que as primeiras constituem matéria substancial para o desenvolvimento das últimas duas. São tratadas as principais definições e dinâmica de interação entre essas tecnologias.

Visando pontuar os aspectos positivos da tecnologia para os negócios, discute-se os principais ganhos advindos da apropriação de processos de *Big Data* nas organizações, incluindo novas formas de gerar valor, personalizar ofertas, compreender melhor as necessidades dos consumidores, prever demandas, identificar riscos e aprimorar a experiência do usuário.

Todavia, são abordados os riscos e aspectos negativos relacionados aos processos de *Big Data*, principalmente no que tange à privacidade de dados e aos

aspectos éticos da extração por meio de empresas de tecnologia, abordando, principalmente, conceitos de datificação. Também são observadas as esferas de monitoramento e vigilância que são fomentadas por meio dessa ferramenta.

Por fim, no intuito de perceber como os estudantes da Faculdade de Comunicação (FAC/UnB) compreendem, lidam e enxergam os processos de *Big Data*, foi realizada uma pesquisa, com viés de sondagem de opinião e conduzida com amostra supracitada, para posterior análise e estabelecimento de relação, com as temáticas abordadas durante o desenvolvimento do trabalho.

2.1 DEFINIÇÃO DE PROCESSOS DE BIG DATA

O conceito de processos de *Big Data* é amplo e não há uma definição precisa, havendo referências de vários autores. Entretanto, há um conjunto de características que o definem e que são comuns a todas as descrições. De acordo com Mayer-Schonberger e Cukier (2013), inicialmente esses processos foram referenciados pelo fenômeno no qual o volume de informações havia crescido tanto que a quantidade a ser examinada não cabia mais na memória que os computadores utilizavam para fazer análise.

Mayer-Schonberger e Cukier (2012) também complementam afirmando que *Big Data* se refere a trabalhos em grande escala que não podem ser feitos em escala menor, com o intuito de extrair novas ideias e criar novas formas de valor de maneiras que alterem os mercados, as organizações, a relação entre cidadãos e governos etc.

Para Davenport (2014), processos de *Big Data* referem-se a dados grandes demais para caber em um único servidor, não estruturados demais para caberem em um banco de dados de linha e coluna ou que fluem de forma muito contínua para um armazém de dados estático.

Mayer-Schonberger e Cukier (2013) inicialmente utilizaram três “V” para definir os processos de *Big Data* (volume, variedade e velocidade). Na sequência, houve a adição de mais dois Vs por parte de teóricos da área (veracidade e valor).

Assim, é possível afirmar que a primeira característica comum aos processos de *Big Data* é seu volume descomunal de dados.

Atualmente, vivemos com uma quantidade sobrenatural de dados: postagens no *Twitter*, transações financeiras, pesquisas em mecanismos de busca, interações em redes sociais, compras de passagem, tráfego em *websites*, dentre outros

exemplos. O *World Economic Forum* (DESJARDINS, 2019) demonstra que todo o universo digital deve atingir 44 *zettabytes* até 2020, além de pontuar que, atualmente, a cada dia 500 milhões de *tweets* e 294 milhões de e-mails são enviados; 4 *petabytes* de dados são criados no *Facebook*; 65 bilhões de mensagens são enviadas no *WhatsApp*; 5 bilhões de pesquisas são feitas; e 4 *terabytes* de informação são criados por carro conectado.

Mayer-Schonberger e Cukier (2013) apontam que quando a quantidade de dados é enorme, a exatidão, muitas vezes, não é o objetivo, mas sim a percepção de tendências gerais. Torna-se necessário, então, aprender a tratar os dados como imperfeitos e imprecisos. O aforismo “quantidade tem uma qualidade própria” nunca foi tão verdadeiro quanto na era digital. (MARR, 2018, on-line)

Outro aspecto marcante em relação aos processos de *Big Data* diz respeito à sua variedade, já que se associa a diversas fontes de dados existentes, em formatos distintos, estruturados ou não. Observa-se que os dados se encontram em formato de vídeo, texto, som, imagem, dentre outros.

De acordo com MIT Sloan Management Review (2016), durante os últimos anos, as mil empresas da *Fortune*, em seu período de exploração de processos de *Big Data*, constataram que o que gera bons resultados a partir dessas ferramentas não diz respeito à capacidade de processar e gerenciar grandes volumes de dados, mas sim de integrar cada vez mais fontes diferentes: dados novos e antigos, estruturados ou não; dados de mídias sociais; dados comportamentais; e dados herdados.

De acordo com Sterne (2017), a grande variedade de dados está se expandindo em uma taxa fenomenal. A seguir, estão os principais tipos de dados coletados para categorização:

Quadro 1 - Tipos de informações coletáveis

Tipos de Informações Coletáveis		
Tipo	Direcionador	Exemplos
Identidade	Quem são eles?	Nome, gênero, raça, endereço, telefone, digital, peso, número de identidade.
História	O que está no passado deles? O que eles já fizeram ou alcançaram?	Educação, carreira, histórico criminal, exposição na imprensa, publicações,

		prêmios, questões jurídicas, divórcio, pontuação de crédito, empréstimos.
Tendências	O que os atrai? São liberais ou conservadores? Do que eles gostam?	Preferências, hobbies, entretenimento, grupos sociais, afinidade com marcas, vocações, partido político, histórico de <i>browser</i> .
Posses	O que eles têm? Seja comprado, adquirido, fabricado ou encontrado?	Renda, casa, carro, joias, coleções, roupas, investimentos, dispositivos, assinaturas, relacionamentos, associações.
Atividades	O que eles fazem e como fazem?	Teclas, gestos, rastreamento ocular, localização, endereço IP, postagens sociais, jantares fora de casa, frequência cardíaca ao longo do tempo.
Crenças	Como eles se sentem e como se posicionam sobre determinadas questões?	Religião, valores, doações, partido político, introvertido/extrovertido, generoso/misericordioso, adaptável/inflexível, agressivo/passivo, opinião, humor.

Fonte: Sterne, 2017

A velocidade também corresponde a outro “V” caracterizador de processos de *Big Data*. Esse aspecto relaciona-se diretamente à rapidez no processamento dos dados, especialmente no quesito de análise em tempo real. Fomenta uma grande vantagem competitiva para as empresas, ao passo que o modelo de consumo atualmente é caracterizado por um perfil de consumidor exigente, com baixa disponibilidade de tempo e inúmeras opções. Assim, qualquer minuto a mais pode ser suficiente para que ele opte pelo concorrente. O ideal de processos de *Big Data* é que os dados sejam analisados em tempo real, sem necessidade de armazenamento.

A veracidade também é relevante, considerando, basicamente, se uma informação é verdadeira ou não. Caso não haja veracidade, não há valor agregado. Segundo um estudo conduzido no Canadá pela Revista Exame (AFP, 2019), 86% das pessoas entrevistadas admitiram já ter acreditado em pelo menos uma notícia falsa com a qual se depararam em Redes Sociais. Não obstante, a pesquisa realizada

também pontua que 77% dos usuários do *Facebook* consultados viram circular informação total ou parcialmente falsa. Já no caso dos usuários do *Twitter*, essa taxa corresponde à 62%.

Por fim, é importante ressaltar o valor dos dados adquiridos: “Qual é o valor efetivo das ferramentas de *Big Data* para o negócio?” (AFP, 2019, on-line). Considerando este questionamento, é importante avaliar o retorno obtido com os investimentos em arquitetura técnica para comportar a grande volumetria de dados, o esforço de análise e a coleta.

Atualmente, considerando que houve um barateamento da parte tecnológica, é possível montar um ambiente de forma rápida, por exemplo com os *Softwares as a Service*, conforme apontado por Rebecca Barros, PhD em *Economic Analytics* (informação verbal)². Assim, os retornos para as organizações que adotam a tecnologia tornam-se mais evidentes.

A mudanças ocasionadas por processos de *Big Data* não são restritas à esfera organizacional, mas também fomentam um novo modelo de relacionar-se com o mundo e de uma linha lógica de compreensão.

Segundo Mayer-Schonberger e Cukier (2013):

A era do *Big Data* desafia a maneira como vivemos e interagimos com o mundo. Mais importante, a sociedade precisará conter um pouco da obsessão pela causalidade e trocá-la por correlações simples: sem saber o porquê, apenas o quê. Essa mudança subverte séculos de práticas consagradas e desafia nossa compreensão mais básica de como tomamos decisões e compreendemos a realidade. (2013, p. 4)

A tendência natural dos seres humanos que compõem a sociedade em que vivemos é buscar entender causas e consequências para fenômenos. No entanto, as ferramentas de *Big Data* propõem uma inédita conjectura, na qual a linha lógica causal internalizada pelos indivíduos precisará ser substituída pela aceitação das correlações, sem justificativa óbvia ou aparente para as mesmas.

Na esfera dos negócios, é preciso adotar cautela ao tratar das correlações estatísticas que podem ser percebidas por meio de ferramentas de *Big Data*. De acordo com Sterne (2017, p.35), “as correlações são uma pista maravilhosa, mas devem ser tratadas como pistas e não como resultados. Correlações são o estímulo para procurar uma causa, não o fim da história. ”

² BARROS, op. cit.

Sterne (2017) também aborda a importância da comunicação correta acerca dos *insights* obtidos por meio de dados, pontuando que não deve ser transferida muita certeza em relação aos resultados, uma vez que “você não recebeu a tarefa de somar uma linha de números e entregar a resposta. Em vez disso, você foi solicitado a examinar os dados e verificar se há algo que possa ser direcional.” (STERNE, 2017, p. 35)

Considerando o dinâmico e cada vez mais acelerado processo de avanço tecnológico no mundo, as ferramentas de *Big Data* se posicionaram como um novo alicerce ao negócio de inúmeras organizações, além de afetarem a forma como nos relacionamos enquanto indivíduos e em relação às organizações.

Segundo Davenport (2014), o *Big Data* não muda apenas os processos de tecnologia e gerenciamento, mas também as orientações básicas e culturais nas organizações. Simplesmente não se pode mais pensar nos negócios da mesma maneira com este novo recurso.

Conforme será tratado adiante, os benefícios e as vantagens de processos de *Big Data* são inúmeros para as organizações, ao passo que sua utilização possibilita a identificação de padrões de comportamento do público por meio da análise de quantidades descomuns de dados, que podem fornecer *insights* determinantes.

2.2 PROCESSOS DE *BIG DATA* E INTELIGÊNCIA ARTIFICIAL

A coleta de dados é uma prática extremamente antiga em nossa sociedade. Os egípcios, por exemplo, consolidaram cuidadosamente as informações que tinham sobre armazenamento de grãos. O que mudou foi a forma de lidar com essa coleta e o caráter dos dados coletados, quebra de conjectura ocasionada pelas tecnologias que surgiram nos séculos XIX e XX. Tornou-se possível, então, acessar dados em tempo real, oriundos dos mais diversos dispositivos e máquinas, a partir de um tráfego que aumenta de forma acelerada e que, portanto, permite uma quantidade exorbitante de dados (MILLS, 2018).

Na era contemporânea, os dados se tornaram um dos tipos mais importantes de “matéria-prima”, constituindo um novo paradigma, que está estimulando a desestruturação do modelo antigo das empresas fazerem negócio, criarem serviços e produtos e, por consequência, gerarem valor.

“*Data is the new oil*” é uma analogia bastante utilizada atualmente, e se refere à pulsão no mercado gerada por empresas que produzem e utilizam dados no século XX, exatamente da forma como era gerada pelas empresas que produziam e utilizavam petróleo no século XX.

Há muitos anos, a Exxon e a Apple trocaram sequencialmente a posição de maior valorização. E há apenas dez anos, o mercado de ações continha apenas três empresas de tecnologia entre as dez principais empresas mais valorizadas; cinco eram gigantes do petróleo. Hoje, a tecnologia é responsável por sete das dez principais, com apenas uma importante petroleira ainda nessa lista. (Exxon). (MILLS, 2018, on-line).

Dentre outros fatores, a disponibilidade dos processos de *Big Data* impulsionou e viabilizou o desenvolvimento da Inteligência Artificial, que utiliza essas ferramentas como ativo para suas abordagens. De acordo com Sterne (2017, p.14):

[...] a atual onda de progresso e entusiasmo pela IA começou por volta de 2010, impulsionado por três fatores que se baseiam em cada outro: a disponibilidade de *Big Data* de fontes, incluindo comércio eletrônico, negócios, mídia social, ciência e governo; que forneceu matéria-prima para abordagens e algoritmos aprimorados de Machine Learning.

Assim como o conceito de *Big Data*, o conceito de Inteligência Artificial possui diversas frentes. Porém, de forma generalista e comum a todos aqueles que a definem, a Inteligência Artificial (IA) possibilita que as máquinas aprendam com a experiência, se ajustem a novas entradas e realizem tarefas humanas. A maioria dos exemplos de IA que você ouve hoje - de computadores jogando xadrez a carros autônomos - depende muito de aprendizado profundo e processamento de linguagem natural. Usando essas tecnologias, os computadores podem ser treinados para realizar tarefas específicas, processando grandes quantidades de dados e reconhecendo padrões nos dados (ARTIFICIAL..., 2019). De acordo com Sterne (2017), *Big Data* é o principal ativo para abordagens de marketing baseadas em Inteligência Artificial.

Sterne (2017) acrescenta que a Inteligência Artificial também consiste em um próximo passo lógico da computação, sendo um programa que pode descobrir as coisas por si mesmo. É um programa que pode reprogramar-se. O autor define esta tecnologia por meio de três “D”. O primeiro é “detectar”, e faz referência à capacidade intrínseca da IA descobrir quais elementos são mais preditivos, dentro de uma variedade enorme de dados, podendo então prestar mais atenção em um tipo específico ou ignorá-lo. Na sequência, o segundo é “deliberar”, e diz respeito às regras

que podem ser utilizadas por IA em relação aos dados, no intuito de decidir e considerar a relevância de cada um, de forma a atribuir mais peso aos aspectos preditivos. Por fim, o terceiro é “desenvolver”, e compreende o crescimento e amadurecimento da Inteligência Artificial com cada iteração.

O *Machine Learning*, por sua vez, configura uma vertente específica de Inteligência Artificial, que se resume ao aprendizado das máquinas por meio dos dados. Trata-se de um método de análise de dados que automatiza a construção de modelos analíticos (MACHINE..., 2019b).

Sterne (2017, p. 4) pontua a diferença entre ambas as tecnologias ao afirmar que “a Inteligência Artificial é uma máquina que finge ser humana. Já o *Machine Learning* é uma máquina que finge ser um programador estatístico”. Logo, a grande diferença está na capacidade de aprendizado. O autor também apresenta o *Machine Learning* como a “maneira de um computador usar um determinado conjunto de dados para descobrir como executar uma função específica por tentativa e erro.” (STERNE, 2017, p. 4).

O *Machine Learning* opera por meio de algoritmos. Alguns casos famosos são a *Netflix* - principal rede televisiva na internet do mundo -, e a *Amazon* - empresa multinacional de comércio eletrônico.

A *Netflix* investe no aprendizado de máquina para aprimorar a experiência do usuário e otimizar o serviço fornecido pela plataforma. A personalização é, segundo a própria empresa, a área mais conhecida. O *Machine Learning* potencializa o algoritmo de recomendação de conteúdo (filmes, séries, documentários etc) para que ele seja aderente às preferências que o usuário já tenha demonstrado por materiais consumidos anteriormente. A organização também utiliza o *Machine Learning* para aprimorar seu catálogo, por meio da identificação dos aspectos fundamentais que tornam um conteúdo bem-sucedido. Assim, os estúdios originais *Netflix* potencializam suas produções por meio do conhecimento acerca dos tipos de conteúdo que estão em ascensão no gosto dos usuários (MACHINE..., 2019).

Na *Amazon*, a tecnologia de *Machine Learning* é utilizada pela equipe de recomendação de produtos, visando realizar também previsões acerca dos bens. A Inteligência Artificial e a *Machine Learning* embasam três produtos famosos da *Amazon*: *Alexa* (*chatbot* de assistente virtual), *Amazon Go Store* (loja sem caixa) e mecanismo de recomendação (oferta de produtos com base em preferências individuais). (MORGAN, 2018).

Quando se fala de Inteligência Artificial e de *Machine Learning*, é essencial pontuar a importância dos algoritmos. De acordo com Gillespie (2013), os algoritmos desempenham um papel cada vez mais importante na seleção de qual informação é mais relevante para o usuário, uma característica crucial da participação na vida pública.

Ainda segundo Gillespie (2013), os algoritmos são agora uma chave lógica que governa os fluxos de informação dos quais dependemos. O autor pontua que os algoritmos de recomendação mapeiam nossas preferências, gerenciam nossas interações em sites sociais e, além disso, são projetados para refinar o que é “quente” e tendência, destacando para os usuários apenas o essencial de uma quantidade infinita de informações que são ofertadas.

A *Foxtrot Systems*, startup de tecnologia em logística, busca ajudar seus clientes a realizarem entregas com o menor tempo, com a menor distância e respeitando parâmetros de pontualidade. Segundo Fillipe Santos, *Head of Operations* e *Data Analysis* da Foxtrot (informação verbal)³, a empresa utiliza algoritmos de *Machine Learning* para identificar, por exemplo, qual é a maior probabilidade de tempo de entrega dos motoristas em cada cliente/ponto de venda. É possível identificar, por exemplo, que a entrega em um cliente “X” às 15 horas, leva 30 minutos; e às 18 horas, nesse mesmo cliente “X”, leva 2 horas. Logo, é melhor realizar a entrega às 15 horas.

Dentro de *Machine Learning*, há três tipos de algoritmos de aprendizado de máquina: supervisionado, não supervisionado e por reforço. No primeiro caso, os dados utilizados para treinar a máquina contêm a resposta desejada. Diz-se, então, que os dados são “anotados” com as respostas ou classes a serem previstas. Já quando se trata de algoritmos não supervisionados, não se indica a resposta desejada por ser muito custoso ou até impossível. O computador terá que descobrir “sozinho”. Finalmente, a abordagem por esforço implica que a máquina tente aprender qual é a melhor ação a ser tomada, dependendo das circunstâncias nas quais a atividade será executada (HONDA; FACURE; YAOHAO, 2017).

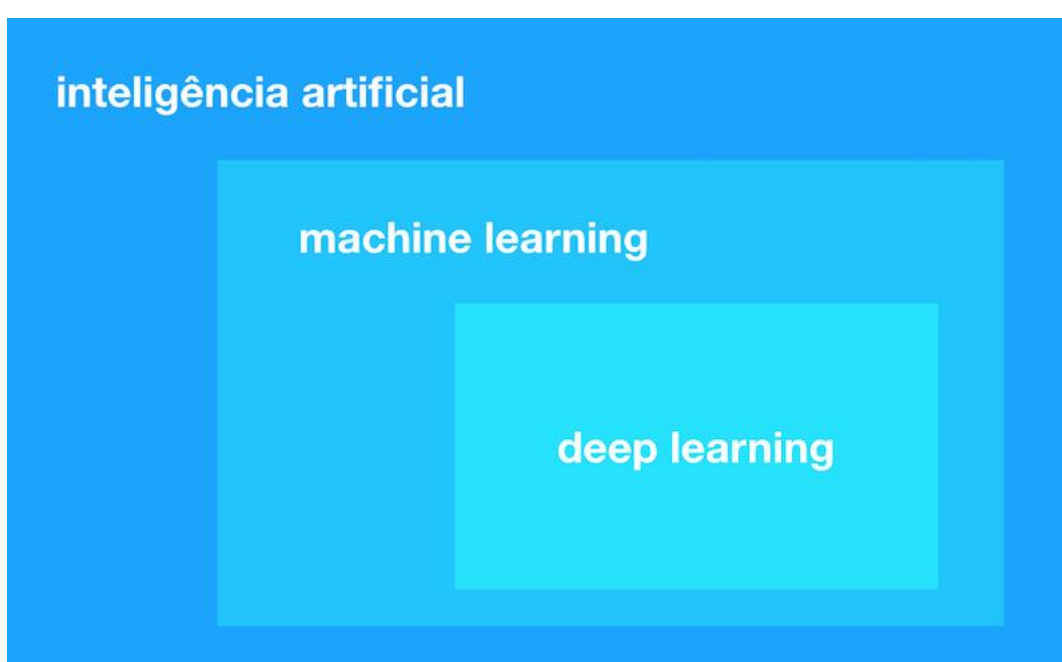
No caso de *Deep Learning*, segundo Sterne (2017), refere-se a camadas cada vez mais profundas de neurônios trabalhando no seu problema, que dão origem a redes neurais dinâmicas nas quais a informação pode fluir de forma muito menos

³ Entrevista concedida por SANTOS, Fillipe. [12 de set. 2019]. Entrevistador: Laura Pereira Hijdra van Hagen. Brasília, 2019. 1 arquivo docx. A entrevista na íntegra encontra-se transcrita no Apêndice B desta monografia.

controlada, permitindo que a máquina construa contextos e chegue a conclusões de forma ainda mais rápida. No *Deep Learning*, os dados são submetidos a várias camadas de processamento não lineares que simulam a forma de pensar dos neurônios (MACHINE..., 2018). Logo, o funcionamento é semelhante ao de *Machine Learning* (por meio de redes neurais), contudo, é muito mais profundo.

A imagem abaixo retrata a relação das três tecnologias e demonstra como o *Deep Learning* é um tipo mais especializado e de alto nível do *Machine Learning*, sendo que este está contido dentro dos mecanismos de Inteligência Artificial:

Figura 1 - Relação entre as tecnologias



Fonte: Salesforce, 2018.

Os processos de *Big Data*, então, demonstram-se fundamentais para a execução dessas tecnologias, ao passo que a observação de ambientes (essência de funcionamento de *Machine Learning* e *Deep Learning*) é realizada por meio do processamento em tempo real de dados, como os de sensores no caso dos carros autônomos, ou por meio do processamento de dados coletados ao longo do tempo e armazenados em bancos de dados.

2.3 ASPECTOS NEGATIVOS DA DATIFICAÇÃO E DATAÍSMO

No que tange aos principais riscos envolvendo os processos de *Big Data*, torna-se importante salientar os aspectos problemáticos que concernem à datificação e ao dataísmo. A datificação é originada e possui como pré-requisito a extração de dados

on-line. O conceito do fenômeno, de acordo com Mayer-Schoenberger e Cukier (2013), corresponde à transformação da ação social em dados quantificados on-line, permitindo rastreamento em tempo real e a análise preditiva. Segundo Mayer-Schoenberger e Cukier (2013, p. 54), “dataficação um fenômeno é colocá-lo num formato quantificado de modo que possa ser tabulado e analisado”.

Atualmente, as informações pessoais dos usuários são extraídas de várias formas e oriundas de diversas fontes no ambiente digital. Todas as interações que são realizadas fomentam uma amostra de comportamento humano que pode ser estudada. E é em cima disso que se baseia a datificação.

A partir do advento do dado como recurso revelador do comportamento humano, a datificação resume-se, basicamente, em quantificar os dados no intuito de analisá-los. A ação de datificação promove o que pode se chamar de dataísmo, que de acordo com Dijck (2014, p. 198), “mostra características de crença generalizada na quantificação objetiva e o potencial monitoramento de todos os tipos de comportamento humano e de sociabilidade, por meio de tecnologias de mídia on-line”.

De forma sucinta, a datificação é a ação de quantificar fenômenos colocando-os na forma de dados estruturados e analisá-los (hoje, comumente obtidos por meio de ambientes digitais), e o dataísmo se refere à crença no potencial de estudos a partir dessa amostra.

Segundo Dijck (2014), o dataísmo é caracterizado, em primeiro lugar, pela crença na objetividade da quantificação e no potencial de rastrear todos os tipos de comportamento humano e da socialidade por meio de dados on-line. Em segundo lugar, os metadados seriam colocados como matéria-prima que pode ser analisada e processada em algoritmos preditivos sobre o comportamento humano futuro.

Dijck (2014) também aponta que, na última década, o *datafication* cresceu e se tornou um novo paradigma aceito para a compreensão da socialidade e do comportamento social. Com o advento da *Web 2.0* e seus sites de rede social em proliferação, muitos aspectos da vida social que nunca haviam sido quantificados antes – amizades, interesses, conversas, pesquisas de informações, expressões de gostos, respostas emocionais etc – foram codificados.

A datificação e o dataísmo apresentam riscos no que se refere ao monitoramento e controle social, ao passo que fomentam um panorama aparentemente completo do comportamento individual e coletivo, o qual pode ser utilizado de forma a obter vantagens por meio do uso desses comportamentos como

instrumento de manipulação. Dijck (2014, p. 199) afirma que “os entusiastas da informação também assumem uma relação evidente entre dados e pessoas, posteriormente interpretando dados agregados para prever comportamentos individuais”.

Zuboff (2015) aborda a captura de pequenos dados dos indivíduos, de ações mediadas por computador e de enunciados em busca de vida efetiva, como a constituição do *Big Data*. A autora expõe como nada é “trivial demais” para a colheita: curtidas de *Facebook*; buscas no *Google*; e-mails; textos; fotos; músicas; vídeos; localização; padrões de comunicação; redes; compras; movimentos; cliques; palavras com erro de ortografia; visualizações de página; entre outros.

Dijck (2014) aponta que dados e metadados selecionados do *Google*, *Facebook* e *Twitter* são geralmente considerados impressões ou sintomas do comportamento ou humor real das pessoas, enquanto as próprias plataformas são apresentadas apenas como facilitadores neutros. Entretanto, segundo Zuboff (2015, p. 79), “esses dados são adquiridos, dataficados, abstraídos, agregados, analisados, empacotados, vendidos, mais analisados e vendidos novamente”.

Dijck (2014) também aborda a crença de que os metadados extraídos das plataformas digitais refletem o comportamento humano como ele é. No entanto, os algoritmos utilizados por mídias sociais possuem um viés naturalmente manipulatório e seletivo. Um exemplo disso seria o conteúdo personalizado que aparece no *feed* de cada usuário (extremamente adaptado àquele indivíduo em específico), bem como recomendações em sites de vendas digitais e *E-commerce*. Segundo Dijck (2014, p. 200), “uma mentalidade de *Big Data* parece também favorecer a paradoxal premissa de que as plataformas de mídia social, ao mesmo tempo, medem, manipulam e monetizam o comportamento humano on-line”.

Assim, é possível afirmar que a natureza dos processos de *Big Data* permite um alto nível de segmentação do público, poder tal que viabiliza a identificação de “fraquezas” e “desejos” urgentes dos usuários, utilizando-se dos mesmos para gerar capital para as empresas por meio do atingimento da pessoa certa, com a mensagem correta, no melhor momento possível.

Zuboff (2015), em sua obra, trata da mudança no relacionamento das organizações com a sociedade. Segundo a autora, o método de produção dos processos de *Big Data* por meio dessa pequena quantidade de dados extraídos demonstra a indiferença formal na relação das empresas com sua população de

“usuários”. Zuboff (2015) afirma que essas populações são a fonte da extração de dados e o local onde os objetivos finais que os dados produzem estão.

Nesse panorama, as empresas utilizam dados dos usuários de forma unilateral, obtendo benefícios que não são garantidos igualmente para a população. Uma vez que o dado é considerado matéria-prima, o valor é agregado apenas para uma das partes (grandes entes, como organizações privadas e públicas).

Dijck (2014, p.198) também explana a extração de dados pessoais de usuários pelas empresas, apontando que “hoje há uma notável tolerância ao Big Brother e ao Big Business acessando rotineiramente as informações pessoais do cidadão, também conhecidas como *Big Data*. ” Tal “aceitação” pode ser percebida na concessão proposital de dados por parte dos indivíduos no intuito de receber ofertas personalizadas de produtos e serviços, ou visando obter algum tipo de desconto, como os que são disponibilizados frequentemente em farmácias, por exemplo. Entretanto, também é possível atribuir esse comportamento à falta de conhecimento acerca do formato no qual essa extração e utilização de dados pessoais por parte de empresas se dá.

A questão da datificação no meio acadêmico também é abordada por Dijck (2014), que reforça que o entusiasmo exacerbado dos pesquisadores pela datificação como paradigma neutro precisa ser substituído por uma análise mais rigorosa e crítica, visando impedir perdas na integridade dos estudos acadêmicos a longo prazo. Assim:

[...] os pesquisadores de *Big Data* precisam identificar as perspectivas parciais dos dados que analisam; mais do que manter as alegações de neutralidade, eles precisam avaliar o contexto no qual os conjuntos de dados são gerados e associar as metodologias quantitativas com indagações qualitativas. (DIJCK, 2014, p. 206).

Além de monitoramento e controle, é válido discorrer acerca do tema de vigilância de dados. Segundo o dicionário de *Oxford*, o conceito se refere à prática de monitorar dados digitais relacionados a detalhes pessoais ou a atividades on-line (DATAVEILLANCE, 2019).

Mayer-Schoenberger e Cukier (2013, p. 105) afirmam que “com o *Big Data* prometendo valiosas ideias para aqueles que os analisam, todos os sinais parecem apontar para uma nova onda de coleta, armazenamento e reutilização de nossos dados pessoais”.

Dijck (2014) indica que a vigilância de dados – monitoramento dos cidadãos a partir de seus dados on-line – difere da vigilância na medida em que realiza o

monitoramento contínuo de metadados sem objetivos especificados, enquanto a vigilância monitora para fins específicos. Segundo a autora (2014, p. 205):

[...] a descoberta das rotineiras táticas de data vigilância ameaçou causar sério dano não apenas na confiança das pessoas nas agências estatais ou nas corporações individuais, mas no suporte ao institucional do dataísmo como um todo.

É assim que, de acordo com Dijck (2014), inicia-se a luta por credibilidade e confiança por parte de grandes corporações, e também de agências governamentais, que precisam recuperar o status de plataformas mediadoras neutras e demonstrar uma devolutiva de transparência para os cidadãos, especialmente após casos como as acusações de práticas ilegais de manutenção de *logs* de dados de usuários que foram direcionadas a empresas como o *Facebook*.

Os aspectos negativos dos processos relacionados à privacidade de *Big Data* podem ser refletidos no caso do *Facebook*, que viralizou no início de 2018 com o período eleitoral norte-americano e ocasionou a descoberta do compartilhamento de dados sensíveis de usuários com outras empresas, especialmente de tecnologia.

O *New York Times* (*NYT*), jornal diário estadunidense mais reconhecido no globo e fundado em Nova York no ano de 1851, conduziu uma investigação sobre uma suposta de venda de dados pessoais de usuários do *Facebook* sem consentimento para gigantes de tecnologia, tais como *Microsoft*, *Spotify*, *Netflix* e *Amazon*. O compartilhamento isentou as organizações parceiras de regras habituais padrão de privacidade.

Segundo o *New York Times*:

[...] o *Facebook* permitiu que o mecanismo de pesquisa Bing da *Microsoft* visse os nomes de praticamente todos os amigos dos usuários do *Facebook* sem consentimento, segundo os registros, e deu ao *Netflix* e ao *Spotify* a capacidade de ler as mensagens privadas dos usuários do *Facebook*. (LAFORGIA; ROSENBERG; DANCE, 2019, on-line).

A prática só foi possível devido à quantidade extraordinária de dados do *Facebook*, o qual possui uma base com aproximadamente 2,2 bilhões de usuários.

O *Facebook*, maior rede social digital do mundo e componente marcante do Vale do Silício, sofre com escândalos de privacidade e compartilhamento indevido de dados desde março de 2018, quando ocorreu uma denúncia de utilização imprópria de seus dados pela *Cambridge Analytica*, consultoria em comunicação estratégica

política e responsável pela condução da campanha eleitoral do presidente norte-americano Donald Trump, em 2016.

O *New York Times* baseou sua pesquisa em milhares de documentos e entrevistas com 50 funcionários do *Facebook*, entretanto, esses materiais não foram compartilhados. No total, o jornal estadunidense constatou que

[...] a rede social tinha acordos especiais com mais de 150 empresas para compartilhar os dados pessoais de seus membros. A maioria deles, segundo a empresa, era de outras empresas de tecnologia, mas a lista também incluía varejistas on-line, fabricantes de automóveis e organizações de mídia, incluindo o próprio NYT, entre outros. (FACEBOOK'S..., 2018, on-line).

De acordo com o *The New York Times*, houve benefícios para algumas das empresas a partir do acesso às informações sensíveis do *Facebook*:

Quadro 2 – Ganhos de empresas a partir do acesso aos dados compartilhados do Facebook

Empresas	Ganho com acesso aos dados do Facebook
- Bing da Microsoft - Pandora (serviço de streaming de música) - Rotten Tomatoes (plataforma crítica de filmes)	Acesso às informações de amigos do Facebook sem consentimento, a fim de personalizar os resultados apresentados.
- Apple	Acesso aos números de contato dos usuários e entradas na agenda dos usuários (mesmo quando o usuário desativa o compartilhamento).
-Netflix -Spotify - Royal Bank of Canada	Ler, escrever e deletar mensagens privadas dos usuários; Ver todos os participantes em um tópico de bate-papo.
-Yandex (provedor de pesquisa russo)	Indexar as identidades dos usuários de páginas e postagens públicas para melhorar seus resultados de pesquisa.
-Yahoo	Visualizar <i>feeds</i> ao vivo de postagens dos amigos.
-Sony -Microsoft -Amazon	Acesso aos endereços de e-mail de seus membros por meio dos amigos.
-Black Berry -Huawei	Extração de dados do Facebook para alimentar seus próprios aplicativos de mídia social.

Fonte: Elaborada pela autora com base no *The New York Times*, 2018.

O ocorrido, em específico, reforça a necessidade dos cuidados para com a privacidade no mundo *Big Data*. É possível concluir que, no caso do *Facebook*, os dados on-line da plataforma foram utilizados como matéria-prima passível de ser quantificada e analisada para compreender o comportamento dos usuários, proporcionando, assim, *insights* milionários para empresas.

3. PROCESSOS DE *BIG DATA* NA INICIATIVA PRIVADA

Neste capítulo, é delineado um panorama geral acerca da inserção de processos de *Big Data* nas organizações privadas e como essa utilização afeta os negócios; quais são os desafios enfrentados durante a jornada; a repercussão que é gerada no comportamento do consumidor; bem como os aspectos positivos e negativos do uso dessas ferramentas, abrangendo aspectos essenciais relativos à privacidade de dados e à ética, especialmente no ambiente digital.

Na sequência, são descritas as principais aplicações de processos de *Big Data* na iniciativa privada, envolvendo os setores que utilizam a ferramenta de forma mais acentuada, suas formas e os objetivos de aplicação, questões relacionadas à cultura organizacional, bem como um mapeamento dos principais desafios para sua implementação nas organizações, ora por questões técnicas, ora por questões relacionadas à cultura organizacional.

Assim, após o panorama geral de utilização de processos de *Big Data* no setor privado, é possível discorrer acerca dos impactos desse uso no comportamento do consumidor, indicando as formas possíveis de influência por meio da natureza do dado, setor do mercado e fenômenos de datificação e *mass customization*, os quais possuem em comum o contato com dados que revelam o comportamento do consumidor.

3.1 PROCESSOS DE *BIG DATA* NOS NEGÓCIOS

Os processos de *Big Data* estão sendo utilizados de forma crescente pelas empresas, especialmente naquelas mais avançadas em termos de transformação digital. Hoje em dia, tornou-se quase impensável não aderir à essa ferramenta, que se encontra presente em várias indústrias.

Para tratar de processos de *Big Data* nos negócios e na esfera privada, é de extrema relevância abordar o conceito de Indústria 4.0. Segundo a *Delloite* (INDUSTRY 4.0, 2019), empresa multinacional especializada em serviços de Auditoria, Consultoria, Assessoria Financeira, *Risk Advisory*, Consultoria Tributária e serviços relacionados, o termo diz respeito à promessa de uma nova revolução industrial – a qual combina técnicas avançadas de fabricação com a Internet das Coisas para criar sistemas que não são apenas interconectados, mas que também

comunicam, analisam e usam as informações para impulsionar ações inteligentes de volta ao mundo físico.

A *Boston Consulting Group (BCG)*, consultoria global em gestão empresarial, por sua vez, define Indústria 4.0 como uma transformação que possibilita a coleta e análise de dados entre máquinas, permitindo processos mais rápidos, flexíveis e eficientes para produzir mercadorias de maior qualidade a custos reduzidos (EMBRACING INDUSTRY 4.0..., 2019). Segundo a empresa de consultoria, essa revolução da fabricação aumentará a produtividade, mudará a economia, promoverá o crescimento industrial e modificará o perfil da força de trabalho - mudando, em última análise, a competitividade das empresas e regiões.

A BCG afirma que há nove tipos de tendências tecnológicas que formam um alicerce à Indústria 4.0 (EMBRACING INDUSTRY 4.0..., 2019):

Figura 2 – Nove Tecnologias da Indústria 4.0



Fonte: Boston Consulting Group, 2019.

A Indústria 4.0 também possui alguns princípios como, por exemplo, a capacidade de operação em tempo real, que consiste na extração e no tratamento de dados de forma praticamente instantânea, permitindo a tomada de decisão imediata. É dessa forma, principalmente, que os processos de *Big Data* se encontram inseridos nesse contexto e fomentam um importante pilar na dinâmica industrial 4.0.

Também pode-se retratar outras diretrizes da nova conjectura, tais como a virtualização (rastreadabilidade e monitoramento remoto de processos por meio de sensores espalhados), descentralização (módulos de fábrica inteligente trabalhando de forma descentralizada a fim de aprimorar os processos de produção), orientação a serviços (utilização de arquiteturas de software orientadas a *Internet of Services*) e modularidade (produção sob demanda, envolvendo o acoplamento ou desacoplamento de módulos de produção – flexibilidade para alternar tarefas de máquinas). (SILVEIRA, 2019).

Dessa maneira, o *modus operandi* de trabalho está sendo ressignificado pelas tecnologias. Por consequência, nota-se um reflexo no modo de atuar frente ao mercado. A KPMG (INDUSTRY 4.0 EXPLAINED, 2017), uma das maiores empresas do mundo de serviços de *Audit, Tax e Advisory*, afirma que a Indústria 4.0 é sobre conectividade, sendo uma oportunidade de mudar radicalmente a maneira como o setor industrial responde às necessidades da sociedade. Segundo a organização, as oportunidades para disrupção são enormes, e aqueles que forem deixados para trás sentirão a perda de forma aguda.

No que se refere aos processos de *Big Data Analytics*, presentes de forma acentuada no modelo 4.0, as oportunidades para os negócios são inúmeras: de processos de tomada de decisão estratégicos e segmentação ultra efetiva de público à otimização da contratação pela área de recursos humanos.

Segundo Mckinsey (MCKINSEY GLOBAL INSTITUTE, 2011, on-line), “os dados digitais agora estão em todo o lugar – em todos os setores, economias, organizações e em todo usuário da tecnologia digital”. A capacidade de armazenar, processar e cruzar esses dados em tempo real oferece grande valor para a iniciativa privada.

De acordo com Boyd e Crawford (2012, p. 663), “*Big Data* se trata muito mais da capacidade de pesquisar, agregar e cruzar grandes conjuntos de dados do que apenas grandes dados em si”. Assim, é necessário pensar não somente no grande volume dos dados em si, mas em como fazer combinações e analisá-los de forma inteligente visando extrair real valor dos mesmos.

O enorme volume de dados permite melhorias e *insights* capazes de mudar direcionamentos significativos das empresas, aprimorando desde a sua estratégia organizacional, até a sua operação. “No cenário atual, um *Big Data* devidamente configurado e utilizado, dentro de um departamento de Análise e Inteligência de

Dados, é capaz de desenvolver estratégias extremamente eficazes e implementar melhorias de maneira inteligente e otimizada para empresas”, aponta André Ferraz, CEO da in Loco (informação verbal).⁴

A apropriação de ferramentas de *Big Data*, assim como as demais tecnologias relacionadas mencionadas neste trabalho, está aumentando o valor de mercado das empresas e gerando um nível de concorrência inédito entre elas. Segundo *Mckinsey Global Institute* (2011, on-line):

[...] a quantidade de dados em nosso mundo está explodindo e a análise de grandes conjuntos de dados - os chamados *Big Data* - se tornará uma base essencial da concorrência, sustentando novas ondas de crescimento da produtividade, inovação e excedentes do consumidor.

De acordo com Schmarzo (2013, p. 25):

[...] o *Big Data* está permitindo que as empresas respondam perguntas que não podiam responder anteriormente e tomar decisões mais oportunas em um nível de fidelidade mais refinado do que antes, produzindo novos *insights* que podem oferecer aos negócios diferenciação e novas eficiências operacionais.

A Mckinsey Global Institute (2011, on-line) também afirma que

[...] na maioria dos setores, concorrentes estabelecidos e novos entrantes aproveitarão estratégias baseadas em dados para inovar, competir e capturar valor a partir de informações profundas e em tempo real.

A capacidade intrínseca dos processos de *Big Data* em realizar correlações e identificar padrões viabiliza que as empresas trabalhem com a análise preditiva. Ou seja, a partir de previsões podem atuar de forma a maximizar os resultados de seu negócio.

Nesse sentido, as organizações privadas vêm apresentando uma postura mais proativa no sentido de compreender primeiro as demandas de seus consumidores, para depois atendê-las. Isso torna a oferta de produtos e serviços demasiadamente mais assertiva e só é possibilitada por meio da análise preditiva.

Schmarzo (2013), em sua obra, reitera o potencial transformador do *Big Data* no que tange aos processos de criação de valor dentro de uma organização.

⁴ Entrevista concedida por FERRAZ, André. [11 de out 2019]. Entrevistador: Laura Pereira Hijdra van Hagen. Brasília, 2019. 1 arquivo docx. A entrevista na íntegra encontra-se transcrita no Apêndice C desta monografia.

Figura 3 - Big Data: criação de valor para organização



Fonte: Schmarzo, 2013 - “Big Data Powers the value creation process”

O autor pontua o tipo de valor agregado para cada um dos setores:

- a) Finanças: identificação de quais linhas de operação de negócios e categorias de produtos são mais eficientes para gerar lucratividade;
- b) Compras: identificação de quais fornecedores são mais econômicos na entrega de produtos de alta qualidade no prazo;
- c) Desenvolvimento de Produtos: identificação de *insights* acerca do uso de produtos para acelerar o desenvolvimento e servir de referência para lançamento de novos produtos;
- d) Manufatura: sinalização de variações de máquinas e processos que podem ser indicadores de problemas de qualidade;
- e) *Marketing*: identificação de que tipos de campanhas são mais eficazes em impulsionar *leads* e vendas;
- f) Distribuição: quantificar níveis ideais de estoque e atividades da cadeia de suprimentos;
- g) Experiência do Cliente: promover um engajamento do consumidor mais relevante e mais personalizado, que gera lealdade e lucratividade a longo prazo;
- h) Operações: otimização de preços de mercadorias;
- i) Recursos Humanos: identificação de características dos funcionários mais eficazes.

A eficiência operacional ocasionada por processos de *Big Data* proporciona análise da produção, *feedbacks* e devolutivas de clientes, assim como antecipação de demandas e subsídio positivo para tomada de decisão. Os tomadores da decisão da *Procter & Gamble*, por exemplo, sabem que os dados, trazidos para um contexto de negócios compreensível, estão corretos e podem ajudar a tomar decisões mais rápidas e melhores, respondendo efetivamente às mudanças velozes do ambiente. (RIJMENAM, 2013).

Samuel Corado, *Big Data Strategist* da *AIS Digital*, afirma que as previsões e predições obtidas por meio de processos de *Big Data* afetam estratégias de marketing, vendas e relacionamento com o consumidor ou cliente (*business to business* - B2B):

Primeiro pela velocidade em que a informação chega para a tomada de decisão e a capacidade de processamento para grandes volumes de dados. Segundo, pela capacidade em cruzar informações estruturadas e não estruturadas gerando insights de forma rápida. Terceiro por democratizar o acesso à informação, fazendo com que outros departamentos participem do processo exploratório de dados. (informação verbal) ⁵

Os processos de *Big Data* também podem ser utilizados para identificar tendências sociais e para usar a identificação de padrões de comportamento no alcance de objetivos. Segundo Raíssa Arriola, *Data Specialist* do *Twitter*:

A rede social visa conectar pessoas por meio de interesses. Isso é realizado por meio da classificação de palavras-chave nos *tweets* para identificar sobre o que as pessoas estão falando. Assim, o foco deixa de ser no usuário e direciona-se à temas, conversas, assuntos (informação verbal) ⁶

As ferramentas de *Big Data* promovem um entendimento mais aprofundado sobre o comportamento do público-alvo das organizações, sejam elas B2B (transações *business to business*) ou B2C (transações *business to consumer*). Por exemplo, a partir dos *insights* gerados pela frente de *analytics* é possível observar tendências de vários temas, no caso do *Twitter*, da moda ao campeonato de futebol Brasileiro (informação verbal). ⁷

⁵ Entrevista concedida por CORADO, Samuel. [19 de set. 2019]. Entrevistador: Laura Pereira Hijdra van Hagen. Brasília, 2019. 1 arquivo docx. A entrevista na íntegra encontra-se transcrita no Apêndice E desta monografia.

⁶ Entrevista concedida por ARRIOLA, Raíssa. [16 de set de 2019]. Entrevistador: Laura Pereira Hijdra van Hagen. Brasília, 2019. 1 arquivo docx. A entrevista na íntegra encontra-se transcrita no Apêndice D desta monografia.

⁷ Idem.

No caso de empresas de rede social, como *Twitter*, *Facebook*, *Instagram*, e *LinkedIn*, a tecnologia de *Big Data* pode atuar identificando padrões de preferência e aproximando pessoas com interesses comuns, ou aproximando pessoas e temas que as interessam. Tal forma de utilização pode ser exemplificada pelos filtros presentes nos *feeds* do *Facebook* e *Instagram*, que não expõem tudo que é publicado pela rede de amigos e seguidores, mas sim publicações de pessoas as quais o algoritmo da plataforma entende que estão mais próximas no seu ciclo social ou mais ligadas aos seus interesses.

Segundo Arriola (informação verbal) ⁸, “o *Twitter* visa revigorizar a comunidade em que está inserido e busca emplacar a responsabilidade social a partir do empoderamento de grupos de diversidade na plataforma, e depois dentro da sociedade”. Assim, no caso do *Twitter*, também é possível utilizar os processos de *Big Data* para criar uma ponte entre minorias dentro da rede social e empoderar esses grupos.

Outro uso comum da tecnologia refere-se à detecção de fraudes. Ela se baseia na identificação de padrões para prever e, por consequência, prevenir comportamentos criminosos. A segurança cibernética tem se tornado cada vez mais relevante no universo digital no qual estamos inseridos hoje. Assim, a análise comportamental examina ações nas redes virtuais em tempo real, identificando anomalias que podem indicar vulnerabilidades e ameaças (ANÁLISES...,2019).

Algoritmos de sistema podem armazenar, processar e analisar milhões de dados, gerando avaliações em tempo real, diminuindo assim a quantidade de operações fraudulentas (CANABARRO, 2017).

É bastante comum o uso de *Big Data* no *E-Commerce*, no que tange à análise de fraudes na internet. Segundo a *E-commerce Brasil* (CANABARRO, 2017), atualmente há diversos indicadores que podem auxiliar na análise de risco no ambiente digital, como tempo de navegação até o momento da compra; tipo de produto adquirido, dispositivo utilizado; endereço de entrega; última utilização do cartão etc.

A manutenção preditiva em indústrias é outra aplicação possível, podendo ser realizada a partir de previsões de falhas mecânicas com base em dados estruturados ou não estruturados. A partir dessa utilização, as empresas podem implementar

⁸ Idem.

processos de manutenção mais econômicos e maximizar o tempo de vida útil de produtos (O QUE É..., 2019).

Por fim, um aspecto importante a ser mencionado corresponde ao *data monetization*. Segundo Schmarzo (2013), a fase de monetização de dados se refere ao momento em que as organizações utilizam os processos de *Big Data* para obter novas oportunidades de receita líquida. Segundo o autor, algumas das iniciativas relacionadas são:

- a) “Empacotamento” de *insights* acerca de clientes, produtos e marketing e venda para outras organizações;
- b) Integração de análises diretamente nos produtos para criar “produtos inteligentes”;
- c) Aproveitamento de *insights* acionáveis e recomendações personalizadas com base em comportamentos e tendências do cliente, visando melhoria do relacionamento e da experiência do usuário.

Os tipos de ações descritos acima possuem um elo comum: todas envolvem o comportamento do consumidor como premissa e apresentam como objetivo final impactá-lo.

3.2 COMPORTAMENTO DO CONSUMIDOR

Os processos de *Big Data* consistem em uma importante fonte de entendimento do comportamento do consumidor cujo conceito, de acordo com (FRIESNER, 2019), corresponde ao estudo dos consumidores e dos processos que os mesmos utilizam para selecionar, consumir e descartar produtos e serviços. Dessa forma, o comportamento do consumidor incorpora ideias de várias ciências, incluindo psicologia, biologia, química e economia, além de considerar aspectos culturais e sociais.

Utilizando conceitos da datificação, é possível realizar uma análise comportamental do público por meio de seus dados digitais. De acordo com Mayer-Schoenberger e Cukier (2013), o fenômeno corresponde à transformação da ação social em dados quantificados on-line, permitindo o rastreamento em tempo real e a análise preditiva.

Tal utilização pode gerar resultados extremamente positivos na medida em que facilita a compreensão mais profunda das empresas acerca do público que se

pretende atingir. Dessa forma, o desenvolvimento de produtos e serviços, abordagem de venda, tom do discurso, promoções e atendimento de desejos e necessidades do consumidor podem adotar um direcionamento cada vez mais assertivo.

De acordo com Sterne (2017), a promessa do *Big Data* é a capacidade de cruzar as mais diversas referências acerca do comportamento do consumidor, incluindo informações relativas a visitas na *web*; respostas por e-mail; central de atendimento; interações; compras; devoluções; registros de mídia social (gostos, amigos, *tweets*, comentários, “qual princesa da Disney você é?”); anúncios; redes (comportamento entre editores/plataformas); dados públicos (casamento, registros, ações judiciais, registros de votação, dados do censo); e dados de terceiros (informações sobre emprego, pontuação de crédito, propriedade de animais de estimação, gastos habituais), a fim de elaborar exatamente a mensagem certa, no momento certo e no dispositivo certo para cada indivíduo.

Em complemento, Schmarzo (2013) pontua que a oportunidade do *Big Data* na análise comportamental está na combinação de transações detalhadas do consumidor (como vendas, devoluções, comentários e cliques na *web*) com novas mídias sociais e os dados móveis. O autor afirma também que o objetivo desse cruzamento é obter *insights* que podem otimizar o engajamento do consumidor com alguns processos do ciclo de vida, como lucro, segmentação, aquisição, ativação, retenção, dentre outros.

Para André Ferraz, CEO e Cofundador da *In Loco*,

[...] entender o comportamento do consumidor e impactá-lo no momento certo e com contexto é uma das formas que obtemos insights valiosos dos nossos usuários e no qual entregamos valor para ele. Por exemplo, imagine um consumidor que está desembarcando em um aeroporto tarde da noite e precisa ir para casa ou hotel. Neste exato momento, ele recebe uma oferta de desconto de um aplicativo de transporte. Isso não faz total diferença para ele? (informação verbal)⁹.

O ponto acima é interessante considerando os aspectos da conveniência e da oferta personalizada. O estudo e a análise do comportamento do consumidor para aprender a melhor forma de abordá-lo também gera valor para o cliente, visto que economiza seu tempo e liga sua necessidade à solução do problema com rapidez e eficácia.

⁹ FERRAZ, op. cit.

Os processos de *Big Data* podem ser utilizados de várias formas. Abaixo, encontram-se mapeadas algumas das principais formas de utilização de processos de *Big Data* e que possuem impacto no comportamento do consumidor:

3.2.1 Análise Preditiva

A análise preditiva é, possivelmente, a grande “essência” dos processos de *Big Data*. É através dela que a maior parte das organizações privadas de diversos setores utilizam aquela tecnologia no intuito de impactar o comportamento do consumidor em favor de seus negócios. O marco inicial está na identificação de padrões e na criação de previsões. De acordo com Schmarzo (2013), a integração de análises preditivas ou avançadas nos principais processos de negócios mantém o potencial de transformar todas as perguntas que os usuários empresariais estão tentando responder, e todas as decisões que estão tentando tomar.

Envolvendo transações bancárias, é possível exemplificar o caso da *Visa* e do *MasterCard*. Mayer-Schoenberger e Cukier (2013, p. 89) pontuam que “ao atender vários bancos e varejistas, eles obtêm mais transações em sua rede e podem usá-las para prever o comportamento do consumidor.”

A análise preditiva pode envolver também a antecipação de demandas e a criação de modelos preditivos para produtos e serviços, visando garantir o sucesso comercial dos lançamentos a partir da análise de enormes quantidades de dados produzidas por consumidores.

Segundo o site DataFlog (RIJMENAM, 2013), a *Procter & Gamble*, empresa multinacional de bens de consumo com uma receita líquida em torno de US\$ 84 bilhões por ano, realiza análises de simulação para projetar novos produtos. A análise de simulação ajuda a garantir o desempenho ideal do produto, levando em consideração variáveis distintas, criando e alterando virtualmente diferentes modelos ou projetos.

Mayer-Schoenberger e Cukier (2013) exemplificam também, em sua obra, a utilização de *Big Data* em dados médicos. Os autores abordam a iniciativa realizada pelo *MedStar Washington Hospital Center*, em Washington, D.C., junto com a *Microsoft Research* e com o uso do software *Amalgama*, da *Microsoft*. Foram analisados dados demográficos de pacientes, bem como seus exames, diagnósticos, tratamentos, entre outros, no intuito de reduzir as taxas de readmissão e infecções. A

ação resultou em uma lista de doenças que criavam oportunidades para que pacientes liberados fossem internados novamente um mês depois, fomentando então uma previsão acerca dos consumidores do serviço médico (pacientes) reincidentes.

Também é possível efetuar análises preditivas no setor de recursos humanos. Pode-se, por exemplo, identificar a probabilidade de um funcionário deixar a empresa.

A Evolv, empresa americana especializada em software de análise de grandes volumes de dados, foi contratada pela Xerox e descobriu que funcionários que fazem parte de uma ou mais — mas não mais de quatro — redes sociais têm menos chances de deixar o emprego e, portanto, diminuem a rotatividade. (JR, 2014, on-line).

A otimização de campanhas de marketing também pode se dar por meio da análise preditiva. Schmarzo (2013) indica que toda empresa gasta dinheiro em marketing, e partes crescentes desse montante estão sendo gastas em canais de mídia altamente mensuráveis. Classificar o nível de eficácia dos gastos em cada tipo de canal, seja ele on-line ou off-line, é um grande desafio. Assim, organizações que são capazes de quantificar e atribuir crédito a cada tipo de canal, além de realizar ações de marketing que impulsionam os negócios e o desempenho das vendas, estão mais bem posicionadas.

Além da identificação do melhor canal para investir, o grande potencial de segmentação de processos de *Big Data*, por si só, já é um grande diferencial para as áreas de *marketing* nas organizações, haja vista que propicia um entendimento mais exato acerca do público. Não obstante, a visão das atividades on-line dos consumidores é precursora de *insights* determinantes que podem ser utilizados a favor do negócio.

3.2.2 Mass Customization e Personalização

Tratando do tema de processos de *Big Data* e impactos no comportamento do consumidor, é de suma importância abordar os conceitos de *mass customization* e personalização.

O *mass customization* corresponde a uma segmentação mais avançada, na qual o consumidor é atingido de maneira mais específica para com seus anseios. Neste modelo, segundo a *Harvard Business Review*, “em vez de focar em mercados homogêneos e ofertas médias, os profissionais de *mass customization* identificaram as dimensões em que seus clientes diferem em suas necessidades.” (PINE, GILMORE, 1997, on-line).

Frank T. Piller e Mitchell M. Tseng (2010) determinam que, como paradigma de negócios, a customização em massa fornece uma proposta comercial atraente para agregar valor, abordando diretamente as necessidades do cliente. E, nesse meio tempo, utilizam recursos de maneira eficiente, sem incorrer em custos excessivos.

B. Joseph Pine II (1992), em sua obra, segregou os tipos de *mass customization* em quatro diferentes categorias. São elas:

- a) Customização colaborativa: formato no qual as empresas, em conjunto com os clientes, constroem produtos e serviços adequados e exclusivos para cada um;
- b) Customização adaptativa: formato no qual as empresas produzem produtos padronizados, mas personalizáveis pelo usuário final;
- c) Customização transparente: formato no qual as empresas fornecem produtos exclusivos para clientes individuais, mas sem declarar abertamente que esses produtos são personalizados;
- d) Customização cosmética: formato no qual as empresas produzem produtos padronizados, mas os vendem e comercializam de modos diferentes para clientes distintos.

Segundo o CEO da In Loco, André Ferraz:

[...] é impossível imaginar um mundo sem uso de dados para personalização de serviços e entrega de produtos daqui para frente. Entender a jornada das pessoas no mundo offline, possibilita que as empresas tracem estratégias mais assertivas, trabalhando o marketing omnichannel para impactar o consumidor certo, com mensagem relevante, no canal certo, dentro do contexto ideal. E o resultado é a fidelização, fluidez e entrega de conveniência para as pessoas e a mudança imediata no ponteiro de vendas das empresas (informação verbal).¹⁰

A personalização também pode ser considerada uma grande aliada na fidelização de clientes que, especialmente na customização colaborativa, sentem que estão construindo seu produto/serviço ideal junto com a empresa. E de fato estão. Esse compartilhamento e essa voz ativa do usuário, em que ele possui o poder de dizer suas preferências e conduzir a produção daquilo que ele mesmo irá consumir, são relevantes no sentido de estreitar o relacionamento com a marca.

¹⁰ FERRAZ, op. cit.

A personalização também possui direta ligação com o impacto que as mídias digitais apresentaram na relação entre usuários e marcas. No modelo antigo de comunicação, as marcas falavam e os usuários apenas ouviam, assumindo postura passiva. Hoje em dia, nota-se uma nova horizontalidade nesse relacionamento. Os *feedbacks* são constantes e, não obstante o monitoramento de comentários, *tags* e repercussões de forma geral nas mídias sociais fomentam grandes indicadores da qualidade de um produto ou serviço ou de como os mesmos foram recebidos pelo público. Tal dinâmica impulsiona, de forma significativa, a melhoria contínua por parte das empresas.

3.2.3 Mecanismos de Recomendação

Os Mecanismos de Recomendação estão bastante presentes em algumas das plataformas mais utilizadas atualmente, como o *Netflix*, a *Amazon* e o *Spotify*, sendo operados por algoritmos programados para identificar preferências individuais. Nesse contexto, Gillespie (2013) aborda em sua obra a capacidade dos algoritmos em identificarem o que é importante para cada usuário.

Popularmente, os sistemas de recomendação são reconhecidos pelos “também pode te interessar”, “pessoas que você talvez conheça” e “pessoas que adquiriram esse item também buscaram” de algumas plataformas digitais.

De acordo com o *LinkedIn* (VEENEMAN, 2018), as abordagens existentes para utilização de sistemas de recomendação são:

- a) Filtragem colaborativa: metodologia que considera a base de usuários para fazer previsões automáticas sobre preferências individuais;
- b) Filtragem baseada em conteúdo: metodologia que analisa semelhança entre itens como produtos, filmes, músicas ou livros. O objetivo é fazer previsões automáticas sobre itens semelhantes;
- c) Soluções híbridas: combinação das duas metodologias acima.

Um dos principais intuits da utilização de mecanismos de recomendação é aprimorar a experiência do usuário. A partir deles, além da facilitação, durante a busca, de produtos interessantes, é possível gerar valor para o usuário por meio de “descobertas”.

No caso do *Spotify*, usuários que baixaram o aplicativo provavelmente possuem interesse por música e *podcasts*. Logo, é interessante conhecer mais músicas de

gêneros preferidos. A mesma dinâmica funciona no caso do *Netflix*, que apresenta para seus usuários filmes e seriados com base em materiais audiovisuais assistidos previamente. Isso ocorre após um mapeamento de preferências. A *Amazon*, por fim, também segue a mesma linha, apresentando livros e outros produtos semelhantes ao que está sendo adquirido pelo consumidor. Comumente, tais descobertas possuem valor e aumentam o nível de satisfação do usuário com a plataforma.

O impacto dos mecanismos de recomendação no comportamento do consumidor pode ser evidenciado pela influência que os mecanismos possuem nos hábitos de compra, de navegação e do conteúdo em geral consumido pelos usuários. Os sistemas de recomendação aceleram o processo de aquisição de produtos e serviços por encurtarem o caminho percorrido pelo consumidor, além de induzirem os usuários a terem contato com novos materiais (novas músicas, filmes e itens). Assim, pode-se dizer que a jornada do usuário dentro dessas plataformas é “moldada”.

3.2.4 Precificação Dinâmica

O *pricing*, ou precificação dinâmica, é caracterizado, de maneira geral, pela “adaptação” de preços de produtos a depender do público e de acordo com alguns aspectos relacionados ao comportamento do consumidor, que podem considerar, por exemplo, o dispositivo que está sendo utilizado para efetuar a compra; o quão disposto o usuário está a comprar o produto; o nível de necessidade e interesse do consumidor; bem como a urgência da compra. Essa estratégia pode ser utilizada também para incentivar o consumidor a adquirir o item mais rapidamente (LARGHI, 2019).

O intuito da técnica é otimizar as decisões envolvendo a precificação, de forma a potencializar a geração de receita por parte da empresa. Para Mckinsey (BAKER; KIEWLL; WINKLER, 2014, on-line):

[...] para cada produto, as empresas devem ser capazes de encontrar o preço ideal que um cliente está disposto a pagar. Idealmente, eles levariam em consideração insights altamente específicos que influenciam o preço - o custo do próximo melhor produto competitivo versus o valor do produto para o cliente, por exemplo - e, então, chegam ao melhor preço.

Em pesquisa do *Valor Econômico* (2019), a utilização da precificação dinâmica por parte do *E-commerce* é reforçada. A partir de um monitoramento de duas semanas por parte da *Valor Investe*, foi percebido que a *Magazine Luiza* e a *Casas Bahia*, por

exemplo, demonstram variações ao longo do dia a partir da lógica do “preço dinâmico”, sendo que essa adaptação estaria mais ligada ao estoque da empresa.

A precificação e a busca pelo “preço ótimo”, especialmente em empresas grandes que possuem uma extensa gama de produtos e não conseguem dar vazão à precificação individual de cada produto, vêm se mostrando bastante eficiente na medida em que expandem as possibilidades do processo de precificação, não restringindo-o apenas à comparação com produtos concorrentes, um modo bastante antigo e tradicional. De acordo com a *Mckinsey* (BAKER; KIEWLL; WINKLER, 2014, on-line), “o aproveitamento da enxurrada de dados disponíveis nas interações com os clientes permite que as empresas tenham preços apropriados - e colham os frutos. ”

Entretanto, o novo modelo baseado em processos de *Big Data* vem gerando algumas discussões e críticas negativas. A onda de questionamentos é oriunda do aumento da percepção do usuário e da consciência digital. Sterne (2017, p. 187) afirma que as tentativas de precificação dinâmica já estão incomodando os consumidores. Segundo o autor:

[...] sites de viagens têm tido problemas em apresentarem preços maiores para usuários de MAC book em comparação a usuários de PC. A Amazon, por sua vez, obteve suas técnicas flagradas por consumidores que realizavam compras em paralelo, por meio de dispositivos diferentes. Também foi percebido que as passagens áreas diminuem de preço quando os cookies são retirados.

Este tipo de “flagra” vem ocorrendo não somente com a precificação dinâmica, mas com uma série de outras utilizações de processos de Big Data, tais como personalização, mecanismos de recomendação e aplicativos. Os anúncios que surgem em tom de coincidência rapidamente adquirem visibilidade dos usuários, que passam a questionar “como a máquina descobriu exatamente o que eu desejava?”.

4. ESTUDO DE CASO: CARACTERIZANDO USO POR EMPRESAS E RECEPÇÃO DE ESTUDANTES UNIVERSITÁRIOS

Neste capítulo, são descritos os parâmetros metodológicos utilizados para o desenvolvimento da monografia, os quais se basearam em pesquisa bibliográfica, estudos de caso, entrevistas em profundidade semiestruturadas com profissionais de *Big Data* de empresas da iniciativa privada e sondagem de opinião com estudantes da Faculdade de Comunicação (FAC) da Universidade de Brasília (UnB), buscando identificar o nível de compreensão dos mesmos acerca das tecnologias de *Big Data*.

Na sequência, são abordados exemplos de experiências das empresas entrevistadas com processos de *Big Data*, envolvendo as principais aplicações da ferramenta no âmbito de cada uma e de seus respectivos setores de atuação, além da demonstração dos ganhos percebidos para os negócios. Não obstante, também são pontuados os principais desafios enfrentados por organizações durante suas trajetórias de implantação de processos de *Big Data* e percepções sobre o *status* do mercado no que tange a utilização da tecnologia de *Big Data*. Também são trabalhados os principais riscos envolvendo a utilização da mesma, principalmente no que diz respeito à privacidade de dados no ambiente digital.

Por fim, apresenta-se os resultados obtidos por meio da pesquisa com os estudantes de Comunicação da Universidade de Brasília, expondo gráficos com valores percentuais, acompanhados de análises.

4.1 PARÂMETROS METODOLÓGICOS

Este estudo foi orientado por abordagem qualitativa, assim como apresenta objeto de estudo também qualitativo. Foram utilizados os seguintes procedimentos para coleta de dados:

a) Pesquisa bibliográfica/documental

Análise de materiais diversos, artigos, livros, pesquisas relacionadas a *Big Data*, *Machine Learning*, Algoritmos, Inteligência Artificial, *Marketing*, Datificação, Comportamento do Consumidor, dentre outros. Esta seção fomenta conhecimento geral acerca dos principais conceitos e ideias defendidos por autores do tema, assim como apresenta debates atuais envolvendo *Big Data* e sua repercussão.

A revisão bibliográfica é utilizada como base e referência para expor resultados das entrevistas com profissionais de empresas privadas que utilizam a tecnologia e pesquisa de sondagem de opinião com os estudantes da Faculdade de Comunicação (FAC) da Universidade de Brasília (UnB).

b) Estudos de caso

Esta etapa consiste na realização de uma análise da aplicação de processos de *Big Data* em empresas conhecidas na iniciativa privada, com vistas a apontar formas de utilização, benefícios percebidos, desafios e *cases* de sucesso, de forma a gerar entendimento sobre a utilização da nova tecnologia com viés prático e atual.

c) Entrevistas em profundidade semiestruturadas com profissionais de *Big Data* de organizações privadas

Seleção de quatro organizações privadas e uma PhD da área para realizar entrevistas, no intuito de mapear o entendimento dos mesmos em relação aos processos de *Big Data*; às formas de utilização; aos impactos decorrentes de seu uso; às áreas envolvidas; aos desafios técnicos de implementação; aos resultados percebidos; e à percepção das empresas entrevistadas em relação à forma que a questão ética afeta a utilização de ferramentas de *Big Data*.

Essa técnica objetivou captar a percepção de profissionais de mercado que ocupam posições de liderança e estão à frente de seus negócios e/ou atuam diretamente com processos de *Big Data*, especialmente no que tange aos impactos no comportamento do consumidor e do público-alvo dessas organizações, sejam elas *B2B* (*business to business*) ou *B2C* (*business to consumer*).

Visando abranger a utilização de processos de *Big Data*, foram realizadas entrevistas com o *Head of Operations and Data Analysis* da *Foxtrot Systems*, empresa de *Software as a Service* com base em São Francisco (CA); a *Data Specialist* do *Twitter*, empresa de rede social e *microblogging*; o *Big Data Strategist* da *AIS Digital*, empresa de tecnologia e *software* especializada em produtos digitais; e CEO e Cofundador da *In Loco*, empresa de tecnologia que fornece inteligência com base em dados de localização.

Além disso, também foi realizada entrevista em profundidade com Rebecca Barros, PhD em *Economic Analytics* e atuante no mercado com Transformação Digital.

d) Entrevista estruturada / Sondagem de Opinião com estudantes da Universidade de Brasília

Esta etapa compreende a aplicação on-line de questionários para alunos da Faculdade de Comunicação da Universidade de Brasília no intuito de identificar: se a amostra compreende de fato o que são processos de *Big Data*; como estes podem ser aplicados no seu cotidiano; a abordagem do tema por parte da Universidade dentro da esfera dos cursos de comunicação; e a percepção dos estudantes respondentes, em relação aos aspectos de privacidade discutidos, a partir desse novo recurso tecnológico.

A amostra de estudantes de Comunicação independe do semestre cursado e habilitação (Comunicação Organizacional, Publicidade, Audiovisual ou Jornalismo).

Os resultados obtidos por meio dos parâmetros metodológicos adotados são demonstrados durante o desenvolvimento da monografia, uma vez que as entrevistas realizadas com as organizações privadas abordadas previamente e a pesquisa de sondagem de opinião com alunos da Faculdade de Comunicação (FAC) da Universidade de Brasília (UnB) possuem notável sinergia com o referencial teórico utilizado.

Tal modelo foi definido no intuito de focar nos assuntos correlatos ao tema de processos de *Big Data*, procurando explicar, da forma mais completa possível, todos os tópicos abordados em linha lógica. Ademais, o resultado das entrevistas realizadas com empresas supracitadas encontra-se consolidado no tópico “Utilização de Processos de *Big Data* nos negócios, ganhos obtidos e impactos no comportamento do consumidor”, de forma a adotar um enfoque nos principais resultados positivos obtidos por essas organizações e como a utilização da tecnologia em questão afetou seu desenvolvimento.

4.2 BIG DATA E EXPERIÊNCIAS DE EMPRESAS

No intuito de ilustrar os resultados obtidos por meio da adoção de ferramentas de *Big Data*, serão explorados abaixo os exemplos de experiências das empresas entrevistadas para o desenvolvimento desta pesquisa. Será abordado, com base nos diálogos realizados, como utilizam essa tecnologia em seus negócios e os principais ganhos percebidos a partir desse uso.

4.2.1 Foxtrot Systems

Setor: Tecnologia aplicada à Logística

Produto ou serviço: Digitalização, automação e otimização logística por meio de operações inteligentes

Entrevistado: Fillipe Santos (*Head of Operations and Data Analysis*)

Em um contexto de evolução desenfreada da Indústria 4.0, os processos de *Big Data Analytics* são a essência da *Foxtrot Systems*. O processo de distribuição logística com o qual a organização se deparou em seus clientes era extremamente engessado, pré-definido e não considerava fatores externos bastante relevantes e que mudam toda a dinâmica do trânsito como, por exemplo, o sentido de vias, as condições de tempo, os congestionamentos, dentre outros.

Assim, o produto inicial da *Foxtrot* consistia em um roteirizador em tempo real, que funcionava a partir da integração de inúmeras bases de dados, ou seja, contemplava um volume descomunal de dados. Algumas dessas bases são oriundas do *Waze*, *Google API's*, *Weather Channel*, *Open Street Maps*, *Inrix*, dentre outras.

O “coração” da empresa consiste na seguinte equação:

$$F(X) = Tempo + Distância + Pontualidade$$

Ou seja, seu objetivo primordial é auxiliar seus clientes a fazerem entregas com o menor tempo possível, com a menor distância possível e buscando atingir a maior pontualidade possível.

Com o passar do tempo, alguns algoritmos de *Machine Learning* foram incorporados aos produtos da *Foxtrot*. Um deles, criado quando a empresa ainda não possuía base de dados própria, dizia qual era o melhor fator externo a ser considerado, ou seja, qual seria a base ideal a ser adotada para definir a rota.

Outro algoritmo da *startup*, por sua vez, constrói uma rota ideal, levando em consideração três fatores principais: menor custo, menor distância e respeito à pontualidade. É válido ressaltar que este já considera os horários de funcionamento dos estabelecimentos, visto que é necessário que a entrega seja realizada dentro desse período.

O algoritmo referente ao padrão de recebimento também configura uma importante ilustração do valor agregado dos dados. Caso o horário de funcionamento

de um estabelecimento seja das 9h às 18h, é possível identificar e aprender em qual horário exatamente dentro desse período a entrega é mais rápida em determinado estabelecimento. Por exemplo, às 10 horas, o motorista leva 30 minutos para entregar, já às 15 horas, ele leva 2 horas. Nesse caso, é melhor realizar a entrega nesse local específico sempre em torno das 14 horas no intuito de otimizar ao máximo a operação logística do cliente.

Esses são apenas alguns exemplos que demonstram o valor que pode ser extraído a partir do cruzamento de diversas fontes de dados, em grande volume e velocidade, e da importância dos processos de *Big Data* enquanto “alimentadores” da engenharia de *Machine Learning*, principal tecnologia da empresa, que necessita desse insumo para operar e aprender.

Atualmente, a *Foxtrot Systems* possui sua própria base de dados que, por toda a sua atuação e experiência com inúmeros pontos de vendas, tornou-se o principal ativo da empresa.

4.2.2 In Loco

Setor: Tecnologia e Dados

Produto ou serviço: Plataforma de Inteligência de Dados

Entrevistado: André Ferraz (CEO)

A *In Loco* é uma empresa de tecnologia que fornece inteligência com base em dados de localização. Suas soluções oferecem relevância aos clientes, garantindo o anonimato dos dados. Segundo informações disponíveis no site, a empresa possui 60 milhões de dispositivos em seu banco de dados, mais de 1.8 bilhão de visitas a locais físicos gravados mensalmente, 25 mil aplicativos utilizando o serviço da *In Loco* e mais de 16 Terabytes de dados entrando em servidores internos diariamente.

Apenas as informações supracitadas já demonstram o valor dos processos de *Big Data* para essa organização, uma vez que tamanho volume de dados possui capacidade indubitável de gerar *insights* que refletem o comportamento humano e, consequentemente, de um grande universo de consumidores.

O serviço funciona da seguinte maneira, de acordo com o site da organização: inicialmente os dados são coletados de dispositivos móveis por meio do *SDK* (*Software Development Kit*) instalado nos aplicativos parceiros. É válido ressaltar que

os últimos são obrigados a apresentar a Política de Privacidade da *In Loco* em seus próprios Termos e Condições de Uso e Política de Privacidade, além de informar que alguns dados pessoais dos usuários podem ser coletados pela *In Loco*.

Após o usuário aceitar a Política de Privacidade, o aplicativo solicita as permissões necessárias para usar as funcionalidades de localização no dispositivo. Uma vez autorizado, a *In Loco* começa a coletar os dados de maneira segura e sem identificações pessoais.

Com a tecnologia de localização de dispositivos ativa, a empresa pode detectar a presença em locais como lojas, shoppings, parques, praças públicas etc., desassociados da identidade do usuário. Dados de visitas a locais sensíveis, como templos religiosos, hospitais, partidos políticos, locais de entretenimento adulto e outros que possam ser usados para fazer inferências confidenciais não são coletados.

Os dados de localização coletados são agrupados em *clusters* - grupos de consumidores não identificados, agregados por comportamento semelhante. Assim, é possível criar métricas sobre o fluxo de visitas nas lojas, segmentar anúncios de acordo com o perfil do dispositivo, propiciar o envio de mensagens de aplicativos relevantes para o dispositivo, por meio da tecnologia de notificações *push*, dentre outros. “Nossa inteligência de localização usa dados anonimizados em soluções de negócio para influenciar fluxo em estabelecimentos e aumentar a relevância e usabilidade dos apps de grandes marcas por seus usuários. Os dados coletados são criptografados e agrupados em clusters de padrões de comportamento. Esses dados podem ser usados, então, para impactar grupos de consumidores com mensagens mais relevantes para seu contexto”, pontua o CEO da *In Loco*, André Ferraz.¹¹

Esse tipo de utilização de processos de *Big Data* por parte da *In Loco* ilustra perfeitamente os impactos da tecnologia no comportamento do consumidor, tendo em vista que ele é estudado e o entendimento alcançado por meio da análise gera a idealização de ações extremamente personalizadas para cada usuário.

“Entender o comportamento do consumidor e impactá-lo no momento certo e com contexto é uma das formas que obtemos *insights* valiosos dos nossos usuários e no qual entregamos valor para ele. Por exemplo, imagine um consumidor que está desembarcando em um aeroporto tarde da noite e precisa ir para casa ou hotel. Nesse exato momento, ele recebe uma oferta de desconto de um aplicativo de transporte.

¹¹ FERRAZ, op. cit.

Isso não faz toda a diferença para ele? Além disso, conseguimos entender o comportamento dos seus consumidores e engajá-los com notificações *push* no momento correto - uma empresa pode enviar um cupom de desconto ou “ativar o consumidor enquanto ele visita o concorrente. Porém, o maior caso de uso da nossa tecnologia é quando facilitamos o processo de abertura de contas digitais, apenas com a informação da localização, sem necessidade de comprovante de endereço, sem necessidade de CPF, RG ou informações que consideramos sensíveis”, destaca o CEO.¹²

A empresa também oferece a opção de validar endereços residenciais automaticamente durante o registro em aplicativos que usam a tecnologia da *In Loco*.

Um ponto que chama atenção na empresa refere-se à preocupação com a privacidade e confidencialidade de dados. “A *In Loco*, atualmente, processa mais de 16 terabytes de dados por dia. É um número enorme e, por isso, desde o nosso início da *In Loco* em 2011, ainda na faculdade, temos um compromisso com a privacidade. Ainda quando não era nem discutida aqui no Brasil. Nossa missão é resolver o problema da privacidade na internet e mostrar que é possível oferecer tecnologia e conveniência aliada à privacidade de dados”, disse Ferraz.¹³

“Como a privacidade é o cerne do nosso negócio, os nossos times de *Business Intelligence* têm maior interface com os *Big Datas*. Atualmente, temos uma equipe de quase 20 analistas de dados que possuem grande interação com os *Big Datas*. Além do time de *insights*, temos times de segurança, anonimização e privacidade que estão em constante contato com os dados para que essas informações estejam criptografadas e anonimizadas para que jamais seja possível chegar a dados identificáveis ou sensíveis dos mais de 60 milhões de dispositivos que possuem nossa tecnologia embarcada”, afirmou o CEO.¹⁴

4.2.3 AIS Digital

Setor: Tecnologia

Produto ou serviço: Produtos digitais (UX, UI, *Machine Learning*, APIs, *Big Data*, dentre outros)

¹² Idem.

¹³ Idem.

¹⁴ Idem.

Entrevistado(a): Samuel Corado (*Big Data Strategist*)

Segundo informações disponíveis no site, a AIS Digital é uma empresa de tecnologia especializada na construção de produtos digitais para clientes exigentes. As equipes da organização imaginam e desenvolvem interfaces digitais adaptadas às necessidades de cada cliente, envolvendo soluções de design, inovação e tecnologia. A empresa se concentra em três principais segmentos: automotivo, serviços financeiros e seguros, com sede no Brasil e clientes nas Américas e Europa. Seu modo de operar baseia-se em três pilares:

- Tecnologia: combinação de práticas de desenvolvimento de software de ponta com as mais recentes tecnologias, visando desenvolver produtos digitais escaláveis e de alta qualidade utilizados por milhões de usuários;
- Design: processos de UX baseados em estratégia de negócios, empatia e jornadas do usuário, no intuito de promover o alcance dos objetivos do usuário por meio da melhor experiência possível;
- Metodologia Agile: estrutura operacional baseada em metodologias ágeis como *Scrum* e *Kanban*, que garantem que o valor máximo seja entregue ao cliente no final de cada *sprint* (período de tempo definido durante o qual um trabalho específico deve ser concluído e preparado para revisão).

Samuel Corado, *Big Data Strategist* da AIS, apontou o valor dos *insights* obtidos pelos processos de *Big Data*, bem como informações que já forneceram acerca do comportamento de seu público: “foi de grande valia, primeiro, trazendo informações antes não apresentadas pelas limitações da tecnologia e grande volume de dados. Visão 360 do cliente, possibilitando a estratégia de *Customer Centric*”.¹⁵

No que tange à análise preditiva e às estratégias de *marketing*, vendas e relacionamento com clientes, Corado discorre sobre os impactos positivos percebidos: “primeiro, pela velocidade em que a informação chega para a tomada de decisão e a capacidade de processamento para grandes volumes de dados. Segundo, pela capacidade em cruzar informações estruturadas e não estruturadas gerando *insights* de forma rápida. Terceiro, por democratizar o acesso à informação, fazendo com que outros departamentos participem do processo exploratório de dados.”¹⁶

¹⁵ CORADO, op. cit.

¹⁶ Idem.

Um dos setores em que há maior utilização de tecnologia de *Big Data* e *Analytics* para clientes, de acordo com o site da organização, corresponde ao de serviços financeiros. A empresa pontua que bancos e instituições financeiras estão evoluindo de um modelo centralizado e fechado para um modelo descentralizado e aberto. Os recém-chegados estão pressionando a indústria para oferecer melhores serviços e experiências aos seus clientes. Assim, a *AIS Digital* oferece soluções para seus clientes que envolvem novas experiências e serviços personalizados, centrados no cliente por meio desse ferramental.

O outro refere-se à área de seguros. Segundo a empresa, as companhias de seguro estão evoluindo para negócios voltados a serviços, e deixando de ser apenas empresas de sinistro que somente cobriam riscos. Agora, elas passam a trazer mais segurança para seus clientes. A AIS ajuda então algumas das maiores companhias de seguros a criarem novas formas de engajar e interagir com seus usuários, valor muitas vezes gerado por processos de *Big Data* na medida em que esses fornecem *insights* acerca de preferências, tendências e prioridades dos consumidores, além da capacidade de aprimorar significativamente a experiência do usuário durante o contato com o serviço financeiro prestado.

4.2.4 Twitter

Setor: Rede Social e *Microblogging*

Produto ou serviço: Plataforma de Rede Social

Entrevistado(a): Raíssa Arriola (*Data Specialist*)

O *Twitter* armazena dados de milhões de usuários pelo mundo. A empresa oferece serviço de rede social e servidor para *microblogging*. Assim, a utilização de *Big Data Analytics* torna-se um instrumento extremamente efetivo para a empresa no sentido de mapear tendências sociais, assuntos mais comentados do mundo por meio de palavras-chave nos *tweets* (o que está *trending*), opinião pública, dentre outros.

Raíssa Arriola, *Data Specialist* do *Twitter*, reforça a mudança do posicionamento e da atuação da plataforma: “hoje em dia o *Twitter* é um local de *news*. Não mais de ver o que o coleguinha está postando. Causa negra é um assunto. Tudo

é um assunto. *Twitter* já foi pessoal. E agora está mais por interesse”.¹⁷ Assim, há um grupo na empresa responsável por entender determinados tipos de usuários, voltado a interesse. O foco passou a ser, então, tópicos de conversas e temas gerais, desconsiderando informações do perfil do indivíduo.

Em seu site, a empresa afirma que “o público do *Twitter* é inclinado, influente e apaixonado por uma variedade de interesses. Portanto, ser capaz de segmentar o público com base em vários parâmetros - interesses, dados demográficos, sistema operacional, marcas que eles seguem e uma variedade de outras variáveis - significa que seu conteúdo é direcionado para as pessoas certas. Isso também garante que as campanhas ressoem, o conteúdo tenha um alto envolvimento e os anúncios tenham um desempenho brilhante nos momentos mais importantes.”¹⁸

É dessa maneira que a empresa utiliza processos de *Big Data* para entender o comportamento do consumidor. Eles são agrupados por interesses e temas comuns, gerando a conexão e aproximação de usuários dentro do ambiente da plataforma. Tal utilização é extremamente proveitosa no mercado de clientes do *Twitter*, que podem utilizar informações da rede para entender melhor o que está sendo dito sobre sua marca.

O site do *Twitter* também afirma que o *dashboard* de *Audience Insights* do *Twitter* inclui informações valiosas sobre o público geral do *microblogging*, seus seguidores e as pessoas que se envolveram com seus *tweets*. É possível identificar novos públicos-alvo relevantes para as próximas campanhas, por exemplo. O *website* indica também que o painel ainda fornece informações agregadas sobre dados demográficos, interesses, estilo de vida e comportamentos de compra do usuário - tudo proveniente de dados fornecidos pelo usuário, dados fornecidos pelo aplicativo, modelos internos do *Twitter* e dados gerados por parceiros. “As pessoas não são unidimensionais - sua publicidade também não deve ser”, finaliza o *Twitter* em seu site.¹⁹

Raíssa Arriola exemplificou o “*Share of Voice (SOV)*”, sob responsabilidade do grupo interno de *Audience Insights*, que identifica, por exemplo, a marca mais mencionada na copa do mundo - ocorre em eventos e no mercado de clientes. É possível constatar qual é o *fast food* mais falado, a cerveja mais comentada, territórios

¹⁷ ARRIOLA, op. cit.

¹⁸ Idem.

¹⁹ Idem.

de maior interesse, dentre outros. Quando ocorre um evento ou situação específica, como a Copa do Mundo, a equipe de pesquisa e *stakeholders* já ficam atentos aos resultados da plataforma após a análise das menções e do conteúdo dos *tweets*.

Além disso, a *Data Specialist* comentou que o time de *Audience Insights* realiza estudos sobre públicos como, por exemplo, os *millennials*. Por conseguinte, nesse caso, é possível perceber que a segmentação ultra efetiva da audiência é um dos ganhos a partir da utilização de processos de *Big Data* e *Analytics*.

Arriola reforça, ainda, a importância dos *insights* obtidos por meio das tecnologias de *Big Data*, explicando o funcionamento da tecnologia dentro da plataforma do *Twitter*. “Há um time de data que consegue misturar tópicos, por exemplo, todos os dados da conversa do brasileiro, maior menção do jogo, de jogador, etc)”, explica.²⁰

O *Twitter* apresenta uma funcionalidade *Analytics*. Segundo o site da organização, é possível analisar os *tweets* e entender seus seguidores. Cada palavra, foto, vídeo e seguidor pode ter um impacto. São oferecidas duas funcionalidades principais:

- *Account Home*: consiste em um boletim do *Twitter*, com estatísticas de alto nível rastreadas mês a mês. Também atua como uma galeria de sucessos, indicando os *tweets* com melhor desempenho e principais influenciadores da rede;
- *Tweet Activity Dashboard*: refere-se às métricas de cada *tweet*. São apresentadas informações, tais como quantas vezes usuários do *Twitter* visualizaram, retweetaram, curtiram e responderam a cada *tweet*.

Outra funcionalidade interessante da plataforma, que possui impacto direto no comportamento do consumidor, refere-se ao *Ad Targeting*, direcionado a organizações que utilizam o *Twitter*. O site da empresa afirma que ao anunciar no *Twitter*, é possível utilizar recursos robustos de segmentação para colocar sua empresa na frente do público certo. A rede oferece várias formas de segmentação para o alcance de metas de negócio:

- Segmentação por idioma (alcance de pessoas que entendem determinado idioma);

²⁰ Idem.

- Segmentação por gênero (alcance de homens ou mulheres);
- Segmentação por interesse (alcance de usuários cujos interesses se alinham amplamente aos negócios da organização precursora da campanha);
- Segmentação por seguidor (alcance de pessoas que provavelmente irão se interessar em consumir seu conteúdo por meio da segmentação de seguidores de contas relevantes);
- Segmentação por dispositivo (alcance de usuários com base no dispositivo móvel específico que estão utilizando para acessar o *Twitter*);
- Segmentação por comportamento (alcance de públicos de alta intenção no *Twitter* com base em padrões de compras e gastos);
- Segmentação por público-alvo personalizado (alcance de grupos específicos de usuários do *Twitter* por meio das listas de CRM);
- Segmentação por palavra-chave (atuação nos sinais de intenção, de forma a entregar mensagens oportunas a usuários com base no que eles *tweetaram* ou se envolveram recentemente nos *tweets*);
- Segmentação por geográfica (alcance de público global ou alcance apenas em um país, região ou até cidade específica).

A extensa gama de possibilidades de segmentação oferecidas pelo *Twitter* demonstra a valia do exorbitante volume de dados da plataforma, e como os processos de *Big Data* podem ser utilizados por organizações privadas, por meio da plataforma *Twitter*, para enviar a mensagem certa para o público certo, de maneira a atingi-lo em alinhamento com os objetivos de negócio.

4.3 PRINCIPAIS DESAFIOS PARA IMPLEMENTAÇÃO DE PROCESSOS DE *BIG DATA* NAS EMPRESAS

Schmarzo (2013), em sua obra, afirma que um dos principais pilares para uma jornada bem-sucedida de *Big Data* consiste no alinhamento entre os negócios e as partes interessadas de Tecnologia da Informação (TI), de forma a utilizar cases que agregam valor suficiente ao negócio e, ao mesmo tempo, possuem alta probabilidade de sucesso. O autor pontua que, apesar do valor comercial atraente do uso de *Big*

Data e analytics avançado, muitos casos possuem baixa probabilidade de êxito devido aos seguintes aspectos:

- a) Indisponibilidade de dados precisos e oportunos;
- b) Falta de experiência com novas fontes de dados (mídias sociais, dispositivos móveis etc);
- c) Ciência de dados, recursos de *analytics* avançado ou habilidades insuficientes;
- d) Falta de experiência com novas tecnologias, como *Hadoop*, *MapReduce* e mineração de texto;
- e) Limitações arquitetônicas e tecnológicas ao gerenciar e analisar dados não estruturados e ingestão de *feeds* de dados em tempo real;
- f) Relacionamento de trabalho fraco entre equipes de negócio e equipes de TI;
- g) Falta de firmeza e apoio gerencial.

Mayer-Schonberger e Cukier (2013, p. 10) apontam que:

[...] o grande passo na direção de um gerenciamento mais eficiente dos dados aconteceu com o advento da digitalização, que tomou as informações analógicas compreensíveis a computadores, o que também facilitou e barateou o armazenamento e processamento.

De acordo com Fillipe Santos, *Head de Operations e Data Analytics de Foxtrot*²¹, o principal desafio técnico encontrado na implementação técnica de processos de *Big Data* reflete-se na maturidade da base de dados de cliente.

Além disso, foi pontuada pelo executivo a necessidade de gestão da mudança. Segundo ele, é necessário fazer com que os funcionários entendam o real valor da nova tecnologia, que costuma ter um impacto nas organizações mais tradicionais. (informação verbal).

Sterne (2017, p. 37) também trata da necessidade de “educar” e engajar o público de uma organização acerca das possibilidades atreladas à utilização de processos de *Big Data*.

A melhor forma de conquistar o coração e a mente daqueles que mais podem se beneficiar com o seu talento para a descoberta de dados é educá-los. Quanto mais pessoas em sua organização entenderem as formas e significados, bem como os riscos e recompensas envolvidos, mais eles irão te procurar para respostas, te incluir em sessões de

²¹ SANTOS, op. cit.

planejamento e suportar suas solicitações para mais dados, pessoas e ferramentas.

Outro ponto cabível relaciona-se à formação de profissionais capacitados, conforme pontua Rebecca Barros, PhD em *Economic Analytics* ²²:

No momento, há dificuldade de recursos e alta demanda. Está faltando profissionais no mercado. As empresas estão focadas ainda em um tipo específico de skill, que é cientista de dados. Mas na verdade a grande carência é para engenheiro de dados.

Para Rebecca Barros, outro desafio relacionado a implementação dos projetos de *Big Data* é criar uma cultura *data driven*, além de garantir a viabilidade econômica dos projetos e os resultados reais para o negócio dos clientes. (informação verbal).

É possível abordar também as dificuldades técnicas de armazenamento do enorme volume de dados, e a necessidade de capacidades técnicas para lidar com essa quantidade (extração de dados, montagem de base *analytics*), de acordo com Rebecca Barros.

Davenport (2014) aponta que, embora seu tamanho receba grande atenção, o aspecto mais difícil do *Big Data* refere-se à sua falta de estrutura. Rebecca Barros, PhD em *Economic Analytics*, indica que “70 a 80% de um projeto de *Big Data* é o esforço de preparar os dados”. (informação verbal) ²³

O CEO da In Loco, André Ferraz ²⁴, também pontua que “o desafio é encontrar profissionais que lidem com um volume enorme de dados, tenham familiaridade com estatística, entendam de computação e tenham a capacidade de automatizar tarefas”. (informação verbal). Sterne (2017) diz ainda que o melhor analista deve ter a capacidade de manipular os dados de várias maneiras para obter informações relevantes, a partir de uma internalização dos objetivos da organização, e gerando *insights* tangíveis que podem ser utilizados para tomadas de decisão na empresa.

É indubitável o barateamento dos custos de armazenamento e processamento dos dados, que tornou tecnologias embasadas em processos de *Big Data* mais sustentáveis e lucrativas dentro de um negócio.

No que se refere ao valor dos processos de *Big Data*, para Davenport (2014), a questão não é se surpreender com o volume de dados, mas analisá-los e convertê-los em *insights*, inovações e valor comercial.

²² BARROS, op. cit.

²³ Idem.

²⁴ FERRAZ, op. cit.

Em complemento, Rebecca Barros, PhD em *Economic Analytics* afirma que é importante se preocupar com a implementação de um projeto de *Big Data*. Ela pontua, dessa forma, o que chama de “tradutores”, que seriam profissionais responsáveis por questionar como a iniciativa pode, de fato, agregar valor para o negócio (informação verbal).²⁵

Sterne (2017, p. 37) também aborda o papel dos denominados tradutores:

Você conhece seus dados por dentro e por fora, mas os consumidores de suas ideias, que devem depender de suas recomendações, não. Para eles, seus dados são tão legíveis quanto um cristal bola ou uma sequência de cartas de tarô. (...) portanto, sua responsabilidade é informar sem confundir, encorajar sem mistificar e tranquilize sem recorrer a truques de mão.

Dessa maneira, é possível evidenciar a relevância da “educação” para os dados, tornando os *insights* obtidos por meio dos mesmos mais palpáveis e compreensíveis para o público ao qual se destina. A atenção depositada nesse fator também pode ser essencial na medida em que demonstra o valor gerado por meio dos processos de *Big Data* que podem, muitas vezes, soar demasiadamente técnicos de forma a não dar visibilidade a todo o modelo de funcionamento e esforço por trás.

4.4 RISCOS RELACIONADOS À PRIVACIDADE E ÉTICA NA UTILIZAÇÃO DE DADOS

Apesar de todos os seus inquestionáveis benefícios para os negócios da iniciativa privada, é imprescindível abordar alguns dos riscos envolvidos nessa utilização. A questão relativa à privacidade dos dados tem se tornado cada vez mais relevante e polêmica. Contudo, outros aspectos são preocupantes tais como o monitoramento social que se torna viável por meio do acompanhamento e análise em tempo real de volumes exorbitantes de dados, os quais possuem viés pessoal.

O excesso de informações pessoais coletadas pelas grandes empresas de tecnologia torna o público vulnerável e mais suscetível à manipulação, considerando também que, normalmente, não há plena ciência sobre quais, quantos, quando ou como estes dados estão sendo extraídos.

Segundo Boyd e Crawford (2012, p. 664):

[...] por um lado, o *Big Data* é visto como uma ferramenta poderosa para lidar com vários problemas sociais, oferecendo o potencial de

²⁵ BARROS, op. cit.

novas ideias em áreas tão diversas quanto pesquisa sobre câncer, terrorismo e mudanças climáticas. Por outro lado, o *Big Data* é visto como uma manifestação preocupante do *Big Brother*, permitindo invasões de privacidade, diminuição das liberdades civis e maior controle estatal e corporativo.

É válido mencionar também o potencial de processos de *Big Data* em ampliar a disseminação de desinformação (informações que não correspondem à realidade, mas são divulgadas como verídicas, comumente denominadas *fake news*). Isso ocorre pela capacidade intrínseca da ferramenta em segmentar determinados públicos-alvo, tornando as campanhas de marketing ultra efetivas. Assim, muitos especialistas apontam os processos de *Big Data* como questionáveis em termos éticos, visto que podem manipular massas.

Outro aspecto negativo envolvendo o uso de processos de *Big Data* refere-se à invasão da privacidade individual a partir da coleta e análise de dados que usuários disponibilizam em sites, aplicativos, redes sociais, lojas etc., muitas vezes sem terem o conhecimento de que suas informações pessoais estão sendo armazenadas naqueles locais. Segundo Schoenberger e Cukier (2013), os dados podem não parecer informações pessoais explícitas, mas, com os processos de análise, podem facilmente dizer a quem se referem ou facilitar a dedução de detalhes íntimos da vida de uma pessoa.

Atualmente, com a percepção acerca da personalização de serviços e produtos ofertados individualmente em redes sociais e outras plataformas *web*, está ocorrendo um certo despertar por parte dos usuários no que diz respeito à privacidade de seus dados pessoais e sobre como as empresas estão captando essas informações.

De acordo com o CEO da In Loco, André Ferraz ²⁶:

[...] existe uma parcela de consumidores que é mais ligada ao digital, a geração touch, que já nasceu com acesso às telas touch screen e que entende como o mundo digital funciona e como o processo de captação de dados influencia a personalização de produtos e serviços. Mas esse movimento não pode ser considerado mainstream, ele está acontecendo aos poucos. As gerações que nasceram antes da ascensão do digital, tem um pouco mais de dificuldade de entender esses processos, mas estão buscando informações para compreender melhor como os seus dados são utilizados no mundo on-line. Isso é perceptível quando falamos sobre privacidade. Os consumidores estão ficando cada vez mais conscientes sobre o volume de dados que estão cedendo para acessar serviços personalizados. Eu acredito que vivenciamos um momento em que é preciso discutir mais sobre privacidade (informação verbal).

²⁶ FERRAZ, op. cit.

Algumas empresas demonstram sua preocupação com a privacidade dos dados de seus usuários. De acordo com Raíssa Arriola ²⁷, *Data Specialist* do *Twitter*:

[...] essa questão parte muito do posicionamento da empresa. Tabelas existem, dados estão aí. A empresa é que é responsável por barrar e determinar que esse comportamento não é condizente com ela. Qualquer um que tem a skill consegue. É só criar a linguagem e dar play, e isso é muito perigoso (informação verbal).

Segundo Arriola, no *Twitter*, “todos os acessos aos dados precisam passar por uma aprovação do Legal. É tudo muito controlado”. (informação verbal) ²⁸

Na *In Loco*, por exemplo, André Ferraz afirma:

[...] além do time de *insights*, temos times de segurança, anonimização e privacidade que estão em constante contato com os dados para que essas informações estejam criptografadas e anonimizadas para que jamais seja possível chegar a dados identificáveis ou sensíveis dos mais de 60 milhões de dispositivos que possuem nossa tecnologia embarcada. (informação verbal) ²⁹

Questões relacionadas à privacidade também podem ser preocupações internas na empresa, a depender do tipo de negócio e da relação com seus colaboradores.

Os motoristas da *Foxtrot Systems*, por exemplo, utilizam um aplicativo que captura informações de localização por meio de georreferenciamento. Segundo Fillipe Santos ³⁰, *Head de Operations e Data Analytics* da *Foxtrot Systems*, “a preocupação atual são os dados do motorista (localização, tempo utilizando o aplicativo, movimentação)”.

A orientação dada pela empresa é de que os motoristas apenas façam *login* no aplicativo ao chegarem no trabalho, para evitar que a empresa tenha acesso à informação pessoal. “Se o motorista não fizer o processo corretamente, a empresa acaba pegando muito dado dele”, afirma Fillipe. ³¹

Atualmente, na iniciativa privada, tem ocorrido uma mudança no sentido de aumentar os processos que garantem a privacidade de dados e de usuários, especialmente com as novas legislações vigentes. No entanto, ainda há um longo

²⁷ ARRIOLA, op. cit.

²⁸ Idem.

²⁹ FERRAZ, op. cit.

³⁰ SANTOS, op. cit.

³¹ Idem.

caminho a ser percorrido. Rebecca Barros ³², PhD em *Economic Analytics*, reforça que ainda existe muita manipulação de dados em ambientes não protegidos.

Para Raíssa Arriola³³, *Data Specialist* do *Twitter*, “com a leis está ficando mais difícil. Está mudando, dois, três anos atrás, qualquer um fazia”.

Além de questões de privacidade, é importante salientar o monitoramento e controle social propiciados por meio da vigilância de dados. Para Mayer-Schoenberger e Cukier (2013, p. 118), “o *Big Data* permite maior vigilância de nossas vidas ao mesmo tempo que torna obsoletos alguns meios legais de proteção à privacidade”. Assim, o autor ressalta as mudanças pragmáticas necessárias para viabilizar a adaptação da sociedade à era do *Big Data* e o seu lado negro. Para proteção da privacidade, é necessário que os usuários da tecnologia tenham mais responsabilidade sobre suas ações, ao passo que a sociedade terá que redefinir sua noção sobre justiça para garantir a liberdade de agir e formar novos tipos de profissionais que possam compreender os algoritmos e que, por conseguinte, defendam as pessoas que podem ser atingidas por eles.

4.5 ESTUDO DE PERCEPÇÃO: ALUNOS DE COMUNICAÇÃO DA UNIVERSIDADE DE BRASÍLIA

Visando mapear as principais questões relativas ao entendimento dos alunos da Faculdade de Comunicação da Universidade de Brasília (FAC/UnB) foi realizada uma pesquisa on-line de sondagem de opinião.

A amostra utilizada para condução do estudo contemplou 143 (cento e quarenta e três) estudantes da Faculdade de Comunicação (FAC), das quatro habilitações do departamento. Foram 41 respondentes de Comunicação Organizacional (de 417 alunos), 38 respondentes de Jornalismo (de 275 alunos), 33 de Publicidade e Propaganda (de 287 alunos) e 31 de audiovisual (de 218 alunos).

O questionário desenvolvido buscou, dentre outros objetivos, mapear o conhecimento geral sobre processos de *Big Data* por parte dos estudantes, com vistas a compreender como os discentes percebem o fenômeno, assim como identificar se a amostra possui ciência da presença dessa tecnologia em seu cotidiano. No intuito de avaliar a validade do que foi medido, foi também solicitado que os respondentes

³² BARROS, op. cit.

³³ ARRIOLA, op. cit.

assinalassem os serviços que utilizam – todos possuem ferramentas de *Big Data* envolvidas.

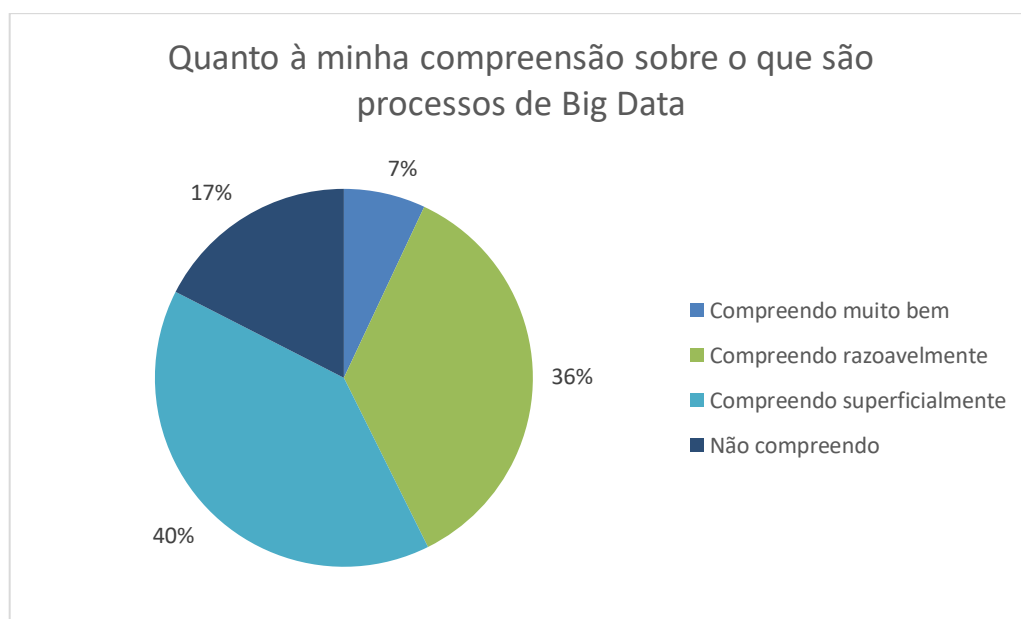
Essa questão também é medida na pergunta relacionada à abordagem ou não desse tema na UnB, seja em momentos universitários seja disciplinas formais.

Não obstante, também buscou-se analisar como a amostra lida com a privacidade de dados. Foi questionado, por exemplo, a frequência com a qual os respondentes informam seu CPF para obterem descontos ou quaisquer vantagens na compra de produtos ou serviços, bem como o hábito cultivado ou não de ler os termos de privacidade ao baixar aplicativos.

Por fim, buscou-se compreender como os alunos de Comunicação da UnB se sentem ao receberem ofertas de produtos e serviços muito personalizadas, com base em seus dados de navegação. Analisou-se, pois, se eles se sentem confortáveis ou não com esse comportamento por parte das empresas, e em caso negativo, se são tomadas atitudes para prevenir isso.

O primeiro resultado obtido diz respeito à compreensão dos estudantes analisados acerca do que são processos de *Big Data*. Mais de 82% da amostra afirmou compreender o conceito em torno dessas ferramentas, independentemente do nível (muito bem – 7%, razoavelmente – 36%, superficialmente – 40%).

Gráfico 1 - Compreensão sobre processos de *Big Data* (%)

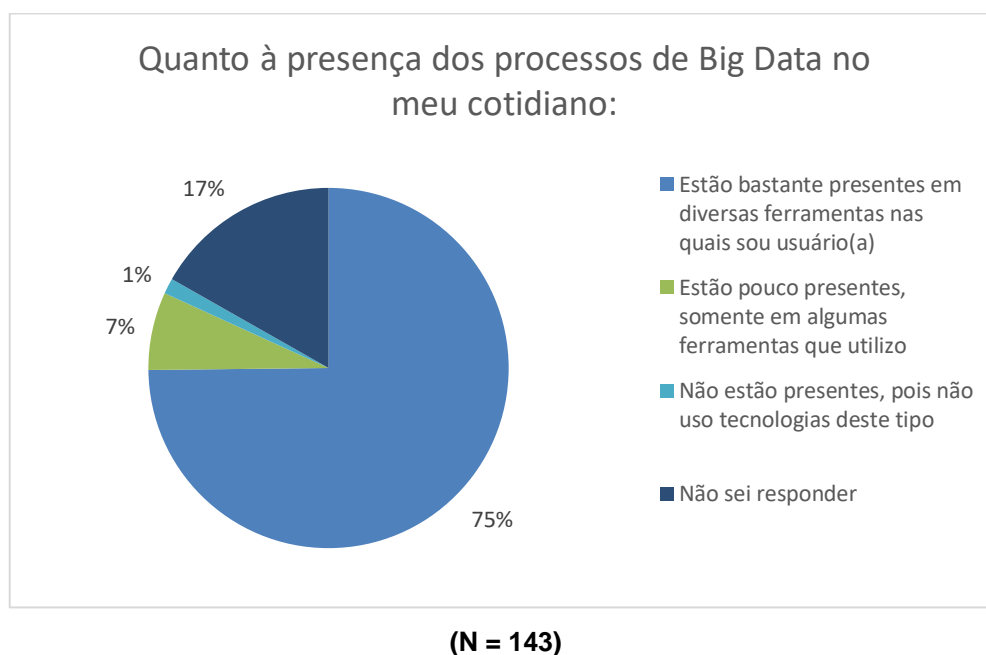


(N = 143)

Fonte: elaborado pela autora, 2019.

Ademais, quase 75% dos respondentes demonstrou ter alto grau de consciência acerca da presença de processos de *Big Data* em seu cotidiano por meio de ferramentas utilizadas com recorrência e que possuem a tecnologia. A maioria dos respondentes indicou que o conhecimento é raso (compreendo razoavelmente e compreendo superficialmente), no entanto, há entendimento suficiente para compreender a presença significativa de processos de *Big Data* no dia a dia dos indivíduos atualmente.

Gráfico 2 - Presença dos processos de *Big Data* no cotidiano (%)



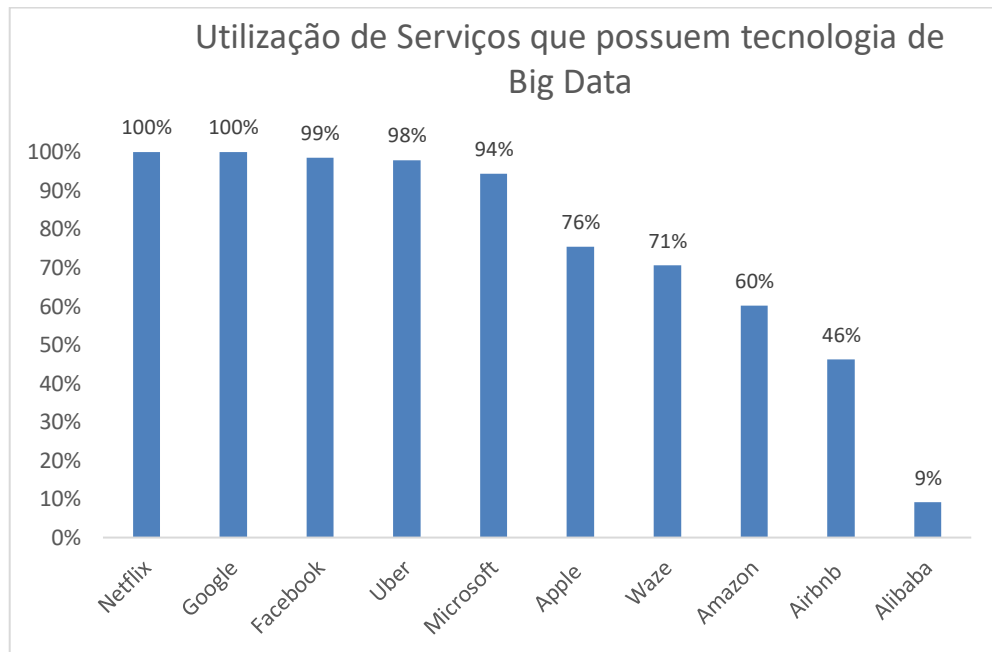
Fonte: elaborado pela autora, 2019.

O gráfico acima demonstra que 74.8% dos respondentes afirmaram que os processos de *Big Data* estão presentes em diversas ferramentas das quais são usuários. Conforme abordado durante o desenvolvimento deste trabalho, a empresa *AIS Digital* afirmou em entrevista realizada que a geração atual possui mais conhecimento sobre o que são esses processos, apresentando então maior facilidade para entender como a extração e a análise de dados é realizada pelas empresas.

Os componentes da geração Z representam os nascidos entre a década de 80 e início dos anos 2000. Essa faixa demográfica vivenciou a revolução da internet. Dessa forma, a alta conectividade de pessoas, produtos e serviços é bastante naturalizada pelo grupo, que foi criado já em um mundo conectado e globalizado.

O gráfico abaixo reforça o alto nível de utilização de serviços que possuem a tecnologia de *Big Data* por parte de estudantes de Comunicação, que compõem tal geração. A grande incidência de uso está acompanhada da consciência sobre a presença da tecnologia e como ela se insere em algumas das plataformas mais utilizadas hoje em dia por parte da amostra respondente.

Gráfico 3 - Utilização de serviços que possuem *Big Data* (%)

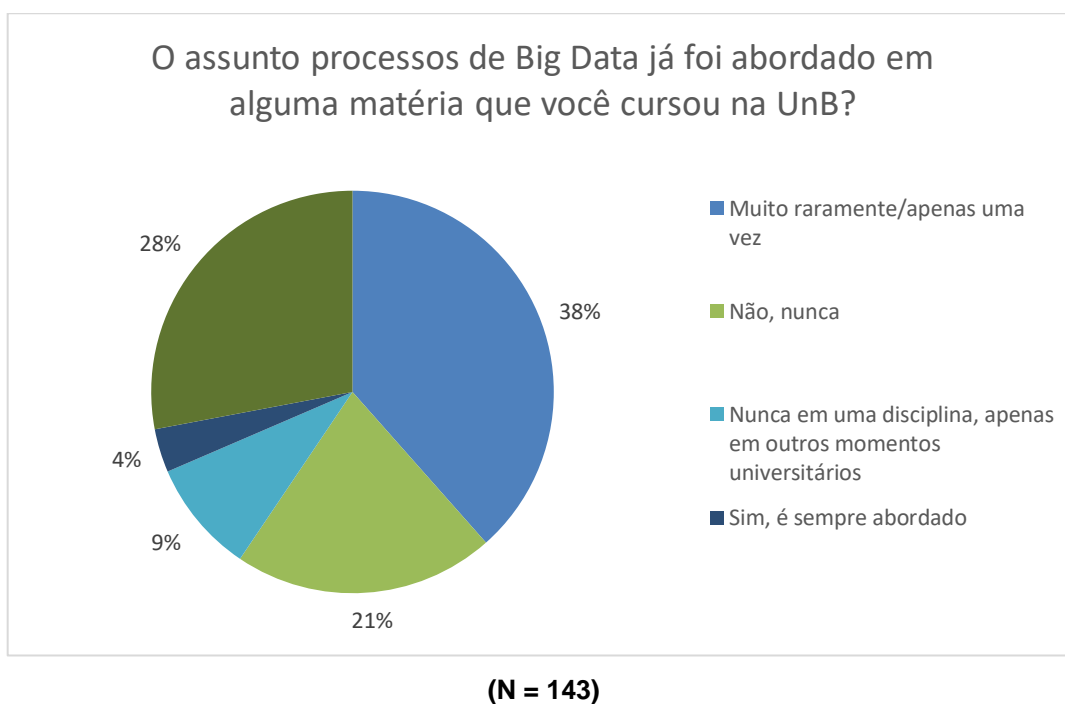


(N = 143)

Fonte: elaborado pela autora, 2019.

O momento tecnológico atual é caracterizado por grande popularidade em torno de Inteligência Artificial, *Machine Learning*, processos de *Big Data* etc. Assim, também é possível atribuir a consciência da amostra acerca da presença de ferramentas de *Big Data* em seu cotidiano à tal popularização e ao fato de que esses conceitos são uma tendência do momento. Comumente, também são levantados aspectos que tangem o novo modelo de mercado trabalho e as novas profissões emergentes, que contarão, em sua vasta maioria ou até exaustão, com a utilização de tecnologias da Indústria 4.0.

Esse é um tema frequentemente debatido e apontado por empresas e estudos. O gráfico abaixo demonstra como a amostra obteve contato com o assunto de processos de *Big Data* dentro da Universidade de Brasília (UnB).

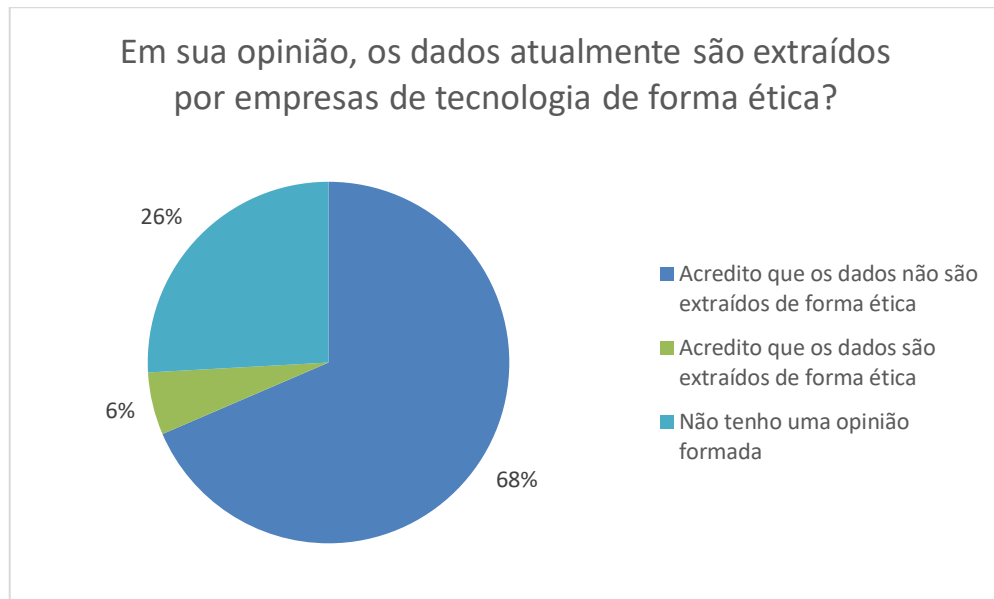
Gráfico 4 - Processos de *Big Data* nas matérias cursadas na UnB (%)

Fonte: elaborado pela autora, 2019.

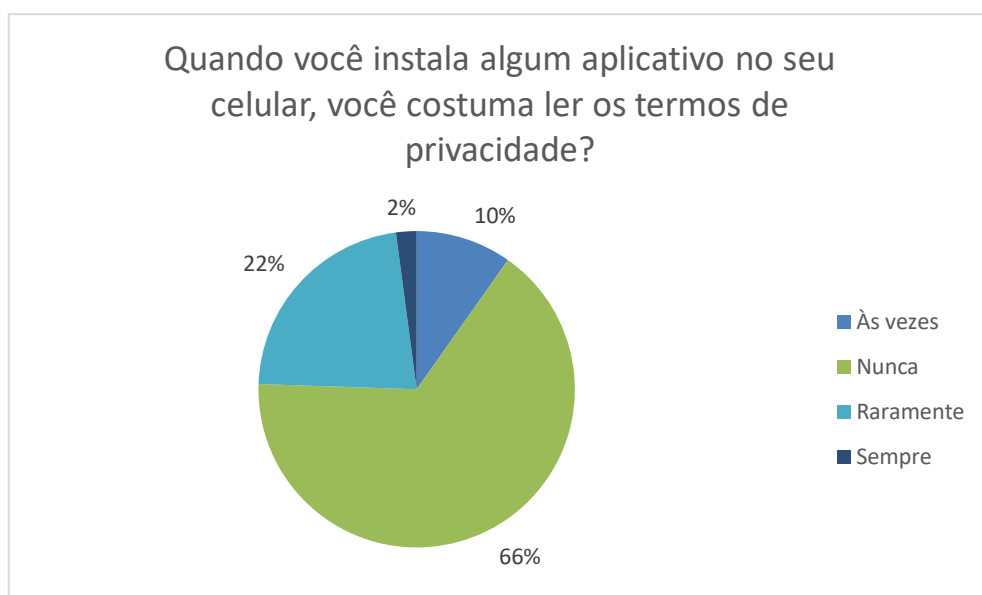
Nota-se que apenas 30 respondentes de um universo de 143 (21%) afirmam que nunca tiveram contato com o tema dentro do ambiente universitário. No entanto, também é interessante observar que, no âmbito dos cursos da Faculdade de Comunicação (FAC), 55 dos respondentes afirmaram que só entraram em contato com o tema “muito raramente/apenas uma vez”. Tal dado sugere que as disciplinas das habilitações da FAC possivelmente estão abordando o tema em menor quantidade.

No que se refere à percepção do aspecto ético presente na extração de dados pessoais por empresas, 68.5% dos respondentes acreditam que os dados não são extraídos de forma ética, ao passo que mais de 25% afirmam não ter uma opinião formada e apenas 3.5% acreditam que os dados são sim extraídos de maneira íntegra.

O resultado sugere que há uma significativa desconfiança da amostra para com o modelo de coleta de dados pessoais por parte de empresas de tecnologia.

Gráfico 5 - Percepção sobre extração dos dados (%)**(N = 143)****Fonte:** elaborado pela autora, 2019.

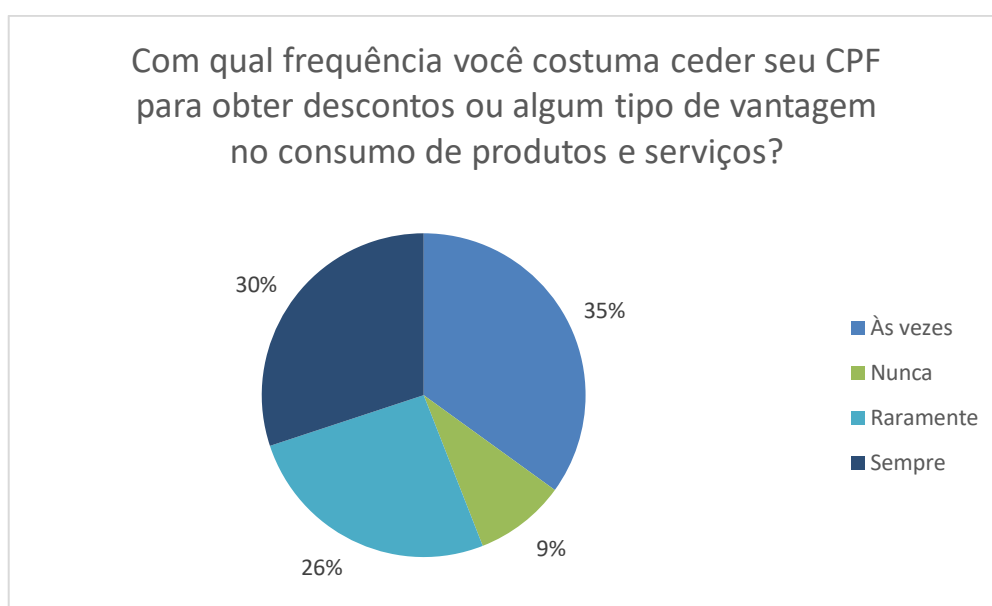
No entanto, curiosamente, 66% dos alunos afirmaram que nunca leem os termos de privacidade ao baixar um aplicativo em seu celular. Isso configura o repasse de informações muito valiosas para companhias, no entanto, não há clareza sobre a função e o papel desse tipo de contrato para a maioria das pessoas.

Gráfico 6 - Termos de privacidade dos aplicativos (%)**(N = 143)****Fonte:** elaborado pela autora, 2019.

Em entrevista realizada com Raíssa Arriola ³⁴, *Data Specialist* do *Twitter*, mencionada acima no desenvolvimento deste trabalho, a profissional comentou a fragilidade dos usuários em relação aos dados, que pode ser verificada por meio do fato exposto no gráfico. A postura pode ser explicada devido à falta de conhecimento técnico sobre como os dados são, de fato, extraídos em dispositivos móveis. Além disso, também pode-se atribuir o resultado aos próprios aplicativos que, muitas vezes, não estimulam os usuários a lerem tais termos de privacidade.

No que se refere à concessão de informações pessoais como CPF para obter descontos em lojas e em demais circunstâncias que implicam no consumo de produtos e serviços, apenas 9% da amostra afirmou nunca fornecer esse tipo de dado para obter vantagens.

Gráfico 7 - Fornecimento do CPF (%)



(N = 143)

Fonte: elaborado pela autora, 2019.

Um dos aspectos interessantes percebidos a partir dos resultados dessa pesquisa corresponde à aparente incoerência dos usuários no que tange a sua noção de privacidade e disposição para fornecer dados pessoais: 68% dos respondentes disseram que acreditam que as empresas não extraem dados de forma ética, contudo, 65,7% afirmaram que nunca leem os termos de privacidade quando baixa um

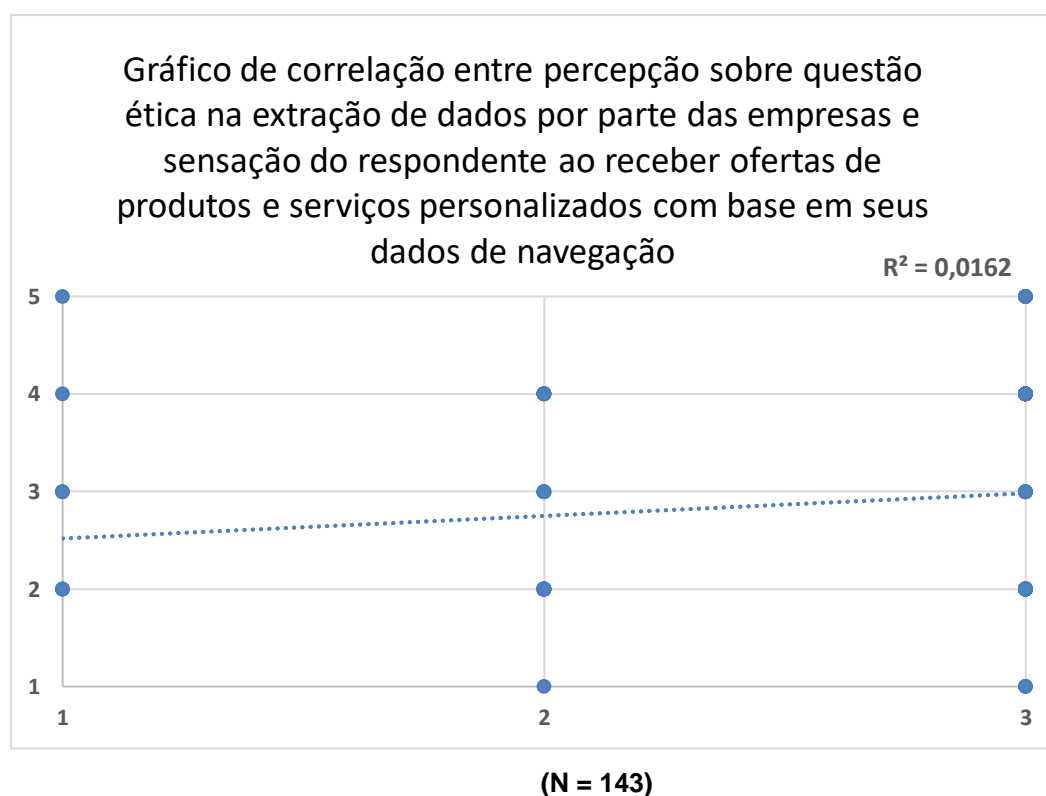
³⁴ ARRIOLA, op. cit.

aplicativo, e 65% responderam que às vezes ou sempre concede seu CPF para obter descontos ou vantagens no consumo de produtos e serviços.

Tal fato pode ser relacionado com o estudo de Boyd e Crawford (2012), que trata do *Big Data* como sendo uma manifestação de *big brother* que permite invasões de privacidade e também Dijck (2014, p. 198), que pontua “uma notável tolerância ao *Big Brother* e ao *Big Business* acessando rotineiramente as informações pessoais do cidadão, também conhecidas como *Big Data*. ”

O gráfico abaixo tem por objetivo demonstrar a correlação entre as respostas para as perguntas “Em sua opinião, os dados atualmente são extraídos por empresas de tecnologia de forma ética?” e “Quando alguma empresa utiliza seus dados de navegação para te ofertar seus serviços de forma mais personalizada você considera isso:”

Gráfico 8 - Correlação entre variáveis



Fonte: elaborado pela autora, 2019.

Na estatística descritiva, o R de Pearson corresponde a um coeficiente de correlação linear entre variáveis. Trata-se de um índice adimensional com valores situados entre -1,0 e 1,0.

$r = 1$ Significa uma correlação perfeita positiva entre as duas variáveis.

$r = -1$ Significa uma correlação negativa perfeita entre as duas variáveis - isto é, se uma aumenta, a outra sempre diminui.

$r = 0$ Significa que as duas variáveis não dependem linearmente uma da outra.

De acordo com estatística, então, quanto mais próximo de 1,0, mais próximo está de haver correlação entre duas variáveis.

É possível constatar que não há nenhum tipo de correlação entre ambas as perguntas realizadas no questionário, pois no caso ilustrado acima, R-Quadrado é igual a 0,0162.

Essa é uma questão interessante a ser analisada, uma vez que não há coerência entre considerar uma postura de extração de dados antiética por parte de empresas de tecnologia (primeira pergunta), mas não afirmar que o uso de dados de navegação por parte de empresas para realizar ofertas personalizadas seja inaceitável (segunda pergunta).

Uma justificativa possível para tal incoerência pode consistir no fato de que, apesar da extração ser antiética aos olhos dos respondentes, as ofertas personalizadas e demais conclusões que as empresas chegam sobre seu consumidor geram, sim, valor para o cliente final. Dessa maneira, nota-se uma postura ambígua por parte dos usuários, que permitem o *big brother* mesmo discordando ou, no mínimo, tendo desconfiança sobre o modelo de coleta de dados pois, no final, eles também obtêm ganhos com isso.

5. CONSIDERAÇÕES FINAIS

Primeiramente, esta monografia foi idealizada vislumbrando discutir as principais características de processos de *Big Data* e como o desenvolvimento desta tecnologia está sendo refletido no mercado privado, seja com ganhos quantitativos ou qualitativos, nível de competitividade, trajetória, desafios ou projeções futuras. Em um segundo momento, considerando a quebra de paradigmas anteriores e criação de novas conjecturas de negócio, buscou-se compreender, sob uma ótica mais específica, como o comportamento do consumidor está sendo impactado e influenciado neste novo cenário.

Não obstante, a visão do próprio consumidor também foi analisada, no intuito de compreender o posicionamento dos usuários diante do novo relacionamento construído entre organizações e marcas com seus públicos. Não é raro, por exemplo, notar a desconfiança e curiosidade dos indivíduos quando percebem que as ofertas realizadas por marcas durante a sua navegação em sites, corresponde exatamente ao que estavam buscando no *Google* no dia anterior. Aplicativos de controle de saúde feminina sabem antes mesmo da própria mulher sobre uma gravidez. Este tipo de acontecimento configura um modelo inédito de abordagem com o consumidor.

Dessa maneira, como conclusão do trabalho realizado, é imprescindível mencionar os avanços da tecnologia com a ascensão da Inteligência Artificial, *Machine Learning* e Internet das Coisas dentro de um contexto de Indústria 4.0, além de ressaltar a posição dos processos de *Big Data* como insumo e alicerce para que a execução de tais ferramentas seja possível.

As ferramentas de *Big Data* consistem em uma mudança de paradigmas não somente para os negócios, mas também no que se refere às formas de compreender e ler os dados, buscando o desapego das relações de causalidade e abraçando a correlação, que é subsidiada, muitas vezes, por padrões sem motivo aparente, óbvio, ou justificável.

Na iniciativa privada, seu uso tornou-se fundamental para o avanço dos negócios em termos de valor de mercado e competitividade, sendo comumente utilizada por organizações para o desenvolvimento de produtos e serviços, campanhas e ações de *marketing*, prevenção contra fraude, gerenciamento de riscos, precificação dinâmica, mapeamento de tendências, dentre outros.

O impacto dos processos de *Big Data* no comportamento do consumidor é indubitável, na medida em que funções de análise preditiva, *mass customization*, mecanismos de recomendação e precificação dinâmica buscam compreender ao máximo diversos aspectos do consumidor no intuito de utilizar tais percepções em função do alcance de metas nos negócios, buscando maximizar seus resultados.

Nessa conjectura, os termos de datificação e dataísmo são grandes norteadores, uma vez que pregam a utilização de dados on-line para compreender o comportamento dos usuários e a crença de que tais amostras quantitativas são verídicas e proporcionam conclusões alinhadas à realidade.

A utilização de ferramentas de *Big Data* apresenta ganhos exponenciais para as empresas e possibilita a visão de aspectos que nunca foram percebidos antes de tais tecnologias. A inovação também permitiu soluções jamais pensadas anteriormente, muitas decorrentes da quantidade de dados passíveis de serem adquiridos em quantidades elevadas e de inúmeras fontes. O poder de escuta dos dados tornou-se uma competência ultra valorizada, especialmente quando a análise é feita de forma correta por profissionais capacitados e traduzida da melhor forma possível para o negócio, de forma a realmente gerar valor para as organizações e seus respectivos clientes.

Para implantar os processos de *Big Data*, as empresas precisam enfrentar desafios relacionados à criação de uma cultura *data driven* (orientada a dados), contratação de profissionais com as competências necessárias para o desenvolvimento de atividades relacionadas à essa tecnologia e tradução apropriada das soluções para os negócios, de forma que os *insights* obtidos por meio dos processos de *Big Data* sejam devidamente utilizados na realidade e esfera de atuação de cada organização.

Entretanto, é importante mencionar os riscos advindos do uso de ferramentas de *Big Data*. O mais relevante e conhecido deles refere-se à invasão de privacidade por meio do acesso a dados pessoais de usuários por parte das empresas. Muitas vezes, o contato com esse tipo de informação é realizado sem consentimento, ou com consentimento, mas sem explicação clara ou devido destaque. Outro aspecto negativo refere-se à utilização do poder preditivo dos processos de *Big Data*, bem como conceitos de datificação e dataísmo para monitoramento e controle social.

No intuito de observar a percepção de alunos de Comunicação da Universidade de Brasília (UnB), foi realizada uma pesquisa de sondagem de opinião. A amostra

demonstrou bastante consciência em relação à presença de processos de *Big Data* em seu cotidiano, bem como a preocupação com aspectos éticos da extração de dados por empresas de tecnologia. Entretanto, foi percebida certa ambiguidade. Os resultados demonstraram que a grande maioria dos respondentes não lê termos de privacidade de aplicativos ao baixá-los em seu celular, bem como fornecem dados pessoais como CPF para obter vantagens no consumo de produtos e serviços. Além disso, uma significativa parcela dos alunos afirmou que considera aceitável quando empresas utilizam seus dados de navegação para realizar ofertas personalizadas.

Futuramente, o desenvolvimento de pesquisas relacionadas à postura do consumidor e usuário frente às novas tecnologias seria interessante, buscando compreender o seu comportamento em relação à utilização de seus dados pessoais por empresas e agências governamentais, além dos novos moldes de relacionamento entre indivíduo e organização dentro de uma conjectura digital e de Indústria 4.0. Além disso, o viés sociológico também é interessante de ser analisado a partir de uma perspectiva de mudança nas relações interpessoais e contratos sociais, que passariam a ser regidos por monitoramento social digital.

REFERÊNCIAS

AFP. Pesquisa global revela que 86% dos internautas já acreditaram “fake news”. **Exame**, 12 jun. 2019. Disponível em: <https://exame.abril.com.br/brasil/pesquisa-global-revela-que-86-dos-internautas-ja-acreditaram-fake-news/>. Acesso em: 15 nov. 2019

ANÁLISES Preditivas: o que são e qual sua importância? **SAS**, 2019. Disponível em: https://www.sas.com/pt_br/insights/analytics/analises-preditivas.html. Acesso em: 27 out. 2019.

ARTIFICIAL Intelligence: what is and why it matters? **SAS**, 2019. Disponível em: https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html. Acesso em: 15 nov. 2019

BAKER, W.; KIEWLL, D.; WINKLER, G. Using big data to make better pricing decisions. **McKinsey&Company**, jun. 2014. Disponível em: <https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/using-big-data-to-make-better-pricing-decisions>. Acesso em: 27 out. 2019.

BEAN, R. Variety, Not Volume, Is Driving Big Data Initiatives: For large corporations, data variety trumps volume when looking for insights. **Mit Sloan Management Review** (Org.), 28 mar. 2016. Disponível em: <https://sloanreview.mit.edu/article/variety-not-volume-is-driving-big-data-initiatives/>. Acesso em: 09 out. 2019.

BIG DATA: os cinco Vs que todo mundo deveria saber. **Canaltech**. Disponível em: <https://canaltech.com.br/big-data/Big-Data-os-cinco-Vs-que-todo-mundo-deveria-saber/>. Acesso em: 04 nov. 2019.

BOYD, D.; CRAWFORD, K. Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication & Society**, [s.l.], v. 15, n. 5, p. 662-679, jun. 2012. Informa UK Limited. Disponível em: <http://dx.doi.org/10.1080/1369118x.2012.678878>. Acesso em: 4 nov. 2019.

CANABARRO, T. Como Utilizar a Tecnologia de Big Data Para Barrar Fraudes no E-Commerce? **E-Commerce Brasil**, 12 jan. 2017 Disponível em: <https://www.ecommercebrasil.com.br/artigos/como-utilizar-tecnologia-de-big-data-para-barrar-fraudes-no-e-commerce/>. Acesso em: 27 out. 2019.

DANCE, G. J.X; LAFORGIA, M.; CONFESSORE, N. As Facebook Raised a Privacy Wall, It Carved an Opening for Tech Giants: Internal documents show that the social network gave Microsoft, Amazon, Spotify and others far greater access to people’s data than it has disclosed. **The New York Times**, Nova Iorque, 18 dez. 2018. Disponível em: <https://www.nytimes.com/2018/12/18/technology/facebook-privacy.html?module=inline>. Acesso em: 08 out. 2019.

DAVENPORT, T.H. **Big Data at Work**: Dispelling the Myths, Uncovering the Opportunities. Harvard Business Review Press, 2014.

DATAVEILLANCE. **Oxford Dictionary**. Lexico, 2019. Disponível em: <https://www.lexico.com/en/definition/dataveillance>. Acesso em: 26 out. 2019.

DEEP learning: o que é e qual sua importância? **SAS**, 2019. Disponível em: https://www.sas.com/pt_br/insights/analytics/deep-learning.html. Acesso em: 20 out. 2019.

DESJARDINS, J. How much data is generated each day? **World Economic Forum**, 17 abr. 2019. Disponível em: <https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/>. Acesso em: 09 out. 2019.

DIAS, C.R. Qual a relação entre Deep Learning e Machine Learning? **IGTI Blog**, 26 jan. 2018. Disponível em: <http://igti.com.br/blog/relacao-entre-deep-learning-e-machine-learning/>. Acesso em: 20 out. 2019.

EMBRACING INDUSTRY 4.0 and Rediscovering Growth. **Boston Consulting Group**. 2009. Disponível em: <https://www.bcg.com/pt-br/capabilities/operations/embracing-industry-4.0-rediscovering-growth.aspx>. Acesso em: 01 nov. 2019

FACEBOOK'S data-sharing deals exposed. **BBC**, 19 dez. 2018. Disponível em: <https://www.bbc.com/news/technology-46618582>. Acesso em: 08 out. 2019

GILLESPIE, T.; BOCZKOWSKI, P.J.; FOOT, K.A. **The Relevance of Algorithm**. In: GILLESPIE Tarleton; BOCZKOWSKI. Pablo J.; FOOT. Kirsten A. **Media Technologies, Essays on Communication, Materiality and Society**. Elsevier, 2013. P.168-193

HONDA, H; FACURE, M; YAOHAO, P. Os Três Tipos de Aprendizado de Máquina. **Laboratório de Aprendizado de Máquina Aplicado a Finanças e Organizações. (LAMFO/UnB)**, 27 jul. 2017. Disponível em: <https://lamfo-unb.github.io/2017/07/27/tres-tipos-am/>. Acesso em: 20 out. 2019.

INDUSTRY 4.0. **Deloitte Insights**, 2019. Disponível em: <https://www2.deloitte.com/us/en/insights/focus/industry-4-0.html>. Acesso em 01 nov. 2019.

INDUSTRY 4.0 EXPLAINED. **KPMG Belgium**, 2018. Disponível em: <https://home.kpmg/be/en/home/insights/2018/01/industry-4-0-explained.html>. Acesso em: 01 nov. 2019.

JR, S.T. O futuro do RH está no Big Data. **Exame**, 16 jun. 2014. Disponível em: <https://exame.abril.com.br/carreira/o-futuro-do-rh-esta-no-big-data/2/>. Acesso em: 27 out. 2019.

KING, T. 80 Percent of Your Data Will Be Unstructured in Five Years. **Solutions Review**, 28 mar. 2019. Disponível em: <https://solutionsreview.com/data-management/80-percent-of-your-data-will-be-unstructured-in-five-years/>. Acesso em: 09 out. 2019.

LAFORGIA, M.; ROSENBERG, M.; DANCE, G. Facebook Data Deals Are Under Criminal Investigation. **The New York Times**, 13 mar. 2019. Disponível em: <https://www.nytimes.com/2019/03/13/technology/facebook-data-deals-investigation.html>. Acesso em: 08 out. 2019.

LARGHI, N. Preço no comércio eletrônico varia de acordo com celular. **Valor**, São Paulo, 7 out. 2019. Disponível em: <https://valor.globo.com/financas/noticia/2019/10/07/preco-no-comercio-eletronico-varia-de-acordo-com-celular.ghtml>. Acesso em: 12 out. 2019.

MACHINE Learning e Deep Learning: aprenda as diferenças. **Salesforce Brasil**, 30 abr. 2018. Disponível em: <https://www.salesforce.com/br/blog/2018/4/Machine-Learning-e-Deep-Learning-aprenda-as-diferencas.html>. Acesso em: 20 out. 2019.

MACHINE Learning: learning how to entertain the world. **Netflix**, 2019. Disponível em: <https://research.netflix.com/research-area/machine-learning>. Acesso em: 19 out. 2019.

MACHINE Learning: o que é e qual sua importância? **SAS**, 2019b. Disponível em: https://www.sas.com/pt_br/insights/analytics/machine-learning.html. Acesso em: 20 out. 2019.

MAYER-SCBONBERGER, V.; CUKIER, K. **Big Data**: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana. Editora Elsevier, 2013.

MAYER-SCBONBERGER, V.; CUKIER, K. **Big Data**: A Revolution that will transform how we live, work and think. Editora Elsevier, 2012.

MANYIKA, J. et al. The McKinsey Global Institute (mgi). Big data: The next frontier for innovation, competition, and productivity. **McKinsey Global Institute**, mai. 2011. Disponível em: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier>. Acesso em: 06 out. 2019.

MCKINSEY GLOBAL INSTITUTE. Big data: The next frontier for innovation, competition, and productivity. **McKinsey Global Institute**, 2011. Disponível em: https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.ashx. Acesso em: 28 out. 2019

MARR, B. The Key Definitions of Artificial Intelligence (AI) That Explain Its Importance. **Forbes**, 14 fev. 2018. Disponível em:

<https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#1e6c70e44f5d>. Acesso em: 19 out. 2019.

MILLS, M.P. Apple Is The New Exxon And Data Is The New Oil: The Path To The First \$10 Trillion Company. **Forbes**, 9 ago. 2018. Disponível em: <https://www.forbes.com/sites/markpmills/2018/08/09/apple-is-the-new-exxon-data-is-the-new-oil-and-the-path-to-the-first-10-trillion-company/#d027ab51526f>. Acesso em: 18 out. 2019.

MORGAN, B. How Amazon Has Reorganized Around Artificial Intelligence And Machine Learning. **Forbes**, 16 Jul. 2018. Disponível em: <https://www.forbes.com/sites/blakemorgan/2018/07/16/how-amazon-has-reorganized-around-artificial-intelligence-and-machine-learning/#d4170fe73618>. Acesso em: 15 nov. 2019

NETTO, A.S.D.P; MORO, E.P.; FOLLY, F. Os 5 V's do Big Data. **Grupo de Teleinformática e Automação (GTA/UFRJ)**. Disponível em: https://www.gta.ufrj.br/grad/15_1/bigdata/vs.html. Acesso em: 20 out. 2019.

O QUE É Big Data? **Oracle**, 2019. Disponível em: <https://www.oracle.com/br/big-data/guide/what-is-big-data.html>. Acesso em: 27 out. 2019.

PILLER, F.T.; TSENG, M.M. (2010). **Handbook of Research in Mass Customization and Personalization**. 1ª Edição, Singapore, World Scientific Publishing Co. *E-book*. Disponível em: https://books.google.com.br/books?hl=pt-BR&lr=&id=fjRkDQAAQBAJ&oi=fnd&pg=PR5&dq=difference+between+mass+customization+and+personalization&ots=YYnWjxHmlh&sig=hDC_Z8LN7Y-DwH4fXHAqzUMTVLM#v=onepage&q=difference%20between%20mass%20customization%20and%20personalization&f=false. Acesso em: 28 out. 2019

PINE II, B.J.; GILMORE, J.H. The Four Faces of Mass Customization. **Harvard Business Review**, jan-fev. 1997. Disponível em: <https://hbr.org/1997/01/the-four-faces-of-mass-customization>. Acesso em: 27 out. 2019.

PINE II, B.J. Mass Customization: The New Frontier in Business Competition. **Harvard Business Review Press**, 1992. Disponível em: <https://www.investopedia.com/terms/m/masscustomization.asp>. Acesso em: 27 out. 2019

RIJMENAM, M.V. How P&G Uses Big Data To Turn Diapers Into Insights. **Datafloq**. Disponível em: <https://datafloq.com/read/pg-big-data-turn-diapers-insights/312>. Acesso em: 12 out. 2019.

SCHMARZO, B. **Understanding How Data Powers Big Business**. Indianapolis: John Wiley & Sons, Inc., 2013.

SILVEIRA, C. O que é Indústria 4.0 e Como Ela Vai Impactar o Mundo. **City Systems**. Disponível em: <https://www.citisystems.com.br/industria-4-0/>. Acesso em: 01 nov. 2019.

STERNE, J. **Artificial Intelligence for Marketing, Practical Applications**. 1 ed. Wiley: 2017.

THE WORLD'S most valuable resource is no longer oil, but data. **The Economist**, 6 mai. 2017. Disponível em: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>. Acesso em: 17 out. 2019.

VAN DIJCK, J. In data we trust? The implications of datafication for social monitoring. **Universiteit Utrecht**, Utrecht, Países Baixos. Amsterdam, 2017.

VAN DIJCK, J. (Surveillance and Society) Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. **University of Amsterdam**, Países Baixos. Amsterdam, 2014.

VEENEMAN, M. How do recommendation engines work? **LinkedIn**, 23 ago. 2018. Disponível em: <https://www.linkedin.com/pulse/how-do-recommendations-engines-work-mourad-veeneman>. Acesso em: 28 out. 2019.

ZUBOFF, S. Big other: surveillance capitalism and the prospects of an information civilization. **Harvard Business School Emerita**, Boston, Estados Unidos. Berkman Center for Internet and Society, Cambridge. DOI:10.1057/jit.2015.5. 2015.

FRIESNER, T. What is Customer Behavior. **Marketing Teacher**. Disponível em: <https://www.marketingteacher.com/what-is-consumer-behavior/>. Acesso em: 16 dez. 2019

APÊNDICES

APÊNDICE A – Entrevista com Rebecca Barros, PhD em *Economic Analytics*

1. Quais setores da organização possuem maior interface com os processos de Big Data utilizados?

Não se aplica.

2. Quais foram os principais desafios técnicos para implementação dos processos de Big Data nas organizações? Eles costumam ser completamente superados?

O principal problema é a criar uma cultura data driven, garantir a viabilidade econômica dos projetos e os resultados reais de projetos para os negócios. Os projetos que começaram só com dados e infra tecnológica não funcionam. Analytics – elementos: pessoas, equipamentos hardwells e softwares, processos, projeto para resultado, ingredientes (dados).

70 a 80% de um projeto de Big Data é o esforço de preparar os dados. “Dado é que nem petróleo – só tem valor quando é refinado.”

A extração dá muito trabalho (arquitetura – data supply chain)

A parte tecnológica barateou nos últimos tempos – é possível montar um ambiente muito rápido, especialmente no modelo de Software as a Service (SaaS)

No momento, há dificuldade de recursos e alta demanda. Está faltando profissionais no mercado (gap). As empresas estão focadas ainda em um tipo específico de skill, que é cientista de dados. Mas na verdade a grande carência é para engenheiro de dados (implementação de resultado para o negócio)

Engenheiro – extrai dados, monta uma base analytics. Já o cientista de dados – implementação de algoritmo (clientes). É necessário o tipo de profissional que é o “tradutor” – O que eu vou fazer com isso? Iniciativa, reversão? Como aplicar no negócio de fato?

Big Data é meio, e não é fim. Há muitos equívocos nesse sentido

Big data tem um problema de armazenamento – por isso a importância do engenheiro de dados (quais equipamentos, tecnologias usar, qual segurança).

3. Como são obtidos insights pelos processos de Big Data? Que tipo de informação eles fornecem acerca de comportamento de público-alvo?

Formalmente isso é analytics – big data é fazer analytics. Arquitetura tecnológica completamente diferente.

Telecom é indústria que mais utiliza este tipo de tecnologia e analytics (bem desenvolvida nesse sentido). Indústrias de varejo com canais digitais também utilizam muito. Entretanto, qualquer empresa B2C precisa estar se preparando. Preocupação em trazer valor efetivo para o negócio. Tem muitos projetos analíticos e de big data que a empresa não implementou aquilo na prática

4. Como você acredita que as previsões e predições obtidas por meio do Big Data afetam as estratégias de marketing, vendas e relacionamento com o consumidor ou cliente (B2B)?

Já respondido acima.

5. Como você vê questão ética nos processos de Big Data que visam influenciar o comportamento do consumidor? Como isso é tratado pelas organizações?

As organizações estão heterogêneas no sentido de privacidade. A legislação chegou e obrigou muitas empresas a se readaptarem. Ainda estão correndo nesse sentido.

O time de TI não acompanhou a evolução digital (privacidade, segurança, volumetria). Há uma discussão bastante recorrente sobre “levar ou não para nuvem? Não significa que tem mais ou menos segurança, é questão de cultura.

Ainda há muita manipulação de dados em ambientes não protegidos.

APÊNDICE B – Entrevista com Filipe Santos, *Head of Operations and Data Analysis* da Foxtrot Systems

1. Quais setores da organização possuem maior interface com os processos de Big Data utilizados?

Filipe como Head de Analytics é responsável por entender necessidades. Depois, por levar as ideias, sugestões de algoritmos, para a equipe responsável. A execução do algoritmo é feita em São Francisco. Inclusive, estão desenvolvendo uma área de vendas lá.

Em relação aos clientes externos, é realizado um contato com a área da empresa que contrata o serviço. Por exemplo, setores de logística dos clientes, que interagem bastante.

Em recursos humanos também ocorre essa utilização (horário de entrada e saída de funcionários, hora extra, gestão de pessoas, garantindo por exemplo que o funcionário faça horário de almoço).

A área financeira também pode ser citada, com controle de renda física, forma de pagamento, quanto está entrando em cash, etc.

Contudo, costumam ser as áreas fim: logística e comercial. Mas todos os setores das empresas clientes consomem em algum nível. Quando você entra com um processo, todas as áreas participam. Sempre há então interface com áreas que não sejam fim.

2. Quais foram os principais desafios técnicos para implementação dos processos de Big Data na organização? Eles foram completamente superados?

O principal foi o estágio de maturidade das principais organizações, que é inferior ao que imaginávamos encontrar em termos de base de dados.

Todos têm bases de dados estruturadas, mas é sempre uma bagunça e falta informação. Dados errados, falta de limpeza dos dados, etc. É preciso, então, tornar o produto agnóstico para não depender de input da empresa.

Por ser uma startup, é necessário fazer uma grande de gestão da mudança. Quebrar paradigmas, fazer com que os funcionários antigos entendam o valor daquilo. Esse é o impacto das tecnologias em empresas mais tradicionais. É preciso conscientizar para o ganho de eficiência, e tomar cuidado com o que as áreas que vão falar.

Um dos desafios da Foxtrot é que, por ela fornecer um produto muito tecnológico, não há muita visibilidade de tudo o que está sendo feito por trás do que aparece.

Outro ponto é a linguagem com o cliente final, ele tem que entender. A user experience do caminhoneiro precisa ser simples. É necessário também pensar na construção de um app com uma boa UX, amigável, intuitivo, etc.

3. O quão valioso são os insights obtidos pelos processos de Big Data? Que tipo de informação eles já te forneceram acerca do comportamento do seu público alvo?

A ideia da Foxtrot era, inicialmente, um roteirizador em tempo real (inúmeras fontes de dados de trânsito, mapa, tempo, sentido da via). O processo de distribuição era engessado antes: predefinido, ignorando fatores externos que mudam completamente a dinâmica do trânsito. A Foxtrot então pensou: como podemos otimizar a rota do motorista então? A grande parte dos clientes da Foxtrot é de comidas e bebidas.

Fizemos contratos para consumir algumas bases de dados: APIs da google, Inrix (mapas e sentido de vias), waze em caso de licença, etc. Um dos desafios é que empresas com base de velocidade de via utilizam crowd sourcing – veículos não adequados para o tipo de público alvo da empresa.

O que precisa ser melhorado no produto Foxtrot: fazer com que o tempo de chegada seja mais realista e, ao mesmo tempo, oferecer uma experiência mais legal para o cliente final, por exemplo, avisando-que a entrega está chegando, proporcionando um acompanhamento, pois gera valor para ele também. Que nem sabemos que o motorista do Uber está chegando à nossa localização.

O coração da Foxtrot é a seguinte equação: $F(X) = \text{tempo} + \text{distância} + \text{pontualidade}$ (pesos diferentes, algoritmo também pode ser por demanda) -> exemplo de empresa sabin, que coleta excretos humanos. O mais importante para ela é a pontualidade.

Alguns dos algoritmos da empresa referem-se à:

- No começo só tinha uma base externa (weather channel, google, waze, open street maps). Este algoritmo então dizia qual base de dados seria melhor usar, qual seria a mais ideal
- Indicava como montar a rota ideal, levando em consideração três fatores principais: menor custo, melhor rota (distância, tempo – prever horário, pontualidade) – respeitar restrições de horário dos pontos de venda

Um grande gargalo encontrado foi o processo normal das empresas “vendedor chega, faz cadastro do cliente, joga no Google com latitude e longitude e erra 10 metros.” Isso gera uma perda enorme de tempo na distribuição

O principal gargalo encontrado então por parte dos clientes foi erro em base de dados.

Há um outro algoritmo também de horário de entrega -> o aplicativo fica capturando dados do motorista o tempo inteiro (quanto tempo parado, andando, celular ligado ou desligado, etc). Mas somente em horário de trabalho para garantir a privacidade dos motoristas.

Pegamos por exemplo então o tempo que o motorista levou para entregar as cervejas. Num horário X, foram 30 minutos. Em outro horário y, foram 3 horas. Sabe-se onde está localizado, onde recebe. Assim, ocorre uma identificação de padrão de recebimento (esse local é melhor entregarmos entre 12h e 14h por exemplo, apesar de seu horário de funcionamento ser de 9h às 18h).

A Foxtrot então passou a gerar sua própria base de dados -> geram valor para os clientes então em termos de pontualidade

- Curva de aprendizado de horários de entrega com sucesso

- Input para roteirização (algoritmo de sequência de roteirização) -> é preciso também que a rota seja viável.

Possuímos os seguintes modelos de Machine Learning

- Machine Learning 1 (Padrão de Recebimento)

O mais valioso é pegar vários tipos de informação e trabalhar com elas (aprendizado de máquina nesse caso por meio do histórico de vendas). O mais valioso na Foxtrot é conseguir não ter apenas um algoritmo, mas ter a integração de vários conversando entre si Para solucionar isso, a equipe da Foxtrot encontrou uma forma de corrigir os problemas de base de dados das empresas.

- Machine Learning 2 (Tempo de Parada)

Ponto de parada -> interpretação

Motorista informa no aplicativo “Parei” -> mas a Foxtrot viu que o motorista deu o aviso de que entregou com o veículo em movimento – então descarta. Foi criado então um barramento de informações erradas para que elas não sujassem o aprendizado da máquina. Este tipo de comportamento de motoristas pode acontecer e comprometer a base de dados.

Ações e comportamentos do motorista que eles detectam:

- Entregou ou não (sinalização do motorista);
- Tempo de atendimento do motorista naquele cliente (a foxtrot tem que conhecer o cliente – premissa de que o motorista está parado e verificação se ele está na cerca de georreferenciamento)
- A cerca contempla imprevistos – cruza com o aviso de entregue por parte do motorista, se está perto então faz sentido

A Foxtrot, então, captura todos os dados do motorista por meio de app e cruza com o “ok” do motorista de que entregou

O algoritmo pega a localização das entregas/ quanto mais visitas o cliente faz, maior a precisão do endereço

- Machine Learning 3 (tempo de atendimento)

Quanto tempo o motorista vai demorar para entregar? – Aprendendo o tempo do cliente, variável de acordo com a quantidade dos produtos.

Todo algoritmo começa com média e mediana (base histórica). Clientes mandam a base de produto. É importante considerar os horários em que os estabelecimentos abrem

4. Como você acredita que as previsões e predições obtidas por meio do Big Data afetam as estratégias de marketing, vendas e relacionamento com o consumidor ou cliente (B2B)?

Feature de desenvolvimento de forecast de demanda para o vendedor é o que os próprios clientes pediram. Há clientes que pedem do início ao fim do processo, e só um pedaço do processo já é extenso. A parte de vendas preditivas possui oportunidades de venda muito altas. Foco em grandes empresas que sofrem com esse gargalo.

80% das empresas de parceria da Foxtrot querem uma solução na linha de venda preditiva para o cliente final, que na verdade não seja só computador, mas que esteja na mão do vendedor.

Todo mundo quer a experiência Rappi, Uber, Loggi. Ligar o cliente final para que ele tenha a experiência B2C. Experiência legal, compras simples, facilitação do processo das empresas. É necessário fazer com que o cara tenha visibilidade na entrega, definir horário de atendimento, etc.

5. Como a empresa vê a questão ética nos processos de Big Data que visam influenciar o comportamento do consumidor? Como isso é tratado pela organização?

A preocupação atual são os dados do motorista (hoje em dia usa-se a localização, tempo utilizando app, movimentação).

A empresa não quer pegar dados depois do horário de trabalho, por ética com o motorista e própria empresa. Também há riscos de processos, a empresa precisa se proteger de possíveis processos judiciais envolvendo clientes B2B. Assim, a política da Foxtrot é bem clara nesse sentido.

Visão de futuro: Sugestão de restaurantes que estão próximos do motorista: aumentar a experiência dele, e somente em seu horário de trabalho.

Se o motorista não faz o processo correto, a empresa pega muito dado dele.

Em relação aos clientes, há um contrato sobre compartilhamento de dados entre empresas concorrentes. Há Ids únicos para cada uma. As informações são paralelas para as empresas, mesmo em casos que a Foxtrot poderia atuar de forma benéfica para as duas. No caso de, por exemplo, dois caminhões indo ao mesmo tempo para o mesmo estabelecimento, a Foxtrot sabe disso, mas não pode fazer nada. Há muita sobreposição de clientes em estabelecimentos também, especialmente em mercados. Possuímos grande base de dados dos clientes finais, o que tem muito valor.

APÊNDICE C – Entrevista com André Ferraz, CEO e Cofundador da In Loco

1. Quais setores da organização possuem maior interface com os processos de Big Data utilizados?

R: Atualmente, a In Loco processa mais de 16 terabytes de dados por dia. É um número enorme e, por isso, desde o nosso início da In Loco em 2011, ainda na faculdade, temos um compromisso com a privacidade. Ainda quando não era nem discutida aqui no Brasil. Nossa missão é resolver o problema da privacidade na internet e mostrar que é possível oferecer tecnologia e conveniência aliada a privacidade de dados. No cenário atual, um Big Data devidamente configurado e utilizado, dentro de um departamento de Análise e Inteligência de Dados, é capaz de desenvolver estratégias extremamente eficazes e implementar melhorias de maneira inteligente e otimizada para empresas. Como a privacidade é o cerne do nosso negócio, os nossos times de Business Intelligence têm maior interface com os Big

Datas. Atualmente, temos uma equipe de quase 20 analistas de dados que possuem grande interação com os Big Datas. Além do time de insights, temos times de segurança, anonimização e privacidade estão em constante contato com os dados para que essas informações estejam criptografadas e anonimizadas para que jamais seja possível chegar a dados identificáveis ou sensíveis dos mais de 60 milhões de dispositivos que possuem nossa tecnologia embarcada.

2. Quais foram os principais desafios técnicos para implementação dos processos de Big Data na organização? Eles foram completamente superados?

R: Há dois anos, era muito comum ouvir as pessoas falarem que big data era a profissão do século. Acho que esse discurso foi repaginado para Inteligência Artificial (AI). Naquela época não havia uma formação formal voltada para os profissionais de Big Data. Hoje, também não temos uma que seja específica para AI.

O desafio é encontrar profissionais que lidem com um volume enorme de dados, tenham familiaridade com estatística, entendam de computação e tenham a capacidade de automatizar tarefas. Ou seja, a gente está falando de um perfil que mescla diferentes áreas do conhecimento humano e, conseqüentemente, as empresas precisam treinar esses profissionais.

Como são profissões novas e não há formações completas para isso, há um gap em encontrar profissionais. Porém, na In Loco temos analistas de dados incríveis. Muita gente que começou conosco ainda na universidade e que hoje são especialistas no assunto.

3. O quão valiosos são os insights obtidos pelos processos de Big Data? Que tipo de informação eles já te forneceram acerca do comportamento do seu público alvo?

R: Entender o comportamento do consumidor e impactá-lo no momento certo e com contexto é uma das formas que obtemos insights valiosos dos nossos usuários e no qual entregamos valor para ele. Por exemplo, imagine um consumidor que está desembarcando em um aeroporto tarde da noite e precisa ir para casa ou hotel. Neste exato momento, ele recebe uma oferta de desconto de um aplicativo de transporte. Isso não faz toda a diferença para ele? Além disso, conseguimos entender o comportamento dos seus consumidores e engajá-los com notificações push no momento correto - uma empresa pode enviar um cupom de desconto ou ativar o consumidor enquanto ele visita o concorrente. Porém, o maior caso de uso dos nossos da nossa tecnologia é quando facilitamos o processo de abertura de contas digitais, apenas com a informação da localização, sem necessidade de comprovante de endereço, sem necessidade de CPF, RG ou informações que consideramos sensíveis.

Nossa inteligência de localização usa dados anonimizados em soluções de negócio para influenciar fluxo em estabelecimentos e aumentar a relevância e usabilidade dos

apps de grandes marcas por seus usuários. Os dados coletados são criptografados e agrupados em clusters de padrões de comportamento. Estes dados podem ser usados, então, para impactar grupos de consumidores com mensagens mais relevantes para seu contexto.

4. Como você acredita que as previsões e predições obtidas por meio do Big Data afetam as estratégias de marketing, vendas e relacionamento com o consumidor ou cliente (B2B)?

R: É impossível imaginar um mundo sem uso de dados para personalização de serviços e entrega de produtos daqui pra frente. Entender a jornada das pessoas no mundo offline, possibilita que as empresas traçam estratégias mais assertivas, trabalhando o marketing omnichannel para impactar o consumidor certo, com mensagem relevante, no canal certo, dentro do contexto ideal. E o resultado é a fidelização, fluidez e entrega de conveniência para as pessoas e a mudança imediata no ponteiro de vendas das empresas.

5. Como a empresa vê a questão ética nos processos de Big Data que visam influenciar o comportamento do consumidor? Como isso é tratado pela organização?

R: Sob a perspectiva de oferta de valor, buscar insights a partir do conhecimento sobre o público é essencial. Entender o que o consumidor precisa para envelopar um produto é primordial para a saúde de qualquer negócio.

Há várias formas de buscar insights. O uso de dados como ferramenta para auxiliar a tomada de decisão é uma delas. Para facilitar a compreensão, vamos imaginar que um aplicativo de delivery identifica que o consumidor pede uma pizza sempre que chega em casa. Nesse caso, a Inteligência Artificial (AI) traz o insight identificando o hábito do consumidor. E, a partir dessa informação, é possível automatizar um processo de fidelização desse cliente. Mas o responsável por tomar essa decisão ainda é o ser humano.

Existe uma parcela de consumidores que é mais ligada ao digital, a geração touch, que já nasceu com acesso às telas touch screen e que entende como o mundo digital funciona e como o processo de captação de dados influencia a personalização de produtos e serviços.

Mas esse movimento ainda não pode ser considerado mainstream, ele está acontecendo aos poucos. As gerações que nasceram antes da ascensão do digital, tem um pouco mais de dificuldade de entender esses processos, mas estão buscando informações para compreender melhor como os seus dados são utilizados no mundo online.

Isso é perceptível quando falamos sobre privacidade. Os consumidores estão ficando cada vez mais conscientes sobre o volume de dados que estão cedendo para acessar

serviços personalizados. Eu acredito que vivenciamos um momento em que é preciso discutir mais sobre privacidade.

APÊNDICE D – Entrevista com Raíssa Arriola, Data Specialist do Twitter

1. Quais setores da organização possuem maior interface com os processos de Big Data utilizados?

Data Ops and Content classification (classificar o que os tweets estão falando para armazenar dados). Ajuda a entender como impactar o usuário e obter insights através do que as pessoas estão falando. Os dados são salvos em categorias, caixinhas. O pessoal de vendas trabalha com Twitter ADs (conversas categorizadas. Produto, usuário final, marketing, vendas (vai diretamente para o ADs)

Um exemplo de algoritmo é aquele que categoriza o que os usuários falaram sobre um jogo por exemplo, ou sobre jogadores. Por meio de uma Landing Page, essas pessoas recebem o convite para entrar na timeline desse assunto.

Também há Identificação de palavras-chave (sistema: semantic core – só utilizamos semântica, atualmente. No futuro, queremos categorizar imagens, vídeos. Hoje em dia utilizamos o @, tudo o que uma conta fala, hashtags, palavras chave)

2. Quais foram os principais desafios técnicos para implementação dos processos de Big Data na organização? Eles foram completamente superados?

É um time comprado de fora, o Twitter compra bastante startup. Ainda há problemas, como por exemplo não trabalhar com vídeos e imagens. Esse é o principal desafio a ser superado.

O Twitter, como muitas empresas, do vale do silício. Tem dificuldade em se internacionalizar. Globalmente não é tão bom. Categorizam muito bem os tweets do USA e indonésia, mas tudo fica muito fechado ao HQ. É preciso se adaptar e crescer.

3. Os quão valiosos são os insights obtidos pelos processos de Big Data? Que tipo de informação eles já te forneceram acerca do comportamento do seu público alvo?

Os dados são armazenados. O time de Data consegue misturar tópicos (todos os dados da conversa do brasileiro, maior menção do jogo, de jogador, etc). Há uns três anos atrás era mais fácil, hoje em dia tudo precisa da aprovação do Legal.

A Questão legal precisa ser destacada – Twitter não consegue mais buscar tudo, antes era muito fácil. Depende muito do CEO da empresa.

Os Dados ADS não dão a localização exata, com base nos últimos três meses o algoritmo só dá uma sugestão. Construção da parte de privacidade estava em

vendas. O caminho veio do time de engenharia com produto. Time de Audience Insights (time de pesquisa), conduz por exemplo estudo sobre millennials.

Plataforma de ADS. Pesquisa por tweet promovido. Em Ad – voluntário, o usuário se dispõe a responder algo para nós. Ajuda internamente, e ajuda externamente. Assim, é possível melhorar a plataforma e melhorar o negócio interno.

As tabelas de dados que usamos são todas aprovadas pelo legal interno.

Surgiu um time que tem o objetivo de entender um determinado tipo de usuário, voltando a interesse. O foco é no tópico das conversas, temas – mas tira perfil do usuário.

4. Como você acredita que as previsões e predições obtidas por meio do Big Data afetam as estratégias de marketing, vendas e relacionamento com o consumidor ou cliente (B2B)?

Trass release. Ajuda muito com os eventos.

O Twitter tem muitas facetas, muitas características – a empresa preza muito pelos movimentos sociais. Conectar as pessoas pelos interesses. Investir nesses dados para que vocês possam ajudar. Emplacar responsabilidade social é muito importante – BGRs (grupos de diversidade – coletivo feminista, LGBTQA, negros) – esses grupos se empoderaram sobre como ajudar dentro do Twitter, e depois dentro de suas comunidades.

O objetivo do Twitter é revigorizar a comunidade em que está inserido

O movimento negro é muito forte no Twitter, política, LGBT – conexão é muito mais fácil. Você segue interesse por interesse, segue quem você quiser.

Hoje em dia, o Twitter é um local de News. Não mais de ver o que o coleguinha está postando. Causa negra é um assunto. Tudo é um assunto. Não é uma interconexão pessoal, é por interesse. Twitter já foi pessoal. E agora tá mais interesse! Não só conta flamengo, mas conversa flamengo. Essa é a novidade. Isso objetiva gerar mais pessoas ativas, conversas e movimento.

Share of Voice (SOV) – Marca mais mencionada na copa do mundo (funciona em eventos – funciona no mercado dos clientes). Fast food mais falado, cerveja mais falada, território que tem muito interesse. O pessoal de pesquisa e stakeholders já ficam preparados quando sabem que vai acontecer um evento.

Um exemplo são as conversas relacionadas a marcas na copa do mundo – coca cola sobre todas as de marca – x%

O grupo de pesquisa entra depois para gerar análises.

Vendas pedem os insights, que é subsídio para as vendas – pedem pelo número de usuários na plataforma – muito forte em conversas e eventos, tendências, etc.

Dados não das pessoas, mas de conversa. E sempre dados de fácil acesso.

5. Como a empresa vê a questão ética nos processos de Big Data que visam influenciar o comportamento do consumidor? Como isso é tratado pela organização?

No Twitter todas as tabelas passam por aprovação do legal, é preciso aprovação para ter acesso. Esse assunto parte muito do posicionamento da empresa. Tabelas existem, dados estão aí. A empresa é que decide o que vai fazer com isso. Qualquer um que tem a skill consegue. A empresa tem que virar e falar “não, isso não condiz com a gente”. Vai muito do CEO.

Com as leis, está ficando mais difícil, está mudando. Há 2, 3 anos atrás qualquer pessoa conseguia fazer. Se você tem habilidade, é muito fácil. Criar linguagem e dar play. E isso é algo muito perigoso.

O Twitter não tem dado de renda, por exemplo. A tabela que existe contém – gênero, idade, localização. Língua se descobre por três palavras no tweet, ou como a pessoa setou no celular.

APÊNDICE E – Entrevista com Samuel Corado, Big Data Strategist da AIS Digital

1. Quais setores da organização possuem maior interface com os processos de Big Data utilizados?

Marketing, Comercial e Relacionamento com o Cliente.

2. Quais foram os principais desafios técnicos para implementação dos processos de Big Data na organização? Eles foram completamente superados?

Conhecimento sobre as novas tecnologias, Convivência dos processos de BI vs Big Data e Limitação sobre os recursos de tecnologia;

Foram superados, mas não na sua totalidade. A convivência entre os processos de BI e Big Data se mantiveram onde exigiu uma maior governança entre esses outros dois mundos.

3. Os quão valiosos são os insights obtidos pelos processos de Big Data? Que tipo de informação eles já te forneceram acerca do comportamento do seu público alvo?

Foi de grande valia, primeiro, trazendo informações antes não apresentadas pelas limitações da tecnologia e grande volume de dados. Visão 360 do cliente, possibilitando a estratégia de Customer Centric.

4. Como você acredita que as previsões e predições obtidas por meio do Big Data afetam as estratégias de marketing, vendas e relacionamento com o consumidor ou cliente (B2B)?

Primeiro, pela velocidade em que a informação chega para a tomada de decisão e a capacidade de processamento para grandes volumes de dados. Segundo, pela capacidade em cruzar informações estruturadas e não estruturadas gerando insights de forma rápida. Terceiro, por democratizar o acesso à informação, fazendo com que outros departamentos participe do processo exploratório de dados.

5. Como a empresa vê a questão ética nos processos de Big Data que visam influenciar o comportamento do consumidor? Como isso é tratado pela organização?

Ao nosso ver, as empresas que partem para uma estratégia de Bigdata, são orientadas (na sua maioria) por desenhos de jornada, pesquisa de campo e outras iniciativas do digitais, o que tornam as questões de éticas, mas simples de serem tratadas. Em todos os casos que participamos, o objetivo era conhecer melhor o consumidor/cliente para melhor atendê-lo e não para ações que possam influenciar de forma intrusiva o comportamento do consumidor.

APÊNDICE F – Questionário aplicado para estudantes da Faculdade de Comunicação (FAC) da Universidade de Brasília (UnB)

Percepção de estudantes de Comunicação da UnB sobre processos de Big Data

Este formulário tem como objetivo compreender o nível de familiaridade dos alunos de Comunicação da Universidade de Brasília (UnB) com o tema de processos de Big Data, bem como identificar a percepção dos mesmos em relação ao aspecto ético desta utilização e da concessão de dados para aprimorar serviços de vendas.

É válido lembrar que as respostas desta pesquisa são confidenciais.

**Obrigatório*

1. Qual é a sua habilitação em Comunicação Social na UnB? *

Marcar apenas uma oval.

- Comunicação Organizacional
- Publicidade
- Jornalismo
- Audiovisual

2. Marque se você já utilizou algum dos serviços abaixo (quantos forem necessários). *

Marque todas que se aplicam.

- Netflix
- Amazon
- Google
- Facebook
- Apple
- Microsoft
- Waze
- Alibaba
- Airbnb
- Uber

3. Quanto à minha compreensão sobre o que são processos de Big Data: *

Marcar apenas uma oval.

- Compreendo muito bem
- Compreendo razoavelmente
- Compreendo superficialmente
- Não compreendo

4. Quanto à presença dos processos de Big Data no meu cotidiano: *

Marcar apenas uma oval.

- Não estão presentes, pois não uso tecnologias deste tipo
- Estão bastante presentes em diversas ferramentas nas quais sou usuário(a)
- Estão pouco presentes, somente em algumas ferramentas que utilizo
- Não sei responder

5. O assunto processos de Big Data já foi abordado em alguma matéria que você cursou na UnB? *

Marcar apenas uma oval.

- Sim, é sempre abordado
- Sim, em algumas disciplinas específicas
- Muito raramente/apenas uma vez
- Nunca em uma disciplina, apenas em outros momentos universitários
- Não, nunca

6. Em sua opinião, os dados atualmente são extraídos por empresas de tecnologia de forma ética? *

Marcar apenas uma oval.

- Acredito que os dados são extraídos de forma ética
- Não tenho uma opinião formada
- Acredito que os dados não são extraídos de forma ética

7. Com qual frequência você costuma ceder seu CPF para obter descontos ou algum tipo de vantagem no consumo de produtos e serviços? *

Marcar apenas uma oval.

- Sempre
- Às vezes
- Raramente
- Nunca

8. Quando você instala algum aplicativo no seu celular, você costuma ler os termos de privacidade? *

Marcar apenas uma oval.

- Sempre
- Às vezes
- Raramente
- Nunca

9. Quando alguma empresa utiliza seus dados de navegação para te ofertar seus serviços de forma mais personalizada você considera isso: *

Marcar apenas uma oval.

- Muito aceitável. Gosto que me ofereçam serviços de forma personalizada

- Aceitável. Os serviços oferecidos de forma personalizada às vezes me ajudam
- Indiferente
- Inaceitável, mas não tomo atitudes para prevenir isso
- Inaceitável, e tomo atitudes para prevenir isso