

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO FACULDADE DE  
ADMINISTRAÇÃO E CIÊNCIAS CONTÁBEIS - FACC

MARIANA HERMES DA FONSECA DE LOSSIO E SEIBLITZ

**A IMPORTÂNCIA DO *BIG DATA* EM FINANÇAS E CONTABILIDADE**

Rio de Janeiro

2021

**Mariana Hermes da Fonseca de Lossio e Seiblit**

**A IMPORTÂNCIA DO *BIG DATA* EM FINANÇAS E CONTABILIDADE**

Trabalho de conclusão de curso apresentado como requisito para a obtenção do título de Bacharel em Ciências Contábeis pela Faculdade de Administração e Ciências Contábeis da Universidade Federal do Rio de Janeiro.

Orientadora: Prof. Marcia Revoredo

Rio de Janeiro

2021

**Mariana Hermes da Fonseca de Lossio e Seiblit**

**A IMPORTÂNCIA DO *BIG DATA* EM FINANÇAS E CONTABILIDADE**

Trabalho de conclusão de curso apresentado como requisito para a obtenção do título de Bacharel em Ciências Contábeis pela Faculdade de Administração e Ciências Contábeis da Universidade Federal do Rio de Janeiro.

---

Marcia Revoredo

---

---

**Rio de Janeiro, 22 de Fevereiro de 2021**

## RESUMO

Com a identificação do enorme volume de dados produzidos por máquinas e seres humanos nos últimos anos, com inúmeras transações financeiras feitas na internet, além das informações geradas pelos dispositivos conectados pela internet das coisas (conectividade e interação entre diversos dispositivos e sensores conectados pela internet), se torna altamente interessante o estudo de como essa nova era tecnológica afeta o mundo de contabilidade e finanças. O estudo encontra evidências de que o *Big Data* será ainda mais importante para o mundo financeiro, trazendo uma perspectiva de novos horizontes para análises mais sofisticadas e eficientes de transações e lançamentos contábeis, riscos financeiros e operacionais, de forma quase simultânea com a ocorrência de atos e fatos contábeis, permitindo uma possível auditoria contínua, considerando também o desafio de lidar com a questão da segurança e privacidade dos dados manipulados.

**Palavras-chave:** *Big Data*. Contabilidade. Finanças. *Analytics*.

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	6
<b>2. OBJETIVOS E JUSTIFICATIVA</b> .....	8
<b>3. REVISÃO DA LITERATURA</b> .....	9
3.1 EXPLICANDO <i>BIG DATA</i> .....	9
3.2 DE ONDE VÊM TANTOS DADOS .....	14
3.3 USOS DO <i>BIG DATA</i> .....	15
3.4 FERRAMENTAS E TECNOLOGIAS DO <i>BIG DATA</i> .....	17
3.5 PRIVACIDADE E SEGURANÇA .....	19
3.6 RELAÇÃO DE <i>BIG DATA</i> COM CONTABILIDADE E FINANÇAS.....	22
<b>4. METODOLOGIA</b> .....	26
<b>5. ESTUDOS DE CASO</b> .....	27
5.1 CONTABILIDADE FORENSE .....	27
5.2 <i>BIG DATA</i> NA AUDITORIA GOVERNAMENTAL .....	29
<b>6. CONSIDERAÇÕES FINAIS</b> .....	31
<b>7. REFERÊNCIAS</b> .....	32

## 1. Introdução:

Até 2018 tínhamos 33 zettabytes<sup>1</sup> de dados no mundo digital, porém é esperado que até 2025 isso chegue a 175 zettabytes (REINSEL; GANTZ; RYDNING, 2018). Isso é, em suma, o *Big Data*, uma gigante quantidade de dados disponíveis produzidos não só pelo homem, como por sistemas automatizados e sensores (TAURION, 2013).

São geradas novas informações a todo momento por cada pessoa por meio de redes sociais, em formato de áudio, texto, vídeo, fotos, entre outros, também por meio de sensores, como relógios inteligentes, que registram dados de batimentos cardíacos, passos e distâncias percorridas, qualidade do sono, além de pesquisas feitas por pessoas em ferramentas de busca. Esses dados, podem ser extremamente úteis tanto para a vida pessoal de indivíduos tentando entender melhor sobre si, quanto para empresas buscando visões de mercado e suas tendências, preferência de clientes, entre outros elementos que podem auxiliar decisões estratégicas, ajudando a melhorar a performance de vendas e rentabilidade das mesmas com a aplicação de ações que tornem as atividades das companhias mais eficientes e eficazes.

O grande desafio observado na era da abundância de dados possibilitados por *Cloud Computing* (computação com armazenamento de dados, redes e aplicações em ambientes compartilhados de maneira integrada através da internet), pela rapidez da geração de novas informações em qualquer hora e lugar do mundo por dispositivos móveis, televisões *smart* (capazes de acessar a internet e de rodar aplicativos de terceiros), e computadores cada vez mais poderosos, é a capacidade de analisar tantos elementos ao mesmo tempo, de modo a torná-los úteis para seus usuários e grupos de interesse. Dada dificuldade cria a necessidade de novas ferramentas tecnológicas capazes de organizar, estruturar e prover percepções, assim como a de profissionais capazes de lidar com essas tecnologias e auxiliar na tomada de decisão.

---

<sup>1</sup> 1 Byte equivale a 8 bits, 1 Kilobyte a 1024 bytes, 1 Megabyte a 1024 kilobytes, 1 Gigabyte a 1024 megabytes, 1 Terabyte a 1024 gigabytes, 1 Petabyte a 1024 terabytes, 1 Exabyte a 1024 petabytes, 1 Zettabyte a 1024 exabytes e 1 Yottabyte a 1024 zettabytes. Alecrim, 2007.

Nesse cenário surgem os profissionais especializados em lidar com o *Big Data*, chamados de *Data Scientists* (cientistas de dados), capazes de gerar valor a partir de dados, que passa a ser considerada agora como uma nova moeda de troca (TAURION, 2016). Esse novo segmento procura auxiliar na organização e filtragem de dados estruturados (padronizados em formato, como em um banco de dados) e não estruturados (mais flexíveis, disponíveis em diversos formatos e disposições), pois um grande gasto para as organizações atualmente é a má qualidade de informações disponíveis.

No novo âmbito da criação do profissional cientista de dados, se torna importante não só o foco na área de TI (Tecnologia da Informação), como também na combinação da função do especialista em tecnologia com o financeiro e contábil, que é uma área já com grande quantidade de dados, com acelerado crescimento e que pode se beneficiar de maneira a aproveitar ferramentas de análise para ampliar o escopo dos estudos realizados por seres humanos atualmente. Assim, o chamado *Analytics* (Inteligência Analítica) pode ser essencial para a melhoria da qualidade de informação vinda de contadores e especialistas de finanças, pois pode auxiliar em análises mais profundas sobre as demonstrações financeiras, dados do mercado e internos da companhia, como pode também criar modelos preditivos capazes de indicar tomadas de decisão pelos gestores de forma mais certa.

Deste modo, o profissional de contabilidade e finanças pode mudar seu papel principal de gerador de informação para analista, o que será feito em uma velocidade nunca vista antes. Isto com a ajuda de profissionais especializados em lidar com *softwares* (programas computacionais) capazes de processar dados e gerar *insights* (compreensão humana intuitiva precisa e profunda de uma pessoa ou coisa) com mais qualidade e velocidade. Porém, é determinante que se tenha cautela em relação aos dados considerados para análise sob o quesito de privacidade e segurança, algo que pode ser muito vulnerável na internet.

Áreas que provavelmente sofrerão drásticas mudanças no novo âmbito tecnológico são a contabilidade gerencial, auditoria e contabilidade financeira, sendo estas as quais o profissional contábil poderá se especializar em gerar ainda mais valor para os tomadores de decisão (CEOLATO, 2019). Porém, para que mudanças ocorram na área contábil, ainda serão necessários grandes

investimentos em habilidades técnicas para os atuantes do setor para que possam trabalhar com novas ferramentas e tipos de dados. Isto se dá devido ao foco inicial de esforços em ferramentas e desenvolvimento de habilidades para áreas relacionadas diretamente aos consumidores, e não de forma interna nas companhias (GOES apud COCKROFT e RUSSEL, 2018).

Assim, o estudo conta com a base teórica para situar os conhecimentos básicos sobre *Big Data* para que sejam considerados no ambiente contábil e de finanças, para por fim, serem vistos estudos de caso para entendermos os usos práticos dos novos recursos, ainda em estágio embrionário.

## **2. Objetivo e Justificativa:**

Com o acelerado crescimento da geração de dados por diversas fontes espalhadas pela internet, mais de 2.5 quintilhões de bytes de dados diários criados (MARR, 2018), o equivalente a aproximadamente 215 milhões de vezes o filme Titanic (1997) na resolução de 1080p HD, ferramentas e técnicas de análise de dados inovadoras estão surgindo para se dar conta da quantidade de dados estruturados e não estruturados existente.

O presente estudo pretende identificar os usos das novas tecnologias que surgiram com o advento do fenômeno do *Big Data* e *Analytics* e como essa onda de inovação impactará as áreas de finanças e contabilidade, tendo como base literaturas e estudos de casos que nos trazem uma visão ou uma expectativa prática mais apurada do que nos espera em um futuro próximo, procurando observar também, como consequência, as mudanças as quais o profissional da área obrigatoriamente deverá absorver.

O Tema se mostra interessante por ser um ponto que pode mudar drasticamente o trabalho do contador e especialista financeiro, pois há um potencial em transformar seu foco em analisar informações ao invés de produzi-las.

Até o presente momento não foram encontrados muitos estudos específicos relacionando *Big Data* e *Analytics* com finanças e contabilidade, sendo eles, do ponto de vista teórico, os de Bhimani e Willcocks (2014), de Cockroft e Russel (2018), de Shi, Sun e Zhang (2019), de Lindell, (2018), e de Neves (2019), além dos estudos em casos de uso prático de Costa e Dutra



(2014), de Lima, Serrano e Cupertino (2020). Deste modo, é possível identificar a importância do início da pesquisa sobre o tema, principalmente no Brasil, onde se encontra em um estágio primordialmente inicial.

### 3. Revisão da Literatura:

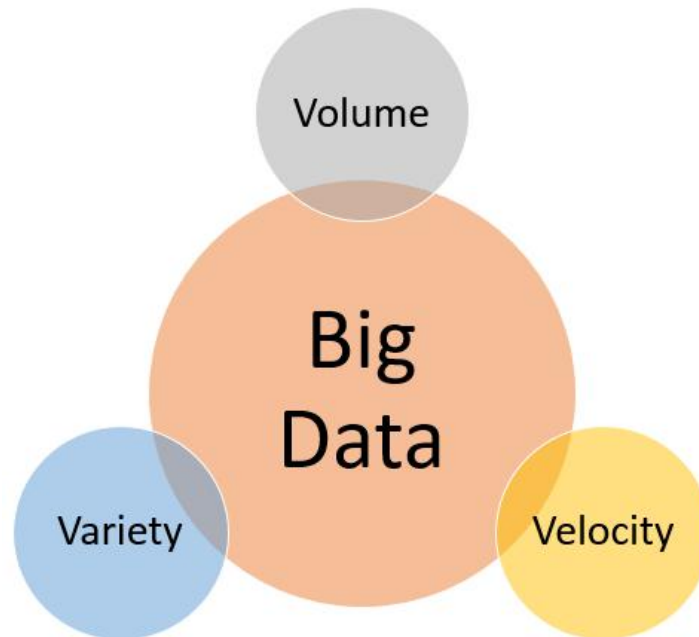
#### 3.1 Explicando *Big Data*:

O termo *Big Data* foi utilizado uma das primeiras vezes em 1997, em um cenário onde se discutia a falta de capacidade de hardwares e softwares de processarem e armazenarem a grande quantidade de dados que estavam surgindo durante o período de disseminação do uso da internet (COX; ELLSWORTH, 1997). Porém, com o desenvolvimento tecnológico e computação em nuvem, atualmente estes não são considerados mais problemas de armazenamento, mas sim de processamento de uma quantidade de dados imensamente maior a cada dia, exigindo uma análise e respostas em tempo real (TAURION, 2013).

Para Taurion (2013, p. 29), muitos dados não eram considerados e podem ser utilizados para melhoria de processos e tomada de decisão, e assim, "o que o microscópio foi para a medicina e a sociedade, o *Big Data* também será para as empresas e a própria sociedade", e seria uma evolução da antiga gestão por Excel, com mais complexidade e rapidez. Para ele, *Big Data*:

Não é apenas um produto de software ou hardware, mas um conjunto de tecnologias, processos e práticas que permitem às empresas analisarem dados a que antes não tinham acesso e tomar decisões ou mesmo gerenciar atividades de forma muito mais eficiente (TAURION, 2013, p.32).

Frente a esse cenário, existem diversas definições do que é *Big Data*. Segundo o Gartner (2020), é uma intensa variedade de ativos informacionais, em alta velocidade e volume, que exigem formas inovadoras de analisar informações. Desse modo, os componentes volume, velocidade e a variedade, que afetam o âmbito dos dados, são conhecidos como os 3 Vs do *Big Data*, e foram introduzidos pelo analista da Gartner, Doug Laney em 2001 (AHMED; PATGIRI, 2016).

Figura 1: 3Vs do *Big Data*

Fonte: Ahmed; Patgiri, 2016. Figura Produzida pela autora.

Para a Oracle (2020), são informações de um conjunto mais denso e complexo, vindos de fontes diferentes, o que exige novas opções de software para se trabalhar, porém abre a possibilidade de resolver problemas das empresas de uma forma nunca possível antes. Tal conjunto de dados estruturados (mais organizados e padronizados) e não estruturados (mais dispersos e com diferentes formatos como texto, vídeo e fala), segundo a IBM (2020), não se encontra nas limitações dos bancos de dados relacionais tradicionais, e sua análise auxilia os tomadores de decisão das empresas a tomá-las com mais rapidez e com maior escopo do que antes. Tudo isso se torna possível com avançadas análises de texto, *Machine Learning* (aprendizado de máquina, ramo da inteligência artificial baseado na ideia de que sistemas podem aprender com dados, identificar padrões e tomar decisões com o mínimo de intervenção humana), análise preditiva, mineração de dados, estatística e processamento de linguagem natural, de forma a criar uma percepção de negócios diferenciada.

Taurion (2013) traz uma nova abordagem dos 3Vs do *Big Data* de Laney, introduzindo mais dois elementos para esta definição, e resumindo em uma fórmula que soma o volume, velocidade, variedade e veracidade, de uma forma que se agrega valor (Figura 2).

Figura 2: 5Vs do *Big Data*



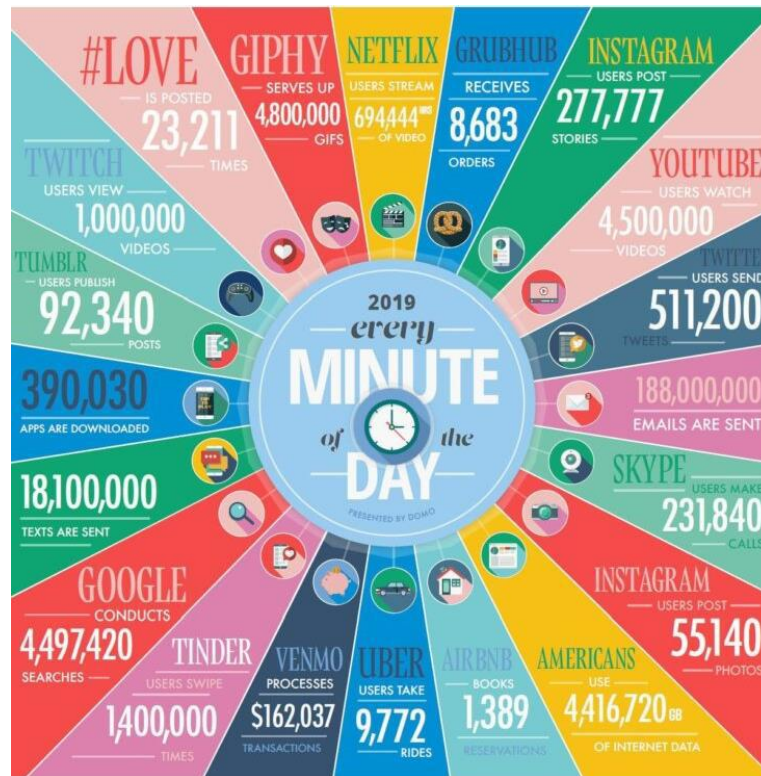
Fonte: Taurion, 2013. Figura produzida pela autora.

A variedade se dá devido as diversas formas de dados disponíveis, podendo eles serem estruturados, semiestruturados ou não estruturados, advindos de redes sociais, e-mails, documentos eletrônicos, em vídeos ou escritos, de sensores voltados para humanos ou ambientes, entre outros. O bom uso dessas informações com tantas diferenças estruturais e as vezes pouca relação entre si, é que a sua combinação pode trazer visões preditivas trazendo muito mais eficiência para o usuário (TAURION, 2013) quando são tratados previamente para que gerem significado (ORACLE, 2020).

O volume é um dos mais evidentes, com as grandes quantidades de dados gerados todos os dias a partir de múltiplas fontes e plataformas diferentes, e é estimado que elas dobrem a cada ano e meio (TAURION, 2013). Tais plataformas que podem ser de multimídia, redes sociais e a internet das coisas, são consideradas grandes fatores de um crescimento exponencial de dados a serem observados no futuro (MCKINSEY GLOBAL INSTITUTE, 2011).

Na figura abaixo, podemos observar a quantidade de dados gerados apenas em cada minuto de 2019 em diversas plataformas da internet.

Figura 3: Dados gerados por minuto.



Fonte: Domo (2019)

Outro aspecto é a velocidade, que se dá devido a necessidade observada de ação praticamente em tempo real (TAURION, 2013). Essa rápida resposta analítica exigida pode ser desafiadora, porém quando feita da forma correta, com auxílio das ferramentas adequadas, pode trazer lucro para a área de finanças e até salvar vidas quando aplicada no setor da saúde, por exemplo (GUO; YU, 2016).

Um importante ponto que se tornou uma preocupação é a veracidade das informações coletadas. Em uma era em que se tornou abrangente a criação das *Fake News* (notícias falsas), se torna essencial que seja garantida a confiabilidade e autenticidade dos dados (TAURION, 2013). A má qualidade de dados produzidos ou capturados é um grande fator na diminuição da vantagem competitiva das organizações, atingindo companhias em uma média de 15 milhões de dólares estadunidenses de custo financeiro ao ano (GARTNER, 2018).

Por fim, todos os dados utilizados por uma organização devem ter valor, para que a empresa que esteja investindo em *Big Data* tenha um retorno sobre

o capital investido (TAURION, 2013). Até 2017, apenas 43% do setor público extraía valor de seus dados, 33% do setor automotivo, e apenas 26% e 13% quando se trata dos setores de serviços financeiros e hospitalidade e lazer (PWC, 2018), Esse ambiente pode se transformar para níveis de eficiência mais elevados com um melhor aproveitamento do *Big Data Analytics* acompanhado pelo uso da Inteligência Artificial (IA). Isto é reforçado com a visão de especialistas que a IA irá auxiliar pessoas a fazerem melhores decisões de forma mais rápida e até mais barata (PWC, 2016a).

Para Davenport (2014), o *Big Data* é representado em suma por uma quantidade muito grande de dados para caberem em apenas um servidor, muito desestruturados para *databases* relacionais e que se movem com muita velocidade para as tradicionais *data warehouses*, dificultando muito as tentativas de controle e gerenciamento, o que só pode ser feito por meio de novas tecnologias que estão surgindo e manipuladas por cientistas de dados.

Segundo a SAS (2020a, não paginado), "*Big Data* não gira em torno da quantidade de dados que você tem, mas do que você faz com eles", e com sua análise é possível "reduzir custos; economizar tempo; desenvolver novos produtos e otimizar ofertas". Essa quantidade de dados tratados pode, por meio de técnicas avançadas, ser uma fonte valiosa para que se descubram fraudes, erros processuais e se calcule risco praticamente em tempo real.

Tais técnicas são conhecidas como *Analytics*, e são o meio pelo qual podemos gerar *insights* que nunca pensamos em sermos capazes de ter, e os principais métodos analíticos utilizados atualmente são a estatística descritiva, análise preditiva e análise prescritiva (SAS, 2020b). Em português a chamada inteligência analítica é capaz de determinar parâmetros de negócio futuros como o preço correto de um item, o melhor local para uma fábrica e os juros que um cliente deve pagar (SAS, 2020b). Tais técnicas ganharam força com o *Big Data*, pois a maioria das ferramentas de mineração de dados ou análises estatísticas costumam ser adaptadas para grandes quantidades de informações, inclusive pois geralmente quanto maior a amostra, melhor é a acuracidade da estatística e da análise (RUSSOM, 2011).

Essas novas possibilidades introduzidas na vida das empresas devido ao *Big Data* e *Analytics* abre aos tomadores de decisão melhores possibilidades, que segundo uma pesquisa da PWC (2016b) sobre os aspectos

de velocidade e sofisticação, foi descoberto que todos querem que as tomadas de decisão sejam mais rápidas, principalmente nos setores bancário, de seguros e saúde, mas entendem que ainda há muito trabalho para ser feito no quesito sofisticação. Com mais sofisticação, significaria que se pode calcular o impacto em receita e lucro quando se simula entradas em novos mercados com seus indicadores ao invés de olhar apenas para o tamanho de suas metas, de forma a gerar muito mais valor (PWC, 2016b).

### **3.2 De onde vêm tantos dados?**

A quantidade de dados gerada diariamente na casa dos petabytes (TAURION, 2013) é evidente, e a necessidade de entender e agregá-los também. Porém, as fontes de dados são extremamente diversas, por isso uma das dificuldades encontradas atualmente é o tratamento desses dados, que podem ser em formato de texto, vídeo, áudio, imagens entre outros, gerados por máquinas cada vez mais automatizadas como sensores e softwares com IA e *Machine Learning*, além dos gerados por humanos, que passam a maior parte de seus dias conectados por meio de tablets, computadores de mesa ou laptops, televisões e relógios inteligentes, smartphones e outros. Um exemplo são as informações criadas por sensores de tráfego que devem ser ligadas aos veículos circulando e seus proprietários (TAURION, 2013).

O notável aumento de dados gerados pode ser explicado também pela criação de informações por fontes não tradicionais de dados sociais e econômicos pesquisados na internet, transações, difusão de dados, e interações sociais e não deliberadas (BLAZQUEZ e DOMENECH, 2017).

Com a perceptível falta de padrão de dados, é notada a necessidade de tratamentos específicos para cada formato existente, podendo ser feita uma separação por categoria e critérios de validação, segurança e aceitação, que variam de acordo com suas origens (TAURION, 2013). Isso demonstra com grande relevância quando percebemos que empresas, por exemplo, não necessitam apenas de dados operacionais advindos de uma base de dados relacional, pois passam a ter a necessidade de utilizar dados desestruturados

como informações de clientes e mídia social de diversas formas diferentes, com dados internos e externos à organização (HURWITZ et al, 2013).

Um dos outros grandes responsáveis pela criação de novos dados foi a Internet das Coisas, com mais dispositivos conectados entre si, criando informações sobre padrões de uso do cliente e desempenho de produtos. E isso tudo está apenas em expansão, pois o horizonte que a computação em nuvem nos traz é realmente inimaginável (ORACLE, 2020).

Todo esse novo volume de dados gerados de tantas origens diferentes trás novos desafios, como a capacidade de sistemas que suportem grandes quantidades de usuários em uma base consistente, como sistemas de *e-commerce* transacional ou aplicativos de *call center* (HURWITZ et al, 2013).

Sob o ponto de vista contábil e financeiro, dados diferentes dos tradicionais *ledgers* (sistemas de registros) advindos do sistema de partidas dobradas criados por Frei Luca Pacioli, em 1494, e sistemas internos de manejo da operação de empresas passaram por uma evolução, com dados que a empresa ainda não utiliza e analisa, podendo gerar uma visão de reporte, previsões e análises capazes de inferir impactos financeiros na organização (LINDELL, 2018).

### **3.3 Usos do *Big Data***

Como forma de possibilitar a materialização dos impactos do *Big Data* na vida real, e demonstrar exemplos que comprovam o funcionamento de propostas oferecidas pelas técnicas de análise e uso de dados de forma a tomar proveito de seu valor, é interessante analisar casos de uso nas organizações já funcionais de forma bem sucedida atualmente.

Um interessante caso de sucesso de uso do *Big Data* é o caso na política da campanha de reeleição do presidente Barak Obama, dos Estados Unidos, quando informações coletadas e analisadas foram utilizadas por grupos de voluntários como forma de orientar as estratégias de cativação de eleitores com respostas em tempo real (MORAES, 2012 apud AMARAL, 2017). A eficiência do atingimento de pessoas se deu pelo bom manejo de dados, onde a comunicação era melhor e permitiu que tempo não fosse desperdiçado em

busca de alvos já perdidos. Um exemplo é quando foi previsto 15 dias antes do primeiro debate de Obama que seu desempenho seria ruim, o que permitiu que fosse bolada uma ação on-line que iria ao ar uma hora após o programa (SLABY, entrevista para a Revista Isto É, 2013 apud AMARAL, 2017).

Outro exemplo de um caso de bom uso do *Big Data* é a Netflix, que por ser focada em dados, toma muitas de suas decisões estratégicas depois de analisá-los, obtendo informações como histórico de transmissão de vídeos e recomendações de terceiros que ficam armazenados em nuvem, dos próprios clientes ou de companhias especializadas, como a Nielsen. Porém, o início de seu interesse em dados advém desde a época na qual era apenas uma empresa de aluguel de DVDs, quando criava recomendações de conteúdo baseadas nas notas de cada vídeo assistido. Seguindo essa cultura de análise de dados, para garantir uma melhor performance frente aos seus concorrentes, a Netflix investiu na produção de conteúdo original a partir do uso de tecnologias *Big Data*. Uma das grandes análises feitas foi para a antiga série da BBC *House of Cards*, onde foi feita uma previsão do sucesso da série por meio do banco de dados de seus clientes, combinado com a popularidade dos atores envolvidos, o que fez com que comprassem as duas primeiras temporadas da série por 100 milhões de dólares. Ademais, ainda foram feitos 10 trailers para serem divulgados para os assinantes de acordo com suas preferências, onde eram mostradas cenas capazes de cativá-los de acordo com seus gêneros preferidos de televisão e cinema (NETFLIX, 2016 apud PEREIRA, 2016).

Outro exemplo, já adentrando uma empresa da área financeira, foi a Serasa Experian, que já focava em dados desde sua fundação, porém começou a incluir o *Big Data* não só no seu portfólio de produtos como na melhoria de seus processos. Nesse caso, a empresa realizou estratégias como o investimento de 7 milhões de reais em tecnologia Hadoop, que é um projeto que oferece uma solução para problemas relacionados à *Big Data*, tendo em seu núcleo duas partes essenciais: o Hadoop Distributed Filesystem (HDFS), que é um sistema de arquivos distribuído e confiável, responsável pelo armazenamento dos dados, e o próprio Hadoop MapReduce, responsável pela análise (TAURION, 2013), com servidores Teradata, pela sua capacidade de análise de dados variados e volumosos, sendo até 10 vezes mais rápida em



relação a métodos tradicionais. Um dos projetos de destaque foi o Tetris, que passou a se chamar "Dados não Estruturados" mais tardiamente, sendo capaz de ir além da análise de dados estruturados, as variáveis comuns da análise de crédito, mas considerando e cruzando também variáveis advindas de sites da internet, blogs, CRM logs (Customer Relationship Management - se refere ao conjunto de práticas, estratégias de negócio e tecnologias focadas no relacionamento com o cliente) e informações públicas de portais de notícias englobando momento e estilo de vida de indivíduos. Essas ações foram possíveis unindo a capacidade de uso de dados não estruturados de fontes externas, com tecnologias Hadoop, Spark (ferramenta de TI para *Big Data* construído com foco em velocidade, facilidade de uso e análises sofisticadas) e a linguagem Python (linguagem de programação de alto nível) para uma infraestrutura de TI com o apoio de especialistas da área de análise de crédito, como matemáticos, estatísticos e físicos (CAMPOS, 2015).

### **3.4 Ferramentas e Tecnologias do *Big Data***

A Era do *Big Data* traz a necessidade de ferramentas avançadas de análise por meio de algoritmos, de forma a unir homem e máquina, permitindo que executivos tomem decisões baseadas em julgamento humano, porém auxiliadas pelo que dizem os dados. Dessa forma, poderão ser tomadas decisões mais rápidas, melhores e mais baratas, além de livre de vieses, pois a tecnologia, diferente de um líder de negócios, não escolhe uma informação baseada em algum ponto de vista específico, nem ignora fatos, o que nos permite enxergarmos a realidade de uma forma diferente (PWC, 2016b).

Para Taurion (2013), os antigos modelos tecnológicos das empresas não previram dados não estruturados e nem para um volume tão alto, e por isso há uma necessidade de foco em *Analytics*, visto que é o fator que transformará o ativo informacional e trará valor para a entidade. Sob essa ótica, as tecnologias para lidar com o *Big Data* estão divididas em duas categorias: as que tratam de *Analytics*, como Hadoop e MapReduce, que é um modelo de programação para rodar no ambiente Hadoop, desenhado para processar grandes volumes de dados em paralelo, dividindo o trabalho em um conjunto de tarefas

independentes, e as de infraestrutura, sendo elas o ambiente onde se tratam as informações, com foco nos bancos de dados NOSQL (banco de dados projetado para pesquisas não relacionais, fundamentais para o *Big Data*), para que os dados possam ser armazenados e processados na escala de petabytes.

Outras tecnologias importantes para a adaptação das empresas ao *Big Data* são: o *In-memory Processing* (nova tecnologia para processamento de dados onde não é mais suficiente o processamento usando pesquisas de dados relacionais e SQL em um universo ínfimo quando comparado as atuais nuvens de informação, e performa *Analytics* realizados dentro do *database*), *Machine Learning*, para rápidas gerações de modelos e testagem, processamento de linguagem natural, *Analytics* feito em nuvem e softwares de *Analytics visual* (DAVENPORT, 2014).

Para Davenport (2014), temos três fases de *Analytics*, sendo a primeira o tradicional, com análises e reportes advindos de dados internos, para decisões internas e lentas, a segunda, já encobrimdo a era do *Big Data*, com dados mais complexos e não estruturados sobre clientes, englobando novas capacidades computacionais e analíticas, com o surgimento dos profissionais especializados nessas tecnologias, os cientistas de dados, com um maior fluxo de transações e negócios ocorrendo online, e por fim a terceira, com uma mistura do *Analytics* tradicional com o *Big Data*, de forma mais integrada ao negócio, com um modelo de entregas e visões de negócio rápido e ágil, de forma a utilizar os dados ao nível de decisão.

Esses tipos de tecnologias trazem perguntas respondidas em ciclos muito mais curtos, pois trabalham a informação antes mesmo dela ser otimizada, relacionada ou organizada de forma racional, isso tudo por meio de análises avançadas que permitirão as empresas que delas aproveitam garantirem sua sustentabilidade no mercado (TAURION, 2013).

Porém, para que tudo isso seja posto em prática ainda existem alguns desafios, como os gerentes de produtos ou internos que não compreendem os produtos de dados e o *Analytics*, além de não concordarem com decisões advindas de um lugar desconhecido, além de trabalhadores de primeira linha que não querem dados e sistemas ensinando a eles como performarem seus trabalhos, e clientes e parceiros que pensam ser donos dos dados (DAVENPORT, 2014). Mesmo com tudo isso, as vantagens do *Big Data*

demonstram uma importante decisão para se ter em mente, considerando os obstáculos e as mudanças na forma de trabalhar, para que as empresas aproveitem da potencial vantagem competitiva advinda dessas novas possibilidades de tomada de decisão.

### **3.5 Privacidade e Segurança**

Com uma maior quantidade de dispositivos conectados à internet, que crescem progressivamente a cada dia, a questão da privacidade e segurança do acesso, armazenamento e uso de dados vem ganhando uma grande importância, inclusive estão sendo criadas leis em vários países voltadas para o tema. Uma grande inspiração para países latino americanos, Austrália e Estados Unidos adotarem leis para tratarem de proteção de dados foi o Regulamento Geral de Proteção de Dados (ou a sigla em inglês GDPR), criado na União Europeia, que está em vigor desde maio de 2018. Tal regulamentação inspirou a criação da Lei Geral de Proteção de Dados (LGPD) no Brasil, que entrou em vigor em 2020 e nos aproxima mais da tendência mundial de termos controles para resguardarmos dados pessoais (SERPRO, 2020a).

Com o costume de pessoas divulgarem amplamente seus dados pessoais em formulários da internet, redes sociais, compras online e outros meios, cresce amplamente a vulnerabilidade dessas informações, que são usadas por empresas para fazer todo o tipo de análise de seus consumidores, de estratégia de marketing até uma análise de crédito, e com isso as empresas começam a sofrer uma pressão para que esses dados sejam muito bem protegidos. Nesse cenário, a LGPD se torna um marco com a padronização de procedimentos e normas, de forma a garantir a segurança jurídica dos dados pessoais e ao mesmo tempo a igualdade de benefícios e competição para o empresariado, e podemos ver suas principais atribuições na imagem abaixo (SERPRO, 2020b).

Figura 4: Infográfico LGPD



Fonte: SERPRO, 2020c.

Tal lei traz diversas mudanças para os donos dos dados em relação ao acesso, correção, eliminação, portabilidade e revogação do consentimento, de forma a empoderar o consumidor pela garantia de indenização em caso de danos. As empresas coletoras de dados deverão seguir os 10 princípios descritos na lei de forma que as informações pessoais adquiridas sejam usadas de forma a atender uma finalidade de negócio válida, utilizando apenas uma quantidade mínima, correta e de qualidade. Essas informações deverão estar mantidas em segurança pela empresa que as coleta, e em determinados casos, qualquer incidente de segurança deverá ser comunicado ao titular, e relatórios de riscos de privacidade podem ser solicitados pelo órgão regulador, o que no caso de descumprimentos detectados, podem ser penalizados com multa (PWC, 2020).

Nesse cenário, a questão ética é muito discutida nos fóruns sobre uso da informação, pois é essencial que sejam definidas a identidade do proprietário de informações e sua privacidade, a propriedade de dados e a reputação originada pelos dados analisados. Isso tudo ainda traz muitas discussões

importantes para que os pontos pendentes sejam compreendidos profundamente, pois já existem citações que consideram os dados pessoais até como um novo ativo econômico (TAURION, 2013).

A preocupação observada na sociedade se dá pelas possibilidades que surgiram nos últimos anos de cruzar dados que sozinhos não teriam tanto significado, porém que combinados trazem comportamentos realizados na internet por pessoas, que podem ser utilizadas em poderosas estratégias de marketing direcionado especificamente para certos indivíduos. Uma possível solução para minimizar esses riscos às informações pessoais são as ferramentas que permitem que as pessoas não sejam identificadas, como a criptografia e o uso de pseudônimos, para que nomes de indivíduos não sejam identificados ou sejam substituídos em bases de dados (BRETERNITZ e SILVA, 2013).

Sob outro ponto de vista, ainda considerando que é necessário um melhor desenvolvimento de discussões técnicas e éticas, devemos também considerar a necessidade de acesso a dados para que sejam de fato úteis para tomadas de decisão, e com isso as organizações passam por desafios da cultura social, que ainda enxergam o compartilhamento de dados como uma perda de poder de barganha ou fraqueza. Para isso, é preciso uma melhor compreensão das entidades sobre o não esgotamento de dados digitais, que na verdade não são uma ameaça, e sim uma oportunidade (ZUPPO et al, 2013 apud CAMPOS, 2015).

Nesse ambiente, a maior preocupação em relação a dados sensíveis é de consumidores em relação aos seus dados de saúde e financeiros, o que pode ser utilizado para o desenvolvimento de produtos personalizados de muita utilidade, além de tratamentos médicos inovadores, porém é evidente que é essencial que seja feito um balanceamento entre privacidade e a utilidade das informações (MCKINSEY GLOBAL INSTITUTE, 2011). Com isso, um dos desafios de fornecedores é garantir que os usuários tenham o acesso limitado e necessário pra criar análises, o que na era do *Big Data* é feito por meio da infraestrutura de nuvem, e seu modelo de segurança deve garantir que o usuário acesse apenas o que precisa, visto que ela é a plataforma de originação e armazenamento das informações (GUO; YU, 2016).

### 3.6 Relação de *Big Data* com Contabilidade e Finanças

Desde o período colonial, o Brasil começou a desenvolver-se no tema da contabilidade, quando era necessário um maior controle da situação patrimonial devido ao crescimento das atividades nas províncias (SILVA, 2015 apud SANT'ANA, 2019), o que também estava relacionado a um crescimento do foco nos gastos públicos. Com isso, foi sendo criado um cenário no qual as escriturações contábeis só pudessem ser realizadas por quem tivesse estudo contábil, conforme foi decretado ainda no período do Brasil Império (BACCI, 2002 apud SANT'ANA, 2019).

Desde então, a contabilidade mudou de apenas um registro de bens, direitos e obrigações patrimoniais e informações com objetivo fiscal para uma poderosa ferramenta de análise de desempenho e planejamento para a tomada de decisão de uma entidade.

E tudo isso, com a popularização do uso de *Big Data* chegando na área de finanças e contabilidade, está prestes a mudar mais uma vez, melhorando seu nível de relevância e qualidade de informação, isso por meio da criação e aprimoramento de normas contábeis a fim de garantir uma informação útil num cenário de economia em tempo real (WARREN ET AL, 2015 apud NEVES, 2019).

Seguindo o avanço das normas contábeis, cada vez mais há uma constante busca por parte dos profissionais da área por formas de automatizar processos repetitivos já consolidados há um tempo considerável e focarem em análises do significado dos números. Ceolato (2019) encontrou pesquisas sugerindo que com as novas tecnologias e sistemas, o tempo antes gasto com tarefas manuais pode ser utilizado para outras que exigem análise de dados, pensamento estratégico e tomada de decisão, de forma a gerarem valor para os gestores da empresa, e grandes aliadas dessas mudanças será a inteligência artificial, *Machine Learning* e técnicas de *Big Data*. O estudo também sugere que as áreas que mais sofrerão alterações serão as de contabilidade gerencial, auditoria e contabilidade financeira com as mudanças tecnológicas no horizonte.

Nesse cenário, podemos observar também que mesmo com a contabilidade ainda em déficit de habilidades técnicas desenvolvidas para funcionar com *Big Data* (GOES apud COCKROFT e RUSSEL, 2018), e muitas empresas de serviços financeiros ainda estarem focadas em *Big Data* apenas para melhorias em relação ao consumidor, o tema ainda é de grande interesse do mundo financeiro e da contabilidade. A internet das coisas (ou a sigla em inglês, IOT) será também um aliado nessa jornada para o uso de dados financeiros, principalmente em um momento onde os smartphones estão substituindo cartões de crédito, e os aplicativos ao redor dessa tecnologia, tendo assim um potencial de revolucionar a indústria financeira (TURBAN ET AL apud COCKROFT e RUSSEL, 2018). Além disso, grandes quantidades de importantes informações advindas de diferentes partes e origens de dados da empresa irão poder ser unificadas e analisadas em conjunto com o melhor poder de processamento e métodos estatísticos e algoritmos sofisticados, para que sejam aproveitadas as possibilidades de reportes com comparações de performances de dados financeiros e não financeiros da companhia (AITE apud COCKROFT e RUSSEL, 2018).

O papel do contador pode ser um dos grandes afetados ou até o escolhido para tratar a função de *Big Data* na empresa, pois são confiáveis pelas companhias, capacitados para entender com clareza o significado dos números, o que pode fazer com que foquem menos em controle de custos como uma forma de melhoria de performance organizacional (COCKROFT e RUSSEL, 2018), e se direcionem para análise de tendências ao invés de apenas dados históricos para guiar a tomada de decisão (BHIMANI e WILLCOCKS, 2014). Com isso, podem entender as demonstrações financeiras e correlacioná-las com a estratégia da empresa e garantir sua veracidade (RICHINS ET AL, 2016 apud NEVES, 2019), tudo isso por meio do aproveitamento da melhor capacidade de processamento, qualidade de algoritmos e análises estatísticas capazes de melhorar reportes e a conexão de dados financeiros e não financeiros (BHIMANI e WILLCOCKS, 2014).

O impacto para o tipo e conteúdo de trabalho contábil tende a variar fortemente, com a inclusão de mecanismos de feedback em tempo real na direção de sistemas de avaliação de performance, métricas contábeis e financeiras, todas em formato a oferecer informações qualitativas e

quantitativas sobre o universo do contador, de forma gráfica e interativa. Isso demonstra que a demanda por informação contábil e financeira tende a manter um modelo de operação eficiente e eficaz a se tornar mais rápido e exigente na identificação de tendências possíveis de guiarem decisões de controle da companhia (BHIMANI e WILLCOCKS, 2014).

Para os gestores financeiros, o ambiente organizacional está mudando pouco a pouco, e com isso a forma que a informação é acessada, analisada e reportada também, o que demonstra a necessidade de mudanças também nos serviços oferecidos. Com isso, a análise de informações não financeiras pode levar a um desenvolvimento de inteligência financeira com impactos na organização da gestão de custos, precificação e decisões operacionais de controle. Outro ponto que pode ser afetado nesse sentido com o uso da análise exaustiva da trilha digital gerada por consumidores é a oportunidade de novas fontes de receita como anúncios direcionados a grupos específicos de pessoas por meio de estratégias de marketing (BHIMANI e WILLCOCKS, 2014).

Sob o ponto de vista de auditoria, surgem novas possibilidades em relação ao profissional, que no caso interno pode prever e evitar custos de falhas de segurança, e no caso interno e externo provavelmente terá que se capacitar de ferramentas e técnicas para o trato da qualidade de dados (ALLES apud COCKROFT e RUSSEL, 2018). Um exemplo é a possibilidade de auditores terem evidências de ativos fixos registrados com informações relevantes em vídeo, imagem, áudio e texto de forma a facilitar a avaliação dos mesmos. Além disso, também pode ser uma oportunidade no julgamento da natureza de ativos intangíveis, tudo isso por meio de algoritmos de extração de dados que serão compilados, processados e analisados de acordo com os indicadores fundamentais do ativo (WARREN ET AL, 2015 apud NEVES, 2019).

Outro ponto que a auditoria pode ser impactada, até talvez de forma revolucionária, é na atual limitação de testagem por amostragem, o que com mecanismos de mineração de dados permitiria que dados externos como redes sociais para avaliações de risco de fraude e controles internos fossem analisados, além de técnicas de IOT como tecnologias de transmissores de radiofrequência que permitem contagens automáticas de inventário em qualquer localização, além de poderem contar com serviços de GPS para confirmações de entregas de vendas de algum produto (CAO ET AL apud



NEVES, 2019). Tudo isso pode não só reduzir o custo de obtenção de dados como melhorar as informações que irão além de dados tradicionais, aumentando o valor operacional gerado desses dados, que podem ser o rastreamento de um bem desde o momento de fabricação até a venda de forma simplificada, além de detalhes sobre a obsolescência do bem em questão (VASARHELYI ET AL; KRAHEL E TITERA, 2015 apud NEVES, 2019).

Com isso, a auditoria tem a capacidade de se desenvolver em ritmo de se tornar contínua por meio de análises automatizadas em tempo real, com uma ação proativa ao invés de reativa, identificando erros à medida que ocorrem (CHAN E VASARHELYI, 2011 apud NEVES, 2019). Uma tecnologia que ajudará muito nesse sentido é o Blockchain, que viabiliza a confirmação imediata das transações e de forma segura, o que faz com que a integridade seja garantida para a análise de auditores, que por conseguinte terão um menor papel de verificação, porém maior no seu julgamento e visão (DAI E VASARHELYI, 2017 apud NEVES, 2019).

Algumas importantes áreas da contabilidade e finanças já podem ser consideradas pouco exploradas em termos de pesquisa, e elas são: privacidade e segurança, visualização de dados, manejo e qualidade de dados e análise preditiva (COCKROFT e RUSSEL, 2018), sendo a última uma ótima aliada à análise comportamental para detecção e prevenção de fraude (FERGUSON apud COCKROFT e RUSSEL, 2018).

Uma evidência dos avanços no mercado para tecnologias de *Big Data* e *Analytics* são as patentes relacionando *Big Data* e *Analytics* registradas por empresas, onde podemos identificar em liderança a Amdocs Development, do ramo de *software* e serviços, seguida pela chinesa Huawei, a americana IBM e a alemã SAP (LIMA et al, 2019).

Além disso, muitos desafios ainda estão por vir na adaptação de *Big Data* para finanças e contabilidade, e entre eles estão em encontrar um modelo de negócios de *Analytics* de *Big Data* para finanças, organizar os dados de forma eficiente e eficaz, integrar informações de variadas fontes, garantir a segurança dos sistemas financeiros para proteção da privacidade individual, e além disso, também lidar com as tradicionais dificuldades da área financeira como alta quantidade de transações, risco de crédito, análise financeira e gerenciamento

de riscos e regulamentação, tudo de forma criativa direcionada pelo *Big Data* (SHI, SUN e ZHANG, 2019).

#### 4. Metodologia:

Quanto aos seus objetivos, a pesquisa é classificada como exploratória, visto que busca trazer mais conhecimento sobre o tema, fazendo com que se torne mais explícito, e esse tipo toma geralmente a forma de pesquisa bibliográfica ou estudo de caso (GIL, 2002). Nesse caso, são utilizados “exemplos que estimulam a compreensão” (SELLTIZ et al., 1967, p 63), o que aliado à parte bibliográfica, pode trazer “um novo enfoque para o assunto” (FREITAS; PRODANOV, 2013, p 52).

Em relação à abordagem do problema, é qualitativa, onde a fonte de dados é o próprio ambiente, e não há manipulação feita de forma proposital pelo agente de pesquisa. É um tipo de pesquisa descritiva, na qual o autor analisa os dados de forma indutiva, sem o uso de métodos e técnicas estatísticas (FREITAS; PRODANOV, 2013).

Quanto aos procedimentos técnicos utilizados, está presente a bibliográfica, por meio de buscas em livros, jornais e periódicos pela internet e o estudo de caso, incluindo os mesmos tipos de fontes pela internet, além de notícias e entrevistas, visando as buscas de casos reais sobre o tema. O primeiro, advindo de informações já existentes e compartilhadas publicamente, apresentam uma grande vantagem por possibilitar que o pesquisador acesse e consolide um grupo muito maior de fatos e ocorrências, que estariam espalhados pelo espaço (GIL, 2002). Já o segundo, consta no exaustivo e profundo estudo de forma a encontrar grande conhecimento (YIN, 2001), sendo analisado um indivíduo, ou algum tipo de grupo para que se entendam diversos pontos sobre o mesmo (FREITAS; PRODANOV, 2013).

A seleção de artigos, periódicos e matérias utilizadas como base na presente pesquisa se deu primordialmente com a busca por palavras e expressões chaves utilizando a plataforma do google acadêmico e scielo, sendo elas *Big Data*, *Big Data* e contabilidade, *Big Data* em contabilidade e finanças, *Big Data* e *Analytics*, *Analytics* para finanças e *Analytics* e contabilidade. Foram dados enfoques nos resultados em materiais encontrados

de instituições renomadas em pesquisa como Harvard Business Review, USP, UFPE, além de autores considerados referências no tema de *Big Data* como Davenport e Taurion, sendo grande parte dos trabalhos encontrados ainda em inglês, pelo tema não estar muito desenvolvido atualmente no Brasil. Foram pesquisadas também, com as mesmas palavras chave, em todos os períodos possíveis, em bases de dados dos periódicos contábeis Revista de Gestão, Finanças e Contabilidade (com publicações de 2012 a 2019); Revista Contemporânea de Contabilidade (com publicações de 2004 a 2021) e Pensar Contábil (com publicações de 2001 a 2020), encontrados pelo site do Conselho Regional de Contabilidade, e não foram encontrados materiais relacionados ao escopo do presente trabalho. Outras buscas realizadas utilizaram da ferramenta do google para pesquisas com as mesmas palavras e expressões chaves anteriores, porém em empresas ou organizações consolidadas há anos como especialista e consultoras de TI e do mercado de finanças como a PWC, IBM, Serpro, Oracle e Sas.

## **5. Estudos de Caso:**

### **5.1 Contabilidade Forense:**

Um exemplo de uso real de técnicas de *Big Data* e *Analytics* é o apresentado no estudo sobre contabilidade forense no combate à lavagem de dinheiro do congresso internacional da USP sobre contabilidade (LIMA, SERRANO e CUPERTINO, 2020). No trabalho, é discutida uma inovação tecnológica para tratar de forma diferente o tradicional trabalho da contabilidade forense de solucionar processos judiciais fazendo uso de habilidades específicas e investigativas, demandados cada vez mais por setores corporativos e governamentais. Por meio de grafos criados com o software IBM i2 Analyst's Notebook, considerando a legislação penal adequada nacional e internacional, foram analisadas grandes quantidades de transações para a verificação da possibilidade de identificação de transações suspeitas de lavagem de dinheiro, considerando que ferramentas de *Big Data*, *Machine Learning* e análise de dados, como mineração de dados podem dar o apoio tecnológico para a melhoria do trabalho forense.

É demonstrada a necessidade de formas mais avançadas e eficientes de análise de extratos bancários por contadores forenses visto que os mesmos processam até milhões de lançamentos em uma única investigação, enquanto utilizam técnicas de direito, auditoria, contabilidade, finanças, economia, sociologia e criminologia para tratar problemas judiciais da esfera civil e privada para identificação de atos ilícitos e sua devida punição. Tal necessidade se dá pelas diversas técnicas de ocultar movimentações ilícitas por meios de se passar indetectável pelos órgãos de controle. E, uma das formas de analisar tantos dados em um mundo mais focado em entender grandes volumes de dados, se torna possível o uso de grafos no processamento de dados para a análise facilitada de redes sociais pelo contador forense, tudo isso devido à possibilidade dessas grandes massas de dados revelarem padrões escondidos sobre atos ilegais.

As técnicas ilícitas podem ser ações que dissimulam a origem, destino, de responsáveis ou beneficiários finais de transações bancárias, fragmentações de valores movimentados para que o valor total não seja identificado, depósitos e saques atípicos em relação à atividade principal do cliente, contas de passagem (*polling accounts*) e uso de intermediários (*straw men*) entre outros.

Para tais dados o tratamento foi dado em 3 etapas, sendo elas limpeza, padronização e concatenação, e durante esses processos, os grafos criados foram capazes de identificar milhares de entidades de pessoas físicas ou jurídicas. Um exemplo de grafos analisados nos traz um alto volume de transações em espécie, de saques e depósitos de recursos, não identificadas e vinculadas a quatro pessoas sendo vistas como investigadas, o que indica uma possível evidência de algum crime. Também foi identificado que os grafos podem ajudar no rastreamento de recursos localizando relações entre duas entidades, revelando assim conexões entre pessoas ou organizações investigadas.

Dessa forma, foi encontrado que os grafos são um meio aplicável de reprimir crimes de lavagem de dinheiro, pois possibilitam a identificação facilitada de transações suspeitas como *polling accounting* e *straw men*, e o caminho das mesmas, auxiliando também o rastreamento de recursos no sistema bancário. Com essa ferramenta, foi possível identificar também, além de relações financeiras, relações entre sócios, comerciais, familiares e outros,

o que de forma combinada por meio de seus dados, pode deixar mais sofisticados os resultados da contabilidade forense e sua consequência penal.

## **5.2 *Big Data* na Auditoria Governamental:**

Outro exemplo onde é possível identificar a discussão de implicações práticas do uso de *Big Data* na área financeira é o estudo sobre *Big Data* para a auditoria financeira para avaliação e resposta a riscos em demonstrações financeiras do Governo Federal, publicado na revista do Tribunal de Contas da União (COSTA E DUTRA, 2014). No dado estudo, é investigado os possíveis desdobramentos da era do *Big Data* na auditoria de demonstrações governamentais, que consiste em certificar-se, de forma limitada, sobre dados agregados com base na abordagem baseada em riscos, cortes materiais e amostras estatísticas, tudo isso de forma anual, normalmente.

A inovação levantada é em relação a mecanismos chamados *audit Analytics* e auditoria contínua, onde seria possível analisar as demonstrações de forma contínua, com menor intervalo de tempo entre cada análise, de forma integral e com melhor tempestividade e confiabilidade nas opiniões.

Na auditoria financeira, é essencial que seja dada uma garantia razoável da ausência de erro ou fraude, o que deve ser feito com o melhor nível de transparência e credibilidade possível, para que as decisões tomadas com base nas demonstrações sejam tomadas de forma segura. E com isso, há o risco de auditoria, quando há uma possibilidade de que seja considerado que as informações são confiáveis, e na realidade não são, visto que são feitos testes por amostras. Com isso, surge a importância do *Big Data* e de suas ferramentas inovadoras de análise para que sejam possibilitadas automações de processos, mudando mecanismos de comunicação de resultados para uma periodicidade muito menor do que a anual obrigatória, podendo ser feita por meio da internet quase que ao mesmo tempo em que ocorrem as transações, tornando mais eficiente a dinâmica da economia e governos.

As possíveis mudanças trazidas por desenvolvimentos tecnológicos podem afetar tanto o setor privado quanto público em relação à utilidade das informações, que devem ser confiáveis para tal. Um ponto importante é que para que a confiança seja garantida é necessário que não se perca a qualidade

dos procedimentos, o que é possível com os mecanismos de *Analytics* que surgem na nova era tecnológica, e ainda de forma a reduzir a periodicidade das auditorias com melhorias na operacionalidade do trabalho, permitindo que se chegue a um patamar de auditoria contínua, sem distorções relevantes nas demonstrações.

Com uma auditoria contínua por meio de uso de ferramentas específicas, se torna factível que testes sejam aplicados à grandes volumes de dados financeiros, em diversos níveis da organização, aposentando a necessidade de cortes com base em amostras e materialidade e melhorando a abordagem baseada em risco. Tudo isso é possível pois com novas tecnologias, é possível focar em itens mais relevantes, ou seja, de maior impacto nas demonstrações, como informações agregadas e contas específicas, com testes estatísticos também mais aperfeiçoados, além de monitorar tais itens praticamente simultaneamente ao momento que surgem.

Tais mudanças tecnológicas são aplicáveis ao Sistema Integrado de Administração Financeira (Siafi), que está ligado ao Orçamento Fiscal e a Seguridade Social, além de outros sistemas gerenciadores de recursos federais, o que produz movimentações ao redor de 30 milhões de lançamentos contábeis de entrada de dados em um ano (2013), sendo uma parte de registros manuais. Nele, já é possível identificar certos testes aproximados ao monitoramento contínuo como análises de demonstrativos já automatizadas, verificações sobre integrações com outros sistemas, também com formas de se identificar erros e monitorá-los até sua correção. E tais inovações de controles internos beneficiam diretamente o órgão que audita externamente tais demonstrações, que é o Tribunal de Contas da União.

Além das tecnologias de *Analytics* já existentes para a auditoria financeira, também temos um outro facilitador na estruturação de dados, que é o Plano de Contas Aplicado ao Setor Público (PCASP), que padroniza as contas públicas de forma consolidada, como indica a Lei da Responsabilidade Fiscal (LRF). E assim, com uma evolução de recursos tecnológicos, gestão de risco, governança e com pessoas capacitadas, será possível atingir um novo nível de eficiência da auditoria financeira governamental, com até uma melhor confiabilidade, transparência e menor periodicidade.

## 6. Considerações Finais:

Por fim, podemos concluir que com todos os recursos tecnológicos trazidos junto aos 5 Vs do *Big Data* (TAURION, 2013) e as ferramentas de *Analytics*, frente ao grande volume de dados, temos um crescimento nunca antes visto. Com isso, podemos observar uma possível nova revolução industrial, com base no que dizem os dados, e uma completa mudança na forma de trabalhar, agora indo um passo além da análise de dados históricos para tomar melhores decisões no presente, para o uso desses dados de forma a fazer análises preditivas, para que decisões sejam feitas já pensando em um resultado esperado, e com alto grau de confiabilidade.

Por meio desta pesquisa, também foi possível notar o aspecto promissor que a era do *Big Data* proporciona para o mundo de contabilidade e finanças, com melhorias nas análises de altos números de transações, riscos financeiros e operacionais, de forma rápida, com monitoramento simultâneo com a ocorrência de eventos, além de uma possibilidade de auditoria contínua. Com isso, foi observado também que áreas pouco exploradas e com grande potencial são a de privacidade e segurança, que é um desafio para todas as matérias tocadas por *Big Data*, a de visualização de dados, de manejo e qualidade de dados, além da análise preditiva (COCKROFT e RUSSEL, 2018).

Dessa forma, é possível observar que os estudos e literaturas relacionando diretamente contabilidade e finanças à temática do *Big Data* ainda são escassos, visto que foram encontrados apenas cinco estudos teóricos e dois sobre usos práticos. Isto demonstra que ainda não é possível obter de forma completa uma ideia sobre os novos horizontes de análises, aperfeiçoamento e precisão de processos e reportologias financeiras que tal novo ambiente tecnológico pode oferecer. Portanto, clara a grande importância do *Big Data* para contabilidade e finanças nos anos que estão por vir, e se torna interessante uma maior quantidade de estudos revelando aplicações práticas sobre o tema em processos contábeis e do universo financeiro.

## Referências:

AHMED, Arif; PATGIRI, Ripon. *Big Data: The V's of the Game Changer Paradigm*. In: IEEE HIGH PERFORMANCE COMPUTING AND COMMUNICATIONS, 18., 2016, Assam. Congress. Sydney: National Institute Of Technology Silchar, 2016. p. 1-24. Disponível em: [https://www.researchgate.net/publication/311642627\\_Big\\_Data\\_The\\_V%27s\\_of\\_the\\_Game\\_Changer\\_Paradigm](https://www.researchgate.net/publication/311642627_Big_Data_The_V%27s_of_the_Game_Changer_Paradigm). Acesso em: 10 jul. 2020.

ALECRIM, Emerson. Kilobyte, Megabyte, Gigabyte, Terabyte... e depois? 2007. Disponível em: <https://www.infowester.com/blog/kilobyte-megabyte-gigabyte-terabyte-e-depois/>. Acesso em: 9 fev. 2020.

AMARAL, Brenna Nicole Jurchacks Santos do. *BIG DATA: Um benchmark de grande volume de dados de serviços de internet*. 2017. 55 f. TCC (Graduação) Curso de Curso de Gestão da Informação, Departamento de Ciencia da Informação, Universidade Federal de Pernambuco, Recife, 2017. Disponível em: <https://repositorio.ufpe.br/bitstream/123456789/34941/1/Brenna%20Nicole%20Jurchacks%20Santos%20do%20Amaral.pdf>. Acesso em: 10 jul. 2020.

BHIMANI, Alnoor; WILLCOCKS, Leslie. DIGITISATION, 'BIG DATA' AND THE TRANSFORMATION OF ACCOUNTING INFORMATION. Online, 21 maio 2014. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/00014788.2014.910051>. Acesso em: 03 set. 2020.

BLAZQUEZ, Desamparados; DOMENECH, Josep. *Big Data sources and methods for social and economic analyses*. 2017. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0040162517310946>. Acesso em: 10 jul. 2020.

BRETERNITZ, Vivaldo Jose; SILVA, Leandro Augusto. *Big Data: Trazendo Novas Oportunidades e Desafios*. 2013. 10th International Conference on



Information Systems and Technology Management – CONTECSI. São Paulo, p. 2906-2914. Junho, 2013.

CAMPOS, Fábio Rocha. A gestão da inovação em serviços intensivos em conhecimento: oportunidades e desafios do *Big Data*. 2015. 1 recurso online (124 p.). Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Geociências, Campinas, SP. Disponível em: <<http://www.repositorio.unicamp.br/handle/REPOSIP/286566>>. Acesso em: 15 ago. 2020.

CEOLATO, Renata Varela. Análise bibliométrica de artigos da área de Sistemas de Informação Contábil e suas contribuições relacionadas à aplicação de tecnologias emergentes na contabilidade. 2019. 20 f. TCC (Graduação) - Curso de Ciências Contábeis, Universidade Federal do Rio Grande do Sul (Ufrgs), Rio Grande do Sul, 2019. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/198557/001098050.pdf?sequence=1&isAllowed=y>. Acesso em: 10 set. 2020.

COCKROFT, Sophie; RUSSEL, Mark. *BIG DATA OPPORTUNITIES FOR ACCOUNTING AND FINANCE PRACTICE AND RESEARCH: BIG DATA IN ACCOUNTING AND FINANCE*. Austrália, 01 fev. 2018. Disponível em: [https://www.researchgate.net/publication/322938351\\_Big\\_Data\\_Opportunities\\_for\\_Accounting\\_and\\_Finance\\_Practice\\_and\\_Research\\_Big\\_Data\\_in\\_Accounting\\_and\\_Finance](https://www.researchgate.net/publication/322938351_Big_Data_Opportunities_for_Accounting_and_Finance_Practice_and_Research_Big_Data_in_Accounting_and_Finance). Acesso em: 15 ago. 2020.

COSTA, Gledson Pompeu Corrêa da; DUTRA, Tiago Alves de Gouveia Lins. Auditoria financeira na era do *Big Data*:: novas possibilidades para avaliação e resposta a riscos em demonstrações financeiras do governo federal. Revista do TCU. Brasília, p. 54-61. Set/Dez 2014.

COX, Michael; ELLSWORTH, David. *MANAGING BIG DATA FOR SCIENTIFIC VISUALIZATION*. São Francisco, 1 Maio 1997. Disponível em: [https://www.researchgate.net/publication/238704525\\_Managing\\_big\\_data\\_for\\_scientific\\_visualization](https://www.researchgate.net/publication/238704525_Managing_big_data_for_scientific_visualization)

DAVENPORT, Thomas. Webinars *Big Data* at Work: dispelling the myths, uncovering the opportunities. In: HARVARD BUSINESS REVIEW WEBINAR, ., 2014, Online. Webinar. Online: Harvard Business School Publishing, 2014. p. 1-7. Disponível em: [https://www.sas.com/content/dam/SAS/en\\_us/doc/whitepaper2/hbr-big-data-at-work-107035.pdf](https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper2/hbr-big-data-at-work-107035.pdf). Acesso em: 8 jul. 2020.

DOMO. Data Never Sleeps 3.0. How much data is generated every minute? 2015. Disponível em: [https://web-assets.domo.com/blog/wp-content/uploads/2015/08/15\\_domo\\_data-never-sleeps-3\\_final1.png](https://web-assets.domo.com/blog/wp-content/uploads/2015/08/15_domo_data-never-sleeps-3_final1.png). Acesso em: 08 jul. 2020.

SELLTIZ, Claire;[et al.]. Métodos de pesquisa nas relações sociais. São Paulo: Ed. Herder, 1967.

SHI, Yufeng; SUN, YUNCHUAN; ZHANG, Zhengjun. FINANCE *BIG DATA*: MANAGEMENT, ANALYSIS, AND APPLICATIONS. Online, 06 jan. 2019. Disponível em: <https://doi.org/10.1080/10864415.2018.1512270>. Acesso em: 03 set. 2020.

GARTNER. Gartner Glossary: *Big Data*. 2020. Disponível em: <https://www.gartner.com/en/information-technology/glossary/big-data>. Acesso em: 08 jul. 2020.

GARTNER. How to Stop Data Quality Undermining Your Business. 2018. Disponível em: <https://www.gartner.com/smarterwithgartner/how-to-stop-data-quality-undermining-your-business/>. Acesso em: 08 jul. 2020.

GIL, A. C. Como elaborar projetos de pesquisa. São Paulo: Atlas, 2002.

HARVARD BUSINESS REVIEW. *Big Data* at Work: Dispelling the Myths, Uncovering the Opportunities. 2014. Disponível em: <https://hbr.org/2014/03/big->

data-at-work-dispelling-the-myths-uncovering-the-opportunities. Acesso em: 10 jul. 2020.

HURWITZ, Judith et al. *Big Data for Dummies*. New Jersey: John Wiley & Sons, Inc., 2013.

IBM. Análítica de grandes dados. 2020. Disponível em: <https://www.ibm.com/br-pt/analytics/hadoop/big-data-analytics>. Acesso em: 08 jul. 2020.

LIMA ET AL. A CONTABILIDADE NA ERA DIGITAL: PROSPECÇÃO TECNOLÓGICA PARA UMA ANÁLISE DE TENDÊNCIAS. Maceió, 27 nov. 2019. Disponível em: <https://portalseer.ufba.br/index.php/nit/article/view/33062>. Acesso em: 15 set. 2020.

LIMA, Rafael Sousa; SERRANO, André Luiz Marques; CUPERTINO, Cesar Medeiros. Contabilidade Forense e Grafos no Combate à Lavagem de Dinheiro. XX Usp International Conference In Accounting. São Paulo, p. 1-17. Julho 2020.

LINDELL, Jim. *Analytics and Big Data for Accountants*. New Jersey: Wiley, 2018.

MARGARET ROUSE. 3Vs (volume, variety and velocity). 2013. Disponível em: <https://whatis.techtarget.com/definition/3Vs>. Acesso em: 08 jul. 2020.

MARR, Bernard. How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read: The Mind-Blowing Stats Everyone Should Read. 2018. Disponível em: <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#14bc023260ba>. Acesso em: 07 jul. 2020.

MCKINSEY GLOBAL INSTITUTE. *Big data: The next frontier for innovation, competition, and productivity*. 2011. Disponível em: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>. Acesso em: 08 jul. 2020.

NEVES, Maria Sofia Camacho. *A Utilização do Big Data na Contabilidade*. 2019. 94 f. Monografia (Especialização) - Curso de Mestrado em Contabilidade, Universidade do Minho, Portugal, 2019. Disponível em: [https://repositorium.sdum.uminho.pt/bitstream/1822/64721/1/Sofia\\_Relatorio%20Estagio\\_C.ALTERAC\\_O\\_ES\\_V2.pdf](https://repositorium.sdum.uminho.pt/bitstream/1822/64721/1/Sofia_Relatorio%20Estagio_C.ALTERAC_O_ES_V2.pdf). Acesso em: 28 ago. 2020.

NODE GRAPH. *How much data is on the internet? The Big Data Facts Update 2020*. 2020. Disponível em: <https://www.nodegraph.se/how-much-data-is-on-the-internet/>. Acesso em: 06 jul. 2020.

ORACLE BRASIL. *O Que é Big Data?* 2020. Disponível em: <https://www.oracle.com/br/big-data/what-is-big-data.html>. Acesso em: 08 jul. 2020.

PEREIRA, Vanessa Alves da Silva. *Big Data – Um Estudo em Gestão Empresarial*. 2016. Trabalho de Conclusão de Graduação – Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016. Disponível em: <https://pantheon.ufrj.br/handle/11422/169>. Acesso em: 02 ago. 2020.

PRESS, Gil. *6 Predictions For The \$203 Billion Big Data Analytics Market*. 2017. Disponível em: <https://www.forbes.com/sites/gilpress/2017/01/20/6-predictions-for-the-203-billion-big-data-analytics-market/#726dcdae2083>. Acesso em: 06 jul. 2020.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de, *Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico*, 2ª Ed., Novo Hamburgo - RS, Associação Pró-Ensino Superior em Novo Hamburgo -ASPEUR Universidade Feevale, 2013. Disponível

em: <http://www.feevale.br/Comum/midias/8807f05a-14d0-4d5b-b1ad-1538f3aef538/E-book%20Metodologia%20do%20Trabalho%20Cientifico.pdf>,  
Acesso em: 07 jul. 2020.

PWC. AI will help answer the big question about data. 2018. Disponível em: <https://www.pwc.com/us/en/services/consulting/library/artificial-intelligence-predictions/big-data-roi.html>. Acesso em: 08 jul. 2020.

PWC. Importância do *big data*. 2016b. Disponível em: <https://www.pwc.com.br/pt/setores-atividade/saude/importancia-big-data.html>. Acesso em: 08 jul. 2020.

PWC. LGPD: o que muda na prática com a Lei 13.709/18. 2020. Disponível em: <https://www.pwc.com/br/pt/setores-atividade/saude/lgpd-o-que-muda-na-pratica-com-a-lei-1370918.html>. Acesso em: 15 ago. 2020.  
PWC. PwC's Global Data and *Analytics* Survey 2016 Big Decisions™. 2016. Disponível em: <https://www.pwc.com/us/en/services/consulting/analytics/big-decision-survey.html>. Acesso em: 08 jul. 2020.

PWC. The human factor: Working with machines to make big decisions. 2016a. Disponível em: <https://www.pwc.com/us/en/services/consulting/library/human-factor.html>. Acesso em: 10 jul. 2020.

REINSEL, David; GANTZ, John; RYDNING, John. The Digitization of the World: From Edge to Core. 2018. Disponível em: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>. Acesso em: 07 jul. 2020.

RUSSOM, Philip. TDWI BEST PRACTICES REPORT: *Big Data Analytics*. 2011. Disponível em: [ftp://ftp.software.ibm.com/software/tw/Defining\\_Big\\_Data\\_through\\_3V\\_v.pdf](ftp://ftp.software.ibm.com/software/tw/Defining_Big_Data_through_3V_v.pdf) . Acesso em: 02 ago. 2020.

SANT'ANA, Samuel Gonçalves. A CONTABILIDADE NO BRASIL: o seu início aos dias atuais. 2019. 19 f. TCC (Graduação) - Curso de Faculdade de Ciências

Contábeis, Universidade Federal de Uberlândia - Ufu, Uberlândia, 2019. Disponível em: <https://repositorio.ufu.br/bitstream/123456789/27661/4/ContabilidadeBrasilIn%C3%ADcio.pdf>. Acesso em: 18 ago. 2020.

SAS. *Big Data* O que é e qual sua importância? 2020a. Disponível em: [https://www.sas.com/pt\\_br/insights/big-data/what-is-big-data.html?source=post\\_page-----](https://www.sas.com/pt_br/insights/big-data/what-is-big-data.html?source=post_page-----). Acesso em: 08 jul. 2020.

SAS. *Analytics* O que é e qual sua importância? 2020b. Disponível em: [https://www.sas.com/pt\\_br/insights/analytics/analytics.html](https://www.sas.com/pt_br/insights/analytics/analytics.html). Acesso em: 08 jul. 2020.

SERPRO. MAPA DA PROTEÇÃO DE DADOS. 2020a. Disponível em: <https://www.serpro.gov.br/lgpd/menu/a-lgpd/mapa-da-protecao-de-dados-pessoais>. Acesso em: 15 ago. 2020.

SERPRO. O IMPACTO DA LGPD NOS NEGÓCIOS. 2020b. Disponível em: <https://www.serpro.gov.br/lgpd/empresa/o-impacto-lgpd-nos-negocios>. Acesso em: 15 ago. 2020.

SERPRO. O QUE MUDA COM A LGPD. 2020c. Disponível em: <https://www.serpro.gov.br/lgpd/menu/a-lgpd/o-que-muda-com-a-lgpd>. Acesso em: 15 ago. 2020.

TAURION, Cezar. *Big Data*. Rio de Janeiro: Brasport Livros e Multimídia Ltda., 2013.

TAURION, Cezar. Maturidade em Data Science é determinante. 2016. Disponível em: <https://cio.com.br/maturidade-em-data-science-e-determinante/>. Acesso em: 06 jul. 2020.

YIN, Robert K. Estudo de caso – planejamento e métodos. (2Ed.). Porto Alegre: Bookman. 2001

YU, Shui; GUO, Song. *Big Data Concepts, Theories, and Applications*. Suíça: Pringer International Publishing, 2016.