

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ADMINISTRAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO**

**Fabiano Chiapinotto Saffi**

**INTELIGÊNCIA ESTRATÉGICA ANTECIPATIVA:  
IDENTIFICAÇÃO DE SINAIS FRACOS POR MEIO DO BIG DATA ANALYTICS**

Porto Alegre, Setembro 2020

Fabiano Chiapinotto Saffi

INTELIGÊNCIA ESTRATÉGICA ANTECIPATIVA:  
IDENTIFICAÇÃO DE SINAIS FRACOS POR MEIO DO BIG DATA ANALYTICS

Dissertação de Mestrado apresentada ao PPGA  
da Universidade Federal do Rio Grande do Sul,  
como requisito parcial para a obtenção do título  
de Mestre em Administração.

Orientadora: Profª Drª Raquel Janissek-Muniz

Porto Alegre, Setembro 2020

Fabiano Chiapinotto Saffi

INTELIGÊNCIA ESTRATÉGICA ANTECIPATIVA:  
IDENTIFICAÇÃO DE SINAIS FRACOS POR MEIO DO BIG DATA ANALYTICS

Dissertação de Mestrado apresentada ao PPGA  
da Universidade Federal do Rio Grande do Sul,  
como requisito parcial para a obtenção do título  
de Mestre em Administração.

Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Raquel Janissek-Muniz

Conceito Final: .....

Banca Examinadora:

---

Prof. Dr. Ariel Behr – PPGA/UFRGS

---

Prof<sup>a</sup> Dr<sup>a</sup> Carla Bonato Marcolin – FAGEN/UFU

---

Prof<sup>a</sup> Dr<sup>a</sup> Fernanda Maciel Reichert – PPGA/UFRGS

---

Prof<sup>a</sup> Dr<sup>a</sup> Raquel Janissek-Muniz – Orientadora – PPGA/UFRGS

Porto Alegre, Setembro 2020

## **Agradecimentos**

Tanto no desenvolvimento desta pesquisa, como também durante a caminhada do mestrado acadêmico, com seus trabalhos e disciplinas, foi necessária a disponibilização de tempo. Tempo que não volta, tempo que foi dedicado não apenas por este pesquisador, mas por toda sua família. Minha esposa Fernanda não apenas deu apoio durante estes anos. O incentivo, os conselhos, a compreensão do quanto esta formação é importante para mim, as restrições necessárias nos finais de semana e no dia a dia, a falta de convivência, a paciência, tolerância e o companheirismo só podem ser explicados pelo amor que nos une. Agradeço imensamente e dedico este trabalho a ela, sabendo que este gesto e poucas palavras não são suficientes. Ao meu filho, João Pedro, fica a promessa de compensar este tempo nos próximos anos. Sei que você entende, ainda que tenha apenas um ano, que fiquei um pouco ausente por motivos importantes.

Agradeço aos meus pais Cicero e Vera, ao meu irmão Alexandre, aos cunhados Renê, Eduardo e Tanisa. Agradeço ao meu sogro e sogra José Antônio e Cecy.

Agradeço à minha orientadora, professora Raquel, por todo conhecimento transmitido, pela confiança de oportunizar espaços para eu repassá-los, mesmo que apenas por uns minutos, e entender o quão difícil é esta profissão que tanto respeito.

Agradeço aos professores Ariel, Carla e Fernanda, pelos ensinamentos e por aceitaram participar desta banca.

Agradeço a todos colegas do nosso grupo de estudos IEA Future Lab e do PPGA, pelas ajudas, sugestões, pesquisas, correções e risadas, que ocorreram em diversos momentos nestes anos. Obrigado pela amizade e parceria.

A todos deixo o meu muito obrigado!

## Resumo

Com o ambiente empresarial cada vez mais dinâmico e complexo, seu monitoramento tem sido uma prática obrigatória para as organizações. Dado que nas últimas décadas há um volume importante de informações disponibilizadas por tecnologias digitais, organizações estão investindo em ferramentas *Big Data Analytics* (BDA) com o propósito de compreender o mercado e reduzir a incerteza na tomada de decisão. Ao considerar esse tipo de dados nos processos de monitoramento, analistas possivelmente fortalecem a identificação de Sinais Fracos (sinais antecipativos). Acreditando que tais processos possam ser guiados por dados, buscou-se investigar a seguinte questão de pesquisa: qual o impacto das práticas do BDA na identificação de Sinais Fracos? Objetiva-se, portanto analisar a relação entre o uso de práticas do BDA e a identificação de Sinais Fracos, fonte de informações da Inteligência Antecipativa. Neste intuito, aproximando as disciplinas da área de Sistemas de Informação - *Analytics* e Inteligência Antecipativa - realizou-se duas pesquisas empíricas, a primeira de natureza qualitativa e a segunda quantitativa. Com o método de entrevistas em profundidade, a primeira pesquisa objetiva compreender se os gestores percebem Sinais Fracos por meio das ferramentas BDA, e seu uso potencial na tomada de decisão estratégica. Foram reveladas práticas do BDA que influenciam o monitoramento e a identificação de Sinais Fracos. Apresentou-se evidências de que dados candidatos a Sinais Fracos podem ser identificados a partir do BDA e de que essas ferramentas facilitam o monitoramento do ambiente. Na segunda pesquisa, a fim de medir o efeito das práticas do BDA na identificação de Sinais Fracos, desenvolveu-se hipóteses relacionadas em um modelo conceitual. Para suportá-lo, elaborou-se uma pesquisa do tipo *survey*, com 123 respondentes. As respostas foram analisadas mediante modelagem de equações estruturais (*PLS-SEM*), suportando o modelo e resultando em achados relevantes, indicando que o uso das práticas do BDA tem efeito positivo na identificação de Sinais Fracos. Como contribuições teóricas destaca-se o suporte empírico ao modelo desenvolvido, demonstrando a importância das variáveis que têm impacto significativo na identificação de sinais fracos.

**Palavras-chave:** Inteligência Antecipativa; *Big Data Analytics*; Sinais Fracos; *Foresight*

## **Abstract**

As the business environment is increasingly becoming dynamic and complex, its monitoring has been a mandatory practice for organizations. Given that in the last decades there is an important volume of information available, organizations are investing in Big Data Analytics (BDA) tools, with the purpose of understanding the market and reducing the uncertainty in decision making. Considering this type of data in the monitoring processes, analysts possibly strengthen the identification of weak signals. Believing that such processes could be data-driven, we sought to investigate the following research question: what is the impact of BDA's practices in the identification of weak signs? Thus, the aim of this study is to analyze the relationship between the use of BDA practices and the identification of weak signs, an information source from Anticipative Intelligence. To this end, bringing together the Information Systems area disciplines - Analytics and Anticipative Intelligence - two empirical types of research were carried out, the first had a qualitative nature and the second had a quantitative nature. With the method of in-depth interviews, the first research aims to understand whether managers perceive weak signals through the BDA tools, and their potential use in strategic decision-making. BDA practices that influence the monitoring and identification of weak signals were revealed. Evidence was presented that candidate data for weak signals can be identified by BDA. In the second research, in order to measure the effect of BDA practices in the identification of weak signals, related hypotheses were developed in a conceptual model. To support it, a survey was conducted, with 123 respondents. The responses were analyzed using structural equation modeling (PLS-SEM), supporting the model and resulting in relevant findings, indicating that the use of BDA practices has a positive effect on identifying weak signals. As theoretical contributions, it stands out the empirical support of the model, demonstrating the importance of variables that have a significant impact on the identification of weak signals.

**Keywords:** Anticipative Intelligence, Big Data Analytics, Weak Signals, Foresight

## **Lista de Figuras**

Figura 1: Desenho da pesquisa .....	22
Figura 2: Respostas para P03, dados não estruturados .....	41
Figura 3: Respostas para P04, BDA e insights de eventos futuros.....	42
Figura 4: Respostas para P07, uso dos Outliers.....	42
Figura 5: Respostas para P08, uso do método indutivo.....	42
Figura 6: Respostas para P06, insights e data-driven .....	43
Figura 7: Respostas para P09, uso do BDA para monitoramento do ambiente e identificação de Sinais Fracos .....	43
Figura 8: Fatores do BDA que influenciam na identificação de Sinais Fracos.....	52
Figura 9: Modelo Conceitual de Pesquisa .....	78
Figura 10: Desenho da pesquisa .....	79
Figura 11: Resultados - Modelo Estrutural da Pesquisa.....	91
Figura 12: Tela inicial com apresentação e termo de aceite .....	122
Figura 13: Ponto de triagem .....	122
Figura 14: Cálculo do tamanho da amostra mínima.....	123

## **Lista de Tabelas**

Tabela 1: Resumo da Pesquisa .....	22
Tabela 2: Características dos Sinais Fracos .....	31
Tabela 3: Paralelo entre Sinais Fracos e Big Data.....	35
Tabela 4: Lista e elementos de investigação.....	36
Tabela 5: Roteiro de perguntas semiestruturado .....	39
Tabela 6: Grupo de entrevistados .....	41
Tabela 7: Tipos de análises realizadas .....	44
Tabela 8: Resumo das respostas .....	51
Tabela 9: Roteiro estruturado de perguntas para consulta a especialistas .....	75
Tabela 10: Síntese das respostas dos especialistas para cada construto avaliado.....	77
Tabela 11: Mudanças na nomenclatura dos construtos após consulta a especialistas .....	77
Tabela 12: Roteiro estruturado de perguntas da pesquisa survey .....	83
Tabela 13: Perfil da amostra .....	85
Tabela 14: Análise fatorial exploratória – cargas no bloco .....	86
Tabela 15: Validade convergente e confiabilidade composta .....	86
Tabela 16: Validade discriminante segundo critério Fornell-Larcker .....	87
Tabela 17: Validade discriminante segundo critério HTMT .....	87
Tabela 18: Collinearity statistics – Outer VIF values.....	88
Tabela 19: Collinearity statistics – Inner VIF values .....	88
Tabela 20: Coeficientes de determinação, determinação ajustado e validade preditiva.....	90
Tabela 21: Resultados das hipóteses.....	90
Tabela 22: Gaps, contribuições e estudos futuros .....	102
Tabela 23: Análise de conteúdo, tabela de categorias e códigos .....	120

## Sumário

Agradecimentos .....	4
Resumo .....	5
Abstract.....	6
Lista de Figuras.....	7
Lista de Tabelas .....	8
1 Introdução .....	10
1.1 Problemática e Questão de Pesquisa.....	14
1.2 Objetivos .....	16
1.3 Justificativa .....	16
1.4 Delineamento da Pesquisa e Estrutura do trabalho.....	20
2 Artigo I - Explorando usos potenciais do <i>Big Data</i> para a Inteligência Antecipativa.....	23
Resumo .....	23
2.1 Introdução .....	23
2.2 Referencial Teórico.....	25
2.2.1 <i>Big Data Analytics</i> .....	27
2.2.2 Teoria dos Sinais Fracos e Inteligência Antecipativa .....	29
2.2.3 BDA e Inteligência Antecipativa .....	32
2.3 Método de Pesquisa .....	37
2.4 Análise dos Resultados .....	40
2.5 Discussões.....	44
2.6 Considerações Finais .....	52
2.7 Referências.....	54
3 Artigo II - O efeito do <i>Big Data Analytics</i> na identificação de Sinais Fracos.....	55
Resumo .....	55
3.1 Introdução .....	55
3.2 Referencial Teórico.....	60
3.2.1 Sinais Fracos e a Inteligência Estratégica Antecipativa.....	60
3.2.2 <i>Data-driven Culture</i> .....	62
3.2.3 <i>Big Data Analytics</i> .....	64
3.3 Modelo de Pesquisa .....	74
3.4 Método .....	78
3.4.1 Pesquisa <i>Survey</i> .....	79
3.4.2 Desenvolvimento do instrumento .....	81
3.4.3 Procedimento de coleta de dados .....	82
3.4.4 Análise e Estatística dos Dados.....	84
3.5 Análise dos Resultados .....	85
3.5.1 Amostra .....	85
3.5.2 Modelo de Mensuração .....	85
3.5.3 Modelo Estrutural.....	88
3.6 Discussão dos Resultados .....	92
3.7 Considerações Finais .....	95
3.8 Referências.....	97
4 Considerações Finais .....	98
5 Referências .....	104
6 Anexo I – Tabela de categorias e códigos do Artigo I .....	120
7 Anexo II – Telas iniciais do questionário online do Artigo II.....	122
8 Anexo III – Estimativa do tamanho mínimo da amostra.....	123

## 1 Introdução

O entendimento do **ambiente** empresarial para a **tomada de decisão** se torna cada dia mais complexo (CARAYANNIS et al., 2017; FONSECA; BARRETO, 2011; MÜHLROTH; GROTTKE, 2018; SCHOEMAKER; DAY, 2009), impulsionado pelo ritmo de novos fenômenos e mudanças repentinas, que aumentam sua volatilidade (VECCHIATO, 2015) e imprevisibilidade. A globalização retira barreiras de concorrentes distantes e desconhecidos, e a *internet* potencializa mercados globais, transformando o cenário e criando novos nichos, que surgem e desaparecem, rapidamente (ROHRBECK; BADE, 2012). Em meio a esses inúmeros negócios e mercados, sistemas políticos e tecnológicos interdependentes, novas formas de comunicação possibilitam a troca de um grande volume de informações (MENDONÇA; CARDOSO; CARAÇA, 2012). Como consequência, está mais difícil manter a vantagem competitiva (ROHRBECK; BADE, 2012), o que exige abordagens avançadas para descobrir e explorar oportunidades e ameaças (CONSTANTIOU; KALLINIKOS, 2015).

Com um ambiente em constante mudança, as organizações são obrigadas a estar sempre atentas, observando-o permanentemente (ILMOLA; KUUSI, 2006; LESCA, 2003; ROSSEL, 2009). Ao ter uma capacidade de **monitoramento** instalada, reduz-se o nível de incerteza e melhora-se a qualidade da tomada de decisão, viabilizando a ação rápida, no momento certo (JANISSEK-MUNIZ; LESCA; FREITAS, 2006). A resiliência e o sucesso econômico são viáveis se as organizações estiverem em constante adaptação com base na análise do meio externo, não agindo somente de forma reativa, mas sim proativamente, de maneira **antecipada** (EREVELLES; FUKAWA; SWAYNE, 2016). Para tal, deve-se decidir a partir de percepções prévias de futuras discontinuidades do ambiente (JANISSEK-MUNIZ; BORGES; BORTOLI, 2015), em um processo que exige mudanças radicais na estratégia corporativa, na gestão da inovação e nas suas estruturas internas (ROHRBECK, 2010). Essa “escuta ativa” – aptidão de buscar, armazenar e analisar informações – é imperativa para as organizações (JANISSEK-MUNIZ; FREITAS; LESCA, 2007; KIM; LEE, 2017), pois proporciona a antecipação de fenômenos e o alinhamento ágil e proativo do posicionamento frente às transformações rápidas do mercado (CHOO, 1999, 2001; LESCA et al., 2015; ROSSEL, 2009). Aquelas que não possuem recursos para práticas de monitoramento, tornam-se vulneráveis, perdendo competitividade em relação aos concorrentes (JANISSEK-MUNIZ; BORGES; BORTOLI, 2015). Esta busca por informações não pode ser realizada de forma pontual ou passiva (MENDONÇA;

CARDOSO; CARAÇA, 2012), é necessário desenvolver uma postura proativa, com processos contínuos, sistemáticos e formalizados (CAINELLI; JANISSEK-MUNIZ, 2019; FONSECA; BARRETO, 2011).

Assim, processos de inteligência estratégica – atividade que envolve monitoramento, identificação e interpretação de informações do ambiente – são relevantes para minimizar incertezas, reduzir o risco da tomada de decisão, entregar valor e competitividade (CAINELLI, 2018). Um dos pioneiros a destacar a importância do monitoramento estratégico foi Aguilar (1967). Ansoff (1975), em seguida, sugeriu um sistema de alerta antecipado às mudanças, destacando o fato de que a empresa precisa lidar com transformações não esperadas, para a sua continuidade e resiliência. Interessado não apenas em identificar ameaças, mas também em propor soluções, sugeriu conceitos e ferramentas para lidar com a turbulência do contexto organizacional (ROSSEL, 2012).

Uma dessas ferramentas é **Inteligência Estratégica Antecipativa (IEA)** que define um conjunto de práticas para monitorar o mercado e antecipar mudanças (AGUILAR, 1967; LESCA, 2003). A questão “estratégica” da IEA, simbolizada pela letra “E” da sigla, reforça que essas práticas fornecem informações de importância estratégica aos tomadores de decisão. Para Lesca (2003), eles não podem somente basear suas decisões em modelos comprovados na prática, pois tais situações ainda não ocorreram. Decisões estratégicas são tomadas em contexto de ambiguidade, no qual os gestores devem aceitar o risco envolvido pela incerteza (JANISSEK-MUNIZ; FREITAS; LESCA, 2007). A literatura enfatiza a necessidade de detectar estas mudanças no ambiente corporativo o mais cedo possível (ANSOFF, 1975; ROHRBECK; THOM; ARNOLD, 2015), sendo este, portanto, outro objetivo importante dos processos de inteligência antecipativa (LESCA, 2003).

Como os sinais do meio externo normalmente são encontrados de forma difusa, ambígua e fragmentada, a identificação de informações pertinentes não é simples. Já que são apenas indícios de eventos em formação, essas informações não são apresentadas prontas; elas precisam ser amplificadas, desenvolvidas e validadas (MENDONÇA; CARDOSO; CARAÇA, 2012). Na IEA, estas informações são chamadas **Sinais Fracos (SF)** (ANSOFF, 1975), definidos como indicadores antecipados de fenômenos de mudança e possuem características singulares como: incompletos, ambíguos, incertos e de difícil detecção (JANISSEK-MUNIZ; BORGES; BORTOLI, 2015; LESCA, 2003). Sua interpretação requer um método, preferencialmente coletivo e contínuo, para que possam fornecer, sistematicamente, valor estratégico (JANISSEK-MUNIZ; BORGES; BORTOLI, 2015).

Esta prática implica em ser capaz de identificar informações relevantes imersas em um grande “ruído” de dados (LESCA et al., 2015).

De fato, velocidade e dinamismo do ambiente, aliados ao progresso tecnológico cada vez mais acelerado, provocam uma “torrente” de dados para serem ingeridos pelos processos de inteligência. Devido à avalanche informacional, tomadores de decisão, sem métodos e ferramentas adequados, são incapazes de considerar todas essas informações (KELLER; VON DER GRACHT, 2014), inviabilizando o processo de compreensão e antecipação de eventos (MILOVIDOV, 2018). À medida que os dados se avolumam, tornando o processo de monitoramento mais complexo, a racionalidade limitada (SIMON, 1955) dos analistas restringe a interpretação desse ambiente desconhecido (EREVELLES; FUKAWA; SWAYNE, 2016). Mesmo com a abundância de dados, as informações pertinentes são difíceis de encontrar (DAVENPORT; BARTH; BEAN, 2012), o que interfere na eficiência da tomada de decisão (JIANG; GALLUPE, 2015). Da mesma forma, profissionais estão sendo afetados por este volume de dados para além da capacidade de compreensão e uso de suas organizações (DAY, 2011).

Com o apoio de algum dispositivo, ferramenta ou técnica, as capacidades do observador individual podem ser ampliadas (CONSTANTIOU; KALLINIKOS, 2015; KELLER; VON DER GRACHT, 2014; MENDONÇA; CARDOSO; CARAÇA, 2012). Dentre os motivos pelos quais as organizações não percebem as mudanças disruptivas de forma antecipada, Rohrbeck (2010) indica o excesso de informação que a alta gerência recebe, implicando na redução da capacidade de avaliar o impacto potencial dessas informações, e reforçando a dificuldade que as empresas têm de monitorar o ambiente. Gerenciar esse dilúvio de informações, remover ruídos e distrações indevidas enquanto buscam por SF passa a ser uma questão a ser investigada (JIANG; GALLUPE, 2015; SCHOEMAKER; DAY; SNYDER, 2013).

Por outro lado, autores afirmam que a disponibilidade de **dados em enormes volumes** também potencializa o desempenho da tomada de decisão organizacional (CHEN; PRESTON; SWINK, 2015; GHASEMAGHAEI; EBRAHIMI; HASSANEIN, 2018; GOES, 2014). Observa-se, para este fim, um movimento em direção à utilização de ferramentas de TI mais potentes (JIANG; GALLUPE, 2015). Diversos avanços nas tecnologias digitais permitem que as empresas desenvolvam formas inovadoras de coletar dados (DAVENPORT; DYCHÉ, 2013; DUAN; CAO, 2015) de fontes de informação internas e externas, das mais variadas formas, como dados da *internet*, conteúdo gerado por empresas

e consumidores, mídias sociais e e-mails, dados gerados por dispositivos móveis, por sistemas corporativos e por sensores em tecnologias como *Internet of Things* (IoT) (CHEN et al., 2017; GE; BANGUI; BUHNOVA, 2018; GOES, 2014; KITCHIN, 2013; SIVARAJAH et al., 2017). Estes são apenas alguns exemplos de fontes que disponibilizam uma quantidade sem precedentes de informações econômicas e sociais (GHASEMAGHAEI; EBRAHIMI; HASSANEIN, 2018; MILLER, 2010; MÜHLROTH; GROTTKE, 2018) constituindo o que agora é conhecido como **Big Data** (DUAN; CAO; EDWARDS, 2018).

Nas últimas décadas o *Big Data* consolidou-se como um tema importante na literatura (CHEN; CHIANG; STOREY, 2012; GÜNTHER et al., 2017). O termo, utilizado inicialmente na década de 90 (DIEBOLD, 2018), dá nome a sistemas que armazenam e processam enormes **volumes** de dados amplamente **variados**, que são gerados, capturados e processados em alta **velocidade** (CHEN; CHIANG; STOREY, 2012; GHASEMAGHAEI; EBRAHIMI; HASSANEIN, 2018; GÜNTHER et al., 2017; LANEY, 2001). O *Big Data* envolve grandes quantidades de dados, em diferentes formatos, que podem dar suporte a diversos tipos de decisões (GOES, 2014). A coleta e o processamento desses dados não são viáveis pelas tecnologias computacionais tradicionais (BREUKER et al., 2016; CORBETT; WEBSTER, 2015), pertinentes a bancos de dados relacionais e a pequenas quantidades de dados (CONSTANTIOU; KALLINIKOS, 2015). Entretanto, novas técnicas e ferramentas desenvolvidas permitiram o seu advento, que, em conjunto com a disciplina de ciência de dados (*data science*), se desenvolveu rapidamente na academia e nas empresas.

Para definir esse conjunto de ferramentas, envolvendo dados em grandes quantidades associados ao seu processamento e análise, a literatura utiliza o termo **Big Data Analytics (BDA)** (JANSSEN; VAN DER VOORT; WAHYUDI, 2017). Buscando transformar esses dados em valor, visando responder às mudanças de preferências dos consumidores, às novas tecnologias e aos movimentos dos concorrentes, bem como favorecendo a geração de conhecimento, as organizações estão aumentando seus investimentos em BDA (BARTON; COURT, 2012; CAESARIUS; HOHENTHAL, 2018; CHEN; CHIANG; STOREY, 2012; LAVALLE et al., 2011; PARK; EL SAWY; FISS, 2017; SALLEH, 2016). Ferramentas BDA não tratam apenas da análise, mas também da visualização e interpretação dos dados, com a finalidade principal do apoio à tomada de decisões (SHARMA et al., 2010). São métodos e técnicas analíticas, realizadas por meio de aplicativos complexos, que permitem compilar de forma inteligente dados estruturados e não estruturados (CHEN; CHIANG; STOREY, 2012; DAVENPORT; DYCHÉ, 2013; DUAN; CAO, 2015). Estudos da indústria também têm destacado o desenvolvimento do BDA (JIANG; GALLUPE, 2015). Lavallo

(2011), em sua pesquisa com mais de 3000 CIOs, mostra que organizações de maior performance utilizam *Big Data* cinco vezes mais do que as demais. De fato, muita atenção está sendo dada na literatura acadêmica e profissional ao valor que as organizações podem potencialmente obter com o uso do BDA. Chen et al. (2012), por exemplo, sugere que ele pode ajudar as organizações a entender melhor seu mercado.

### 1.1 Problemática e Questão de Pesquisa

Um benefício importante que o BDA oferece é a geração de *insights* (DREMEL, 2017; EREVELLES; FUKAWA; SWAYNE, 2016; KELLER; VON DER GRACHT, 2014; KITCHENS et al., 2018; LAM et al., 2017; MÜHLROTH; GROTTKE, 2018; SHARMA; MITHAS; KANKANHALLI, 2014; YOON, 2012), definidos como compreensões importantes, percebidas em determinadas situações, que podem melhorar o resultado da organização (JIANG; GALLUPE, 2015; SCHULTE; HOVORKA, 2017). Ao serem transformados em conhecimento mediante interpretações associativas, fornecem benefícios como conhecer melhor os clientes, apoiar a inovação, e reduzir a incerteza na tomada de decisão estratégica (CHEN; PRESTON; SWINK, 2015; DAVENPORT; BARTH; BEAN, 2012; DUAN; CAO, 2015; GÜNTHER et al., 2017; KITCHENS et al., 2018). Porém, em um ambiente socioeconômico dinâmico contemporâneo, não basta somente colher *insights* a partir de padrões e tendências, calculados sobre dados internos. Este tipo de análise e resultado seria suficiente se houvesse apenas a evolução incremental de tendências históricas, mas não é bem-sucedido ao lidar com surpresas, **eventos disruptivos**, fenômenos que podem mudar o mercado rapidamente (ANSOFF, 1975; JANISSEK-MUNIZ; BORGES; BORTOLI, 2015; ROSSEL, 2012).

Neste contexto organizacional de necessidade de monitoramento com disponibilidade de dados em múltiplas fontes, um dos objetivos do BDA é o de ser um sistema de relacionamento com o mercado (DUAN; CAO; EDWARDS, 2018; THORLEUCHTER; VAN DEN POEL, 2015). Ele é um dos meios que proporcionam sua observação (JIANG; GALLUPE, 2015) e compreensão (KAYSER; BLIND, 2017; KELLER; VON DER GRACHT, 2014), introduzindo mudanças em como a informação pode ser utilizada pelas empresas (CONSTANTIOU; KALLINIKOS, 2015) para obter uma visão abrangente (LAM et al., 2017). Para Brynjolfsson e Reichman (2016), o *Big Data* oportuniza monitorar e modelar intenções, preferências e opiniões para prever fenômenos econômicos. Seu desenvolvimento permite não apenas projetar tendências ou obter *insights* a partir de dados estruturados internos, mas também identificar Sinais Fracos a partir dados relacionados a

fenômenos político-econômicos, sociais, mercadológicos e tecnológicos (KIM et al., 2013; MILLER, 2010; MÜHLROTH; GROTTKE, 2018; YOON, 2012).

Neste sentido, interroga-se a respeito da possibilidade de que os métodos de identificação de Sinais Fracos e criação de sentido da inteligência estratégica antecipativa (IEA), que inicialmente prevê a captura de sinais apenas em base qualitativa, possam ser atualizados para comportar igualmente essas novas ferramentas, que, em geral, são mais quantitativas. Busca-se, portanto, nesta pesquisa, explorar a possibilidade de busca de SF a partir das ferramentas BDA. Erevelles, Fukawa e Swayne (2016), em estudo sobre o impacto do BDA em atividades de marketing, afirmam que a análise do consumidor está no centro da revolução dos dados e que a tecnologia permite a coleta de dados importantes sobre os fenômenos. Os autores mostram, em seu *framework* teórico, que empresas podem ser proativas às mudanças no ambiente, capturando SF por meio de ferramentas BDA, potencializando o protagonismo no mercado que permite aumentar a competitividade e reduzir a necessidade de reagir quando uma tendência já estiver sedimentada pelos concorrentes.

Os *insights* obtidos por meio do BDA não emergem automaticamente dos algoritmos desenvolvidos. Geralmente são fruto de um sistema de criação de sentido estruturado, executado por analistas e tomadores de decisão, apoiados ou guiados por dados e ferramentas analíticas (CONSTANTIOU; KALLINIKOS, 2015; JIANG; GALLUPE, 2015; MCAFEE; BRYNJOLFSSON, 2012; SHARMA; MITHAS; KANKANHALLI, 2014). A geração de *insights* é a última etapa do processo de análise dos dados, um passo imprescindível para a antecipação dos cenários futuros, envolvendo a transformação das informações em percepções e decisões, valiosas para a organização se forem implementadas (DUTTA; BOSE, 2015). Estudos sugerem que melhores decisões são tomadas quando gestores estão embasados nos dados e não apenas em intuições e experiências (DAVENPORT, 2006; DUTTA; BOSE, 2015; LAVALLE et al., 2011; LYYTINEN; GROVER, 2017; SHARMA; MITHAS; KANKANHALLI, 2014).

Considerando esse universo de dados, investiga-se se o BDA pode contribuir para processos de inteligência antecipativa mais eficientes e precisos. Assim, a **problemática deste estudo** envolve dois principais elementos: a Inteligência Estratégica Antecipativa e a análise de grandes volumes de dados. A proposta é entender o uso potencial de ferramentas de BDA para a identificação de Sinais Fracos, contribuindo para que elas se encaixem como uma nova engrenagem nos processos de monitoramento de ambiente e de IEA. Pretende-se

estudar novas oportunidades para a área da Inteligência Estratégica Antecipativa e tomada de decisão, alavancadas pelas ferramentas *Big Data Analytics*, compreendendo como o uso destas ferramentas pode facilitar a percepção de SF e promover a tomada de decisão estratégica. Assim, a questão que orienta este estudo é: **qual o impacto das práticas do BDA na identificação de Sinais Fracos?**

## 1.2 Objetivos

O presente estudo tem por objetivo geral **analisar a relação entre as práticas do BDA e a identificação de Sinais Fracos**.

Para atingir o objetivo geral, os seguintes objetivos específicos serão guia deste estudo:

- Analisar a relação existente na literatura entre os Sinais Fracos e os tipos de informações obtidas nos processos de inteligência usando BDA.
- Investigar se as organizações percebem Sinais Fracos em análises utilizando BDA.
- Identificar as variáveis relacionadas ao BDA que impactam na identificação de Sinais Fracos, propondo um modelo de pesquisa.
- Analisar o modelo proposto e suas hipóteses, mensurando o efeito das variáveis relacionadas ao BDA na identificação de Sinais Fracos.

## 1.3 Justificativa

Esta pesquisa é justificada pela **mudança contextual nos processos de tomada de decisão** (ABBASI; SARKER; CHIANG, 2016; FRISK; BANNISTER, 2017; GOES, 2013; MCAFEE; BRYNJOLFSSON, 2012; SHARMA; MITHAS; KANKANHALLI, 2014) e **de inteligência** (CONSTANTIOU; KALLINIKOS, 2015; EREVELLES; FUKAWA; SWAYNE, 2016; JIANG; GALLUPE, 2015; LYCETT, 2013; MÜHLROTH; GROTTKE, 2018), gerada pelo **grande volume de informações disponíveis** aos analistas e gestores, acompanhado pelo **advento do BDA** (CONSTANTIOU; KALLINIKOS, 2015). Há um ambiente em transição na economia digital, que desafia os conceitos tradicionais de gerenciamento dos negócios (GEORGE; HAAS; PENTLAND, 2014). Essa transformação disruptiva está em andamento, principalmente na forma de como as organizações administram sua estratégia (DUTTA; BOSE, 2015; RAGUSEO, 2018), visto que os gestores, cada vez mais, tomam decisões baseadas em dados (DAVENPORT, 2006; DUAN; CAO; EDWARDS, 2018; LAVALLE et al., 2011). De fato, o *Big Data* abriu novas possibilidades para o suporte à decisão (GHASEMAGHAEI; EBRAHIMI; HASSANEIN,

2018; MIAH et al., 2017). Constantiou e Kallinikos (2015) afirmam que, a partir da sedimentação acadêmica e prática do BDA, **as práticas de geração de sentido nos processos de inteligência serão alteradas**. Para os autores, ademais, o *Big Data* introduz importantes **mudanças em como a informação é produzida** e torna-se relevante para as organizações. Essas mudanças apresentam implicações nas práticas de obtenção e uso das informações, ou seja, nos processos de inteligência, a partir dos quais a tomada de decisão é realizada. Assim, o BDA está totalmente ligado à forma de como as organizações percebem e atuam em seus ambientes internos e externos e, então, aos processos de inteligência e estratégia organizacional.

Contudo, a maior parte dos estudos que tratam do BDA concentram-se em capacidades organizacionais táticas (GROVER et al., 2018), pouco abordando sua utilização como apoio à decisão estratégica (VAJJHALA; STRANG; SUN, 2015). A exemplo, a pesquisa de Ghasemaghaei, Ebrahimi e Hassanein (2018) dirimiu um *gap*, operacionalizando e validando o conceito de competência de *data analytics*. Para os autores, não havia estudos anteriores com esta problemática, que avalia o impacto do uso eficaz de *data analytics* (ou BDA) no desempenho da tomada de decisão empresarial.

O foco deste estudo é entender como o uso destas novas ferramentas, suas características e práticas, podem impactar na estratégia e na inteligência antecipativa das organizações. Embora existam pesquisas endereçadas a entender e categorizar BDA, seus métodos de análise e formas de entregar valor (GÜNTHER et al., 2017), poucas os relacionam às noções de Sinais Fracos (ECKHOFF et al., 2014). Esses poucos trabalhos, em sua grande maioria, tratam de aspectos específicos e tecnológicos, sem uma visão, modelo ou *framework* abrangente (ECKHOFF et al., 2014; MÜHLROTH; GROTTKE, 2018). Portanto, a literatura sobre detecção de Sinais Fracos em BDA não oferece uma visão geral de como as tecnologias podem ser usadas (ECKHOFF et al., 2014).

Jiang e Gallupe (2015) sugerem que o BDA está relacionado à obtenção de *insights* e inovações, promovendo e alavancando melhores resultados para a organização. Os autores afirmam que há *gaps* entre a implementação de ferramentas BDA e a entrega eficiente de *insights* estratégicos. Eckhoff et al. (2014), em revisão sistemática na literatura, procuraram casos relevantes descrevendo a detecção de Sinais Fracos por meio da tecnologia. Na sua busca, realizada a partir de 2009, o conjunto de resultados colhido foi pequeno: comparou-se sete abordagens documentadas e apenas cinco ofereceram informações relevantes, descrevendo casos concretos. Os autores consideraram o estudo como um ponto de partida

para *surveys*, a realizar em profissionais e especialistas, para obter dados empíricos a respeito dessas práticas de inteligência. Identifica-se nestas pesquisas, portanto, este *gap* na literatura a respeito da percepção, em um BDA, de possíveis Sinais Fracos, conforme conceituados na tradição de Ansoff (1975).

Neste trabalho será investigado como os processos de inteligência podem ser executados na presença deste “dilúvio” de dados e das ferramentas BDA, considerando o contexto da inteligência e da estratégia organizacional e a forma como as empresas monitoram, decidem e agem. No sentido de buscar respostas, encontra-se a revisão de literatura de Günther et al. (2017), que trata da obtenção de valor a partir do *Big Data*. Foram identificados debates centrais sobre como as organizações percebem o valor retornado. A autora sugere que a literatura tangencia o assunto, não tratando diretamente do relacionamento entre *Big Data* e Inteligência Antecipativa. Adicionalmente, Liang e Liu (2018), em recente estudo panorâmico e bibliométrico, apresentam as principais tendências de pesquisas acadêmicas sobre *Big Data*, identificando como os resultados acadêmicos relacionados a *Big Data* e *Business Intelligence* cresceram e evoluíram na última década. O artigo sugere orientações de pesquisas futuras, contudo, mesmo citando gestão do conhecimento, tomada de decisão e sistemas de apoio a decisão, nenhum tópico comentado nesta ampla pesquisa está relacionando *Big Data* diretamente com inteligência antecipativa e Sinais Fracos.

Chen, Chiang e Storey (2012), no artigo “*Business Intelligence and Analytics: From Big Data to Big Impact*” descrevem a evolução do BDA, ilustrando aplicações e impactos desta área de pesquisa. Neste estudo são apontados trabalhos que tratam de aplicações de *Big Data* em inteligência de mercado e *environmental scanning*, através de técnicas de: *text and web analytics*, *text and network analytics*, *network visualization*, *sentiment and affect analysis* e *relation mining*. O *environment scanning on-line* relatado oferece aos executivos oportunidades de usar a “inteligência da *web*” para desenvolver *insights* sobre fatores socioculturais e político-econômicos, com efeito no aumento das vendas, satisfação do cliente e principalmente na decisão estratégica em fusões e aquisições.

Devido à crescente quantidade de dados disponível, métodos e sistemas auxiliados por computador para identificar Sinais Fracos estão cada vez mais sendo necessários (ECKHOFF et al., 2014; KELLER; VON DER GRACHT, 2014; MÜHLROTH; GROTTKE, 2018; YOON, 2012). A tendência é de que os processos de inteligência antecipativa sejam, de forma crescente, implementados e apoiados por ferramentas baseadas em SI. Eles serão uma das colunas do desenvolvimento da inteligência, tanto para sua

eficiência como para eficácia (KELLER; VON DER GRACHT, 2014). Lycett (2013) argumenta que o BDA permite que analistas e tomadores de decisão se engajem em um processo de criação coletiva de sentido baseado em dados, sendo utilizado como meio para entender os fenômenos do ambiente. É razoável esperar que exista uma relação entre o uso do BDA e melhores *insights* e decisões, mas não está claro como esses resultados podem ser obtidos. Na literatura não são encontrados estudos que demonstrem o que é preciso priorizar em uma implementação de BDA, para que o monitoramento do ambiente e a identificação de Sinais Fracos sejam mais eficientes. Assim, acredita-se que a relevância desta pesquisa esteja neste *gap*, que correlaciona construtos que representam práticas do BDA à essa variável da inteligência antecipativa.

Para obter um maior retorno no uso do BDA, é importante utilizá-lo como apoio para a tomada de decisão (BUMBLAUSKAS et al., 2017; FRISK; BANNISTER, 2017; JANSSEN; VAN DER VOORT; WAHYUDI, 2017). É preciso entender o papel dos dados como um fator crítico de sucesso para decidir com menores riscos e maior discernimento, adquirindo uma cultura orientada por dados (*data-driven culture*) (DREMEL, 2017). Abbasi, Sarker e Chiang (2016), propuseram uma agenda de pesquisa argumentando que pesquisadores devem repensar, em razão do advento *Big Data*, como os dados são coletados, preparados, analisados e apresentados para os processos de tomada de decisão. É necessário, portanto, entender como o processo decisório está sendo transformado e que tipo de mudança cultural precisa ser realizada para que as iniciativas de *Big Data* sejam eficazes. Eles apresentaram as seguintes áreas de investigação para o tópico de pesquisa “natureza da tomada de decisão”: “Como as organizações / indivíduos / grupos realmente tomam decisões no ambiente de *Big Data*? Até que ponto os modelos tradicionais de tomada de decisão se sustentam no novo ambiente?” (ABBASI; SARKER; CHIANG, 2016, p. 12). Os autores também apontam a seguinte pergunta ainda não tratada na literatura, classificando como fundamental, que é abordada neste estudo: “Os processos de tomada de decisão, realizados nos vários níveis da organização, estão sendo transformados devido ao *Big Data*? Em caso afirmativo, como esta transformação ocorre?” (ABBASI; SARKER; CHIANG, 2016, p. 10). Os autores apontam a necessidade de pesquisas que demonstrem o resultado dos investimentos em *Big Data* em relação a diversos aspectos da tomada de decisão.

Assim, este trabalho, tratando da relação do BDA com a Inteligência Estratégica Antecipativa (IEA), contribui com esta literatura, pesquisando possibilidades de retorno de valor para tomada de decisão estratégica. Neste mesmo sentido, Sharma, Mithas e Kankanhalli (2014) citam o seguinte *gap* na literatura: “Como o uso do *business analytics*

influencia os processos de tomada de decisões organizacionais?” (SHARMA; MITHAS; KANKANHALLI, 2014, p. 434). Vajjhala, Strang e Sun (2015), também reafirmam este *gap*, destacando que a maioria dos estudos empíricos em *Big Data* explora comportamentos específicos.

O *Big Data* e o BDA são temas importantes de pesquisas (CHEN; CHIANG; STOREY, 2012), tanto na academia como nas empresas. Contudo, o termo permanece confuso e pode tornar-se uma “*buzzword*” (POSPIECH; FELDEN, 2013), inibindo novos estudos, uma vez que o campo de pesquisa não está bem estruturado. Assim, percebe-se que a demanda de uma base teórica na literatura de SI existe e é importante, já que tópicos relevantes ainda são desconhecidos. Neste sentido, Sharma, Mithas e Kankanhalli (2014) sugerem que, mesmo havendo evidências de que investimentos em BDA podem gerar valor, esta afirmação ainda precisa de uma análise mais profunda. Já para Raguseo (2018), a literatura oferece poucos estudos empíricos com evidências sobre benefícios associados ao uso de *Big Data*. Embora o termo tenha se tornado comum pelos praticantes, há poucos estudos publicados que explorem as possíveis contribuições teóricas e práticas do *Big Data* (GEORGE; HAAS; PENTLAND, 2014).

#### **1.4 Delineamento da Pesquisa e Estrutura do trabalho**

Para cumprir os objetivos desta investigação de natureza exploratória, optou-se por estruturá-la em dois artigos, atendendo sua problemática por meio de abordagens qualitativa e quantitativa na produção do conhecimento. Mesmo possuindo escopos individuais, os artigos foram planejados de modo que, em conjunto, respondam à questão de pesquisa delimitada e atendam os objetivos gerais e específicos deste estudo.

Ambos os artigos tratam dos temas contextualizados no Capítulo 1, discorrendo em relação à identificação de SF por meio do BDA. Portanto, o referencial teórico deste trabalho desenvolveu-se através dos dois artigos, de modo que eles comungam os mesmos tópicos: SF, IEA, *Big Data Analytics*, e as relações entre essas disciplinas da área de Sistemas de Informação. No Artigo II foram detalhados aspectos, ou práticas, do BDA, ilustrando conceitos tratados pela academia, com a finalidade de desenvolver construtos e hipóteses para um novo modelo teórico. São eles: *Data Lakes*, dados não estruturados, *Advanced Analytics*, Análise de *Outliers* e o Método Indutivo. Houve, ainda, a necessidade de elaborar o construto da cultura guiada por dados (*Data-driven Culture*), visto que é um antecedente relevante.

Os artigos possuem abordagens metodológicas diferentes. No primeiro entendeu-se necessário um estudo qualitativo, no intuito de pesquisar de maneira detalhada a viabilidade da identificação de SF por meio do BDA. Inicia-se esta investigação, portanto, com um estudo empírico que parte da revisão de literatura e utiliza o método de entrevistas em profundidade. Ao explorar se os *insights* colhidos no BDA podem ser considerados SF, identificou-se as possíveis variáveis, práticas do BDA, que influenciam o monitoramento do ambiente e a identificação de SF. Foram apresentadas evidências de que dados candidatos a SF podem ser identificados a partir do BDA e de que essas ferramentas facilitam o monitoramento do ambiente.

Após o enfoque qualitativo, no Artigo II optou-se pela abordagem quantitativa. Empregou-se instrumentos de medida para analisar dados abstratos da realidade. Eles formaram o elo entre o modelo e a observação do campo (HOPPEN; LAPOINTE; MOREAU, 1996), permitindo a coleta de dados, mediante de artefatos de pesquisa e ferramentas estatísticas, para suportar hipóteses deduzidas (SAMPIERI; COLLADO; LUCIO, 2006). Assim, foram definidos construtos com o objetivo de medir os impactos de práticas do BDA na identificação de SF. Oriundos da pesquisa qualitativa e do referencial teórico, eles foram submetidos a especialistas de IEA para a sua homologação, a fim de desenvolver o Modelo Conceitual de Pesquisa, objeto da fase quantitativa deste estudo. Na etapa seguinte, desenvolveu-se os artefatos para a pesquisa quantitativa. O propósito, pautado pela fundamentação teórico-conceitual e pela pesquisa exploratória, foi de suportar o Modelo Conceitual indicado e suas hipóteses, apontando e mensurando as variáveis relacionadas ao BDA que impactam na identificação de SF. Dentro do paradigma positivista quantitativo, realizou-se uma pesquisa do tipo *survey*, baseado em um questionário *on-line* estruturado de perguntas. Para analisar a relação entre as práticas do BDA e a identificação de SF, utilizou-se a modelagem de equações estruturais (SEM), com mínimos quadrados parciais de regressão (*Partial Least Squares* ou *PLS-SEM*), que viabiliza examinar relacionamentos entre múltiplas variáveis (HAIR et al., 2009).

Na Tabela 1 apresenta-se um resumo, relacionando os objetivos específicos com a pergunta de pesquisa de cada artigo.

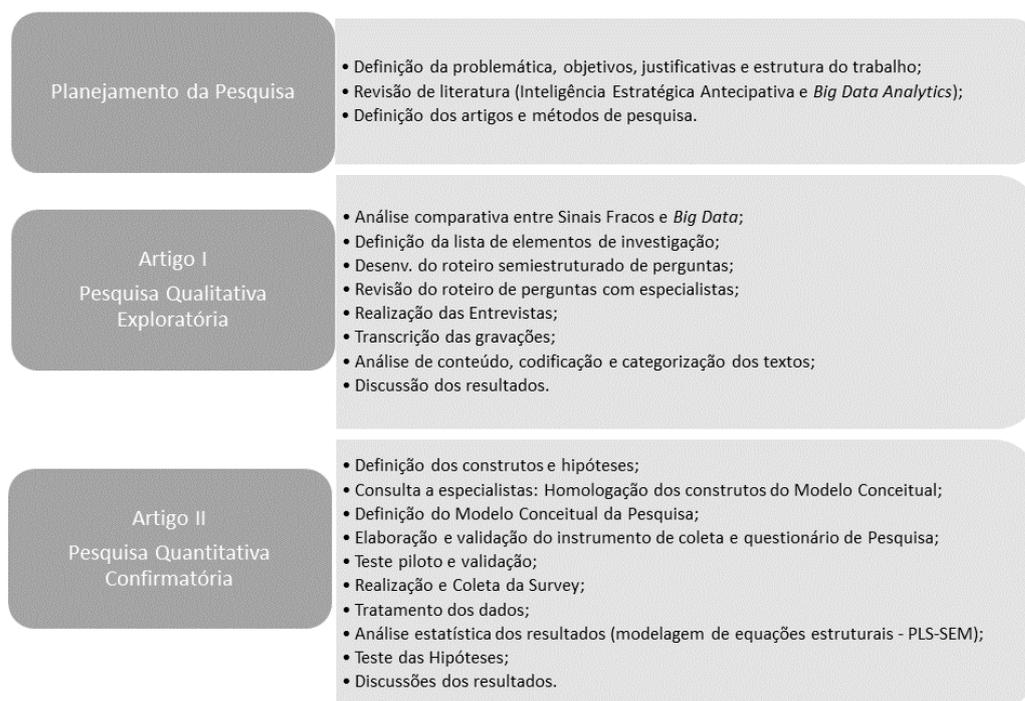
Tabela 1: Resumo da Pesquisa

<b>Título da dissertação:</b>	
Inteligência Estratégica Antecipativa: Identificação de Sinais Fracos por meio do <i>Big Data Analytics</i> .	
<b>Questão de pesquisa:</b>	
Qual o impacto das práticas do BDA na identificação de Sinais Fracos?	
<b>Objetivo Geral:</b>	
Analisar a relação entre as práticas do BDA e a identificação de Sinais Fracos.	
<b>Objetivos Específicos atendidos pelo Artigo I:</b>	<b>Questão de pesquisa Artigo I:</b>
1. Analisar a relação existente na literatura entre os SF e os tipos de informações obtidas nos processos de inteligência usando BDA.	Gestores percebem SF por meio do BDA? Esses sinais são utilizados na tomada de decisão estratégica?
2. Investigar se as organizações percebem SF em análises usando BDA.	
<b>Objetivos Específicos atendidos pelo Artigo II:</b>	<b>Questão de pesquisa Artigo II:</b>
3. Identificar as variáveis relacionadas ao BDA que impactam na identificação de SF, propondo um modelo de pesquisa	Qual o efeito das práticas do BDA na identificação de Sinais Fracos?
4. Analisar o modelo proposto e suas hipóteses, mensurando o efeito das variáveis relacionadas ao BDA na identificação de SF.	

Fonte: elaborado pelo autor

A Figura 1 ilustra o desenho da pesquisa, plano metodológico que guia o desenvolvimento deste trabalho.

Figura 1: Desenho da pesquisa



Fonte: elaborado pelo autor

## 2 Artigo I - Explorando usos potenciais do *Big Data* para a Inteligência Antecipativa

*Fabiano Chiapinotto Saffi*

### Resumo

*Big Data* tem se desenvolvido significativamente na academia e nos negócios, permitindo que empresas adotem formas inovadoras de coletar e analisar grandes volumes de dados, em variadas formas e origens. Para obter maior valor, organizações estão investindo em ferramentas *Big Data Analytics* (BDA), buscando aprimorar sua capacidade de entender e antecipar as mudanças no ambiente. Ao considerar esses dados em práticas de inteligência, usando o BDA como um mecanismo de relacionamento com o mercado, analistas potencializam a identificação de Sinais Fracos (sinais antecipativos). Este estudo empírico, realizado mediante entrevistas em profundidade, objetiva compreender se os gestores identificam Sinais Fracos por meio das ferramentas BDA, bem como, o seu uso na tomada de decisão estratégica. Como resultado, apresentou-se práticas do BDA que influenciam o monitoramento e a identificação de Sinais Fracos, com evidências de que dados candidatos a Sinais Fracos podem ser identificados a partir do BDA e de que essas ferramentas facilitam o monitoramento do ambiente.

**Palavras-chave:** *Big Data*; Inteligência Antecipativa; *Big Data Analytics*; Sinais Fracos.

### 2.1 Introdução

O *Big Data* surgiu como uma área de pesquisa significativa nas últimas duas décadas, com impulso expressivo em comunidades acadêmicas e de negócios (CHEN; CHIANG; STOREY, 2012; GÜNTHER et al., 2017). Este campo de estudos se tornou uma nova fronteira para oportunidades, habilitadas pela tecnologia e permitidas pela revolução da informação (DUAN; CAO, 2015; GOES, 2014). O termo *Big Data* foi utilizado pela primeira vez em 1998, referindo-se ao armazenamento de conjuntos de dados massivos e amplamente variados, que são gerados, capturados e processados em alta velocidade (CONSTANTIOU; KALLINIKOS, 2015; DIEBOLD, 2018; LANEY, 2001).

Avanços em tecnologias emergentes viabilizaram, para as organizações, a coleta destes dados em fontes internas e externas (DUAN; CAO, 2015; GOES, 2014; KITCHIN, 2013; SIVARAJAH et al., 2017). Para transformá-los em valor, elas estão investindo em ferramentas *Big Data* e *Business Analytics* (BARTON; COURT, 2012; LAVALLE et al., 2011), que proporcionam a análise e a visualização, oportunizando a interpretação dos dados

para o apoio à tomada de decisões (SHARMA et al., 2010). São métodos e técnicas analíticas realizadas por aplicativos complexos, que permitem compilar os dados de forma inteligente (CHEN; CHIANG; STOREY, 2012; DAVENPORT; DYCHÉ, 2013). A este conjunto de ferramentas, a literatura usa o termo *Big Data Analytics* (**BDA**) (JANSSEN; VAN DER VOORT; WAHYUDI, 2017), voltando a atenção ao valor potencial criado por meio do seu uso. A exemplo, Chen et al. (2012) sinaliza que o BDA ajuda as organizações a entender melhor seus negócios e mercados dinâmicos, alavancando oportunidades e expondo possíveis ameaças.

Por outro lado, a manutenção da competitividade e da resiliência das organizações depende, em grande parte, da capacidade de entender e antecipar as mudanças no ambiente (ILMOLA; KUUSI, 2006; LESCA et al., 2015; ROSSEL, 2009). Para reduzir a incerteza na tomada de decisão e alinhar a estratégia às transformações rápidas do mercado (CHOO, 2001), é necessário um processo proativo de monitoramento do ambiente. Ele não pode ser realizado de forma passiva (MENDONÇA; CARDOSO; CARAÇA, 2012), para ser efetivo, necessita práticas sistemáticas formalizadas (CAINELLI; JANISSEK-MUNIZ, 2019; FONSECA; BARRETO, 2011). Esta escuta “ativa” oportuniza que as empresas se posicionem considerando as informações encontradas (JANISSEK-MUNIZ; FREITAS; LESCA, 2007; KIM; LEE, 2017). Como elas chegam de forma difusa e fragmentada, a detecção de informações relevantes não é simples, sendo dificilmente apresentadas aos decisores. Os dados precisam ser amplificados e interpretados (MENDONÇA; CARDOSO; CARAÇA, 2012), já que são evidências de fenômenos que ainda estão em formação.

Em Inteligência Antecipativa tais informações são chamadas Sinais Fracos (ANSOFF, 1975). Sinais Fracos (SF) são pequenas mensagens reconhecidas como indícios antecipados de fenômenos de mudança (JANISSEK-MUNIZ; FREITAS; LESCA, 2007) e possuem características singulares, como: são incompletos, ambíguos e incertos, portanto, são difíceis de descobrir (LESCA, 2003). Sua interpretação requer uma análise metodológica e preferencialmente coletiva para que possam potencialmente fornecer valor estratégico (JANISSEK-MUNIZ; BORGES; BORTOLI, 2015). Esta prática implica em ser capaz de coletar e selecionar informações significativas, imersas em um grande “ruído” de dados, para depois analisá-las, adaptando a organização aos sinais encontrados (LESCA; JANISSEK-MUNIZ, 2015). As maneiras pelas quais as organizações capturam sinais ambientais e fomentam sua tomada de decisão são questões importantes de pesquisa e essenciais para a vantagem competitiva (JIANG; GALLUPE, 2015).

Em razão da possibilidade de coleta de dados digitalizados em fontes externas, considera-se o BDA como um mecanismo que aprimora o monitoramento do ambiente empresarial (DUAN; CAO; EDWARDS, 2018; JIANG; GALLUPE, 2015; THORLEUCHTER; VAN DEN POEL, 2015). Ele viabiliza o entendimento do mercado (KELLER; VON DER GRACHT, 2014; MÜHLROTH; GROTTKE, 2018) e a identificação de SF, considerando que os conjuntos de dados coletados contêm informações político-econômicas, sociais, mercadológicas e tecnológicos (MÜHLROTH; GROTTKE, 2018; YOON, 2012). Pretende-se, neste estudo, **investigar se os gestores identificam Sinais Fracos por meio ferramentas BDA, e se estes sinais são utilizados na tomada de decisão estratégica**. Como enquadramento metodológico, optou-se por uma pesquisa qualitativa, coletando dados a partir de entrevistas baseadas em um roteiro de perguntas semiestruturado.

Esta pesquisa é justificada por uma mudança de contexto na área da inteligência antecipativa gerada pelo advento do BDA, ainda não tratada de forma extensa na literatura de Sistemas de Informação (SI). Em contrapartida a este desenvolvimento na academia e nas empresas, BDA também deve ser considerado como um tópico de pesquisa emergente. O termo permanece confuso e pode tornar-se muito genérico, utilizado como uma “*buzzword*” (WHITE, 2012). Percebe-se que a demanda de uma base teórica na literatura de SI existe, já que tópicos importantes e teorias ainda são desconhecidos (POSPIECH; FELDEN, 2013). Embora existam pesquisas endereçadas a entender e categorizar BDA, seus métodos de análise e formas de entregar valor (GÜNTHER et al., 2017), poucas estudos empíricos o relacionam aos Sinais Fracos (ECKHOFF et al., 2014). Depara-se, portanto, com uma lacuna na literatura a respeito da percepção, em um BDA, de possíveis SF.

Este trabalho está organizado de forma a apresentar, após esta introdução, o referencial teórico da pesquisa. Após, descreve-se o método, seguindo com os resultados encontrados e discussões. Por fim, apresenta-se conclusões e considerações finais.

## **2.2 Referencial Teórico**

*Small Data* é qualquer conjunto de dados coletado e analisado sem a velocidade e a exaustividade proporcionada pelas ferramentas *Big Data*, portanto, a partir de técnicas de amostragem (KITCHIN; LAURIAULT, 2015). Para Lam et al. (2017), *Small Data* está no outro lado do espectro dos 3Vs: possui volume muito menor, é gerado vagarosamente e contém pouca variedade. Antes do *Big Data*, a construção do conhecimento se dava a partir dos resultados de experiências ou observações em escala humana, limitados em relação ao seu volume, com pouca variedade e que, geralmente, foram planejados para responder

questões específicas, com foco na causalidade e na compreensão da realidade (FARAWAY; AUGUSTIN, 2018). Estudos baseados em *Small Data* são adequados para responder perguntas ou hipóteses deduzidas (KITCHIN; LAURIAULT, 2015) e para explorar em detalhes, de forma aprofundada, a problemática da pesquisa (LINDSTROM, 2016). Com o emprego de abordagens *Small Data* pode-se focar em casos específicos, estudando histórias individuais, com suas nuances e contextos. *Small Data*, nesta visão, está relacionado a *insights* obtidos a partir de pequenos dados coletados em observações diretas no ambiente. *Small Data* é, portanto, qualquer tipo de dado que não é considerado como *Big Data*. Podem ser dados de uma pesquisa do tipo *survey*, pode ser qualquer pesquisa de mercado. Podem ser dados de uma entrevista qualitativa (KITCHIN; LAURIAULT, 2015). São considerados como *Small Data*, também, *insights* obtidos em visitas, contatos com pessoas, etnografias, percepções (LINDSTROM, 2016). Um censo nacional, por exemplo, possui um grande volume de dados, contudo, também é considerado como *Small Data*, considerando que seus dados são bem estruturados e coletados de forma lenta (KITCHIN, 2014).

Historicamente, os Sinais Fracos são encontrados por meio dos processos de monitoramento do ambiente, em dados do tipo *Small Data*, já que o *Big Data Analytics* se desenvolveu, na academia e na indústria, apenas após a primeira década deste século (WATSON, 2017). Também afirma-se que Sinais Fracos são tipicamente coletados em *Small Data* porque seu processo de identificação inicia na percepção de um estímulo (FONSECA; BARRETO, 2011) registrado na interação da organização com seu ambiente (JANISSEK-MUNIZ; FREITAS; LESCA, 2007). Essa percepção, geralmente, é realizada de forma direta por colaboradores voluntários selecionados, nas suas observações em clientes, fornecedores, parceiros, governos, entre outros (LESCA, 2003). No método *L.E.SCA*<sup>®</sup> que operacionaliza a IEA, estes sinais podem ser identificados na busca ativa por informações em um contexto específico, motivadas por atores e temas que foram modelados na fase anterior do processo, ou em possíveis alvos identificados a posteriori (JANISSEK-MUNIZ; LESCA; FREITAS, 2006). Assim, os Sinais Fracos estão relacionados a *insights* obtidos a partir de pequenos dados coletados – *Small Data* – em observações diretas no ambiente, interno ou externo.

O BDA, por outro lado, ao disponibilizar uma grande quantidade de dados juntamente com suas ferramentas de análise e visualização, proporciona a geração de *insights* (JIANG; GALLUPE, 2015; LAM et al., 2017) guiados por dados, que oferecem vantagens competitivas (CAESARIUS; HOHENTHAL, 2018; SALLEH, 2016), como o aumento a capacidade organizacional para a tomada de decisões estratégicas (CHEN;

PRESTON; SWINK, 2015) e a inovação de novos produtos e serviços (DUAN; CAO, 2015). Portanto, para melhor compreender esta mudança de contexto na área da inteligência estratégica antecipativa, para desenvolver a temática deste estudo e embasar a pesquisa empírica, aprofundou-se o conhecimento em BDA, na Teoria dos Sinais Fracos e na relação entre estas duas áreas. Apresenta-se nas próximas seções as principais referências analisadas.

### **2.2.1 *Big Data Analytics***

O conceito dos sistemas de apoio à decisão em larga escala (*Decision Support Systems - DSS*) surgiu em torno dos anos de 1970, tendo sua adoção difundida nas organizações no início da década de 1980. Esses sistemas davam suporte à algumas tarefas pertinentes à tomada de decisão (WATSON, 2017). Posteriormente foram criados os sistemas de informação executiva (*Executive Information System - EIS*), um conjunto recursos integrado que proporcionaram, entre outros benefícios, relatórios, consultas, modelagem, análises financeiras e estatísticas (SPRAGUE, 1980). Após, nos anos 1990, surgiu o *Data Warehousing*, oportunizando uma nova geração de sistemas classificados como *Business Intelligence (BI)* (CHEN; CHIANG; STOREY, 2012), termo abrangente que inclui aplicativos e práticas. Eles oferecem acesso e análise à informações, com o objetivo de melhorar o resultado e a tomada de decisão (LARSON; CHANG, 2016). Em seguida, nos anos 2000, foi possível o armazenamento de dados em tempo real e, nas últimas duas décadas, despontou o conceito de *Big Data Analytics (BDA)*. Em cada uma dessas fases foram promovidos avanços significativos na oferta de informações para a decisão (DAVENPORT, 2006; PLADS; ROBERT; STREET, 2013; WATSON, 2017).

O objeto deste trabalho são as ferramentas BDA e suas características que estão mudando as práticas da inteligência estratégica. Chen et al. (2012) em seu artigo seminal apresentaram um *framework* que distinguiu a evolução destas tecnologias ao longo do tempo, descrevendo-as em termos das suas características e capacidades. Na literatura de sistemas de informação, BDA foi definido de inúmeras maneiras, por vários autores. É um termo ainda ambíguo, associado à coleta e análise de grandes conjuntos de dados (BUMBLAUSKAS et al., 2017). Chen et al. (2012), porém, mostraram que há um uso comum deste termo. Segundo os autores, BDA é utilizado para conceituar um conjunto de ferramentas que lidam com grandes volumes de dados exigindo armazenamento avançado, gerenciamento, tecnologias de visualização e, principalmente, novas formas de análises estatísticas.

O crescimento exponencial do BDA ocorre principalmente em razão do desenvolvimento de tecnologias e processos, e sua rápida incorporação às práticas organizacionais. Essas novas tecnologias são importantes tanto na geração dos dados, como na sua coleta e processamento (KITCHIN, 2013). O declínio nos custos de processamento computacional, da memória, da capacidade de armazenamento e as novas técnicas de manipulação de dados em arquivos distribuídos como *Hadoop* (VILLARS; OLOFSON; EASTWOOD, 2011), também contribuíram de forma decisiva para a difusão destas ferramentas. Para definir *Big Data* de forma mais específica, Laney (2001) inicialmente determinou os “3Vs” e boa parte da literatura (DIEBOLD, 2018; JEBLE; KUMARI; PATIL, 2018) enfatiza as oportunidades oferecidas por ele.

*Big Data* necessariamente é enorme em **volume**, tornando obrigatório o uso de novas técnicas e ferramentas de coleta, armazenamento e análise. O mundo gerou 1 zettabytes de dados até 2010 e está previsto para gerar 40 zettabytes até 2020 (EREVELLES; FUKAWA; SWAYNE, 2016; VILLARS; OLOFSON; EASTWOOD, 2011). Este imenso volume de dados permite às organizações ir além dos dados históricos internos e obter uma visão abrangente do mercado consumidor (LAM et al., 2017). Devido ao reduzido custo de armazenamento e grande capacidade de guardar dados em um formato nativo, organizações estão mantendo os dados brutos nos chamados *Data lakes*, ficando disponíveis quando forem necessários para análise (KITCHENS et al., 2018; WATSON, 2017). Esses dados passivos, ou agnósticos, não são coletados em razão de uma solicitação específica ou para um propósito definido (CHEN et al., 2016). Esta prática de coleta promove favorece as análises realizadas pela abordagem indutiva e abduativa (CONSTANTIOU; KALLINIKOS, 2015; GÜNTHER et al., 2017; OLBRICH, 2014), e gera o registro completo das informações “captadas” no ambiente. Ambas práticas, tanto as análises indutivas, quanto o armazenamento do que é coletado para exames futuros, são características de processos de inteligência antecipativa. As perguntas com código **P01** são referentes ao volume de dados brutos e sua análise posterior, conforme Tabela 5.

**Velocidade** é a frequência de geração, processamento, armazenamento e publicação das informações. Esta propriedade do BDA impacta na agilidade necessária para armazenar e analisar dados (BHOLAT, 2015) que são produzidos de forma muito rápida (*streaming*, *real-time* ou *near real-time*). Como exemplo, cita-se o Wal-Mart, que lida com mais de um milhão de transações de clientes por hora. Essas informações são importadas em seus enormes bancos de dados, que possuem um volume de mais de 2,5 petabytes (LAM et al., 2017). A velocidade também é essencial em análises que pretendem utilizar dados de

sensores, que normalmente são gerados e transferidos por *streaming*. Há igualmente dados não estruturados de redes sociais, como textos, imagens e vídeos que necessitam, da mesma forma, do processamento imediato. Esta constante atualização em um BDA viabiliza análises e respostas a perguntas de pesquisa em tempo real, processo que é chamado de *nowcasting*. Elas envolvem algoritmos sofisticados, lidando com conjuntos de dados dinâmicos, e podem dar sentido a contextos de mudança ambiental rápida. Essas análises e orientações em tempo real podem ser importantes para negócios sensíveis a ajustes contínuos, que precisam ser constantemente alinhados às tendências de mercado. Exemplos são as indústrias da moda ou da música (CONSTANTIOU; KALLINIKOS, 2015). Organizações como Netflix e Amazon se encaixam neste contexto de *nowcasting*, dado que seus sistemas analisam e apresentam respostas imediatas, conforme *streamings* de dados gerados pelos consumidores. As perguntas com código **P02** referem-se à velocidade viabilizada pelo BDA, conforme Tabela 5.

Tecnologias BDA processam uma **variedade** de dados, como informações de texto, imagens, sensores e da internet (ARUNACHALAM; KUMAR; KAWALEK, 2018). A grande diferença entre os dados de um BDA e os dados tradicionais está na oportunidade de manipular não apenas informações transacionais estruturadas, mas também comportamentais não estruturadas (EREVELLES; FUKAWA; SWAYNE, 2016). Os dados não estruturados podem ser lidos de maneira qualitativa (LEE, 2018), favorecendo a obtenção de *insights*, candidatos a Sinais Fracos. A partir desta possibilidade de analisar dados não estruturados definem-se as perguntas com o código **P03**.

Além dos 3Vs, novas dimensões foram associadas para caracterizar um *Big Data*: exaustividade, valência, resolução, escalabilidade, veracidade, valor, variabilidade, visibilidade e virtude (GÜNTHER et al., 2017; KITCHIN; MCARDLE, 2016; LAM et al., 2017; TIEFENBACHER; OLBRICH, 2015). Kitchin e McArdle (2016), porém, argumentam que o BDA realmente se diferencia quando possui velocidade e exaustividade dos dados coletados. Nem sempre é a quantidade, o volume e a variedade dos dados que são importantes, mas sim o que as organizações fazem com os dados, a sua utilidade.

### 2.2.2 Teoria dos Sinais Fracos e Inteligência Antecipativa

Os conceitos existentes na Teoria dos Sinais Fracos foram propostos pela primeira vez no trabalho seminal de Ansoff (1975) (WARNKE; SCHIRRMESTER, 2016). Ele foi precursor na área de inteligência, monitoramento estratégico e Sinais Fracos, tendo sido o primeiro a destacar o fato de que o ambiente empresarial é turbulento e, portanto, exige recursos

gerenciais apropriados (ROSSEL, 2012). Na mesma linha, Lesca (2003) defende que o dinamismo do ambiente exige uma adaptação proativa constante por parte das organizações. Atualmente, a complexidade do contexto empresarial e da tomada de decisão estratégica tem sido exponenciada pelo volume de informações e fenômenos desconhecidos (MÜHLROTH; GROTTKE, 2018; ROHRBECK; BADE, 2012; WARNKE; SCHIRRMESTER, 2016). Lesca (2003) sugere que, se a taxa de mudança é alta, o melhor enfoque é ter flexibilidade estratégica e decidir quais ações serão viáveis, preparando-se dentro de um escopo determinado pelos Sinais Fracos.

Sinais fracos (**SF**) são sintomas imprecisos e antecipados de futuras ameaças ou oportunidades (ANSOFF, 1975; GHEORGHIU; ANDREESCU; CURAJ, 2015), e possuem um papel importante na inteligência estratégica das organizações, pois mostram que eventos futuros podem ser percebidos antes que existam como um todo. Esta percepção é realizada por meio desses sinais, inicialmente fragmentados, ambíguos e incompletos (HEUPEL; VON JUTERZENKA, 2015; LESCA, 2003). SF são dados ambientais brutos de natureza: incerta, não estruturada, prematura. Eles podem ser refinados e transformados em informações valiosas e em conhecimento estratégico (JANISSEK-MUNIZ; BORGES; BORTOLI, 2015; KIM; LEE, 2017). São antecipatórios, qualitativos e podem vir em vários formatos, a partir de fontes distintas (FONSECA; BARRETO, 2011). De fato, *a priori*, SF são hipóteses de uma futura mudança, em uma combinação de lógica dedutiva de *insights* obtidos em conjunto com outras informações ou experiências. Isoladamente eles têm pouca utilidade, pois não são suficientes em si para estabelecer uma percepção de um evento futuro (JANISSEK-MUNIZ; BORGES; BORTOLI, 2015). São construções interpretativas de indivíduos, baseadas em dados brutos, compilados a partir de um olhar atento e interessado. Para gerar valor, os SF dependem das interpretações realizadas pelos analistas e gestores (BORGES; JANISSEK-MUNIZ, 2018), que devem estar sempre atentos ao que ocorre. Para descrever resumidamente os Sinais Fracos, apresenta-se na Tabela 2 suas principais características.

Tabela 2: Características dos Sinais Fracos

<b>Características</b>	<b>Porque se chamam Sinais Fracos</b>
Fragmentado	A informação em um SF normalmente é incompleta, sendo necessário correlacionar com outros sinais para interpretar e obter sentido.
Disseminado	Um SF em geral é difuso, disperso e misturado com outras informações não úteis, com ruídos que dificultam sua identificação.
Estranho / Inesperado / Surpreendente	É uma informação incomum, com aparência não esperada, não familiar. Encontrada, muitas vezes por acaso.
Ambíguo	Uma informação do tipo SF não é objetiva, é dúbia. Muitas interpretações podem ser realizadas. É pouco clara, ambivalente, equivocada.
Utilidade não percebida	Não possui ligação direta ou clara com um fenômeno, ameaça ou oportunidade atual. A mesma informação pode ser fortemente interperativa para uma pessoa e totalmente sem interesse aparente para outras. Um SF é aparentemente desprovido de significado operacional.
De difícil detecção	Um SF permanece facilmente sem ser percebido. Ele é dissimulado, permanece escondido, em uma quantidade de dados.
Aleatório / Imprevisível	Um SF não participa de uma tendência ou padrão. Ele é apresentado de maneira aleatória, imprevista.
Formatos diversos	Um SF pode assumir diversos formatos, textos, imagens, vídeos, sons, percepções, diálogos, etc.

Fonte: Adaptado de Bortoli, Janissek-Muniz e Borges (2015, p.5)

Há, no ambiente, estímulos a todo instante: informações, notícias, estudos, números, opiniões, etc. Nessa oferta abundante, pode haver algum dado prematuro e não estruturado que aponta para transformações disruptivas (MENDONÇA; CARDOSO; CARAÇA, 2012). A princípio parece ser parte de um “ruído de fundo”, mas, examinando melhor, ou associando a outros sinais (LESCA, 2003), pode tornar-se um sinal forte no futuro, representando um evento que aos poucos se materializa (KIM; LEE, 2017; SCHOEMAKER; DAY; SNYDER, 2013). Sua importância reside, portanto, na possibilidade de correlacionar a outros SF (LESCA, 2003), o que pode levar a novas interpretações e caminhos (MENDONÇA; CARDOSO; CARAÇA, 2012) mediante decisões estratégicas.

Para obter maior vantagem competitiva a partir de uma decisão estratégica proativa, é necessário interpretar essas informações precoces o mais rápido possível, em uma visão dinâmica do futuro. Um dos pré-requisitos para essa percepção é a coleta dos dados, e, para que ela ocorra, é necessário um processo sistemático de exame do ambiente (ILMOLA; KUUSI, 2006; LESCO, 2003). As atividades, portanto, devem ser implantadas de forma contínua (CAINELLI; JANISSEK-MUNIZ, 2019; FONSECA; BARRETO, 2011), como uma capacidade (RHISIART; MILLER; BROOKS, 2015) difundida pela cultura da

organização. Perguntas com código **P04** estão associadas ao conceito e características dos Sinais Fracos.

A coleta de informações do tipo SF exige método e gestão. Já na década de 70, Ansoff (1975) sugeriu uma forma de identificar Sinais Fracos, estabelecendo um sistema de monitoramento do ambiente, valoração do impacto dos sinais percebidos e planejamento com desenvolvimento de ações (FONSECA; BARRETO, 2011), baseado no monitoramento ativo e contínuo do ambiente com três filtros (HOLOPAINEN; TOIVONEN, 2012). O filtro de **vigilância** permite a observação do ambiente pela organização, que precisa ser restrito a alguns domínios, assim como a definição de alvo proposta por Lesca (2003). O filtro da **mentalidade** comporta percepções, interpretações e decisões dos indivíduos, realizados sempre em função do seu estilo cognitivo, da sua formação, conhecimento e experiências anteriores. O filtro de **poder** trata da influência de hierarquias e de processos organizacionais sobre o reconhecimento dos Sinais Fracos (WARNKE; SCHIRRMESTER, 2016). Estes três filtros refletem a percepção de um estímulo, a interpretação com a intenção de criar sentido e valor, e a sua incorporação ao banco de conhecimento e tomada de decisão (FONSECA; BARRETO, 2011). O objetivo é detectar um padrão antes que todas as peças do quebra-cabeça se encaixem, ou antes que seja tarde demais (MENDONÇA; CARDOSO; CARAÇA, 2012). Os filtros de Ansoff motivaram o desenvolvimento das perguntas com o código **P05**.

### 2.2.3 BDA e Inteligência Antecipativa

O uso do BDA pode ser visto como uma capacidade organizacional que facilita a geração de *insights* e de conhecimento a partir da análise de dados (CHEN; PRESTON; SWINK, 2015). Contudo, obter *insights* não é o mesmo que identificar Sinais Fracos. *Insight* refere-se a uma compreensão relevante de parte de um conhecimento, de um fato ou objeto, de modo que ideias relacionadas possam ser compiladas em uma mesma narrativa (SCHULTE; HOVORKA, 2017). É a capacidade de obter uma compreensão intuitiva de determinada situação, que potencialmente beneficia o resultado da organização (JIANG; GALLUPE, 2015). Os *insights* não são apresentados pelo BDA como um resultado pronto, completo (MCAFEE; BRYNJOLFSSON, 2012). Eles decorrem de um processo humano de criação de sentido, individual ou coletivo, que utiliza os dados processados como um meio (CONSTANTIOU; KALLINIKOS, 2015; SHARMA; MITHAS; KANKANHALLI, 2014). Eles podem, portanto, ser considerados como candidatos a Sinais Fracos e ser tratados no processo de inteligência antecipativa, entrando na fase inicial de coleta e percepção de

informações, conforme proposto por Lesca (2003). No processo de criação de sentido orientado por dados (*data-driven*), analistas têm total influência sobre quais elementos serão selecionados para descrever os fenômenos, como por exemplo, quais os *outliers* importantes e quais associações relevantes podem estar contidas nos dados (LYCETT, 2013). *Insights* baseados em dados podem ser usados para criar uma narrativa que dá sentido aos fenômenos estudados, alimentando a tomada de decisão e planos subsequentes (SHARMA; MITHAS; KANKANHALLI, 2014). Perguntas relativas aos *insights* obtidos em BDA para a tomada de decisão possuem o código **P06**.

A busca e a identificação de SF em processos de inteligência antecipativa envolvem desafios de coleta e seleção manual a partir do ambiente (KIM; LEE, 2017), podendo ser tratados por métodos sugeridos e praticados dentro da tradição Ansoff. Essas práticas são baseadas na percepção intuitiva de especialistas e nas perspectivas avaliadas pelos analistas, durante processo de criação de sentido. Embora custosos e demorados, eles proporcionam resultados importantes (LEE, 2018; YOON, 2012). No entanto, com o crescimento no volume de dados disponível no ambiente de negócios atual (JIANG; GALLUPE, 2015), é oportuno buscar pelos SF também a partir das análises sobre os dados (EREVELLES; FUKAWA; SWAYNE, 2016; MÜHLROTH; GROTTKE, 2018). Para tal, evidencia-se a importância dos *outliers* observados estatisticamente em volumes de dados (KIM; LEE, 2017; MENDONÇA; CARDOSO; CARAÇA, 2012). *Outliers* são conceituados como anomalias estatísticas que não se encaixam facilmente no entendimento dos resultados apresentados, e podem ser pontos singulares interpretados como um nicho potencial de inovação. Neste sentido, cenários futuros alternativos gerados por *outliers*, baseados em pesquisas de identificação de SF, também podem ser elaborados. As oportunidades e riscos percebidos nesses diferentes cenários podem ser transformados em uma agenda de decisão e ação (MENDONÇA; CARDOSO; CARAÇA, 2012). Perguntas relativas aos *outliers* possuem o código **P07**.

Com o volume de dados existente, suas associações passam a ser mais complexas, o que torna a capacidade de discernimento dos analistas um limitante importante ao decifrar e interpretar o ambiente (MÜHLROTH; GROTTKE, 2018; SAMMUT; SARTAWI, 2012). Deste modo, para monitorar o ambiente, não basta apenas a adoção de novas tecnologias de análise, uma mudança metodológica é necessária. Para entender os dados, gerar informação e conhecimento, para obtendo melhores resultados no uso do BDA, é necessário repensar ou complementar o método (BHOLAT, 2015; MCABEE; LANDIS; BURKE, 2017). Estrategicamente, recomenda-se não apenas consultar ou manipular os dados para validar

conjecturas pré-concebidas, mas também usá-los para formular novas hipóteses (EREVELLES; FUKAWA; SWAYNE, 2016), no chamado *sistema de informação inverso* (LESCA; JANISSEK-MUNIZ, 2015). No contexto do BDA, visualiza-se esta mudança epistemológica (KITCHIN; LAURIAULT, 2015; TAYLOR; SCHROEDER; MEYER, 2014). A abordagem indutiva, em conjunto com as tecnologias *Big Data*, permite que se identifique padrões, tendências e *outliers* sem desenvolver hipóteses *a priori* (LYCETT, 2013), refletindo uma maneira particular de perceber e entender o ambiente (KITCHIN; LAURIAULT, 2015). Tal tipo de investigação científica requer menos confiança no conhecimento existente, dando maior foco no que é desconhecido. Nesta transformação, deve-se passar de uma visão baseada no conhecimento para uma visão baseada na ignorância (SAMMUT; SARTAWI, 2012). Perguntas relativas ao método indutivo possuem o código **P08**.

Estas considerações sugerem que, além das capacidades de observadores individuais, novos processos guiados por dados são necessários para perceber, identificar e utilizar SF na inteligência antecipativa. Portanto, é importante adotar técnicas de identificação que permitam a percepção de SF partindo do processamento de dados (MENDONÇA; CARDOSO; CARAÇA, 2012), e o BDA pode ser uma delas (EREVELLES; FUKAWA; SWAYNE, 2016; MÜHLROTH; GROTTKE, 2018). Contudo, há uma diferença entre aceitar SF e perceber ou interpretar o que eles significam. Tomadores de decisão possuem o viés de ver o mundo conforme suas experiências e, se os SF apresentados não se ajustarem a este padrão, serão ignorados ou distorcidos (BORGES; JANISSEK-MUNIZ, 2018). Quando há ambiguidade, pode-se facilmente manipular os sinais até obter o que se deseja acreditar (SCHOEMAKER; DAY, 2009). Esta observação reforça a necessidade de um processo coletivo e sistemático de análise das informações obtidas em um BDA. A pergunta sobre como as organizações podem utilizar o BDA para obter informações do ambiente possui o código **P09**.

Inicia-se o comparativo entre Sinais Fracos e *Big Data*, resumindo e relacionando suas principais características na Tabela 3. Essa síntese é o primeiro produto deste estudo, analisando a relação existente na literatura entre os Sinais Fracos identificados nos processos de Inteligência Antecipativa e as informações obtidas nos processos de análise usando o BDA. Neste paralelo aponta-se as semelhanças entre a natureza desses dois tipos de informações. As características dos Sinais Fracos são as mesmas listadas na Tabela 2 (BORTOLI; JANISSEK-MUNIZ; BORGES, 2015).

Tabela 3: Paralelo entre Sinais Fracos e Big Data

Características dos SF	Paralelo em Big Data
Disseminado	Da mesma forma, informações importantes podem estar disseminadas ou dispersas no “dilúvio” de dados de um <i>Big Data</i> . A exemplo, usuários da rede social <i>Twitter</i> geram milhares de mensagens por segundo (KITCHIN; MCARDLE, 2016). Para utilizar essas informações é necessário remover o “ruído” (GÜNTHER et al., 2017), criando metadados e tornando-as legíveis por algoritmos (KITCHIN, 2014). Os dados não estruturados disponíveis são difíceis de lidar, contudo, é possível identificar sinais relevantes (TIEFENBACHER; OLBRICH, 2015). O “ruído” pode correlacionar dados incorretamente, como por exemplo, identidades de pessoas são confundidas ou dados de períodos diferentes são conectados (JANSSEN; VAN DER VOORT; WAHYUDI, 2017).
Estranho / Inesperado / Surpreendente	<i>Outliers</i> encontrados por meio de BDA são amostras inesperadas e surpreendentes, são resultados conceituados como anomalias (KIM; LEE, 2017) ou desvios das médias estatísticas (MÜHLROTH; GROTTKE, 2018), que não se encaixam facilmente no entendimento das informações apresentadas. Em grandes volumes de dados, é provável que os Sinais Fracos tenham baixa frequência absoluta de ocorrências e alta amplitude de flutuação (YOON, 2012). Em geral são sinais que possuem um padrão anormal e que raramente são expostos (LEE, 2018).
Ambíguo, incompleto, incerto, fragmentado	Informações obtidas em análises em <i>Big Data</i> também possuem estas características. Ao utilizar BDA, não é possível confiar apenas nos métodos tradicionais baseados em frequências para guiar as interpretações. Ferramentas estatísticas apresentam facilmente falsas correlações neste contexto. O grande volume de dados ocasiona vieses de confirmação, já que é possível criar inúmeras relações entre as variáveis existentes. Muitas correlações também seriam estatisticamente significativas, tornando o teste de significância de hipótese nula inútil (GEORGE; HAAS; PENTLAND, 2014; MCABEE; LANDIS; BURKE, 2017). Em um <i>Big Data</i> há cenários onde muitas hipóteses podem ser suportadas, já que analistas utilizam leituras específicas dos dados, em visões influenciadas pelo viés de confirmação (BHOLAT, 2015; GÜNTHER et al., 2017), intencionalmente ou acidentalmente (KUSHIRO; MATSUDA; TAKAHARA, 2014).
De difícil detecção	A identificação de informações importantes em um <i>Big Data</i> é tecnicamente desafiadora, pois envolve a modelagem de dados estruturados e não estruturados (DUTTA; BOSE, 2015). As ferramentas estatísticas tradicionais não podem ser utilizadas neste contexto, são planejadas para tratar pesquisas no formato <i>Small Data</i> : amostras de dados estruturados, com propriedades planejadas e conhecidas (KITCHIN; LAURIAULT, 2015). Há uma incapacidade de técnicas para explorar o volume e a velocidade dos dados corretamente (ARUNACHALAM; KUMAR, 2018), tornando a detecção dos <i>insights</i> mais complicada. Muitos desafios existem neste tipo de análise, tornando difícil a conversão de grandes quantidades de dados brutos em informação e conhecimento que permita a tomada de decisão (BUMBLAUSKAS et al., 2017).
Aleatório / Imprevisível Sem hipóteses pré-definidas	Por meio do método de pesquisa indutivo torna-se viável encontrar informações aleatórias e imprevisíveis em <i>Big Data</i> . Esta abordagem <i>bottom-up</i> permite que os analistas obtenham <i>insights</i> sem desenvolver hipóteses <i>a priori</i> (LYCETT, 2013).
Formatos diversos	As tecnologias BDA podem processar uma variedade de tipos de dados diferentes (ARUNACHALAM; KUMAR; KAWALEK, 2018).

Fonte: elaborado pelo autor

Para melhor entendimento da relação entre o referencial teórico e o roteiro de perguntas utilizado na coleta de informações, apresenta-se a Tabela 4 contendo a lista de elementos de investigação, com sua descrição, referências e observações. A lista é um conjunto de características de Sinais Fracos e de elementos que potencialmente influenciam o monitoramento do ambiente e a identificação de Sinais Fracos.

Tabela 4: Lista e elementos de investigação

Elemento	Referências	Observação	Pergunta
Coleta de <b>grandes volumes de dados brutos</b> ( <i>raw data</i> ), para a modelagem e análise posterior.	(CHEN et al., 2016; LAM et al., 2017)	Avaliar se as organizações realizam este tipo de coleta e se este mecanismo influencia o monitoramento digital do ambiente, por meio de grandes volumes de dados, que podem ser modelados e analisados conforme escopos de pesquisa posteriores.	P01
Coleta e análise de dados com <b>velocidade</b> .	(BHOLAT, 2015)	A constante atualização que o <i>Big Data</i> oferece permite respostas em tempo real. Em um ambiente dinâmico e rápido, a velocidade da identificação dos Sinais Fracos é importante (ANSOFF, 1975; ROHRBECK; THOM; ARNOLD, 2015) para que o alinhamento da organização ao ambiente seja ágil, mantendo a competitividade e o protagonismo. Avaliar se este tipo de coleta é realizado e se ele permite a tomada de decisão estratégica rápida.	P02
Coleta e análise de <b>dados não estruturados</b> , como imagens, textos, vídeos e mapas.	(ARUNACHALAM; KUMAR; KAWALEK, 2018; TIEFENBACHER; OLBRICH, 2015)	Dados não estruturados possibilitam a interpretação qualitativa (LEE, 2018), oferecendo métodos mais adequados para a identificação de Sinais Fracos. Avaliar se este tipo de dado é coletado e se ele é utilizado na tomada de decisão estratégica.	P03
<b>Teoria dos Sinais Fracos</b> : eles são antecipativos e permitem a percepção fenômenos futuros	(ANSOFF, 1975)	Verificar se as análises em BDA facilitam a identificação de Sinais Fracos.	P04
<b>Sinais fracos</b> são informações incompletas, incertas, de difícil detecção, fragmentadas e/ou ambíguas	(JANISSEK-MUNIZ; BORGES; BORTOLI, 2015)	Avaliar se as análises em BDA oferecem informações com estas características.	P04
É necessário um <b>processo sistemático</b> e coletivo de criação de sentido para o melhor discernimento dos Sinais Fracos.	(LESCA, 2003)	Examinar se as organizações possuem processos de inteligência para analisar as informações obtidas em um BDA.	P04
<b>Filtros de Ansoff</b> : vigilância, mentalidade e poder	(ANSOFF, 1975)	Avaliar se as informações obtidas em análises BDA passam pelos mesmos filtros.	P05
<b>Insights</b> colhidos em um BDA podem ser utilizados para a tomada de decisão.	(DUAN; CAO, 2015; SHARMA; MITHAS; KANKANHALI, 2014)	Verificar se em organizações <i>data-driven</i> , <i>insights</i> obtidos em BDA podem ser considerados como candidatos a Sinais Fracos e, então, como apoio à decisão estratégica.	P06
Um <b>outlier</b> pode ser uma evidência de sinal fraco, a ser confirmado em análises posteriores	(KIM; LEE, 2017; MENDONÇA; CARDOSO; CARAÇA, 2012)	Entender se <i>outliers</i> são analisados pelas organizações como uma informação importante antecipativa.	P07
Para obter melhores resultados em um BDA, é necessário complementar o método com a abordagem <b>indutiva</b>	(BHOLAT, 2015; MCABEE; LANDIS; BURKE, 2017)	Avaliar se o método indutivo pode ser utilizado na percepção de possíveis Sinais Fracos em organizações.	P08
O BDA pode ser uma ferramenta para <b>identificação de Sinais Fracos</b>	(EREVELLES; FUKAWA; SWAYNE, 2016; MÜHLROTH; GROTTKE, 2018)	Examinar se o BDA pode ser utilizado para monitorar o ambiente, facilitando a identificação de candidatos a Sinais Fracos.	P09

Fonte: elaborado pelo autor

### 2.3 Método de Pesquisa

A fim de explorar o campo, investigando a viabilidade de perceber SF por meio das ferramentas de BDA, entende-se como necessária uma pesquisa qualitativa de natureza descritiva e exploratória, adotando o método de entrevistas individuais em profundidade (FRASER; GONDIM, 2004; GASKELL, 2002; MAY, 2004; POZZEBON; FREITAS, 1998; TRIVIÑOS, 1987). A pesquisa retrata características, métodos, técnicas e análises realizadas em BDA para fins de inteligência, portanto, é classificada como **descritiva**. Ela também é **exploratória**, já que se propõe a apresentar uma visão geral dos novos fenômenos, compreendendo os fatos, mediante o referencial teórico e a análise dos dados empíricos coletados (JANISSEK-MUNIZ, 2016; PETTY; THOMSON; STEW, 2012). Elegeram-se esta abordagem em razão da emergência dos fenômenos estudados, assim, tornam-se necessárias estratégias exploratórias para identificar e compreender os comportamentos dos indivíduos em seus ambientes. Ela serve como uma base para pesquisas futuras, no sentido de que proporciona dados para testar hipóteses desenvolvidas (GASKELL, 2002). Os objetivos desta etapa são: entender se as organizações estão percebendo SF em análises BDA; explorar se os *insights* colhidos no BDA podem ser considerados como possíveis Sinais Fracos.

O método de entrevistas viabiliza a compreensão detalhada das atitudes, opiniões, valores e motivações (GASKELL, 2002), do ponto de vista dos participantes em seus contextos (POZZEBON; FREITAS, 1998), sendo adequado para explorar em profundidade esta temática. Emprega-se este design de pesquisa, uma vez que ele permite explorar e comparar as possíveis práticas – formas de análises – oferecidas pelo BDA, bem como suas influências na inteligência antecipativa e na tomada de decisão, refletindo sobre suas diferenças e semelhanças. As entrevistas foram selecionadas oportunisticamente por conveniência, atendendo a requisitos estabelecidos, através de redes de contatos acadêmicas e profissionais do pesquisador. Os critérios obrigatórios de seleção foram os seguintes:

- Organizações que já possuem áreas ou processos de inteligência estabelecidos;
- Organizações que utilizam ferramentas de BDA por ao menos um ano;
- Os consultores, analistas, desenvolvedores e gestores entrevistados utilizam as ferramentas de BDA direta ou indiretamente, por meio de aplicações ou relatórios gerados pela equipe de inteligência.

A coleta de dados primários foi realizada em outubro e novembro de 2018, a partir de um roteiro semiestruturado (MAY, 2004), com o objetivo de compreender detalhadamente opiniões, ações e motivações dos entrevistados, nas circunstâncias tratadas na pesquisa (GASKELL, 2002). As perguntas foram formuladas para estimular o entrevistado a elaborar os detalhes, para obter clareza e ficar perto da experiência, do fenômeno ocorrido (STARKS; TRINIDAD, 2007). Desenvolveu-se o roteiro semiestruturado (Tabela 5), ou protocolo da entrevista, tomando como base a fundamentação teórico-conceitual, proporcionando a análise descritiva e a percepção de como as organizações estão tratando o BDA para fins de inteligência estratégica antecipativa. A escolha por realizar entrevistas com perguntas abertas ocorreu devido à importância da coleta de dados espontâneos e não previstos, percebidos no momento da interação com o entrevistado (FREITAS; JANISSEK-MUNIZ, 2000; POZZEBON; FREITAS, 1998). Portanto, a entrevista incluiu perguntas questionando os motivos pelos quais as análises em BDA estão sendo realizadas, como elas ocorrem e, também, como o resultado destas análises está sendo utilizado para entender e monitorar o ambiente, identificar Sinais Fracos e para o apoio à tomada de decisão.

Cada entrevista teve uma duração aproximada de 30 a 60 minutos. Elas foram gravadas em áudio e, também, notas foram tomadas pelo pesquisador. As transcrições das gravações em áudio e as anotações realizadas pelo pesquisador foram codificadas, para a análise qualitativa de dados, em um software de *qualitative data analysis* (QDA ou *Computer assisted qualitative data analysis software* - CAQDAS). O software utilizado foi o NVivo Plus, na versão 11, para Microsoft Windows. Como os entrevistados são colaboradores estratégicos e consultores que não possuem autorização formal para divulgação de informações e processos, tomou-se o cuidado de não citar seus nomes, os nomes das suas empresas e de exemplos de produtos na transcrição.

A análise foi realizada utilizando as técnicas de análise de conteúdo (BARDIN, 2016; FREITAS; JANISSEK-MUNIZ, 2000), respondendo e explorando à questão de pesquisa, por meio de correlações de padrões obtidos empiricamente. A análise de conteúdo foi desenvolvida em três fases: pré-análise, exploração do material e o tratamento de resultados, inferências e interpretações (BARDIN, 2016). Para Bardin (2016), na pré-análise são organizados e sistematizados os dados, escolhendo os documentos a analisar, formulando hipóteses e objetivos e elaborando indicadores que fundamentam a interpretação final. A fase seguinte consiste nas operações de codificação, desconto ou enumeração dos textos. Na terceira e última fase, para que os resultados obtidos sejam significativos, é realizado o seu

tratamento. Dados estatísticos permitem estabelecer visões sobre os resultados como diagramas, gráficos ou modelos.

Após as entrevistas realizadas, identificou-se que não houve nova informação relatada, decidindo pela saturação foi alcançada. Ao utilizar dados obtidos a partir de entrevistas semiestruturadas, eles são analisados apenas após a finalização da coleta. A saturação é então obtida por meio da garantia de uma amostra adequada, que forneça dados suficientes para replicar a classificação de dados em cada item (DENZIN, 2018).

O roteiro inicialmente foi desenvolvido com 24 perguntas, sendo realizado primeiramente, para fins de teste piloto, com dois consultores experientes em ferramentas BDA, lotados em grandes organizações, nacionais e multinacionais, que atuam no mercado brasileiro. Após este teste, o roteiro foi revisado por dois especialistas em Inteligência Antecipativa, pesquisadores do PPGA na Universidade Federal do Rio Grande do Sul. As perguntas foram, então, recompiladas e resumidas em apenas 11, facilitando a coleta e priorizando a qualidade das informações. Apresenta-se, na Tabela 5, o roteiro de perguntas semiestruturado utilizado nas entrevistas.

Tabela 5: Roteiro de perguntas semiestruturado

Código e referência	Perguntas
<b>P01</b> BDA: Volume	Como a sua organização coleta e analisa grandes volumes de dados? Eles são coletados de forma bruta para modelar e analisar posteriormente?
<b>P02</b> BDA: Velocidade	Os dados são coletados e analisados rapidamente (por exemplo: <i>streaming</i> , <i>real-time</i> ou <i>nowcasting</i> )? Se sim: Como esses dados possibilitam gerar orientações ou respostas, em tempo real, às mudanças no ambiente?
<b>P03</b> BDA: Variedade	Sua organização coleta dados não estruturados? Se sim, eles decorrem de fontes externas como textos, imagens e vídeos? Como a organização lida com este tipo de dado?
<b>P04</b> Sinais Fracos e Inteligência Antecipativa	Você percebe que as análises de BDA oferecem <i>insights</i> para um melhor exame ou compreensão de eventos futuros e permitem antecipar oportunidades, ameaças e eventos disruptivos?  Na sua avaliação, as análises de BDA se caracterizam por oferecer informações incompletas, incertas, de difícil detecção, fragmentadas e/ou ambíguas sobre o ambiente de negócios?  Ao perceber informações antecipativas, como sua organização as analisa para dar sentido e criar um entendimento do ambiente? Ela decide e planeja ações para aproveitar oportunidades ou mitigar ameaças de uma eventual disrupção ou mudança de paradigma no ambiente de negócios? Esta análise é realizada em um processo sistemático coletivo?
<b>P05</b> Filtros de Ansoff	Você percebe que uma informação obtida via análise de BDA passa por este caminho ou barreiras? Fale sobre isso. a) Percepção e atenção para o monitoramento do ambiente de negócios? b) Interpretação, julgamento e a tomada de decisão? c) Tomada de ação (poder) no ambiente organizacional?
<b>P06</b> <i>Insights e data-driven</i>	As decisões da sua organização estão sendo tomadas baseadas em informações e <i>insights</i> colhidos no DBA? A sua organização é <i>data-driven</i> ? Fale a respeito do processo de análise.
<b>P07</b> <i>Outliers</i>	Como a sua organização analisa os pontos singulares ( <i>outliers</i> , fora da curva) encontrados no BDA? Eles podem ser utilizados para perceber indicadores de fenômenos futuros (por exemplo, nichos de mercado, novos produtos)?
<b>P08</b> Método Indutivo	Considerando o método de análise de grandes volumes de dados em sua organização: a) As análises já possuem perguntas prontas ou hipóteses a serem respondidas e/ou validadas? b) As análises utilizam o BDA para formular novas hipóteses a partir dos dados? Como?

- c) A organização percebe novas correlações, *outliers*, informações de modo a descobrir novas oportunidades futuras apenas a partir dos dados?  
Exemplifique.

**P09**  
Monitoramento do ambiente e identificação de Sinais Fracos

Na sua opinião, como as organizações podem utilizar o BDA para obter informações a fim de monitorar o que está mudando no ambiente externo?

Fonte: elaborado pelo autor

## 2.4 Análise dos Resultados

Os resultados foram obtidos por meio da análise de conteúdo das transcrições das entrevistas. A análise de conteúdo é uma técnica de pesquisa baseada na leitura detalhada das respostas e seu processo de codificação. O objetivo é obter conhecimento a partir dos textos categorizados, resumidos e organizados. Ela propicia a identificação dos motivos e opiniões subentendidas nos fenômenos estudados de forma indireta, visto que observa-se as respostas transcritas dos respondentes (FREITAS; JANISSEK-MUNIZ, 2000).

O foco das entrevistas foram os métodos, técnicas e análises realizadas em BDA pelos praticantes, não entrando detalhadamente nos contextos de cada organização envolvida. O dado empírico foi coletado essencialmente nas análises desenvolvidas, apoiadas, solicitadas ou planejadas por cada respondente, portanto, foi fundamentado na sua experiência. Ao entrevistar consultores que trabalham em diversas organizações, as informações relatadas podem ter sido colhidas em fenômenos ocorridos em mais de uma empresa.

Na Tabela 6 apresenta-se o perfil do grupo de respondentes. Como descrito no Capítulo 2.3, realizou-se as duas primeiras entrevistas para fins de teste, portanto, foram desconsideradas nos resultados deste estudo. Elas foram essenciais para entender se o roteiro de perguntas poderia responder de forma satisfatória a problemática da pesquisa, para aprimorar os itens, deixando-os mais claros e para reduzir a quantidade de perguntas realizadas. Todos entrevistados possuem mais de 5 anos de experiência na área de inteligência, *marketing* ou na direção das organizações.

Tabela 6: Grupo de entrevistados

Entrevistado	Empresa	Indústria	Cargo	Escolaridade
Entrevistado 01	A	TIC	Partner Manager	Pós-Graduação
Entrevistado 02	B	TIC	Arquiteto de Soluções	Pós-Graduação
Entrevistado 03	C	TIC	IT Specialist   Team Leader	Pós-Graduação
Entrevistado 04	D	TIC	Diretor de Negócios	Pós-Graduação
Entrevistado 05	E	TIC	Analista e Arquiteto de Dados	Pós-Graduação
Entrevistado 06	F	TIC	Consultor de Negócios e Soluções	Pós-Graduação
Entrevistado 07	G	Comunicação	Programmatic Business Developer	Pós-Graduação
Entrevistado 08	H	TIC	CEO	Pós-Graduação
Entrevistado 09	I	TIC	Diretor	Pós-Graduação
Entrevistado 10	J	TIC	Big Data and Cloud Services	Pós-Graduação
Entrevistado 11	K	TIC	R&D Senior Manager	Pós-Graduação
Entrevistado 12	L	TIC	Head Of Research	Pós-Graduação

Fonte: dados da pesquisa (2018)

Para explorar os dados, com o objetivo de levantar hipóteses, ilustra-se alguns resultados quantitativos. Esse estudo descritivo é significativo, oportuniza realizar cruzamentos nos dados, correlacionando quantitativamente as informações colhidas. Ele viabiliza a validação de novas ideias, hipóteses e conclusões (FREITAS; JANISSEK-MUNIZ, 2000). Destaca-se nas figuras exibidas abaixo algumas informações relevantes obtidas na análise quantitativa dos dados textuais mediante as categorias e códigos gerados pela análise de conteúdo. Na Figura 2, referente às perguntas P03, os entrevistados afirmam que coletam dados estruturados, embora nem todos os analisem. A Figura 3 mostra a percepção das análises BDA (perguntas P04), e a grande maioria indica que elas oferecem *insights* para compreensão de eventos futuros.

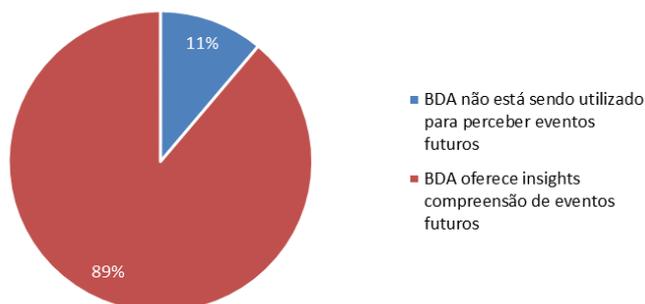
Figura 2: Respostas para P03, dados não estruturados



Fonte: dados da pesquisa (2018)

Figura 3: Respostas para P04, BDA e insights de eventos futuros

BDA permite insights para compreensão de eventos futuros

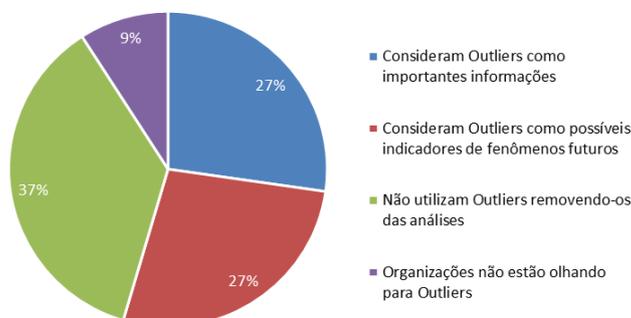


Fonte: dados da pesquisa (2018)

Conforme Figura 4 (perguntas P07), o uso de *outliers* é considerado como importante, e a Figura 5 (perguntas P08) mostra que a abordagem indutiva em análises BDA é utilizada na mesma intensidade que o método dedutivo, embora exista uma predisposição à evolução neste sentido.

Figura 4: Respostas para P07, uso dos Outliers

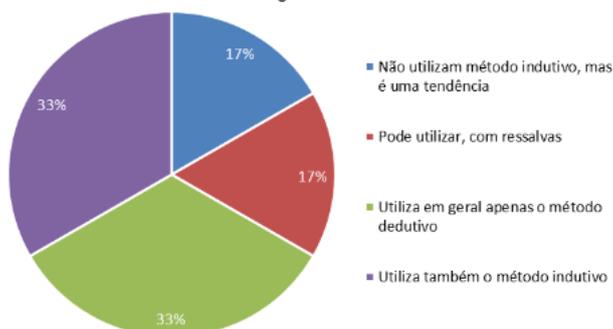
Uso dos Outliers



Fonte: dados da pesquisa (2018)

Figura 5: Respostas para P08, uso do método indutivo

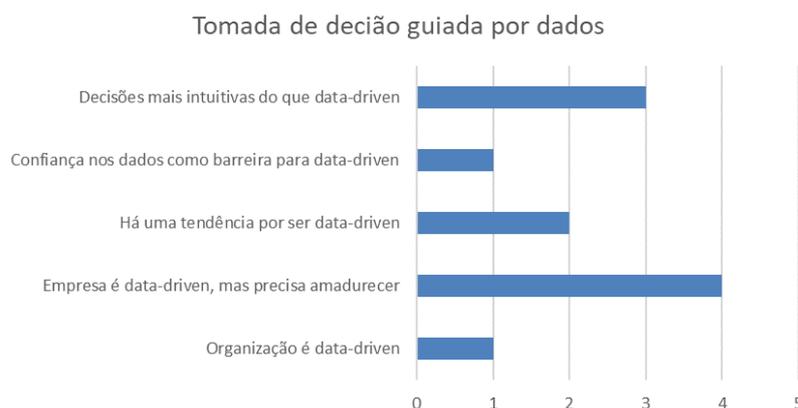
Abordagem Indutiva



Fonte: dados da pesquisa (2018)

A pergunta P06 (Figura 6) avalia se os *insights* obtidos em BDA são utilizados na tomada de decisão estratégica. As respostas mostram que parte das organizações da amostra decidem em geral de forma intuitiva e, mesmo nas que são consideradas pelos respondentes como *data-driven*, ainda há necessidade de amadurecer os processos de decisão guiada por dados.

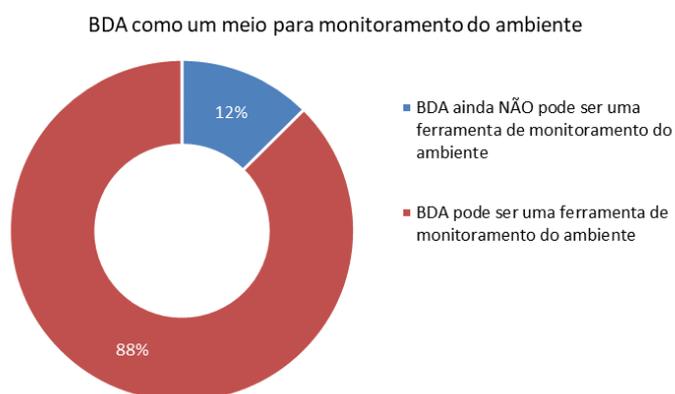
Figura 6: Respostas para P06, insights e data-driven



Fonte: dados da pesquisa (2018)

Por fim, a Figura 7 mostra que, na percepção dos respondentes, o BDA pode ser uma ferramenta de apoio ao monitoramento do ambiente nas suas organizações.

Figura 7: Respostas para P09, uso do BDA para monitoramento do ambiente e identificação de Sinais Fracos



Fonte: dados da pesquisa (2018)

No Apêndice I apresenta-se a Tabela 23 contendo a lista completa de categorias e códigos gerados pela análise de conteúdo, realizada a partir da transcrição das gravações no

software de QDA NVivo Plus. Nele pode-se obter informações descritivas detalhadas das respostas analisadas.

Foram identificadas, em oito entrevistas, vinte referências a exemplos específicos de análises realizadas em BDA. Todos esses exemplos possuem potencial para que, ao examinar seus resultados, os analistas percebam candidatos a Sinais Fracos. Também foram observadas, pela análise de conteúdo, 28 referências a 10 tipos de análises. Elas estão listadas na Tabela 7. Nota-se que a análise descritiva exploratória tem, ainda, o maior número ocorrências.

*Tabela 7: Tipos de análises realizadas*

Tipo de Análise	Referências
Análise de <i>Outliers</i>	2
Análise de sentimento	3
Classificação	4
<i>Clustering</i>	1
Correlações	2
Descritiva exploratória	6
<i>Forecasting</i>	2
Modelos preditivos	4
NLP	1
Reconhecimento de imagem	3

*Fonte: dados da pesquisa (2018)*

## 2.5 Discussões

Buscando compreender como as organizações estão processando e analisando os dados de um BDA, percebe-se que tais práticas dependem do processo em questão, dos motivos da coleta e análise, e, principalmente, da cultura e maturidade da área de inteligência de cada empresa. Na amostra desta pesquisa encontram-se níveis diferentes de maturidade, desde empresas que trabalham apenas com os dados internos estruturados, de forma lenta, até as que trabalham com dados externos, não estruturados e rápidos, exigindo maior complexidade no tratamento da informação (ARUNACHALAM; KUMAR; KAWALEK, 2018; TIEFENBACHER; OLBRICH, 2015). Empresas mais maduras entendem que o dado (ou informações fornecidas por sistemas) é um ativo importante e estratégico (DAVENPORT; DYCHÉ, 2013; DUAN; CAO, 2015; LAVALLE et al., 2011), e procuram armazená-los de forma bruta e em um escopo maior, conforme indicado por Watson (2017). O entrevistado 10 ilustra esta afirmação: “*Já consideram que a informação é um ativo e querem guardar*

*tudo para utilizar porque vai ter valor. Guardam dados brutos de tudo que é possível, sem ter uma pesquisa em andamento.”*

Essas coletas não são realizadas com problemática definida. Normalmente há um projeto pontual que as suportam, mas os analistas tomam o cuidado de aproveitar e armazenar mais dados do que o necessário, sem processá-los *a priori*, para que possam ser utilizados como dados secundários em futuras pesquisas. Assim, as práticas de coleta e análise promovem o monitoramento do ambiente, armazenam informações e conhecimento (KITCHIN; LAURIAULT, 2015), podendo ser usadas em diferentes momentos. Essas práticas também facilitam o uso da abordagem indutiva, um dos caminhos para gerar hipóteses e conhecimentos não deduzidos (LYCETT, 2013), aumentando a chance da percepção de Sinais Fracos. Nas entrevistas, muitos afirmam o uso do método indutivo e também a possibilidade de monitorar o ambiente e perceber Sinais Fracos em análises realizadas em BDA. A coleta de dados externos brutos sem escopo definido também viabiliza o registro histórico do que está sendo alterado no ambiente. É um monitoramento digital, com possíveis alertas em tempo real. Entrevistados 3, 5, 6, 10 e 11 relatam esta característica, como é possível observar na manifestação do entrevistado 6:

*“É o conceito de Data Lake de trazer estes dados, independente de num primeiro momento tu não necessariamente ter uma ideia de uso ou valor, mas trazer isso porque daqui a pouco, no uso de qualquer modelo estatístico, se possa identificar quem possa ter relação, quem faça sentido em um determinado insight, ou necessidade de negócio que venha querer solucionar.”*

Boa parte dos entrevistados relata que há coleta de dados em tempo real do tipo *real-time* ou *streaming*. Alguns sistemas são reconhecidos como *near real-time*, já que a coleta e processamento não é imediata, mas igualmente rápida. Eles informam que esses sistemas possibilitam a resposta automática, ou a decisão rápida, em relação a uma mudança no meio. Porém, em todos os relatos, esta resposta é de natureza operacional ou tática. Não foram percebidos exemplos de *real-time*, ou *near real-time*, que identificam mudanças estratégicas no ambiente, ou que contribuem para antever eventos futuros.

A coleta e análise de dados externos não estruturados (imagens, textos, vídeos, mapas) comportam uma compreensão qualitativa das informações (LEE, 2018), viabilizando métodos mais adequados para a identificação de SF. Alguns entrevistados relatam o uso de uma técnica de processamento de linguagem natural (*Natural Language Processing* - NLP), chamada *análise de sentimento*. Esta técnica é frequentemente adotada para identificar

sentimentos e subjetividades em textos coletados em mídias sociais, como, por exemplo, opiniões de clientes (CHEN; CHIANG; STOREY, 2012). Portanto, ela classifica e categoriza o que a população está sentindo ao escrever textos, em mídias sociais ou processos interativos com a organização no estilo “0800” ou “fale conosco”. Ela utiliza algoritmos de inteligência artificial e pode ser usada em questões operacionais e estratégicas, avaliando como o mercado está assimilando uma marca, por exemplo. O entrevistado 7 afirma que há modelos preditivos usando dados não estruturados que permitem análises e *insights* sobre interação com o consumidor. Mühlroth e Grottke (2018) e Yoon e Kim (2012) apresentam o NLP como sendo uma ferramenta para “minerar” SF em grandes conjuntos de dados.

O E8 afirma que a coleta e análise de dados não estruturados é uma consequência da evolução da aplicação das tecnologias *Big Data*. Muitas organizações estão armazenando dados deste tipo a partir de origens diversas, mas a análise efetiva ainda está no começo. Para analisar estes dados históricos é necessário primeiro um amadurecimento dos processos, com cientistas de dados e algoritmos. Também deve-se coletar dados ao longo do tempo para o treinamento dos modelos, de modo que as informações geradas pelas análises sejam assertivas e tenham melhor qualidade. Uma ferramenta BDA entrega mais valor na medida que obtém mais dados. Quanto maior o volume de informações, maior será a capacidade de aprendizado dos algoritmos, majorando as possibilidades de propor *insights*. Importante observar que parte dos exemplos relatados pelos entrevistados trata da análise de dados não estruturados com o foco em decisões operacionais e táticas, como reconhecimento de imagem, monitoramento de placas de automóveis, entre outros. O uso destes dados para obtenção de *insights* em nível estratégico corporativo, como já dito, deve ser uma evolução natural desses processos, de acordo com o amadurecimento técnico e cultural das organizações. Para que *insights* sejam transformados em conhecimento e possam ser candidatos a SF, visando antecipação e vantagem competitiva (DUAN; CAO, 2015), é essencial que as informações sejam observadas e interpretadas por uma equipe experiente de analistas e decisores (CONSTANTIOU; KALLINIKOS, 2015; SHARMA; MITHAS; KANKANHALLI, 2014). Para Lycett (2013) todo este processo é potencialmente inútil se os resultados não forem aproveitados na tomada de decisão e, para tal, um dos caminhos é o processo de criação de sentido, no qual os dados são a base de entendimento dos fenômenos estudados. Para Sharma, Mithas e Kankanhalli (2014) esse processo de geração de *insights* geralmente envolve vários papéis, de diferentes setores. A composição dessas equipes é resultado de decisões que podem potencializar ou restringir suas capacidades de gerar *insights*.

O entrevistado 11 afirma que a tomada de decisão estratégica é o grande objetivo de um BDA. Ele viabiliza a antecipação de eventos ou comportamentos. Em alguns casos de uso, os entrevistados relatam que os dados brutos possuem informações do ambiente, oportunizando monitorar mudanças e registrar histórico para futuras consultas. O entrevistado 12 ilustra um caso relatando que é possível avaliar mudanças de comportamento, a partir do agrupamento (*clustering*) de consumidores: em um momento, uma análise pode apresentar um pequeno grupo em relação aos demais, um *outlier*, representando um comportamento de determinado segmento da sociedade, sobre uma preferência por um produto, por exemplo. Ao longo do tempo, com dados históricos, é percebido que este *outlier* está crescendo, mas que, dentro do total da população, ainda não é representativo. Esta visão mais analítica talvez permita antecipar fenômenos de comportamento de mercado. Destaca-se um trecho da entrevista do entrevistado 12: “*Com big data e inteligência artificial conseguimos antecipar isso, ter um sinal de mudança. Se será relevante, permanente ou não, é difícil de dizer, mas existe um sinal de mudança.*”

Além do comportamento de consumidores, pode-se analisar o mercado também a partir da visão dos concorrentes, substitutos ou novos entrantes. Observa-se exemplos de uso do BDA para análise da concorrência, que pode gerar *insights* e percepções de possíveis SF sobre a sua atuação e desempenho, permitindo antecipar os próximos passos e eventos futuros. Foram relatados exemplos de análises realizadas a partir da coleta digital de processos jurídicos (entrevistado 6), que disponibiliza o monitoramento do comportamento do consumidor em relação aos produtos e serviços prestados pela organização em questão e seus concorrentes. Ao perceber nos processos jurídicos uma possível ameaça, a organização pode agir evitá-la, de forma proativa. O mesmo pode ser feito investigando dados em sites do tipo “Reclame aqui”, ou em mídias sociais, onde consumidores registram suas críticas e elogios. São exemplos práticos de como o BDA viabiliza este monitoramento do mercado e como é possível identificar candidatas a Sinais Fracos. Outro exemplo relatado é o da coleta de preços da concorrência. Ao analisá-los e apresentá-los como indicadores de valores relativos de mercado, comparando com os praticados pela organização, percebe-se sinais de como se comporta um nicho específico. O entrevistado 7 cita dois exemplos de análise de mercado a partir dos dados. O primeiro exemplo mostra que o BDA viabilizou a decisão de desenvolver e ofertar novos produtos e o cancelamento de outros, que geravam maior custo e menor rentabilidade. No segundo, após uma análise de *marketing* guiada pelos dados, gestores perceberam que havia duas marcas que deveriam ser comunicadas ao mercado como apenas uma. Assim, elas foram unidas em um novo produto digital. Este último caso

mostra, de forma clara, uma previsão e adaptação ao ambiente externo, na percepção de um novo fenômeno disruptivo da indústria e na decisão estratégica ágil antecipativa.

Questionou-se os entrevistados se as análises em BDA oferecem informações com características de SF (fragmentados, ambíguos, incompletos e incertos), conforme indicado pela literatura (BORTOLI; JANISSEK-MUNIZ; BORGES, 2015). Os entrevistados 3, 4, 5, 6, 7 e 8 afirmaram que cada informação obtida em BDA tem seu sentido e se completa em um escopo específico. Porém, para utilizá-la na tomada de decisão, para entender ou perceber a possibilidade de evento disruptivo futuro, deve-se conjugar aquela informação com outras. Assim, em contexto mais amplo, corporativo e estratégico, interno e externo, as análises em BDA proporcionam informações de natureza semelhante à dos Sinais Fracos. Destaca-se esta fala do E7: “[...] os dados, quando eles estão fora de contexto, eles dizem o que tu queres na realidade, e não o que eles querem dizer de fato. O dado não conta nada sozinho.”

Ele relata que é necessário juntar várias informações do BDA para uma análise completa, de um escopo ou contexto de pesquisa. Em alguns casos, além de dados quantitativos, é necessária uma análise qualitativa para obter, detalhar e refinar as conclusões obtidas no monitoramento. O entrevistado 8 apontou exemplos de vieses que podem ocorrer ao considerar apenas uma informação ou dado coletado, portanto, para aumentar a chance da antecipação de um fenômeno, é necessário correlacionar informações, assim como ocorrem com os SF (ANSOFF, 1975).

A maior parte dos entrevistados relata que suas organizações possuem uma área de inteligência, com pessoas dedicadas a analisar o mercado, definir estratégias, explorar hipóteses, trabalhando com BDA e outras ferramentas de apoio. Porém, não possuem processos definidos, tanto para executar as análises quanto para a tomada de decisão. Em geral as demandas têm origem nos gestores e decisores, que são repassadas para a área de inteligência, que as analisa e retorna. Entrevistados afirmam que *insights* obtidos em BDA passam, como os Sinais Fracos, pelos filtros de Ansoff (1975). Contudo, os filtros têm sua importância reduzida em razão da natureza do BDA. O filtro da vigilância é facilitado, pois pode-se estruturar o modelo de análise de forma indutiva, para que sejam acusadas as possibilidades de informações relevantes. Mudanças em um padrão geral de consumo, por exemplo, que utilizam dados internos e dados externos do mercado, podem gerar alertas que não passarão despercebidos pelos analistas. O filtro de poder também é amenizado quando se utiliza BDA, se a organização possui processo de tomada de decisão por meio dos dados (*data-driven*). Se o *insight* é obtido a partir de mensagens identificadas nos dados, ele é de mais fácil persuasão na hierarquia superior da organização. Ao responder sobre o filtro de

poder, alguns entrevistados citaram a palavra “confiança”. É necessário ter confiança nos dados e confiança na equipe de análise para que o filtro de poder seja mais fraco.

Entrevistados afirmam que a tomada de decisão a partir dos dados é uma tendência e está ocorrendo nas organizações que possuem maior maturidade em *analytics* e *data-driven*. De fato, empresas que entendem e confiam nos seus dados já os utilizam, de inúmeras formas, no seu planejamento estratégico. O entrevistado 6 tem a percepção de que, no Rio Grande do Sul, empresas locais estão mudando seus processos neste sentido, a partir da influência de concorrentes, por isomorfismo mimético (HININGS; GEGENHUBER; GREENWOOD, 2018). As entrevistas relatam, porém, que as empresas estão apenas iniciando esta mudança, portanto, em geral as decisões continuam sendo realizadas de maneira intuitiva.

A literatura indica que um *outlier* pode ser uma evidência de sinal fraco, a ser confirmado em análises posteriores (KIM; LEE, 2017; MENDONÇA; CARDOSO; CARAÇA, 2012). As análises em BDA, nas suas diversas abordagens, proporcionam o destaque dos *outliers*, cabendo aos analistas priorizá-los ou não em seus estudos. Nas entrevistas, percebe-se que há organizações que examinam prioritariamente os *outliers*, buscando entender o motivo de eles existirem, porque há pontos fora da curva. Para estas, *outliers* são informações importantes pois propiciam a identificação de ameaças e oportunidades. Se o *outlier* representar uma ameaça de um evento futuro, a organização pode agir, evitando-a ou preparando-se para quando ela surgir. Se for uma nova oportunidade, a estratégia pode ser agir a fim de que este *outlier* entre para a média, aumentando sua amplitude, de modo que ele deixe de ser um *outlier* no futuro e se transforme em algo de grande valor. Destaca-se as seguintes opiniões: “Um ponto fora da curva pode deslocar a média. Então, um outlier pode apresentar uma previsão de uma nova média”; “Sempre que olho para um outlier, descubro como posso adequar minha estratégia para que deixe de ser um outlier.”

Para o entrevistado 4, as organizações não estão olhando para os *outliers* como deveriam. Mesmo assim, ele acredita que o *outlier* representa uma informação estatística que pode ser mais relevante do que o resto da amostra. Afirma que é pelos *outliers* que são obtidos os conhecimentos invulgares, de acordo com a literatura. A análise descritiva estatística sobre os dados sempre deve ser feita, em muitas situações as informações médias são relevantes. Contudo, no volume de dados de um *Big Data*, analisar os *outliers* pode ser ainda mais importante, pois viabiliza uma série de possíveis descobertas: inovações, fenômenos disruptivos ou mudanças estão acontecendo fora das tendências médias

(GEORGE; HAAS; PENTLAND, 2014). O entrevistado 7 afirma que, na análise quantitativa, os *outliers* são removidos, para que os grandes padrões e médias dos dados sejam avaliados e confirmados, antes que os cálculos estatísticos sejam realizados. Porém, como consideram os *outliers* importantes, em alguns casos é realizada uma análise mais profunda e detalhada, para explicar tais ocorrências. Estas análises, em geral, são qualitativas e descritivas, gerando relatórios com as possíveis causas das ocorrências dos *outliers* e, principalmente, com os *insights* obtidos. Esta análise ocorre de forma menos frequente e o objetivo principal é entender quem são os *outliers*, como, por exemplo, clientes com um perfil muito diferente.

A maior parte dos entrevistados afirma que o método indutivo é utilizado nas organizações, colhendo novas informações e *insights* não deduzidos *a priori*. Acredita-se, portanto, que esse método proporcione a identificação de um número maior de candidatos a SF. O entrevistado 6 informa que, para que esta abordagem gere frutos, é necessário definir um norte estratégico. Assim como na inteligência antecipativa há Atores e Temas (JANISSEK-MUNIZ; LESCA; FREITAS, 2006), também é importante definir escopos nas pesquisas indutivas. Sem isso, o método pode levar a vieses, já que existem inúmeras variáveis que podem ser relacionadas. Outro exemplo apresentado pelo entrevistado 7 trata de análises que apresentaram correlações entre o comportamento dos clientes e serviços utilizados. Segundo o entrevistado, estas correlações singulares não poderiam ter sido deduzidas *a priori*. A organização, mediante este *insight*, pôde se adaptar a uma nova realidade de padrão de consumo, percebida antecipadamente pelos analistas.

Quando se questiona se o BDA pode ser uma ferramenta utilizada para monitorar o ambiente, a maioria dos entrevistados afirma que sim. Na literatura, Mendonça et al (2012) evoca a necessidade de dispositivos intelectuais para a detecção de SF e afirma a possibilidade de técnicas de processamentos de dados para este fim. Erevelles et al (2016) e Mühlroth e Grottke (2018) afirmam que BDA pode ser o caminho. Os entrevistados apontam inúmeros exemplos de análises realizadas, que detectaram SF com supervisão humana, em dados internos e externos, estruturados e não, sempre por meio de um grande volume coletado. Na Tabela 8, apresenta-se uma compilação das respostas, realizada a partir dos códigos e categorias da análise de conteúdo.

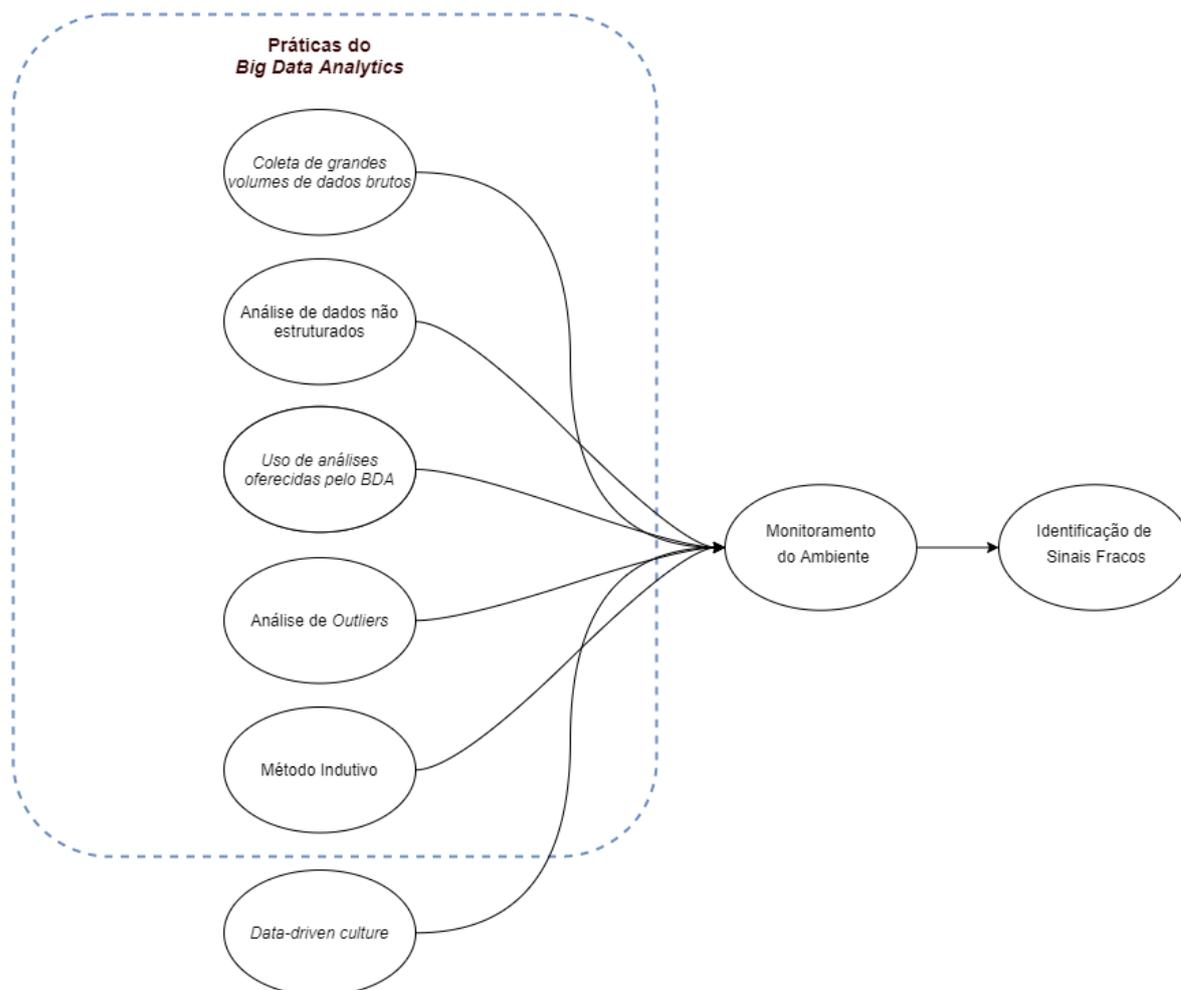
Tabela 8: Resumo das respostas

Perguntas	Resumo das Respostas
P01	A maior parte dos entrevistados (7) informou que suas organizações coletam grandes volumes de dados brutos, para posteriormente modelar e analisar. Apenas dois relatam que tratam, em grande parte, de dados internos e transacionais, portanto, já processados de certa forma pelos sistemas operacionais corporativos e funcionais.
P02	Quatro entrevistados afirmaram que coletam e processam dados rapidamente, com velocidade, em <i>real-time</i> . Estes dados permitem agilidade na percepção de mudanças do ambiente. Suas análises, julgamentos e decisões podem ser caracterizadas como operacionais e táticas.
P03	Seis entrevistados relatam que as organizações coletam dados não estruturados. Apenas um, que trabalha com empresas de porte médio e pequeno, informa que este tipo de coleta não é atualmente uma preocupação. Estas empresas ainda estão criando processos para implementar BI a partir de dados estruturados internos. Algumas empresas estão coletando e analisando dados não estruturados, outras apenas coletando e algumas, com maior maturidade, informam que estas análises possibilitam orientações operacionais e táticas, em tempo real.
P04	Oito entrevistados informam que as análises realizadas em BDA permitem a percepção de fenômenos futuros, oportunizando a antecipação de oportunidades, ameaças e eventos disruptivos. Um entrevistado afirma que não percebe o uso efetivo do BDA para este tipo de análise estratégica.  Seis entrevistados afirmam que em um contexto corporativo e estratégico, as análises em BDA oferecem informações incompletas, incertas, de difícil detecção, fragmentadas e/ou ambíguas.  Cada informação oferecida pelo BDA tem seu motivo e sua completude, mas, para utilizá-la na tomada de decisão, sempre é necessário conjugar ou correlacionar com outras, do próprio BDA ou colhidas em outros contextos.  Cinco entrevistados relatam que as organizações, mesmo com áreas de inteligência, não possuem processos bem definidos, formais e sistemáticos, para coleta e criação de sentido.
P05	Sete entrevistados relatam que os filtros de Ansoff (1975) para SF se aplicam aos resultados das análises em um BDA. Em geral os filtros são amenizados em virtude das características destas ferramentas.
P06	Metade dos entrevistados afirma que suas empresas usam <i>insights</i> colhidos em BDA na tomada de decisão, sendo consideradas como <i>data-driven</i> . Afirmam também, porém, que estão apenas começando este processo, há um longo caminho de maturidade a percorrer. Boa parte das decisões continua sendo realizada de forma intuitiva.
P07	Metade dos entrevistados informa que suas organizações não utilizam <i>outliers</i> , não olham para estas informações e, para realizar uma análise estatística robusta, os <i>outliers</i> são removidos da amostra. Outros afirmam que eles são imprescindíveis, são representações de informações importantes e possíveis indicadores de fenômenos futuros. Um entrevistado pontua que, para estudar padrões eles são descartados, mas, como são importantes, são examinados de forma profunda e detalhada em pesquisas qualitativas.  O entrevistado 01 percebe que <i>outliers</i> são os primeiros dados que são analisados, priorizados, pois podem ser manifestações de oportunidades/ameaças.
P08	Sobre o método de pesquisa, dois entrevistados afirmam que utilizam apenas o método dedutivo. Um enfatizou que não usa na sua empresa o método indutivo, mas é uma tendência. Os demais salientam que o método indutivo é utilizado de alguma forma, obtendo bons resultados, novas informações e novos <i>insights</i> não deduzidos <i>a priori</i> .
P09	Sete entrevistados afirmam que o BDA pode ser uma ferramenta utilizada para monitorar o ambiente, apresentando em suas análises as possíveis mudanças do meio e candidatos a SF. O entrevistado 11, por outro lado, informa que esta atividade depende somente de humanos.

Fonte: elaborado pelo autor

Os resultados revelam que uma maior ênfase nas estratégias de busca de *insights* a partir dos grandes volumes de dados, como a coleta e análise de dados brutos e não estruturados, o estudo de *outliers*, o uso da abordagem indutiva, o uso de técnicas de análise disponíveis no BDA e a promoção da cultura *data-driven*, favorecem a identificação de um número maior de Sinais Fracos. Neste sentido, sugere-se o modelo apresentado na Figura 8, que trata dos fatores relacionados a BDA que influenciam na percepção de Sinais Fracos.

Figura 8: Fatores do BDA que influenciam na identificação de Sinais Fracos



Fonte: elaborado pelo autor

## 2.6 Considerações Finais

O uso de BDA para monitoramento do ambiente e identificação de Sinais Fracos é ainda um tema pouco abordado na literatura de *Big Data* e Inteligência Antecipativa. Este trabalho apresentou um estudo empírico inicial, associando essas disciplinas da área de Sistemas de Informação. Para responder se os gestores percebem Sinais Fracos por meio das ferramentas

BDA e se estes estão sendo considerados na tomada de decisão estratégica, realizou-se uma pesquisa qualitativa coletando dados em 10 entrevistas, com profissionais da área de inteligência com experiência em BDA.

Como resultado, são apresentadas evidências de que dados candidatos a SF podem ser identificados a partir do BDA. Observou-se que algumas organizações percebem SF por meio de dados, mesmo sem possuir processos formais definidos. Apresentando métodos e processos de como o BDA pode ser utilizado para este fim, este trabalho trouxe exemplos de análises já sendo conduzidas nas organizações.

Em termos de contribuições práticas, evidenciou-se que as organizações podem utilizar BDA como uma forma de monitorar o mercado, para antecipar fenômenos, oportunidades e ameaças, ilustrando neste estudo, exemplos e caminhos para este fim. Para tal, necessita-se focar nas seguintes práticas do BDA: a coleta e análise de dados brutos e não estruturados, a análise de *outliers*, o uso da abordagem indutiva e de técnicas de análise disponíveis no BDA. Por fim, para viabilizar a identificação de um número maior de Sinais Fracos por meio do BDA, torna-se importante que a organização da invista na cultura *data-driven*.

Percebe-se igualmente, com os resultados do estudo, que o ambiente social-econômico está ficando a cada dia mais complexo, dado o volume de informações que aumenta exponencialmente. Desta forma, os gestores já não são capazes, sem uma ferramenta, de analisar o que é recebido e perceber o que está mudando no mercado (MÜHLROTH; GROTTKE, 2018).

Como contribuição teórica, ressalta-se a discussão dos dados empíricos e seus pontos convergentes com a literatura estudada. Sugere-se, também, que o BDA seja incorporado como um mecanismo de monitoramento do ambiente, como um meio que permite a coleta, processamento e análise de dados do contexto empresarial, interno e externo. Cita-se, ainda como contribuição para a literatura, o paralelo entre os Sinais Fracos e as informações obtidas nos processos de análise usando BDA, realizado a partir do referencial teórico deste estudo.

Considerada como um dos fatores que justificam esta pesquisa, a mudança na área da inteligência antecipativa gerada pelo advento do BDA é determinante, sendo imprescindível que se passe a colher informações por meio destas novas ferramentas. Tem-se, assim, uma oportunidade de atualização dos processos da inteligência antecipativa a partir da transformação digital. Ao monitorar o ambiente também pelo BDA, em razão do

volume de informações, é possível considerar mais variáveis nas práticas de criação coletiva de sentido.

Cita-se como limitações do estudo o número de entrevistas realizadas. Foram poucas organizações estudadas, de uma mesma região geográfica, o que aumenta a possibilidade de vieses. Da mesma forma, o perfil de empresas é majoritariamente voltado para a indústria de Tecnologia e Comunicação. Entende-se estes limites como oportunidades de estudos futuros. Nesta linha, como sugestão de pesquisa, propõe-se a definição de um modelo contendo elementos de impacto do BDA no monitoramento do ambiente e na percepção de possíveis Sinais Fracos, contribuindo assim para a evolução do processo da inteligência antecipativa.

De fato, percebe-se que o BDA deve colaborar para tornar os processos de inteligência antecipativa mais eficientes e para que o ritmo atual de volume e complexidade de informações seja acompanhado. Sistemas digitais são necessários para ajudar na coleta e análise das informações, proporcionando melhores percepções sobre o ambiente corporativo (ECKHOFF et al., 2014; KELLER; VON DER GRACHT, 2014; PARK; EL SAWY; FISS, 2017).

## **2.7 Referências**

As referências bibliográficas deste artigo encontram-se no Capítulo 5 - Referências.

### 3 Artigo II - O efeito do *Big Data Analytics* na identificação de Sinais Fracos

*Fabiano Chiapinotto Saffi*

#### **Resumo**

Com a finalidade de reduzir a incerteza na tomada de decisão, antecipar eventos futuros, e viabilizar o alinhamento rápido da estratégia às transformações do mercado, o monitoramento ativo do ambiente é uma prática importante para as organizações. Nas últimas décadas encontra-se um volume cada vez maior de dados disponibilizados por tecnologias digitais. Com o propósito de possibilitar a análise desses grandes volumes de dados e utilizá-los no entendimento do comportamento do mercado, organizações estão investindo em ferramentas chamadas *Big Data Analytics* (BDA). O objetivo deste estudo é analisar o impacto do uso de práticas do BDA como fonte de informações para a Inteligência Antecipativa. Pretendeu-se responder a seguinte questão de pesquisa: Qual o efeito das práticas do *Big Data Analytics* na identificação desses sinais? Neste intuito, desenvolveu-se construtos e hipóteses que foram relacionadas em um Modelo Conceitual. Para suportá-lo, realizou-se uma pesquisa quantitativa, do tipo *survey*, com 123 respondentes. Os dados foram analisados por meio de modelagem de equações estruturais (*PLS-SEM*). Os achados indicam que o uso das práticas do BDA tem efeito positivo na identificação de Sinais Fracos. Como contribuições teóricas cita-se o suporte empírico ao modelo desenvolvido e as evidências indicando que o uso das práticas do BDA e a identificação de Sinais Fracos estão positivamente relacionados.

**Keywords:** *Big Data Analytics*; Inteligência Antecipativa; Sinais Fracos.

#### **3.1 Introdução**

O cenário para a decisão e ação gerencial tem se tornado a cada dia mais complexo (MÜHLROTH; GROTTKE, 2018), em uma realidade onde a previsibilidade de eventos futuros tem reduzido sensivelmente (ROHRBECK; BADE, 2012). Inovações disruptivas, provocadas pelo ambiente de negócio incerto (JISSINK; SCHWEITZER; ROHRBECK, 2019), pela competição acirrada, pela rápida mudança no comportamento do consumidor (MÜHLROTH; GROTTKE, 2018) e pelos breves ciclos tecnológicos (KIM; PARK; LEE, 2016), são apresentados de forma cada vez mais acelerada ao mercado. Neste contexto, a compreensão dos fenômenos de forma antecipada se mostra ainda mais importante, pois permite a identificação de oportunidades de negócios antes dos concorrentes (YOON, 2012) e o posicionamento estratégico para redução de riscos (CHOO, 2001; JANISSEK-MUNIZ;

FREITAS; LESCA, 2007). Sustentar e melhorar a competitividade depende, em grande parte, dessa capacidade de identificar mudanças e evoluções no ambiente empresarial (LESCA et al., 2015). Mesmo com acesso a inúmeras fontes e formatos de informações, organizações sentem falta de evidências para a antecipação de eventos incertos (MOREIRA et al., 2015).

Considerando a dinamicidade da época e a demanda por indícios do que pode acontecer, Igor Ansoff (1975), teórico pioneiro do planejamento estratégico, apresentou uma proposição indicando que mudanças disruptivas podem ser previstas através da identificação de pequenos sinais no ambiente, interpretando-os em conjunto e alimentando planos estratégicos para respostas organizacionais. Ansoff chamou estes elementos de Sinais Fracos (SF). Eles são indicadores antecipados de fenômenos (MENDONÇA; CARDOSO; CARAÇA, 2012) e possuem características iniciais como: são incompletos, incertos, de difícil detecção, fragmentados e ambíguos (HEUPEL; VON JUTERZENKA, 2015; LESCA, 2003). É preciso implementar processos de análise e criação de sentido, de forma sistemática e contínua, para que eles forneçam valor (JANISSEK-MUNIZ; BORGES; BORTOLI, 2015). Esta prática implica em ser capaz de perceber, selecionar e coletar pequenos elementos informacionais que estão normalmente imersos em um grande “ruído” de dados, compartilhando-os com interessados, para que, em processos de interpretação coletiva, possam decidir e adaptar a organização aos sinais encontrados (LESCA et al., 2015).

Avanços nas tecnologias digitais, como o rápido desenvolvimento da *internet* e seus artefatos (VEUGELERS; BURY; VIAENE, 2010), têm disponibilizado um número cada vez maior de fontes heterogêneas de dados (KAYSER; BLIND, 2017; KITCHENS et al., 2018). Pessoas, alertas, sensores, dispositivos móveis, documentos, entre inúmeras outras entidades, estão cada vez mais conectados, o que transforma significativamente o método de gerar e analisar as informações (GROVER et al., 2018). Essa velocidade de eventos, aliada ao progresso tecnológico, estão provocando essa “avalanche” constante de dados (KELLER; VON DER GRACHT, 2014), os quais, se adequadamente utilizados para fins de inteligência empresarial, podem oferecer informações pertinentes sobre diferentes ambientes que circundam a organização, como o político, o econômico, o social e o tecnológico (LYCETT, 2013).

Contudo, em razão do volume, para identificar informações relevantes, torna-se necessário um esforço maior de monitoramento (KAYSER; BLIND, 2017). Sem métodos e ferramentas adequados, é inviável a coleta e análise deste volume de informações (KELLER;

VON DER GRACHT, 2014), inviabilizando a antecipação de eventos por meio deles (LEE, 2018; WELZ et al., 2012). À medida que os dados são mais complexos e em maior quantidade, a racionalidade limitada (SIMON, 1955) dos analistas limita sua interpretação (EREVELLES; FUKAWA; SWAYNE, 2016), e as informações desejadas passam a ser difíceis de encontrar (THORLEUCHTER; SCHEJA; VAN DEN POEL, 2014), tornando as decisões ineficientes (JIANG; GALLUPE, 2015). Percebe-se, neste sentido, um aumento na dificuldade de decisão a partir do monitoramento do ambiente.

Gerenciar este dilúvio de informações, remover ruídos e distrações indevidas enquanto buscam por Sinais Fracos, são desafios que têm se demonstrado cada vez mais importantes para a organização (SCHOEMAKER; DAY; SNYDER, 2013). Da mesma forma, profissionais estão sendo afetados por este volume de dados, o qual, pela sua dimensão, está além da capacidade de compreensão (DAY, 2011). Ao invés de monitorar e capturar sinais para alimentar o planejamento da organização, antecipar eventos (EREVELLES; FUKAWA; SWAYNE, 2016) e identificar fenômenos futuros, gestores seguem definindo estratégias sem incorporar os movimentos disruptivos do mercado (DAY, 2011). Embora a literatura já tenha apresentado inúmeros métodos que permitam obter e analisar informações antecipativas (KAIVO-OJA, 2012), eles carecem de ferramentas computacionais associadas que suportem suas etapas (MOREIRA et al., 2015). Neste sentido, há uma lacuna relacionada à tecnologia que permita a detecção de Sinais Fracos (MÜHLROTH; GROTTKE, 2018).

Uma das formas de adquirir, armazenar e processar dados externos se desenvolve por meio de sistemas de informações que tratam grandes volumes de dados, chamados *Big Data* (KOSTIN, 2018). O termo *Big Data* caracteriza sistemas que armazenam e processam **volumes** de dados amplamente **variados**, que são gerados, capturados e processados em alta **velocidade** (CHEN; CHIANG; STOREY, 2012; GÜNTHER et al., 2017; LANEY, 2001). Sua natureza exige abordagens, ferramentas e tecnologias diferentes daquelas disponíveis anteriormente (CORBETT; WEBSTER, 2015). Novas técnicas permitiram seu advento (CORBETT; WEBSTER, 2015), que, em conjunto com a disciplina de ciência de dados, se desenvolveu na academia e nas empresas.

Para definir este conjunto de ferramentas, envolvendo dados, seu processamento e análise, a literatura utiliza o termo *Big Data Analytics* (**BDA**) (JANSSEN; VAN DER VOORT; WAHYUDI, 2017). Estudos da indústria têm destacado o desenvolvimento do BDA (JIANG; GALLUPE, 2015). Lavalley (2011), em pesquisa com mais de 3000 CIOs,

mostra que organizações de maior performance utilizam *Big Data* 5x mais do que as demais. Segundo Barton & Court (2012), empresas que usam BDA em suas operações apresentam taxas de rentabilidade e produtividade 5% ou 6% maiores que seus pares. Muita atenção também está sendo dada na literatura acadêmica ao valor que as organizações podem obter pelo uso do BDA. Chen et al. (2012), por exemplo, sugerem que ele pode ajudar as organizações a entender melhor seus negócios e mercados, alavancando oportunidades e sinalizando possíveis ameaças. A essência do BDA é a transformação de muitos dados, nos seus mais variados tipos, em informação e conhecimento para a tomada de decisão (DUAN; CAO, 2015; GOES, 2014).

O desenvolvimento do BDA tem permitido às empresas coletar e armazenar grandes volumes de dados nos chamados *Data Lakes* (CHEN et al., 2017; KITCHENS et al., 2018; WATSON, 2017) e, a partir da sua análise, identificar Sinais Fracos (MÜHLROTH; GROTTKE, 2018). O BDA está relacionado à forma de como as organizações percebem o que acontece no mercado (CONSTANTIOU; KALLINIKOS, 2015; EREVELLES; FUKAWA; SWAYNE, 2016). Ele aprimora o monitoramento do ambiente (DUAN; CAO; EDWARDS, 2018), oferecendo recursos para conhecer melhor seus clientes (KITCHENS et al., 2018; LAM et al., 2017), provendo a capacidade de prever (BRYNJOLFSSON; GEVA; REICHMAN, 2016; EREVELLES; FUKAWA; SWAYNE, 2016) e, assim, melhor servi-los (CHEN et al., 2017; WATSON, 2017). Neste sentido, e acreditando que os processos para coleta e identificação de Sinais Fracos possam se beneficiar destas novas ferramentas quantitativas, define-se a questão que orienta este trabalho: **Qual o efeito das práticas do BDA na identificação de Sinais Fracos?** Objetiva-se analisar e medir empiricamente o impacto do emprego de práticas do BDA como fonte de informações para a Inteligência Antecipativa. Investiga-se, portanto, os potenciais usos de ferramentas BDA para identificar Sinais Fracos em processos de Inteligência Antecipativa.

Muitas organizações não desenvolvem processos sistemáticos e estruturados de inteligência (CAINELLI; JANISSEK-MUNIZ, 2019), e a cultura organizacional é um dos fatores relevantes para este resultado (SOUZA, 2019). O êxito na implementação de processos de inteligência está fortemente relacionado à cultura organizacional (KONONIUK; SACIO-SZYMAŃSKA, 2015), sendo significativa, portanto, a implementação de uma cultura que os favoreça (JANISSEK-MUNIZ, 2016; SOUZA, 2019). Cultura organizacional é definida como um sistema compartilhado de valores, crenças, suposições e objetivos (CEKULS, 2015), norteado pelos líderes, que compõe ou influencia comportamentos e processos (SOUZA, 2019). Especificamente na temática deste estudo, um

foco apropriado da cultura organizacional quando se estuda a respeito de Inteligência e BDA, é o conceito de cultura da decisão guiada por dados (*Data-driven Culture* ou **DDC**). A DDC refere-se a este sistema compartilhado, difundindo e padronizando a crença da necessidade de usar dados na tomada de decisão. A DDC incentiva, em toda organização, a prática da tomada de decisão baseada em evidências (DUAN; CAO; EDWARDS, 2018; KIRON; SHOCKLEY, 2011). O termo *Data-driven Culture* tem sido utilizado há muitos anos, contudo, obteve maior destaque com o advento do *Big Data* (DUAN; CAO; EDWARDS, 2018). Autores evidenciam a importância da DDC para obter maior vantagem competitiva e maior retorno de valor por meio do BDA (CHEN et al., 2017; DAVENPORT, 2006; GROVER et al., 2018; KIRON; SHOCKLEY, 2011; LAVALLE et al., 2011), visto que ele é usado no apoio à decisão (GÜNTHER et al., 2017; KITCHENS et al., 2018), ao provocar *insights* e gerar conhecimento, reduzindo a incerteza e majorando a capacidade de decidir estratégicas assertivas (CHEN; PRESTON; SWINK, 2015).

A transformação que ocorre nos processos de inteligência e na cultura da tomada de decisão (FRISK; BANNISTER, 2017), impulsionada pelo advento do grande volume de informações digitais (CONSTANTIOU; KALLINIKOS, 2015), denota a relevância do tema deste trabalho. O *Big Data* abriu novas possibilidades para o suporte à decisão (MIAH et al., 2017), porém, a maior parte da literatura sobre o BDA se concentra em como aprimorar capacidades organizacionais táticas (GROVER et al., 2018), pouco abordando sua utilização como apoio à decisão estratégica (VAJHALA; STRANG; SUN, 2015). Há estudos sobre a detecção de Sinais Fracos por meio de *Big Data*, a exemplo de Duan e Cao (2015), afirmando que o BDA influencia diretamente no monitoramento do ambiente. Entretanto, em sua maioria, eles apresentam pesquisas em características funcionais e tecnológicas, não há estudos teóricos propondo modelos relacionando estas duas áreas de pesquisa (MÜHLROTH; GROTTKE, 2018). Não encontram-se estudos sobre detecção de Sinais Fracos por meio de sistemas de informações que apresentem uma visão geral de como as tecnologias podem ser empregadas (ECKHOFF et al., 2014).

Para perseguir os objetivos desta pesquisa, empregou-se uma abordagem dedutiva. Foram criados construtos e hipóteses a partir do referencial teórico e dos resultados do Artigo I, associados aos temas deste estudo, os quais foram relacionados em um modelo que busca medir o impacto das práticas do BDA na identificação de Sinais Fracos. Para suportar o modelo, seus construtos independentes e suas relações, realizou-se uma pesquisa quantitativa, do tipo *survey*, com 123 respondentes. A sequência deste trabalho estrutura-se da seguinte forma: apresenta-se a revisão teórica de literatura, com principais conceitos e

teorias utilizados, seguidos do modelo de pesquisa e do método empregado. Adiante, abordam-se os resultados, suas discussões e considerações finais.

## **3.2 Referencial Teórico**

*Big Data* é um produto da era da informação e potencializa uma nova fronteira de vantagem competitiva para as organizações (CAESARIUS; HOHENTHAL, 2018; SALLEH, 2016). Ele desafia as epistemologias estabelecidas, criando novos paradigmas (KITCHIN, 2014). Suas ferramentas, técnicas e métodos de pesquisa, diferentes dos sistemas tradicionais (LI; ZHANG, 2018), são relevantes entre diversos setores econômicos, empresariais e acadêmicos. Para melhor compreender o efeito de suas práticas na Inteligência Antecipativa e desenvolver os construtos para a pesquisa quantitativa, revisou-se os seguintes temas na literatura de SI: Sinais Fracos e a Inteligência Estratégica Antecipativa, a cultura da decisão guiada por dados (*Data-driven Culture*), o BDA e as práticas que, segundo o Artigo 1, influenciam a identificação de sinais fracos. Nas próximas seções são citadas as principais referências encontradas.

### **3.2.1 Sinais Fracos e a Inteligência Estratégica Antecipativa**

O monitoramento atento do ambiente externo, a coleta e análise de informações relevantes, são atividades de Inteligência que permitem obter maior discernimento do ambiente empresarial, com o foco na tomada de decisão estratégica para a resiliência e o sucesso econômico de longo prazo das organizações (CAINELLI; JANISSEK-MUNIZ, 2019; JANISSEK-MUNIZ, 2016). Segundo Simon (1960), decisão estratégica eficaz é uma habilidade imperativa em mercados voláteis e competitivos e, para que haja um posicionamento assertivo, a compreensão do ambiente deve ser sua base. Foi Aguilar (1967) quem inicialmente defendeu essa ideia de atenção ao ambiente para o planejamento do futuro e determinação das ações a serem tomadas. Ela surgiu pela percepção da incapacidade dos analistas e dos processos de planejamento em identificar antecipadamente as mudanças rápidas no mercado. A organização deveria, para tal, promover um sistema de alerta, observando o ambiente em um processo sistemático (FONSECA; BARRETO, 2011). Criando incentivos e provendo a cultura de atenção, uma organização vigilante é capaz de observar um mercado amplo, explorando e coletando informações relevantes (DAY; SCHOEMAKER, 2009). A Inteligência Estratégica Antecipativa (**IEA**) tem como objetivo oferecer percepções de acontecimentos futuros no contexto organizacional, sendo fundamentada a partir do monitoramento nos ambientes socioeconômicos internos e

externos e na interpretação de pequenos sinais coletados (JANISSEK-MUNIZ; FREITAS; LESCA, 2007; RIOS et al., 2011). A IEA defende que eventos identificados em dados históricos não são suficientes para apontar possíveis eventos disruptivos (ROSSEL, 2009). Deste modo, ao contrário de análises que identificam padrões usando informações do passado, a IEA busca a criação de sentido em cenários representativos definidos por meio de sinais percebidos no ambiente, os quais, pelas suas características e interpretação, podem representar evidências de possíveis fenômenos futuros (MOREIRA et al., 2015). Estes sinais são chamados de Sinais Fracos (SF), conceito proposto pela primeira vez no trabalho seminal de Ansoff (1975).

Pela Teoria dos Sinais Fracos, a antevisão do que pode acontecer deve iniciar por esses sinais (HEUPEL; VON JUTERZENKA, 2015), antecipatórios, qualitativos, que podem chegar em diversos formatos, a partir de fontes distintas (FONSECA; BARRETO, 2011). São informações orientadas para o futuro, que pela sua interpretação permitem antecipar mudanças e fenômenos emergentes (HILTUNEN, 2008). Eles são chamados de “fracos” porque, quando identificados, possuem um baixo volume de emissão, percepção e interpretação (LEE, 2018). Não são considerados como parte de uma tendência, são apenas “rastros” que permitem entender como um evento importante pode ocorrer (LEE, 2018). Seus impactos, a princípio, não podem ser estimados de forma acurada (ANSOFF et al., 1984; THORLEUCHTER; SCHEJA; VAN DEN POEL, 2014), eles são construções interpretativas, baseadas em dados brutos, compilados em *insights* por um “olhar” atento e interessado. *A priori*, SF são apenas hipóteses, uma combinação de lógica dedutiva, *insights* obtidos em conjunto com outras informações (ROHRBECK; THOM; ARNOLD, 2015). Sua importância reside, portanto, na possibilidade de correlacionar a outros SF (LESCA, 2003), o que pode levar a novos caminhos (MENDONÇA; CARDOSO; CARAÇA, 2012) em decisões estratégicas.

No futuro, em novas análises, associando a novos sinais (LESCA, 2003), eles podem ser fortalecidos, quando representam eventos que aos poucos se materializam (KIM; LEE, 2017; SCHOEMAKER; DAY; SNYDER, 2013). Tornam-se sinais fortes quando a interpretação do seu significado fica clara. Sendo assim, perceber estas pequenas “sementes” o quanto antes é uma capacidade valiosa para o desenvolvimento e resiliência das organizações (ROHRBECK; THOM; ARNOLD, 2015). Os SF estão relacionados a uma variedade de possibilidades, desde eventos e problemas moderadamente incertos a outros muito incertos e disruptivos (KAIVO-OJA, 2012). O tratamento desses sinais para geração de conhecimento são processos estratégicos de busca por oportunidades e prevenção de

ameaças (ROSSEL, 2009), permitindo a antecipação de mudanças em sistemas que podem ser direcionados a novos caminhos (HOLOPAINEN; TOIVONEN, 2012; MENDONÇA et al., 2004; MENDONÇA; CARDOSO; CARAÇA, 2012).

Em termos de métodos para operacionalização desses processos, a literatura menciona diversas formas pelas quais os SF são coletados, interpretados, percebidos e identificados (HOLOPAINEN; TOIVONEN, 2012). Em diferentes organizações e contextos sociais, a análise dos SF é realizada por caminhos variados (KAIVO-OJA, 2012; ROSSEL, 2012). Com essas inúmeras formas de observar o ambiente e criar inteligência, em escolas distintas a IEA vem sendo estudada pela academia e praticada pelas empresas ao longo dos anos. Entre os métodos que tratam deste assunto, pode-se citar: *Strategic Issue Management* do próprio Ansoff (1975), *Environmental Scanning* (CHOO, 2001), *Learning Environmental Scanning* (LESCA, 2003), *Strategic Foresight* (ROHRBECK, 2010), *Strategic Radar* (SCHOEMAKER; DAY; SNYDER, 2013), e *Early Warning System* (BEDENIK et al., 2012).

Processos de inteligência antecipativa, que envolvem monitoramento, análise e *insights* para tomada de ação, torna as empresas protagonistas e não reativas, visto que possibilita o alinhamento adiantado entre as circunstâncias do meio e a ação. Tal mecanismo é um complemento ao planejamento estratégico, focado no longo prazo (FONSECA; BARRETO, 2011), considerando constantemente as variáveis ambientais (JANISSEK-MUNIZ, 2016). Um dos pré-requisitos para obter estes *insights* é a existência de um método sistematizado de monitorar o ambiente (ILMOLA; KUUSI, 2006). Sendo um processo, requer atividades implementadas de forma contínua, como uma capacidade difundida pela **cultura** da organização (RHISIART; MILLER; BROOKS, 2015).

### 3.2.2 *Data-driven Culture*

Assiste-se a uma revolução relacionada a possibilidade de obtenção e uso de dados, na qual organizações estão sendo fortemente impactadas, especialmente pela viabilidade de tomar decisões baseadas em grandes volumes de dados (RAGUSEO, 2018). Para lidar com essas importantes mudanças na maneira de gerenciar seus negócios, elas estão investindo na promoção de uma nova cultura, voltada a orientação baseada em dados, fomentando a chamada *Data-driven Culture* (**DDC**) (BARTON; COURT, 2012).

A *Data-driven Culture* (ou *Data-Oriented Culture*) trata de um padrão de comportamentos, crenças e práticas em uma organização, consistentes com a tomada de

decisão analítica (HOLSAPPLE; LEE-POST; PAKATH, 2014) e alinhados à ideia de que obter *insights* a partir dos dados é imprescindível para o sucesso dos negócios (DUAN; CAO; EDWARDS, 2018; KIRON; FERGUSON; PRENTICE, 2013; KIRON; SHOCKLEY, 2011). Esta definição está de acordo com a literatura sobre cultura organizacional (CEKULS, 2015; DUAN; CAO; EDWARDS, 2018; SOUZA, 2019). Com a DDC é possível aplicar o método científico sobre os dados de forma efetiva, partindo das evidências encontradas para a tomada de decisão racional (LI; ZHANG, 2018). Para Kiron e Shockley (2011), a DDC em possui três características: a análise sobre dados é um ativo importante; a alta gestão apoia análises em dados em toda organização; *insights* devem estar disponíveis para quem precisar.

A base da DDC é implementar processos para analisar, compilar e entregar as informações importantes aos tomadores de decisão (VASSAKIS; PETRAKIS; KOPANAKIS, 2018). Essa capacidade, de decidir e agir a partir de práticas guiadas por dados, é caracterizada como um ativo valioso (VASSAKIS; PETRAKIS; KOPANAKIS, 2018), determinante para a competitividade e inovação (MEDEIROS, 2018). Para Duan e Cao (2015), a DDC impacta diretamente na inovação de novos produtos e indiretamente na sua significância. Para Lavallo et al. (2011), organizações que efetivamente adotam a gestão *data-driven* estão no caminho de se tornarem líderes em seus setores. Conforme resultados apresentados no Artigo I, a promoção da cultura *data-driven*, em conjunto com o uso do BDA, facilitam a identificação de um maior número de Sinais Fracos.

É significativo ter métodos eficientes de processamento de grandes volumes de dados para de favorecer a tomada de decisão *data-driven* (SIVARAJAH et al., 2017). Davenport (2006) afirma que organizações que almejam usar o BDA para aumentar sua competitividade devem promover a cultura de medir, testar e avaliar evidências quantitativas. Neste sentido, organizações que possuem uma cultura de decisão orientada por dados reforçam a relevância e o retorno esperado do BDA. Nelas são aprimoradas as previsões e decisões assertivas (DUAN; CAO, 2015; LAVALLE et al., 2011; MEDEIROS, 2018), habilitando o entendimento de problemas de negócios complexos (ARUNACHALAM; KUMAR, 2018). Davenport (2006) e Grover et al. (2018) também enfatizam a importância da cultura de decisões fundamentadas em fatos para majorar os benefícios retornados pelo BDA, afirmando que a falta da DDC é prejudicial para a geração de valor potencial. Para obter vantagem competitiva a partir do BDA, é necessário construir uma cultura que promova suas análises (CHEN et al., 2017). Depende de uma mentalidade analítica, para que gerentes se sintam confortáveis ao tomar decisões baseadas em dados, e não somente em suas

impressões pessoais. Eles devem colocar as evidências na frente de seus julgamentos, para não descartar dados importantes com base em intuições (LYYTINEN; GROVER, 2017). Portanto, em ambientes onde há DDC, o BDA se torna fundamental para a gestão e a tomada de decisão estratégica. Departamentos que tradicionalmente examinam dados para decidir, como finanças e operações, tendem a promover com maior intensidade a adoção de BDA (ABBASI; SARKER; CHIANG, 2016; KIRON; SHOCKLEY, 2011). Nesse sentido, afirma-se que o fator mais importante para ter êxito na implementação do BDA é que a empresa estabeleça a DDC (DREMEL, 2017; DUAN; CAO; EDWARDS, 2018; GROVER et al., 2018; KIRON; SHOCKLEY, 2011; LAVALLE et al., 2011). Uma organização centrada em dados necessita destas inovações culturais, e o alinhamento de uma estratégia BDA com a infraestrutura organizacional exige que as lideranças promovam a DDC.

Da mesma forma, foi a partir do BDA e seu grande volume de dados que a DDC voltou a ser destacada na indústria e na academia (DUAN; CAO; EDWARDS, 2018; KIRON; PRENTICE; FERGUSON, 2012; LAVALLE et al., 2011). Chen et al. (2012), defendem que, para implementar DDC, é preciso ter dados brutos e saber como transformá-los em conhecimento. Para Kiron, Prentice e Ferguson (2012), com a DDC é mais provável que as organizações usem mais dados. Elas tendem a incorporar BDA nos processos, sendo mais eficazes na coleta, análise e disseminação dos *insights* por toda organização (KIRON; PRENTICE; FERGUSON, 2012). Em razão destas observações, é proposta a primeira hipótese dessa investigação, referindo-se ao impacto da DDC no uso de *Data Lakes*, uma das práticas do BDA que antecede à identificação de SF. A literatura conceitua *Data Lakes* como um local onde todos os tipos de dados, estruturados ou não estruturados, são armazenados em seu formato original, para futura utilização (KITCHENS et al., 2018; WATSON, 2017). Este conceito será melhor detalhado no Capítulo 3.2.3.1.

**H1: A cultura *data-driven* tem efeito positivo no uso de *Data Lakes*.**

### 3.2.3 *Big Data Analytics*

O termo *Big Data* é mencionado quando são tratados dados em grandes volumes, de possível formato não estruturado e produzidos com grande velocidade (SOON; LEE; BOURSIER, 2016). Em termos organizacionais, o volume diz respeito a quantidades de informações que possibilitam a ir além de dados históricos internos (LAM et al., 2017); a velocidade trata do ritmo de geração de dados, impactando na agilidade necessária para armazená-los e processá-los (BHOLAT, 2015), e viabilizando a análise em tempo real; a variedade implica na possibilidade de usar diferentes tipos de dados, como os não estruturados: textos,

imagens, áudios e vídeos (ARUNACHALAM; KUMAR; KAWALEK, 2018; KITCHIN; MCARDLE, 2016). O crescimento exponencial do *Big Data* ocorre principalmente em função do desenvolvimento de tecnologias digitais e de processos, bem como sua rápida incorporação às rotinas analíticas organizacionais.

Associando *Big Data e Analytics*, Chen et al. (2012) definiram o termo *Big Data Analytics* (BDA), relacionando conceitos e ferramentas para mineração de dados e análise estatística. A literatura define BDA como tecnologias projetadas para extrair valor de grandes volumes e variedades de dados, permitindo a coleta, o processamento e a análise em alta velocidade (CÔRTE-REAL; OLIVEIRA; RUIVO, 2017), indicando métodos e técnicas analíticas realizadas por meio de aplicativos complexos, que permitem compilar de forma inteligente dados estruturados ou não (CHEN; CHIANG; STOREY, 2012; DAVENPORT; DYCHÉ, 2013; DUAN; CAO, 2015). Historicamente, organizações têm coletado dados estruturados de suas atividades internas visando calcular desempenho e identificar capacidades específicas (CONSTANTIOU; KALLINIKOS, 2015). Com o BDA, elas podem relacionar esses dados internos com dados externos compartilhados, estruturados ou não, proporcionando a identificação de oportunidades (MEDEIROS; MAÇADA; FREITAS JUNIOR, 2019) e melhor conhecimento do seu mercado (CHEN et al., 2017; KITCHIN; LAURIAULT, 2015; LAM et al., 2017).

Em contexto organizacional, o BDA é uma solução de apoio para a geração de *insights* a partir dos dados (LAM et al., 2017), gerando valor nos níveis estratégico e operacional (CHEN et al., 2017; DUAN; CAO, 2015). No nível operacional, pequenas decisões podem ser automatizadas a partir de técnicas de análise específicas. Empresas como a Amazon e Netflix, por exemplo, desenvolveram algoritmos correlacionando buscas de produtos e históricos de compras de clientes, para sugerir e oferecer novos produtos a partir de uma previsão da probabilidade de interesse (CONSTANTIOU; KALLINIKOS, 2015; EREVELLES; FUKAWA; SWAYNE, 2016; JEBLE; KUMARI; PATIL, 2018). No nível estratégico, os *insights* guiados pelos dados decorrem de um processo humano e analítico de criação de sentido, portanto, para este objetivo, o BDA serve como uma ferramenta importante (CONSTANTIOU; KALLINIKOS, 2015; SHARMA; MITHAS; KANKANHALLI, 2014).

Entre as implicações do Artigo I, há as evidências de que a procura de *insights* a partir dos BDA, usando práticas singulares como a coleta e armazenamento de dados brutos (*Data lakes*), a análise de dados não estruturados, o estudo de *outliers*, o uso da abordagem

indutiva e o uso de técnicas de análise disponíveis no BDA (*Advanced Analytics*), contribuem para a identificação de um número maior de Sinais Fracos. Parte-se destes resultados para desenvolver construtos que representam as práticas do BDA que potencialmente impactam direta ou indiretamente na identificação de sinais fracos.

Com o BDA é viável integrar diversas fontes de dados distintas, o que tem grande potencial em gerar valor estratégico. Contudo, há uma série de desafios neste processo, principalmente ao relacionar as informações heterogêneas e não estruturadas. A solução tecnológica atual, para dirimir estas dificuldades na coleta e integração dos dados, é a criação de *Data Lakes* (KITCHENS et al., 2018).

### 3.2.3.1 *Data Lake*

O mundo gerou um *zettabytes* de dados até 2010, com volume multiplicado por 40 em 2020, e previsão de atingir 180 *zettabytes* em 2025 (VASSAKIS; PETRAKIS; KOPANAKIS, 2018). Novas ferramentas, tornaram viável a implementação deste volume de forma segura, escalável e com baixo custo (WATSON, 2017). Neste cenário, muitas organizações capturam e conservam dados em formato nativo, nos chamados *Data Lakes* (CHEN et al., 2017; KITCHENS et al., 2018; WATSON, 2017). Esses dados são chamados de “dados brutos” ou “*raw data*”, já que precisam ser preparados antes das análises (JANSSEN; VAN DER VOORT; WAHYUDI, 2017). São também passivos ou agnósticos, pois não são produzidos/coletados em razão de uma demanda específica, ou para os mesmos propósitos em que são utilizados (CHEN et al., 2016; CONSTANTIOU; KALLINIKOS, 2015; GÜNTHER et al., 2017). Normalmente não se sabe *a priori* quais os requisitos das futuras análises que poderão empregá-los, portanto, define-se um escopo maior no processo de coleta (CHEN et al., 2016).

Guardados na sua forma original, os dados em *Data Lakes* não precisam ser planejados e modelados (WATSON, 2017). De baixo custo, eles ficam disponíveis para futuras análises, a qualquer momento, para toda organização (MITROVIC, 2017; WATSON, 2017). Assim, um *Data Lake* é um repositório que armazena dados brutos (MITROVIC, 2017), local onde todos os tipos de dados (provenientes de fontes internas e externas) de uma organização são depositados, aguardando para serem utilizados (KITCHENS et al., 2018; WATSON, 2017). Os *Data Lakes* estão se tornando cada vez mais difundidos e são essencialmente *Data Warehouses* destinados a servir como um laboratório, ou “*sand boxes*”, para experimentação, descobertas e buscas por *insights* (ABBASI; SARKER; CHIANG, 2016; KITCHENS et al., 2018).

Neste estudo, considera-se *Data Lake* um construto que trata da coleta e armazenamento de grandes volumes de dados brutos para análise posterior (WATSON, 2017). Este processo é desenvolvido em contrapartida à prática de manter apenas o dado já processado, coletado com o objetivo de atender demandas pontuais de informação. Como os dados ficam salvos e facilmente acessíveis (KITCHENS et al., 2018; MITROVIC, 2017; WATSON, 2017), entende-se que eles promovem os demais processos de manipulação e uso deste tipo de informações. Portanto, especificamente nos conceitos estudados neste trabalho, deduz-se um impacto positivo no uso de dados não estruturados, no uso de *Advanced Analytics*, nas análises de *outliers* e no uso do método indutivo. Esses quatro construtos são descritos nos Capítulos 3.2.3.2, 3.2.3.3, 3.2.3.4 e 3.2.3.5.

*Data Lakes* são mecanismos de monitoramento digital do ambiente, já que registram, em um histórico de dados, informações sobre o mercado (KITCHIN; LAURIAULT, 2015). Guardar os dados sem um propósito específico facilita as análises realizadas pela abordagem indutiva e abductiva (CONSTANTIOU; KALLINIKOS, 2015; GÜNTHER et al., 2017; OLBRICH, 2014), de exploração e análise. Essas análises são caminhos para gerar novas hipóteses e conhecimentos não deduzidos (EREVELLES; FUKAWA; SWAYNE, 2016; LYCETT, 2013), majorando a possibilidade da identificação de Sinais Fracos. Para Mendonça (2012), Sinais Fracos podem ser construções baseadas em dados brutos interpretados, o que se aproxima do entendimento de Rohrbeck, Thom & Arnold (2015) de que SF são hipóteses geradas por uma combinação de lógica dedutiva. Em decorrência destas observações, identifica-se as seguintes hipóteses:

**H2a: O uso de *Data Lakes* tem efeito positivo no uso de dados não estruturados.**

**H2b: O uso de *Data Lakes* tem efeito positivo no uso de *Advanced Analytics*.**

**H2c: O uso de *Data Lakes* tem efeito positivo no uso de análises de *outliers*.**

**H2d: O uso de *Data Lakes* tem efeito positivo no uso do método indutivo.**

### **3.2.3.2 Dados não-estruturados**

Tecnologias BDA processam uma variedade de tipos de dados não estruturados (ARUNACHALAM; KUMAR; KAWALEK, 2018), como dados textuais (como mensagens de texto, documentos, livros, artigos acadêmicos) e não textuais (vídeos, imagens e gravações de áudio) (MCABEE; LANDIS; BURKE, 2017). Eles oferecem uma riqueza diversificada de informações, que supera os dados tradicionais estruturados. A literatura classifica da seguinte forma os possíveis dados de um BDA (MCABEE; LANDIS; BURKE,

2017): (1) Estruturados: fornecidos e coletados em formatos fixos; os metadados que os descrevem estão disponíveis e não são dinâmicos; (2) Semiestruturados: possuem uma estrutura organizacional básica para identificar elementos, como os e-mails que incluem o endereço do destinatário, fornecendo mecanismos para classificação e integração (MCABEE; LANDIS; BURKE, 2017); (3) Não estruturados: não há metadados disponíveis descrevendo os dados (TIEFENBACHER; OLBRICH, 2015). A maioria dos dados coletados, criados e armazenados pelas organizações é classificada como não estruturada (BHIMANI, 2015; GROVER et al., 2018). Muitos são obtidos em redes sociais, onde indivíduos compartilham informações pessoais e comportamentais (EREVELLES; FUKAWA; SWAYNE, 2016). Há igualmente oportunidades de analisar um conjunto diversificado de dados históricos, gerados também de forma analógica, mas digitalizados nas últimas décadas (KITCHIN, 2014).

Dados não-estruturados em grandes volumes são analisados de forma qualitativa, por meio de algoritmos, como os baseados no julgamento semântico de palavras-chaves (LEE, 2018) ou *Natural Language Processing* (NLP) (DAVENPORT; DYCHÉ, 2013; YOON, 2012). Eles oferecem *insights* que não são obtidos pelos dados tradicionais, a exemplo da avaliação do comportamento social e econômico do consumidor (DUTTA; BOSE, 2015; EREVELLES; FUKAWA; SWAYNE, 2016; MILLER, 2010; VAJJHALA; STRANG; SUN, 2015), da previsão de tecnologias futuras (ABBAS; ZHANG; KHAN, 2014). Dados textuais são examinados sistematicamente via mineração de texto (*text mining*) (ABBAS; ZHANG; KHAN, 2014; KAYSER; BLIND, 2017; THORLEUCHTER; VAN DEN POEL, 2015), para gerar conhecimento usando fontes como notícias, conferências, patentes, *workshops*, trabalhos acadêmicos, mídias sociais ou qualquer meio de comunicação em massa. Kayser e Blind (2017) citam a mineração de texto como caminho para ampliar a base de informações em análises preditivas (*forecast*). Para Kim et al. (2013), executando essas práticas realiza-se um monitoramento do ambiente por meio de BDA não-estruturado. Eckhoff et al. (2014) estudaram abordagens para identificar Sinais Fracos e afirmam que dados não-estruturados podem ser analisados para este fim, ilustrando com notícias da *web* mineradas, análises de termos emergentes em resumos de conferências. Complementarmente, Abbas et al. (2014) analisaram a detecção de tendências tecnológicas por meio de dados textuais de patentes. Tomando como base estes estudos, apresenta-se a seguinte hipótese:

**H3: O uso de dados não estruturados tem efeito positivo na identificação de Sinais Fracos.**

### 3.2.3.3 *Advanced Analytics*

Em projetos BDA, a análise dos dados é significativamente diferente em comparação com as ferramentas tradicionais de análise, que tratam de conjuntos de dados escassos, estáticos, “limpos” e pouco relacionais, colhidos em amostras e de acordo com premissas e hipóteses existentes (KITCHIN, 2014; MILLER, 2010). O desafio de analisar *Big Data* é lidar com a abundância e variedade dos dados (DUTTA; BOSE, 2015) e com o fato de que a maior parte do que é analisado não tem uma pergunta específica pronta. Seus dados não foram gerados a partir de um escopo estreito definido. Eles, em geral, são secundários, subprodutos de outra pesquisa (KITCHIN, 2014; MILLER, 2010). Portanto, é um erro supor que é apenas necessário adquirir sistemas *Big Data*. Também é indispensável o desenvolvimento de ferramentas de análise específicas (WALLER; FAWCETT, 2013). *Advanced Analytics* é um termo geral que define este conjunto de ferramentas que aplica técnicas analíticas avançadas aos dados, respondendo perguntas de negócios (BOSE, 2009; INTEZARI; GRESSEL, 2017). O objetivo é obter informação, identificando correlações, padrões, tendências e *outliers*, descobrindo os motivos subjacentes que os explicam, por meio do uso extensivo dos dados, da análise estatística e quantitativa (DUAN; CAO; EDWARDS, 2018). *Advanced Analytics* se tornou um ativo decisivo e um elemento importante nos esforços das empresas para melhorar o desempenho (BARTON; COURT, 2012; WALLER; FAWCETT, 2013), aumentando a eficiência e produtividade (DUTTA; BOSE, 2015). Os *insights* obtidos são usados para direcionar, melhorar e automatizar a tomada de decisão (BOSE, 2009) baseada em evidências (DREMEL, 2017), bem como para inovar produtos e serviços (DAVENPORT; BARTH; BEAN, 2012; GÜNTHER et al., 2017).

Duan et al. (2018), Sivarajah, Kamal, Irani e Weerakkody (2017), e Yablonsky (2019), classificam *Advanced Analytics* pelos possíveis modelos de análise:

- A **análise descritiva**, que emprega dados históricos e atuais, é uma importante fonte de *insights*. Ela mostra visões do passado e suas correlações com o que acontece no presente (VASSAKIS; PETRAKIS; KOPANAKIS, 2018), e são práticas do *Advanced Analytics* que oferecem formas de explorar dados sumarizados e detalhados, descrevendo-os por meio de métodos estatísticos (SIVARAJAH et al., 2017). Emprega técnicas ou ferramentas como *Business Intelligence*, mineração de dados (MÜHLROTH; GROTTKE, 2018), análise de *outliers* (MILLER, 2010; MÜHLROTH; GROTTKE, 2018), análise de sentimentos (DUAN; CAO; EDWARDS, 2018), *text mining* (KAYSER; BLIND, 2017;

THORLEUCHTER; VAN DEN POEL, 2015), *web analytics* (DUAN; CAO; EDWARDS, 2018), *sentiment and affect analysis e relation mining* (CHEN; CHIANG; STOREY, 2012), *social monitoring* (CATALDI; CARO; SCHIFANELLA, 2013).

- A **análise preditiva** projeta cenários futuros, no sentido de pensar o que pode acontecer, determinando possibilidades futuras; usando dados, ela é capaz de descobrir padrões e identificar relacionamentos. Aplica modelos estatísticos e aprendizado de máquina (YABLONSKY, 2019), supervisionados ou não supervisionados (SIVARAJAH et al., 2017), como regressões (MÜHLROTH; GROTTKE, 2018) e análise discriminante (BOSE, 2009; DUTTA; BOSE, 2015), árvores de decisão e *clustering* (DUTTA; BOSE, 2015; MÜHLROTH; GROTTKE, 2018), *classification, network analysis e association analysis* (LEE et al., 2009; MILLER, 2010; MÜHLROTH; GROTTKE, 2018), redes neurais (BOSE, 2009; DUTTA; BOSE, 2015), análise de tendências (MILLER, 2010).

- A **análise prescritiva** apresenta planos de ação, apontando o provável resultado de cada caminho, com relacionamentos de causa e efeito, auxiliando os gerentes na sugestão de ações para otimização de resultados (BARTON; COURT, 2012). Exemplos de técnicas da análise prescritiva: otimização, simulação (DUAN; CAO; EDWARDS, 2018), inteligência artificial (YABLONSKY, 2019), *recommendation* (CATALDI; CARO; SCHIFANELLA, 2013).

Inúmeras fontes de dados estruturados ou não estruturados são coletadas, processadas e lidas por ferramentas *Advanced Analytics*, como publicações científicas (KAYSER; BLIND, 2017; LEE, 2018; MÜHLROTH; GROTTKE, 2018), patentes (KAYSER; BLIND, 2017; LEE et al., 2009; VEUGELERS; BURY; VIAENE, 2010; YOON; PARK, 2007), notícias (KAYSER; BLIND, 2017; MÜHLROTH; GROTTKE, 2018; VEUGELERS; BURY; VIAENE, 2010), dados de mídias e redes sociais (KAYSER; BLIND, 2017; MÜHLROTH; GROTTKE, 2018; WELZ et al., 2012), *Websites* (MÜHLROTH; GROTTKE, 2018), relatórios científicos e de estudos futuros (KAYSER; BLIND, 2017).

Considerando este quantitativo de dados, o ser humano tem limites na capacidade de interpretação. Neste sentido, Simon (1955) apresentou a Teoria da Racionalidade Limitada, ponderando que decisões são tomadas com conhecimentos restritos, sem considerar todas as variáveis e alternativas existentes. A lista de escolhas dos decisores não é totalmente conhecida, sendo inviável comparar todos os benefícios (BAWDEN; ROBINSON, 2009). Ao buscar informações para criação de sentido, eles ficam satisfeitos quando julgarem suficiente, mesmo que não sejam as melhores disponíveis. Assim, o BDA tem uma função importante na percepção da correlação entre as informações. Devido à complexidade do

ambiente (JIANG; GALLUPE, 2015), à grandeza dos dados coletados (VILLARS; OLOFSON; EASTWOOD, 2011) e ao ruído de informações existente (VEUGELERS; BURY; VIAENE, 2010), as organizações podem obter valor e vantagem competitiva ao buscar Sinais Fracos também a partir dessas análises em BDA (MÜHLROTH; GROTTKE, 2018). Embora a busca de SF envolva desafios na coleta e análise de dados massivos, o aprendizado de máquina e a mineração de dados podem atenuar significativamente essa barreira (KIM; LEE, 2017; VEUGELERS; BURY; VIAENE, 2010). O uso destas diversas técnicas de análise oferecidas pelo BDA promove o monitoramento do ambiente e a possibilidade de identificar um número maior de Sinais Fracos (KIM; LEE, 2017). Portanto apresenta-se a seguinte hipótese:

**H4: O uso de *Advanced Analytics* tem efeito positivo na identificação de Sinais Fracos.**

#### **3.2.3.4 Análise de *Outliers***

Sinais Fracos são reconhecidos por representar fenômenos estranhos ou irregulares, significativamente distantes dos padrões ou médias encontradas no ambiente empresarial (KUOSA, 2010; MENDONÇA; CARDOSO; CARAÇA, 2012). Para identificá-los, um dos caminhos é explorar e compreender *outliers* coletados estatisticamente em grandes volumes de dados. *Outliers* podem ser evidências de Sinais Fracos, que devem ser confirmadas em análises posteriores (JANISSEK-MUNIZ, 2004; KIM; LEE, 2017; MENDONÇA; CARDOSO; CARAÇA, 2012). *Outliers* são conceituados como anomalias nos resultados estatísticos, que não se encaixam no entendimento dos resultados apresentados. São elementos, encontrados nos dados, que exibem desvios inesperados, quando comparados com o padrão definido. Eles podem ser a manifestação de erros, ou ruídos, que devem ser corrigidos, descartados ou ignorados ao usar métodos estatísticos não afetados por valores discrepantes (LI; ZHANG, 2018; YOON; KIM, 2012). Contudo, também podem representar eventos incomuns, fenômenos que estão iniciando e merecem a atenção adicional (MILLER, 2010). Utiliza-se a busca de *outliers* para identificar dados novos ou desconhecidos (KIM; LEE, 2017), para gerenciar riscos associados a eventos improváveis e que podem gerar grande impacto (ABBASI; SARKER; CHIANG, 2016; BREUKER et al., 2016). Para George, Haas e Pentland (2014), em muitas situações as informações médias são relevantes, porém, analisar *outliers* viabiliza descobrir inovações, fenômenos disruptivos ou mudanças que podem estar acontecendo fora das tendências. Portanto, em processos de criação de

sentido de Sinais Fracos, devem ser elaborados cenários alternativos gerados por *outliers*. Oportunidades e riscos encontrados nestas análises, como por exemplo uma nova orientação de mercado, podem ser transformados em uma agenda de decisão e ação (MENDONÇA; CARDOSO; CARAÇA, 2012).

No contexto do BDA, tecnologias *Outlier Mining Algorithm* são usadas para encontrar anomalias e iniciar a análise dos dados caracterizados como *outliers* (LI; ZHANG, 2018). Como exemplos de buscas de *outliers*, cita-se abordagens para detectar Sinais Fracos em fontes digitais (KIM; LEE, 2017; YOON; KIM, 2012). Yoon e Kim (2012) propõem um método utilizando *outliers* para identificar patentes distintas ou não usuais, que podem ser sinais de disrupções ou oportunidades tecnológicas. Após a coleta e a análise da estrutura sintática do texto usando *Natural Language Processing (NLP)*, são buscados e examinados os *outliers* que representam patentes relevantes, por meio de tratamentos estatísticos. Kim & Lee (2017), no mesmo caminho, evitando vieses e informações não relacionadas com antecipação de fenômenos, também indicam o uso de espaços já segmentados como fontes de informações. Para que o processo se desenvolva de forma guiada por dados (*data-driven*), o método escolhido por Kim & Lee foi aplicar o *Local Outlier Factor (LOF)*, técnica usada para identificar informações desconhecidas, as quais um sistema de aprendizado de máquina entende e detecta como anormal. Como os dados originais podem ser compostos de conteúdos diversos e fragmentados, mesmo tratando apenas de temas futuros, o *LOF* é indicado, já que permite tratar essa heterogeneidade, percebendo anomalias locais, identificando candidatos a SF detectados como *outliers*, exibindo padrões incoerentes ou distribuições não uniformes.

Segundo Hiltunen (2008) e Y. Lee (2018), os dados que têm a possibilidade de serem Sinais Fracos possuem padrões anormais e estão relacionados a tópicos pouco expostos. Yoon (2012) apresentou um método de identificação de SF baseado em mineração de texto e palavras-chaves, medindo quantitativamente a presença desses padrões atípicos. O autor enfatizou como sinais aqueles que são relativamente menos visíveis em volume, portanto fracos, mas que estão sendo difundidos de forma rápida. Ao explorar a frequência de ocorrência das palavras-chaves em documentos digitais, foram percebidos padrões inesperados e apontados como possíveis SF. Classificou-se como Sinais Fracos palavras de baixa frequência e alta taxa de crescimento.

*Outliers* podem formar uma base empírica para identificar novas oportunidades de negócios em inovação de produtos, pesquisa e desenvolvimento e *marketing*

(CONSTANTIOU; KALLINIKOS, 2015; GEORGE; HAAS; PENTLAND, 2014). A avaliação dos *outliers* e suas raridades potencialmente produz novas ideias e soluções alternativas para as possíveis ameaças ou oportunidades (KIM; LEE, 2017). Considerando-se estas observações, apresenta-se a seguinte hipótese:

**H5: O uso de análises de *outliers* tem efeito positivo na identificação de Sinais Fracos.**

### 3.2.3.5 Método de pesquisa Indutivo

Sinais Fracos são incomuns, não familiares. Eles não são esperados, frequentemente são encontrados por acaso (BORTOLI; JANISSEK-MUNIZ; BORGES, 2015). São dados ambientais inadvertidos, não há aviso de que serão encontrados. Há grandes e sistêmicas barreiras para identificá-los. O principal motivo não é a falta de informação, mas os vieses e a rigidez da mentalidade de quem a está analisando. Ao longo do abundante volume de estímulos atualmente existente, é preciso pensar de forma diferente para identificar essas informações incomuns porém importantes (MENDONÇA; CARDOSO; CARAÇA, 2012). Como SF são inesperados, eles podem ser rejeitados imediatamente pelo observador. São originados fora dos sistemas correntes ou das margens do conhecimento convencional, e não possuem uma história, uma sequência temporal, que poderia ser uma base para dedução de hipóteses ou criar modelos (KUOSA, 2010). Na busca pela sua identificação, não é possível contar com perguntas definidas ou hipóteses deduzidas *a priori*. Modelos dedutivos falham ao prever eventos disruptivos, portanto, para obter maior retorno, é necessário complementar o método empregando a abordagem indutiva (BHOLAT, 2015; MCABEE; LANDIS; BURKE, 2017). No método indutivo, utiliza-se informações coletadas para somente então formular novas hipóteses (JEBB; PARRIGON; WOO, 2017; KITCHIN, 2014; STEADMAN, 2013), e não o contrário. Procura-se entender os dados em uma explicação teórica (BHOLAT, 2015). Neste sentido, ao contrário de responder um questionamento, o processo indutivo, sem um direcionamento específico, fornecerá caminhos que um analista humano pode não ter pensado anteriormente (EREVELLES; FUKAWA; SWAYNE, 2016; KITCHIN, 2014). Este método, aplicado em conjunto com as tecnologias de BDA, oferece resultados para diversos objetivos analíticos: a descoberta de relacionamentos empíricos (JEBB; PARRIGON; WOO, 2017), a identificação de correlações, de padrões inusitados (STEADMAN, 2013), de tendências futuras e *outliers* (JEBB; PARRIGON; WOO, 2017).

É possível encontrar informações imprevisíveis em grandes volumes de dados (LYCETT, 2013). O método indutivo reflete uma maneira particular de perceber e entender o ambiente (KITCHIN; LAURIAULT, 2015), que requer menos confiança no conhecimento existente, dando maior foco ao que é desconhecido. Para desenvolver essa abordagem, é necessário mudar para uma visão baseada na ignorância (SAMMUT; SARTAWI, 2012), olhando para o futuro. Janissek-Muniz, Lesca e Freitas (2006) apresentam duas metáforas de formas de estudar o ambiente: a do espelho retrovisor e a do radar. Entende-se que o método indutivo está alinhado com a metáfora do radar, também de acordo com a perspectiva da Inteligência Antecipativa. Assim, além de olhar os dados do passado, respondendo perguntas em um “espelho retrovisor”, também utiliza-se o radar, artefato para captar e obter novas informações não esperadas, em atenção e escuta do ambiente (JANISSEK-MUNIZ; LESCA; FREITAS, 2006). Este é o chamado *sistema de informações inverso* (LESCA; JANISSEK-MUNIZ, 2015), no qual o conhecimento não é obtido por meio de uma investigação com perguntas definidas. Nele, não há requisitos prévios, perguntas ou necessidades específicas. O radar é compatível com esse sistema, escutando o ambiente, percebendo sinais e iniciando um novo ciclo de coleta e interpretação (CAINELLI; JANISSEK-MUNIZ, 2017).

Moreira et al. (2015) citam um exemplo da abordagem *indutiva* em uma análise que emprega o aprendizado de máquina não supervisionado. O fato de ser não supervisionado é importante, pois o algoritmo utilizado, *clustering*, oferece o resultado somente processando os dados, independente de orientação humana. A ideia da solução foi agrupar informações semelhantes obtida em diversas fontes, já que os SF se repetem, mesmo que em quantidade pequena, em diferentes partes dos dados. Assim, apresentar os grupos (*clusters*) pode levar a formulação de novas hipóteses de Sinais Fracos (MOREIRA et al., 2015). Diante do que foi exposto neste capítulo, apresenta-se a seguinte hipótese:

**H6: O uso do método indutivo tem efeito positivo na identificação de Sinais Fracos.**

### 3.3 Modelo de Pesquisa

Tendo em vista o objetivo de medir o efeito das práticas do BDA na identificação de Sinais Fracos, elaborou-se o modelo conceitual com as hipóteses que norteiam este trabalho. Para tal, inicialmente realizou-se uma homologação dos construtos oriundos da fundamentação

teórico-conceitual e dos resultados da pesquisa realizada no Artigo I, expostos respectivamente nos Capítulos 3.2 e 2.5. Eles foram submetidos para a revisão de especialistas em Inteligência Estratégica Antecipativa, por meio de um questionário eletrônico *on-line*, com uso da plataforma *Google Forms*. A Tabela 9 contém as perguntas que foram realizadas nesta homologação, onde cada linha representa um construto a ser examinado pelos especialistas.

Tabela 9: Roteiro estruturado de perguntas para consulta a especialistas

Construto, referências e perguntas	Descrição
<p><b>C01 – Data Lake</b></p> <p>Trata da coleta e armazenamento de grandes volumes de dados brutos (CHEN et al., 2016; LAM et al., 2017; WATSON, 2017)</p> <p>Você concorda que a prática de coletar e armazenar grandes volumes de dados brutos pode impactar no monitoramento digital do ambiente e na percepção de sinais antecipativos?</p>	<p>Este constructo trata da coleta e armazenamento de grandes volumes de dados brutos (<i>raw data</i>), para a modelagem e análise posterior. Este processo pode ser desenvolvido em contrapartida à prática de manter apenas o dado já processado, coletado com o objetivo de atender demandas pontuais de informação.</p> <p>Devido ao custo reduzido de armazenamento e a capacidade de guardar muitos dados em formato nativo, organizações estão mantendo os dados brutos nos chamados <i>Data Lakes</i>. Esses dados passivos não são coletados em razão de uma solicitação específica ou para um propósito definido (CHEN et al., 2016). Desta forma eles ficam disponíveis para possíveis análises futuras (WATSON, 2017).</p> <p>Este mecanismo promove o monitoramento digital do ambiente por meio de grandes volumes de dados brutos, que podem ser analisados conforme escopos de pesquisas futuras. Ele também gera o registro completo das informações monitoradas no ambiente. Ambas práticas são características de processos de inteligência antecipativa.</p>
<p><b>C02 – Dados não estruturados</b></p> <p>(ARUNACHALAM; KUMAR; KAWALEK, 2018; KITCHIN; MCARDLE, 2016; TIEFENBACHER; OLBRICH, 2015)</p> <p>Você concorda que a análise de dados não estruturados pode impactar no monitoramento digital do ambiente e na percepção de sinais antecipativos?</p>	<p>Este constructo trata da análise de dados não estruturados, como informações textuais (mídias sociais, e-mails e mensagens de texto) e dados não textuais (vídeos, imagens e gravações de áudio).</p>
<p><b>C03 – Advanced Analytics</b></p> <p>Trata do uso de técnicas e algoritmos oferecidos pelo BDA</p> <p>(DUTTA; BOSE, 2015; MÜHLROTH; GROTTKE, 2018)</p> <p>Você concorda que o uso de técnicas e algoritmos oferecidos pelo BDA pode impactar no monitoramento digital do ambiente e na percepção de sinais antecipativos?</p>	<p>A análise de dados é significativamente diferente na adoção de projetos BDA, em comparação com projetos de análise tradicionais. Ela é tecnicamente desafiadora, pois envolve a modelagem de um grande volume de dados estruturados e não estruturados (DUTTA; BOSE, 2015).</p> <p>Técnicas quantitativas são usadas nas informações coletadas e pré-processadas para extrair <i>insights</i>. O objetivo é identificar padrões e, principalmente, descobrir motivos subjacentes que os explicam. Para análise de um grande volume de dados estruturados, abordagens estatísticas não tradicionais como regressão, análise discriminante e também de <i>machine learning</i> como árvores de decisão, redes neurais, <i>classification</i>, <i>cluster analysis</i> e <i>association analysis</i>, podem ser usadas. Para a análise de dados não estruturados, a mineração de texto (<i>text mining</i>) pode ser adotada para a determinar variáveis importantes (<i>sentiment analysis</i>) e suas relações.</p> <p>O uso de diversas técnicas de análise disponíveis no BDA é necessário para facilitar a identificação de um número maior de Sinais Fracos e reduzir o viés humano nos processos de inteligência antecipativa e tomada de decisão estratégica. Sistemas para detectar possíveis sinais antecipativos em BDA precisam ser capazes de aprender e acumular conhecimento ao longo do tempo (MÜHLROTH; GROTTKE, 2018).</p>
<p><b>C04 - Análise de Outliers</b></p> <p>(CONSTANTIOU; KALLINIKOS, 2015; KIM; LEE, 2017; MENDONÇA; CARDOSO; CARAÇA, 2012)</p> <p>Você concorda que a Análise de <i>Outliers</i> em grandes volumes de dados pode impactar no monitoramento digital do</p>	<p>Os <i>outliers</i> podem ser conceituados como anomalias nos resultados estatísticos que não se encaixam facilmente no entendimento dos resultados apresentados. Neste sentido, cenários alternativos gerados por <i>outliers</i>, baseados em processos que identificam de Sinais Fracos, também podem ser elaborados.</p> <p>Um <i>outlier</i> pode ser uma evidência de sinal fraco, a ser confirmado em análises posteriores (KIM; LEE, 2017; MENDONÇA et al., 2004).</p> <p>Os <i>outliers</i> ajudam a identificar os comportamentos que se desviam das tendências observadas. Em função do volume de dados utilizados nestas pesquisas, estes <i>outliers</i></p>

ambiente e na percepção de sinais antecipativos?	são representados por números importantes, permitindo indicar fontes de mudanças. Os <i>outliers</i> , então, podem formar uma base empírica para identificar novas oportunidades de negócios em áreas como inovação de produtos, pesquisa e desenvolvimento e marketing (CONSTANTIOU; KALLINIKOS, 2015).
<p><b>C05 - Método de Indutivo</b></p> <p>(BHOLAT, 2015; LYCETT, 2013; MCABEE; LANDIS; BURKE, 2017)</p> <p>Você concorda que utilizar a abordagem indutiva nas análises em <i>Big Data</i> pode impactar no monitoramento digital do ambiente e na percepção de sinais antecipativos?</p>	<p>Em contraste com o raciocínio dedutivo, há o método indutivo (<i>bottom-up</i>), no qual as teorias são formuladas por meio de inferências realizadas a partir dos dados, que generalizam, induzindo a partir de casos particulares. Apesar da popularidade das estratégias dedutivas, o raciocínio indutivo baseado em dados fornece um método viável e importante para a pesquisa (MCABEE; LANDIS; BURKE, 2017).</p> <p>Para obter melhores resultados em um BDA, é necessário complementar o método de pesquisa, analisando os dados também por meio da abordagem indutiva. A abordagem indutiva <i>bottom-up</i>, em conjunto com as tecnologias BDA, permite que identificar informações sem desenvolver hipóteses <i>a priori</i> (LYCETT, 2013).</p> <p>A razão pela qual essa abordagem é defendida por acadêmicos é que os modelos dedutivos dominantes falham ao prever eventos disruptivos (BHOLAT, 2015). Quando tratamos da inteligência antecipativa, ela é ainda mais importante, já que os métodos dedutivos podem projetar fenômenos lineares apenas enquanto as tendências persistem no ambiente de negócios. Em um mundo turbulento, onde eventos podem mudar o mercado a qualquer momento, necessitamos da abordagem indutiva para facilitar identificação e percepção dos Sinais Fracos.</p>
<p><b>C06-A</b></p> <p><i>Data-driven Culture</i></p> <p>(ARUNACHALAM; KUMAR; KAWALEK, 2018; DUAN; CAO, 2015; LAVALLE et al., 2011; SHARMA; MITHAS; KANKANHALLI, 2014)</p> <p>Você concorda que desenvolver uma cultura <i>data-driven</i> pode <b>impactar no</b> monitoramento digital do ambiente e na percepção de sinais antecipativos?</p> <p><b>C06-B</b></p> <p><i>Data-driven Culture</i></p> <p>(DUAN; CAO; EDWARDS, 2018)</p> <p>Você concorda que desenvolver uma cultura <i>data-driven</i> pode <b>moderar o</b> impacto dos demais constructos no monitoramento digital do ambiente e na percepção de sinais antecipativos?</p>	<p>Para obter vantagem competitiva pelo BDA, é necessário desenvolver uma cultura <i>data-driven</i>, na qual as decisões gerenciais dependem de <i>insights</i> baseados em dados (DUAN; CAO, 2015; LAVALLE et al., 2011).</p> <p>A presença da cultura dos dados nas organizações, a possibilidade de se tornar <i>data-driven</i>, permite a tomada de decisão baseada em dados. As organizações que possuem recursos analíticos avançados não extrairão todos seus benefícios se eles não forem aceitos como uma ferramenta na tomada de decisão, se não forem integrados ao processo de planejamento estratégico (ARUNACHALAM; KUMAR; KAWALEK, 2018).</p> <p>A cultura <i>data-driven</i> impacta diretamente na inovação em novos produtos e indiretamente na significância dos produtos, por meio do <i>Environmental Scanning</i> (DUAN; CAO, 2015). Em organizações orientadas por dados, <i>insights</i> obtidos em um BDA podem ser considerados como candidatos a Sinais Fracos.</p>

Fonte: elaborado pelo autor

Após a avaliação de 10 especialistas, acadêmicos e profissionais, todos os construtos foram confirmados por maioria. Optou-se por realizar mudanças nas suas nomenclaturas, de acordo com termos mais comuns, utilizados pela academia, conforme ilustrado na Tabela 11. O modelo de pesquisa inicialmente deduzido também foi revisado. Para simplificá-lo resolveu-se por remover o construto Monitoramento do Ambiente, avaliando diretamente o impacto no construto Identificação de Sinais Fracos. Optou-se, também, por uma estrutura mais detalhada para os caminhos dos efeitos diretos nas variáveis dependentes, alinhando-os com conceitos existentes na literatura. A hierarquia dos construtos *Data-driven Culture* e *Data Lake* foi alterada ao colocá-los como antecedentes, antecipando sua hipotética relação de causa e efeito no modelo. Cogitou-se posicionar o construto *Data-driven Culture* como moderador (tópico C06-B no roteiro de perguntas aos especialistas), contudo, esta opção foi descartada e decidiu-se por considerá-lo como variável independente de efeito indireto na

identificação de SF (C06-A). Na Tabela 10 são resumidas as respostas dos especialistas, com destaque nas discordâncias, e na Figura 9 observa-se o Modelo Conceitual reformado, em relação ao primeiro formato proposto, apresentado na Figura 8 do Artigo I. Os construtos homologados deram origem as variáveis dependentes e independentes, utilizadas no Modelo Conceitual e no questionário de perguntas da *survey*.

Tabela 10: Síntese das respostas dos especialistas para cada construto avaliado

Especialista	C01	C02	C03	C04	C05	C06-A	C06-B
Especialista 1	Concordo	Concordo	Concordo	Concordo	Concordo	<b>Discordo</b>	Concordo
Especialista 2	<b>Discordo</b>	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo
Especialista 3	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo	<b>Discordo</b>
Especialista 4	<b>Discordo</b>	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo
Especialista 5	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo
Especialista 6	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo
Especialista 7	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo
Especialista 8	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo
Especialista 9	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo
Especialista 10	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo	Concordo

Fonte: Dados da pesquisa (2019)

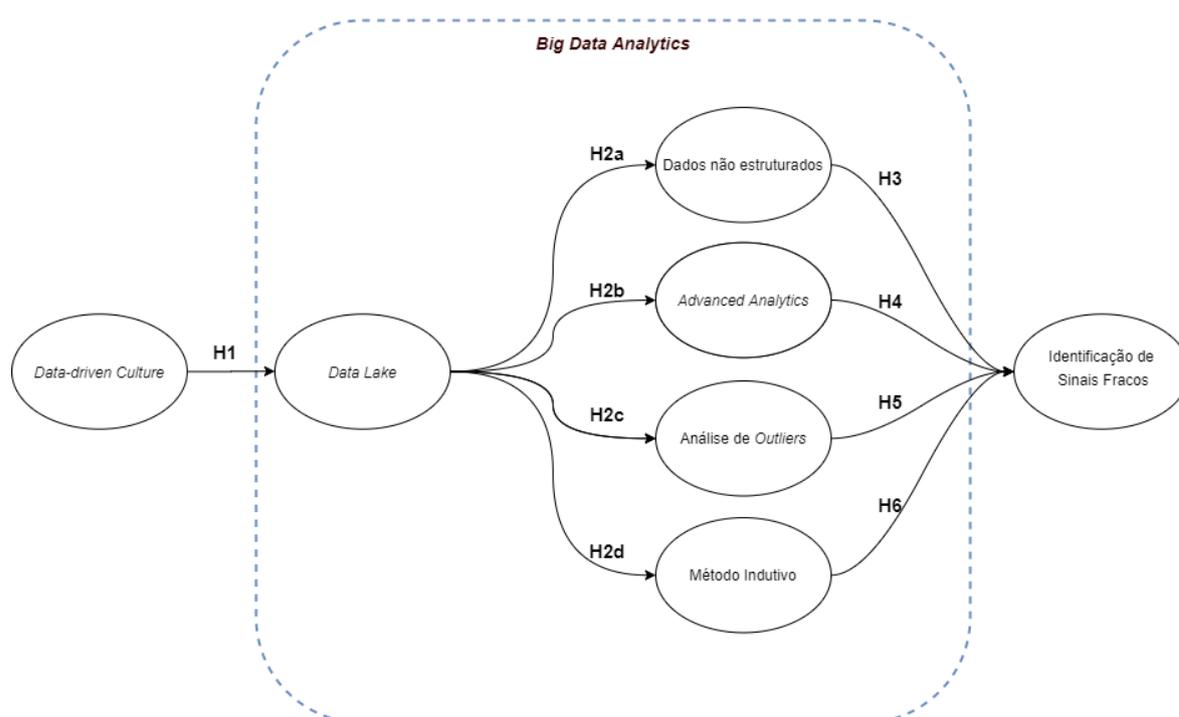
Tabela 11: Mudanças na nomenclatura dos construtos após consulta a especialistas

Nomes dos construtos iniciais propostos	Nomes construtos ajustados
Coletas de grandes volumes de dados brutos	<i>Data Lake</i>
Análise de dados não estruturados	Dados não estruturados
Uso de análises oferecidas pelo BDA	<i>Advanced Analytics</i>
Análises de <i>Outliers</i>	Análise de <i>Outliers</i>
Método Indutivo	Método Indutivo
<i>Data-driven Culture</i>	<i>Data-driven Culture</i>
Monitoramento do Ambiente	Construto removido
Identificação de Sinais Fracos	Identificação de Sinais Fracos

Fonte: elaborado pelo autor

Como consequência da revisão de literatura, também foram apresentadas, na seção 3.2, as hipóteses de pesquisa. Diante dessas hipóteses, e da homologação dos construtos pelos especialistas em Inteligência Estratégica Antecipativa, ilustra-se na Figura 9 o novo Modelo Conceitual de Pesquisa desenvolvido e atualizado a partir da sua versão inicial (Figura 8 do Artigo I).

Figura 9: Modelo Conceitual de Pesquisa



Fonte: elaborado pelo autor

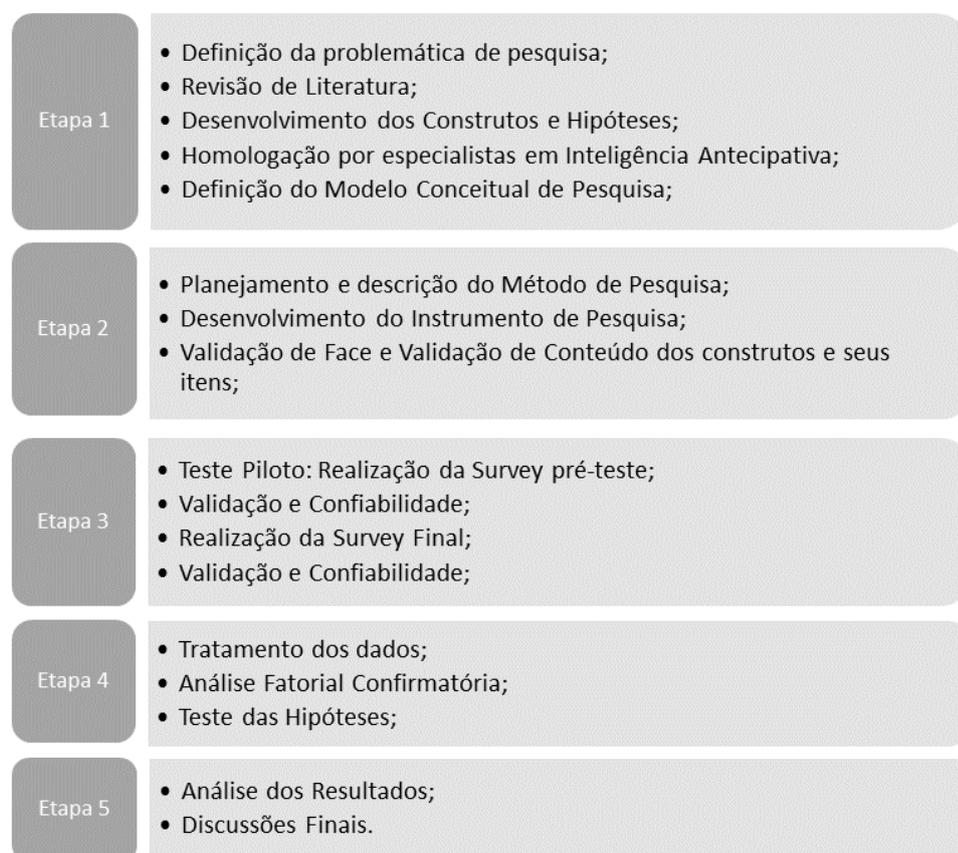
### 3.4 Método

Para atender os objetivos gerais e específicos desta pesquisa, considerou-se necessária a aplicação de um método **quantitativo**. Realizou-se uma coleta amostral de dados, com o objetivo de testar as hipóteses deduzidas, por meio de artefatos de pesquisa e ferramentas estatísticas (SAMPIERI; COLLADO; LUCIO, 2006). Buscando explicar as relações entre múltiplas variáveis, a análise foi realizada seguindo o modelo estatístico de equações estruturais (**SEM**), com estimativa *Partial Least Squares* (**PLS**) (HAIR et al., 2009).

No Capítulo 3.3 foi apresentado o Modelo Conceitual de Pesquisa e sua homologação inicial por dez especialistas em Inteligência Antecipativa. Neste capítulo descreve-se o método de pesquisa e o seu desenvolvimento, bem como a construção e validação do instrumento de pesquisa. A Figura 10 ilustra o desenho da pesquisa. Destaca-se que a homologação dos construtos realizada por especialistas **não** substituiu as validações necessárias ao instrumento de pesquisa e aos modelos analisados, conforme indicado por Koufteros (1999). No desenho da pesquisa constam as etapas estabelecidas pelo autor: validação de face e conteúdo, teste piloto (pré-teste), testes estatísticos para validação e confiabilidade

(fidedignidade) dos resultados obtidos; importantes para avaliar a consistência do instrumento, a confiabilidade das variáveis e a unidimensionalidade dos fatores.

Figura 10: Desenho da pesquisa



Fonte: elaborado pelo autor

### 3.4.1 Pesquisa Survey

Dentro do paradigma positivista quantitativo, elegeu-se por realizar uma pesquisa *survey*, com base em um questionário estruturado. A pesquisa *survey* é um dos métodos de pesquisa quantitativa mais utilizados (HAIR et al., 2009). Ele trata de coletas e análises de informações padronizadas, a respeito dos assuntos que estão sendo pesquisados (PINSONNEAULT; KRAEMER, 1993). A *survey* é caracterizada pelo fato de questionar diretamente a população da amostra, que representa uma população alvo, por meio de um instrumento de pesquisa que, geralmente, é um questionário estruturado de perguntas (HAIR et al., 2009). Em uma *survey* são obtidas informações desta amostra - características, opiniões e ações - que são analisadas e possibilitam inferências (FREITAS et al., 2000).

Há três particularidades importantes que diferenciam a *survey* dos demais métodos: (1) a coleta de informações quantitativas descritivas de algum objeto de estudo da população; (2) as perguntas são realizadas diretamente às pessoas, sem intermediários; (3) estas informações são coletadas em uma amostra, que é apenas parte da população alvo (PINSONNEAULT; KRAEMER, 1993). O presente instrumento tem como objetivo testar um modelo, suportando hipóteses nas suas relações causais, questionando os motivos da sua existência. Classifica-se a pesquisa *survey* deste trabalho como explanatória, em razão do seu propósito principal (FREITAS et al., 2000; PINSONNEAULT; KRAEMER, 1993).

Neste método de pesquisa, quem possui as informações necessárias é a população alvo (HAIR et al., 2009), portanto, sua definição é imprescindível para o bom andamento da pesquisa. A população alvo é a soma dos indivíduos que compartilham o perfil indicado pela pesquisa. A amostra, que é o conjunto de respondentes, é definida como uma parte ou subconjunto dessa população alvo (MALHOTRA, 2012). Para que os resultados da pesquisa não sejam contestados, o pesquisador deve obter uma amostra representativa da população (BRYMAN, 1989). A seleção do conjunto de indivíduos desta pesquisa foi realizada por conveniência, de acordo com os recursos empregados pelo pesquisador (MALHOTRA, 2012). A amostra, portanto, é classificada como não probabilística, já que os respondentes serão escolhidos pela sua disponibilidade (FREITAS et al., 2000). Os participantes foram contatados informalmente, sendo questionados sobre a disponibilidade de responder a pesquisa, já passando instruções de como acessar e responder o questionário eletrônico da *survey*. O questionário foi enviado por diversos meios eletrônicos, e-mail, redes sociais e mensagens eletrônicas, direcionados a profissionais que trabalham com inteligência – gestores, executivos, analistas, desenvolvedores, consultores – cujas organizações possuam grandes volumes de dados.

O *software G\*Power*, na versão 3.1.9.2, foi utilizado para estimar o tamanho mínimo necessário da amostra, de acordo com o Modelo Conceitual de Pesquisa. Para calcular, parametriza-se o aplicativo com a quantidade de preditores da variável dependente do modelo, com as informações de Poder do Teste e Tamanho do Efeito ( $f^2$ ). Recomenda-se indicar o Poder do Teste em 0,80 e o Tamanho do Efeito ( $f^2$ ) em 0,15 (COHEN, 1988; HAIR et al., 2016). O número de preditores da variável dependente, Identificação dos Sinais Fracos, apresentados no Modelo Conceitual de Pesquisa é 4, resultando em uma amostra mínima de **85 participantes**. No “Apêndice III – Estimativa do tamanho mínimo da amostra” apresenta-se a tela do *software* parametrizada e o resultado do cálculo do tamanho mínimo da amostra.

Para Hair et al. (2016), não há um número mínimo de respostas para o modelo de análise de equações estruturais com mínimos quadrados parciais (PLS), utilizado no tratamento estatístico dos dados. Porém, o autor indica que quanto maior o volume, melhor será a assertividade estatística das estimativas, apresentando resultados mais consistentes.

### 3.4.2 Desenvolvimento do instrumento

O Instrumento de coleta de dados deve ser decidido considerando a sua viabilidade, seus custos e a qualidade final das informações (PINSONNEAULT; KRAEMER, 1993). O artefato escolhido para esta *survey* é um questionário de pesquisa estruturado, respondido por um conjunto de participantes selecionados (amostra). Segundo Malhotra (2012) essa técnica estruturada para coletar dados consiste em perguntas, reunidas em um questionário fixo, a ser respondidas pelos integrantes da amostra. O questionário foi desenvolvido a partir de variáveis e seus conceitos, de acordo com o Modelo Conceitual de Pesquisa, a fim de cumprir os objetivos deste estudo. Os conceitos são assim operacionalizados, dando sentido às variáveis neste conjunto de perguntas, o que permite a medição e a avaliação dos resultados (TRIVIÑOS, 1987). O instrumento de coleta deve ser preparado tomando como base referências teóricas (BABBIE, 1999) e, se possível, utilizando itens ou escalas previamente validados, permitindo a absorção e o complemento de conhecimento e a realização de comparações entre os estudos (BOUDREAU; GEFEN; STRAUB, 2001). Esta pesquisa utiliza variáveis e itens estudados e validados em pesquisas anteriores (DUAN; CAO; EDWARDS, 2018; HAJI-KAZEMI; ANDERSEN; KLAKEGG, 2015; ROHRBECK; SCHWARZ, 2013; SCHOEMAKER; DAY, 2009). O construto *Advanced Analytics*, que trata do uso de técnicas e algoritmos oferecidos pelo BDA, bem como o construto que representa a Cultura *Data-driven*, utiliza itens adaptados de Duan et al. (2018). Não foram encontradas, na revisão de literatura, escalas prontas para a variável Identificação de Sinais Fracos. Os itens, portanto, foram desenvolvidos a partir da fundamentação teórico-conceitual, tomando como base escalas dos estudos de Rohrbeck e Schwarz (2013) e Hajikazemi, Andersen e Klakegg (2015). No intuito de responder à questão de pesquisa e os objetivos deste trabalho, também foram formulados, no planejamento deste instrumento, novos itens para os construtos do modelo. As referências para a construção desses itens foram apresentadas no referencial teórico deste artigo, sendo ilustrados na Tabela 12.

Para ter a certeza que o instrumento está representando corretamente os construtos propostos, realizou-se a **validação de face** e **validação de conteúdo** (HAIR et al., 2009),

em seus conceitos e itens (KOUFTEROS, 1999). A validação de face avalia se o questionário apresenta corretamente os conceitos propostos da pesquisa. A validação de conteúdo assegura que os itens operacionalizados de cada variável realmente mede, de forma correta, o que o construto representa (HOPPEN; LAPOINTE; MOREAU, 1996). A validação de face e conteúdo foi conduzida com dez especialistas em Inteligência Estratégica Antecipativa, acadêmicos e profissionais. Foram realizados ajustes nos itens do questionário conforme as observações recebidas, melhorando o texto, deixando-os mais claros e alinhados com seus conceitos, refletindo de forma adequada os construtos.

Para mensurar as variáveis escalares operacionalizadas relacionadas, optou-se por uma escala *Likert*, com cinco pontos, unipolar e graus de concordância de 1 (não concordo) até 5 (concordo plenamente). No questionário, primeiramente realizou-se uma triagem, indagando aos participantes se sua organização possui ou não grandes volumes de dados. A pesquisa finaliza e não encaminha para as perguntas efetivas do questionário estruturado, caso a organização do respondente não possua grandes volumes de dados. Desta forma foram descartadas respostas que não atendem a uma das premissas desta pesquisa. No “Apêndice II – Telas iniciais do questionário online do Artigo II” ilustra-se a apresentação da pesquisa aos participantes, o termo de aceite e o ponto de triagem. Em uma segunda etapa, apresentou-se as perguntas relativas aos construtos, listadas na Tabela 12. Na terceira etapa foram coletadas informações pessoais e profissionais do respondente e sua organização, como: faixa etária, escolaridade, cargo, tempo de empresa, tempo no cargo atual, setor de atividade econômica (ramo de atuação ou indústria), porte e abrangência (nacional, multinacional) da organização.

### **3.4.3 Procedimento de coleta de dados**

A coleta de dados foi realizada em um questionário eletrônico *on-line*. Foi utilizada a plataforma *Google Forms*, para compartilhar, executar as perguntas e guardar seus resultados; trata-se de uma ferramenta gratuita, simples, de fácil utilização, que possibilita a geração de relatórios no formato de planilha eletrônica, para posterior análise.

O questionário foi inicialmente aplicado em um pré-teste, conforme proposto por Koufteros (1999), com 40 respondentes em novembro de 2019. Destes, 32 respostas foram de organizações que trabalham com grandes volumes de dados. Percebeu-se, nesse pré-teste, itens com alto grau de colinearidade, com Fatores de Inflação de Variância (*VIF values*)

maiores que 5. Estes indicadores demonstram respostas com colinearidade em variáveis independentes (HAIR et al., 2016). Os itens pertinentes foram alterados, seguindo a orientação apresentada, antes de realizar a *survey*. Na Tabela 12 exibe-se o roteiro de perguntas estruturado da pesquisa *survey*.

Tabela 12: Roteiro estruturado de perguntas da pesquisa *survey*

	<i>Data Lake</i>	Referências
1.1	Minha organização realiza coletas (ou extrações) de grandes volumes de dados, armazenando-os para futuras análises.	(CHEN et al., 2016, 2017; KITCHENS et al., 2018; LAM et al., 2017;
1.2	Minha organização coleta e armazena dados brutos (não processados) também para utilização futura, independentemente da sua importância imediata em um projeto específico.	MITROVIC, 2017; WATSON, 2017)
1.3	Na minha organização há dados brutos armazenados para processamento, modelagem e análises futuras.	
1.4	Minha organização realiza análises sobre dados brutos coletados e armazenados anteriormente, que não foram buscados especificamente para o propósito da análise em questão.	
<b>Dados não estruturados</b>		
2.1	Minha organização realiza coletas de dados não estruturados.	(ARUNACHALAM ; KUMAR;
2.2	Minha organização realiza processamento e análise de dados não estruturados.	KAWALEK, 2018;
2.3	Na minha organização, os dados não estruturados são fontes de informações que permitem análises qualitativas sobre grandes volumes de dados, como por exemplo a mineração de texto e a análise de sentimento.	DUTTA; BOSE, 2015; KIM et al., 2013; LEE, 2018;
2.4	No ambiente de negócios da minha organização, os dados não estruturados geram informações para a tomada de decisão.	TIEFENBACHER; OLBRICH, 2015)
<b>Advanced Analytics</b>		
3.1	Minha organização utiliza Análise Descritiva para compreender o que aconteceu e o que está acontecendo. Por exemplo, emprega mineração de dados para estudar o contexto e as tendências dos eventos passados e atuais.	(BARTON; COURT, 2012; DUAN; CAO, 2015;
3.2	Minha organização utiliza Análise Preditiva para indicar o que pode acontecer. Por exemplo, usa modelos estatísticos e preditivos para fornecer projeções e explicações de acontecimentos futuros baseados no comportamento passado.	MÜHLROTH; GROTTKE, 2018;
3.3	Minha organização utiliza Análise Prescritiva para indicar o que se deve fazer. Por exemplo, executa otimizações e simulações para recomendar planos de ações, mostrando o resultado provável de cada cenário.	SIVARAJAH et al., 2017;
3.4	Minha organização utiliza Inteligência Artificial. Por exemplo, emprega modelos de machine learning para classificar dados de redes sociais.	YABLONSKY, 2019)
<b>Análise de Outliers</b>		
4.1	Na minha organização, alguns <i>outliers</i> detectados são informações importantes para a tomada de decisão. Por exemplo, eles indicam um novo produto cujo consumo está crescendo em um mercado específico.	(CONSTANTIOU; KALLINIKOS, 2015; HILTUNEN, 2008; KIM; LEE, 2017;
4.2	Minha organização realiza análises detalhadas para compreender melhor os <i>outliers</i> identificados.	MENDONÇA; CARDOSO;
4.3	Na minha organização, alguns <i>outliers</i> são evidências de riscos ou oportunidades.	CARAÇA, 2012)
4.4	Na minha organização, <i>outliers</i> geram insights que facilitam a previsão de fenômenos futuros.	
<b>Método Indutivo</b>		
5.1	Na minha organização, os analistas exploram os dados em busca de novas ideias. Desta forma, NEM SEMPRE os dados são consultados somente para responder perguntas específicas ou para comprovar hipóteses já formuladas.	(BHOLAT, 2015; LYCETT, 2013; MCABEE;
5.2	Na minha organização, os analistas obtêm insights e/ou criam novas hipóteses a partir da exploração dos dados em um escopo amplo de análise.	LANDIS; BURKE, 2017; SAMMUT;
5.3	Na minha organização, os analistas identificam novas correlações, tendências ou padrões, generalizando exemplos encontrados nos dados analisados.	SARTAWI, 2012; STEADMAN, 2013)
5.4	Na minha organização, possíveis oportunidades ou ameaças não conhecidas são percebidas a partir dos dados, sem uma dedução anterior.	
<b>Data-driven Culture</b>		
6.1	Na minha organização, os gestores acreditam que é importante ter, entender e utilizar dados.	(ARUNACHALAM ; KUMAR;
6.2	Na minha organização, novas ideias que desafiam e mudam as práticas atuais são obtidas por meio da análise de dados.	KAWALEK, 2018;
6.3	Na minha organização, a tomada de decisões depende de <i>insights</i> baseados em dados.	BARTON; COURT, 2012; DUAN; CAO, 2015; KIRON;
6.4	Na minha organização, os indivíduos têm a necessidade e/ou a curiosidade de analisar dados.	SHOCKLEY, 2011; MEDEIROS, 2018)

Identificação de Sinais Fracos		
7.1	Na minha organização, as análises apresentam informações caracterizadas como sintomas antecipados de fenômenos futuros.	(ANSOFF, 1975; JANISSEK-
7.2	Na minha organização são coletados dados ambientais que, em um primeiro momento, são incertos, ambíguos e incompletos. Eles podem ser refinados e transformados em informações valiosas.	MUNIZ; BORGES;
7.3	Na minha organização, são identificadas mensagens que fazem sentido apenas em uma análise conjunta e levam a uma melhor compreensão do mercado.	BORTOLI, 2015;
7.4	Na minha organização, são identificadas mensagens que facilitam a percepção de futuras oportunidades e ameaças.	LESCA, 2003)

Fonte: elaborado pelo autor

Na segunda coleta de dados, realizada entre novembro e dezembro de 2019, foram respondidos o total de 167 questionários. As respostas foram tabuladas utilizando a ferramenta Microsoft Excel, descartando registros duplicados (percebeu-se por terem sido registrados pelo mesmo respondente, ao mesmo momento e com as mesmas respostas). Também foram removidas as linhas de empresas que não possuem grandes volumes de dados, em um total de 36. Desta forma, considerou-se nesta etapa 128 respostas válidas.

#### 3.4.4 Análise e Estatística dos Dados

Após esta primeira verificação, os resultados foram analisados de forma descritiva e exploratória, examinando estatisticamente os valores obtidos, utilizando a ferramenta PASW *Statistics* 18. O objetivo foi tratar a base de dados, mantendo somente as respostas relevantes para a verificação dos modelos. Realizou-se, então, a análise de frequência, em uma fase de purificação da amostra, para remover respostas ou itens que não apresentam variância suficiente. Descartou-se 5 questionários por ter mais de 80% de escolhas na mesma alternativa (HAIR et al., 2016), de modo que 123 observações foram consideradas como válidas para esta pesquisa.

O modelo confirmatório e o teste de hipóteses foram realizados por meio de modelagem de equações estruturais, com mínimos quadrados parciais de regressão (*Partial Least Squares* ou *PLS-SEM*) (HAIR et al., 2009; HAIR; RINGLE; SARSTEDT, 2011). A ferramenta utilizada para esta análise foi o *software SmartPLS 3*, indicado para esse tipo de modelagem. O método é apropriado para este estudo, visto que seu objetivo é **suportar um modelo proposto e suas hipóteses**, mensurando o impacto de variáveis independentes na variável dependente “Identificação de Sinais Fracos”. Para Hair et al. (2016), o método *Partial Least Squares* (PLS) pode ser utilizado para suportar hipóteses e para o tratamento de modelos estruturais complexos.

### 3.5 Análise dos Resultados

Após a coleta e tratamento dos dados, apresenta-se nas próximas seções os resultados obtidos a partir das análises realizadas.

#### 3.5.1 Amostra

O total de respondentes válidos considerados nesta amostra é de 123 profissionais. Todos informaram que possuem experiência profissional e trabalham em organizações que utilizam grandes volumes de dados. Quanto às organizações da amostra, são grandes empresas de serviços, em sua maioria. Boa parte dos respondentes possuem mais de quinze anos de experiência e algum tipo de pós-graduação. O descritivo detalhado da amostra é apresentado na Tabela 13.

Tabela 13: Perfil da amostra

<b>Porte da organização</b>	<b>(%)</b>	<b>Segmento de atuação</b>	<b>(%)</b>
Grande	54,5	Serviços	60,2
Médio	28,5	Indústria	19,5
Pequeno	12,2	Comércio / Varejo	10,6
Empresa Individual	2,4	Governo	9,8
Microempresa	2,4		
<b>Experiência profissional</b>	<b>(%)</b>	<b>Escolaridade</b>	<b>(%)</b>
acima de 15 anos	43,9	Pós-graduação, especialização ou MBA	52,0
Até 15 anos	27,6	Graduação	22,0
Até 10 anos	17,1	Mestrado	18,7
Até 5 anos	8,1	Doutorado	4,9
Até 2 anos	3,3	Fundamental, médio ou técnico	2,4
<b>Cargo aproximado</b>	<b>(%)</b>		
Analista, desenvolvedor ou técnico	36,6		
Supervisor ou coordenador	18,7		
Gerente ou superintendente	17,1		
Diretor, C-Level, presidente ou sócio	13,8		
Consultor ou assessor	13,8		

Fonte: Dados da pesquisa (2020)

#### 3.5.2 Modelo de Mensuração

Para avaliar a qualidade do modelo de mensuração, realizou-se inicialmente a análise fatorial exploratória (AFE) com os dados da amostra. O objetivo primário foi de verificar a unidimensionalidade dos fatores, examinando se há construtos latentes subjacentes não percebidos nos itens. Esta análise é importante quando não há escalas prontas na literatura, ou quando analisa-se novos modelos (KOUFTEROS, 1999). Na Tabela 14 observa-se

valores satisfatórios, com cargas que variam entre 0,755 a 0,933, que evidenciam a unidimensionalidade nos construtos (KOUFTEROS, 1999).

Tabela 14: Análise fatorial exploratória – cargas no bloco

<b>Itens</b>	<b>Data Lake</b>	<b>Itens</b>	<b>Método Indutivo</b>
1.1	0,761	5.1	0,858
1.2	0,893	5.2	0,918
1.3	0,834	5.3	0,930
1.4	0,789	5.4	0,854
<b>Itens</b>	<b>Dados não estruturados</b>	<b>Itens</b>	<b>Data-driven Culture</b>
2.1	0,853	6.1	0,853
2.2	0,931	6.2	0,858
2.3	0,881	6.3	0,762
2.4	0,895	6.4	0,840
<b>Itens</b>	<b>Advanced Analytics</b>	<b>Itens</b>	<b>Identificação de Sinais Fracos</b>
3.1	0,755	7.1	0,826
3.2	0,889	7.2	0,876
3.3	0,858	7.3	0,849
3.4	0,843	7.4	0,892
<b>Itens</b>	<b>Análise de Outliers</b>		
4.1	0,932		
4.2	0,913		
4.3	0,931		
4.4	0,933		

Fonte: Dados da pesquisa (2020)

Verificou-se a validade convergente por meio da variância média extraída (AVE), que indica se os itens de cada construto estão correlacionados (HAIR et al., 2009). Observou-se que todos os itens convergem para o conceito da sua variável, já que as validades convergentes são aceitáveis para todos os construtos, com valores AVE superiores a 0,50 (FORNELL; LARCKER, 1981; HAIR et al., 2016; HAIR; RINGLE; SARSTEDT, 2011). A Tabela 15 exibe a variância média extraída para cada construto.

Tabela 15: Validade convergente e confiabilidade composta

	Cronbach's Alpha	Composite Reliability (CR)	Average Variance Extracted (AVE)
<i>Advanced Analytics</i>	0.857	0.903	0.701
Dados não estrut.	0.913	0.938	0.792
<i>Data Lake</i>	0.837	0.890	0.670
<i>Data-driven culture</i>	0.848	0.898	0.687
Método Indutivo	0.913	0.939	0.793
Análise de <i>Outliers</i>	0.946	0.961	0.860
Identificação Sinais Fracos	0.883	0.920	0.741

Fonte: Dados da pesquisa (2020)

A validade discriminante informa se os construtos do modelo são realmente distintos, independentes um dos outros. Para examiná-la utilizou-se o critério *Fornell-Larcker* em cada construto, no qual os valores precisam ser maiores nos elementos da diagonal, que compara a raiz quadrada da AVE de cada construto com a correlação entre eles (FORNELL;

LARCKER, 1981). A Tabela 16 mostra que os valores em negrito são maiores, portanto, confirma-se que há validade discriminante. Também para análise discriminante, adotou-se o teste da razão HTMT (HENSELER; RINGLE; SARSTEDT, 2014), no qual os construtos precisam apresentar valores inferiores a 0,85, limite conservador indicado por Henseler et al. (2014). A Tabela 17 reafirma que nesta amostra os construtos são independentes.

Tabela 16: Validade discriminante segundo critério Fornell-Larcker

Construtos	<i>Advanced Analytics</i>	Dados não estrut.	<i>Data Lake</i>	<i>Data-driven Culture</i>	Método Indutivo	Análise de <i>Outliers</i>	Identificação Sinais Fracos
<i>Advanced Analytics</i>	<b>0.837</b>						
Dados não estrut.	0.476	<b>0.890</b>					
<i>Data Lake</i>	0.407	0.234	<b>0.818</b>				
<i>Data-driven Culture</i>	0.522	0.268	0.355	<b>0.829</b>			
Método Indutivo	0.591	0.441	0.289	0.550	<b>0.890</b>		
Análise de <i>Outliers</i>	0.591	0.350	0.267	0.490	0.558	<b>0.927</b>	
Identificação Sinais Fracos	0.579	0.483	0.302	0.632	0.569	0.474	<b>0.861</b>

Fonte: Dados da pesquisa (2020)

Tabela 17: Validade discriminante segundo critério HTMT

Construtos	<i>Advanced Analytics</i>	Dados não estrut.	<i>Data Lake</i>	<i>Data-driven Culture</i>	Método Indutivo	Análise de <i>Outliers</i>	Identificação Sinais Fracos
<i>Advanced Analytics</i>							
Dados não estrut.	0.531						
<i>Data Lake</i>	0.466	0.268					
<i>Data-driven Culture</i>	0.617	0.304	0.407				
Método Indutivo	0.670	0.474	0.314	0.630			
Análise de <i>Outliers</i>	0.657	0.371	0.299	0.560	0.600		
Identificação Sinais Fracos	0.662	0.528	0.336	0.727	0.629	0.517	

Fonte: Dados da pesquisa (2020)

O coeficiente de confiabilidade, indicado pelo coeficiente de Confiabilidade Composta (CR) e o *alpha de cronbach*, avalia as consistências internas de cada construto (HAIR et al., 2009). Os coeficientes de confiabilidade calculados informam que todos os construtos são consistentes, pois atendem os limites. A CR e o *alpha de cronbach* são maiores que 0,70 (HAIR et al., 2009, 2016; HAIR; RINGLE; SARSTEDT, 2011). A Tabela 15 ilustra estes números.

Constata-se, portanto, que o modelo de mensuração está validado quanto sua consistência interna e confiabilidade composta, validade convergente e discriminante.

### 3.5.3 Modelo Estrutural

Partindo das etapas indicadas por Hair et al. (2016), após a verificação do modelo de mensuração, avaliou-se o modelo estrutural e realizou-se o teste das hipóteses. Iniciou-se a análise do modelo estrutural examinando a colinearidade, calculando o Fator de Inflação de Variância (VIF), apresentado nas Tabela 18 e Tabela 19. O resultado revelou que os valores VIF para todas variáveis (*Inner VIF values*) e seus itens (*Outer VIF values*) são satisfatórios, mostrando que os ajustes realizados na etapa pré-teste foram suficientes e o modelo não foi prejudicado pela colinearidade. Todos valores constam abaixo de 5 (HAIR et al., 2016).

Tabela 18: Collinearity statistics – Outer VIF values

Item	VIF	Item	VIF
1.1	1.580	5.1	2.312
1.2	2.584	5.2	4.406
1.3	1.992	5.3	4.826
1.4	1.720	5.4	2.296
2.1	2.905	6.1	2.180
2.2	4.458	6.2	2.141
2.3	3.071	6.3	1.588
2.4	3.242	6.4	2.043
3.1	1.603	7.1	1.989
3.2	2.574	7.2	2.460
3.3	2.295	7.3	2.212
3.4	2.077	7.4	2.711
4.1	4.374		
4.2	3.650		
4.3	4.532		
4.4	4.528		

Fonte: Dados da pesquisa (2020)

Tabela 19: Collinearity statistics – Inner VIF values

Construtos	Advanced Analytics	Dados não estrut.	Data Lake	Data-driven Culture	Método Indutivo	Análise de Outliers	Identificação Sinais Fracos
Advanced Analytics							1.952
Dados não estrut.							1.363
Data Lake	1.000	1.000			1.000	1.000	
Data-driven Culture			1.000				
Método Indutivo							1.789
Análise de Outliers							1.713
Identificação Sinais Fracos							

Fonte: Dados da pesquisa (2020)

O coeficiente de determinação ( $R^2$ ) representa a quantidade de variância explicada nas variáveis latentes endógenas (HAIR et al., 2012), ou variáveis dependentes, que possuem relação estrutural em sua direção. Os valores  $R^2$  informam o quanto do fenômeno analisado

é explicado pelo modelo de regressão proposto. Eles variam de 0 a 1 e, como o objetivo da abordagem PLS-SEM é esclarecer a variância dessas variáveis, o nível de  $R^2$  dos principais construtos deve ser satisfatório (HAIR; RINGLE; SARSTEDT, 2011). Para Hair, Sarstedt, Ringle, & Mena (2012), o nível aceitável do coeficiente de determinação depende do contexto da pesquisa. Enquanto os resultados de  $R^2$  de 0,20 são vistos como altos para avaliação do comportamento do consumidor, por exemplo, em outras áreas somente valores como 0,75 seriam percebidos desta forma (HAIR; RINGLE; SARSTEDT, 2011). Para Cohen (1988), considera-se os valores  $R^2$  de 0,02 (ou 0,0196), 0,13 e 0,26 como de efeito pequeno, médio e grande, respectivamente. Verificou-se, também, pelo procedimento de *blindfolding*, a relevância preditiva do modelo. Foi calculado, para cada construto dependente, o coeficiente de validade preditiva ou redundância validada cruzada (*cross-validated redundancy* -  $Q^2$ ), o qual necessita ser maior do que zero (HAIR et al., 2016) para indicar que os construtos exógenos têm relevância preditiva sobre os construtos endógenos dependentes (HAIR; RINGLE; SARSTEDT, 2011).

Na Tabela 20 observa-se os coeficientes de determinação  $R^2$  e de validade preditiva  $Q^2$ . O coeficiente de determinação ajustado ( $R^2$  ajustado), é uma modificação do  $R^2$  que considera as variáveis independentes incluídas na equação de regressão (HAIR et al., 2009). Os valores calculados nesta pesquisa indicam um efeito explicado pelo modelo na variância da principal variável latente endógena Identificação dos Sinais Fracos. Com os valores  $R^2$  de 0,45 e  $Q^2$  de 0,30, fica confirmada relevância preditiva satisfatória, já que  $Q^2 > 0$  (HAIR et al., 2016). Se o modelo é propriamente aplicado e analisado, quanto maior o valor de  $R^2$ , maior o poder de explicação da regressão, logo, melhor a previsão da variável latentes endógenas dependentes (HAIR et al., 2009). Desta forma, o modelo proposto apresenta boa capacidade de explicação do construto Identificação de Sinais Fracos, com valores  $R^2$  importantes, uma vez que os demais fatores explicam 45% da variância. Com base na classificação de Cohen (1988), calculou-se  $R^2$  de efeito grande para a variável Identificação de Sinais Fracos. Para as demais variáveis, já era esperado um coeficiente  $R^2$  fraco, em razão de que o modelo explica, de forma significativa, apenas parte da variância dos construtos. Há outros fatores não analisados neste estudo que causam efeito nestas variáveis dependentes. Segundo Cohen (1988), os construtos Dados não estruturados, *Data Lake*, Método Indutivo, Análise de *Outliers* possuem  $R^2$  de efeito pequeno; o construto *Advanced Analytics* possui  $R^2$  de efeito médio.

Tabela 20: Coeficientes de determinação, determinação ajustado e validade preditiva

<b>Construtos</b>	<b>R<sup>2</sup></b>	<b>R<sup>2</sup> ajustado</b>	<b>Q<sup>2</sup></b>
<i>Advanced Analytics</i>	0.166	0.159	0.108
Dados não estruturados	0.055	0.047	0.040
<i>Data Lake</i>	0.126	0.119	0.071
Método Indutivo	0.084	0.076	0.060
Análise de <i>Outliers</i>	0.071	0.064	0.056
Identificação de Sinais Fracos	0.451	0.432	0.304

Fonte: Dados da pesquisa (2020)

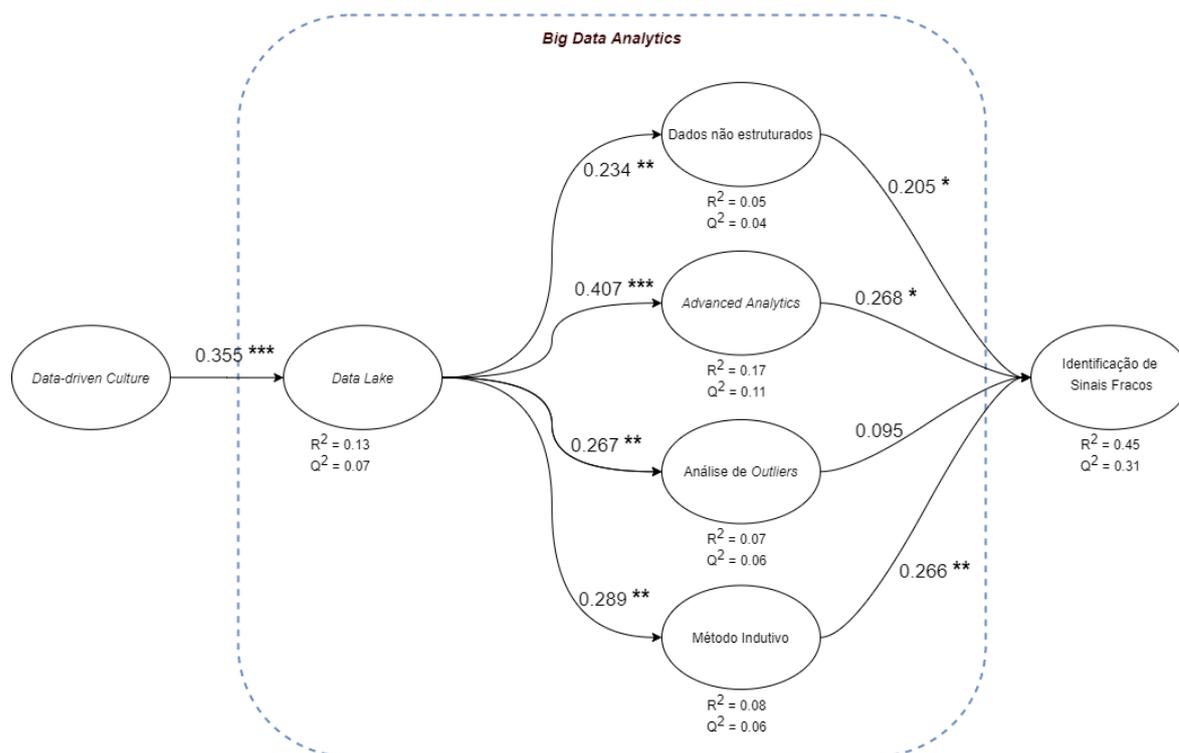
Foi realizado um processo de *bootstrapping* configurado para 5000 amostras, com o objetivo de avaliar a significância estatística dos coeficientes estimados ou das hipóteses a suportar (HAIR et al., 2016). O “teste t” é utilizado para calcular essas significâncias das relações do modelo. Ele é necessário, pois toda a análise é desenvolvida sobre uma amostra da população (HAIR et al., 2009). Os resultados são apresentados na Tabela 21 e na Figura 11, que ilustra o modelo estrutural da pesquisa.

Tabela 21: Resultados das hipóteses

<b>Hipótese</b>	<b>Caminho</b>	<b>Coefficiente (β)</b>	<b>t-Value</b>	<b>p-Value</b>	<b>f<sup>2</sup></b>	<b>Resultado</b>
H1	<i>Data-driven Culture</i> → <i>Data Lake</i>	0.355	4.396	0.000	0.144	Suportada
H2a	<i>Data Lake</i> → Dados não estruturados	0.234	2.768	0.006	0.058	Suportada
H2b	<i>Data Lake</i> → <i>Advanced Analytics</i>	0.407	5.898	0.000	0.199	Suportada
H2c	<i>Data Lake</i> → Análise <i>Outliers</i>	0.267	3.407	0.001	0.077	Suportada
H2d	<i>Data Lake</i> → Método Indutivo	0.289	3.279	0.001	0.091	Suportada
H3	Dados não estrut. → Identificação Sinais Fracos	0.205	2.151	0.032	0.056	Suportada
H4	<i>Advanced Analytics</i> → Identificação Sinais Fracos	0.268	2.454	0.014	0.067	Suportada
H5	Análise <i>Outliers</i> → Identificação Sinais Fracos	0.095	0.998	0.318	0.010	Não Sup.
H6	Método Indutivo → Identificação Sinais Fracos	0.266	2.788	0.005	0.072	Suportada

Fonte: Dados da pesquisa (2020)

Figura 11: Resultados - Modelo Estrutural da Pesquisa



*p-Values* indicam significância: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Fonte: elaborado pelo autor

Para Hair et al. (2009), os valores críticos de significância mais comumente usados são valores-t de 2,58 (nível de 0,01) e 1,96 (nível de 0,05). Assim, por exemplo, ao estabelecer o nível de significância em 0,05, obtém-se o intervalo de confiança de  $\pm 1,96$ , multiplicado pelo erro padrão. Para Koufteros (1999), em geral, valores-t maiores que 2 são considerados como significantes, associados a *p-Values* menores que 0,05. Desta forma a hipótese nula do “teste t” não é suportada, como consequência, as hipóteses propostas são confirmadas.

Observa-se, na Tabela 21 e na Figura 11, que as hipóteses de efeito direto H1, H2a, H2b, H2c, H2d, H3, H4 e H6 foram suportadas. Esses caminhos hipotetizados apresentam valores-t satisfatórios, respectivamente 4.396, 2.768, 5.898, 3.407, 3.279, 2.151, 2.454 e 2.788, e significâncias (*p-Values*) com valores menores que 0,05. A hipótese H5, que trata do caminho “Análise *Outliers* → Identificação Sinais Fracos” não foi suportada por esta análise.

O tamanho do efeito ( $f^2$ ) estima a contribuição para o valor de  $R^2$  em variáveis dependentes. Como referência usou-se para  $f^2$  os seguintes valores: 0,02 para pequeno

impacto na variável endógena, 0,15 para médio e 0,35 para alto (COHEN, 1988; HAIR et al., 2016). Na Tabela 21 são apresentados os valores desta métrica. Todas as hipóteses suportadas apresentam valores de efeitos satisfatórios, sendo superiores a 0,02. A hipótese H5 (Análise *Outliers* → Identificação Sinais Fracos), que não teve significância, também não possui valor de efeito relevante.

Portanto, a modelagem de equações estruturais PLS-SEM deu suporte oito das nove hipóteses deste estudo, com coeficientes de determinação satisfatórios e relevância preditiva. *Data-driven Culture* ( $\beta = 0.355$ ,  $p = 0$ ) tem relação positiva com a implementação de *Data Lakes*. *Data Lake* tem relação positiva com o uso de dados não estruturados ( $\beta = 0.234$ ,  $p = 0.006$ ), *Advanced Analytics* ( $\beta = 0.407$ ,  $p = 0.000$ ), Análise de *Outliers* ( $\beta = 0.267$ ,  $p = 0.001$ ) e do método indutivo ( $\beta = 0.289$ ,  $p = 0.001$ ). Os resultados indicam que uso de dados não estruturados ( $\beta = 0.205$ ,  $p = 0.032$ ), de técnicas de análise classificadas como *Advanced Analytics* ( $\beta = 0.268$ ,  $p = 0.014$ ) e do método indutivo ( $\beta = 0.266$ ,  $p = 0.005$ ), práticas do BDA, tem um efeito positivo na Identificação de Sinais Fracos.

### 3.6 Discussão dos Resultados

Há evidências empíricas que suportam o modelo conceitual proposto nesta pesquisa, com exceção à hipótese H5. O modelo reforça o referencial teórico ao mostrar que a cultura *data-driven* tem um efeito positivo no uso de *Data Lakes* e, conseqüentemente, favorece a identificação de SF. *Data lakes* são recursos que disponibilizam dados previamente coletados, portanto, uma organização que possui DDC necessita e fomenta estas plataformas. Para Dremel (2017), empresas precisam desenvolver recursos técnicos suficientes, aumentando a maturidade da infraestrutura tecnológica, para que os tomadores de decisão aceitem gradualmente os *insights* oriundos dos dados. É evidente que organizações precisam promover esta cultura para obter vantagem competitiva, obtendo visões granulares e panorâmicas para apoiar a tomada de decisão estratégica (KIRON; PRENTICE; FERGUSON, 2012; TEKINER; KEANE, 2013; VASSAKIS; PETRAKIS; KOPANAKIS, 2018). Com processos *data-driven*, organizações melhoram suas previsões e realizam intervenções mais eficazes (VASSAKIS; PETRAKIS; KOPANAKIS, 2018). Para Duan et al. (2018), a cultura orientada a dados influencia positivamente o monitoramento do ambiente, passo imprescindível para a identificação dos SF. Neste sentido, este estudo, que trata de inteligência antecipativa, complementa o de Duan et al. (2018), focado nos efeitos para a inovação. Além disso, *insights* ou candidatos a SF, não são apresentados como um

resultado pronto no uso do BDA. É necessário, primeiramente, desenvolver a cultura empresarial, até o ponto em que as decisões sejam dependentes das percepções baseadas em dados (LAVALLE et al., 2011).

As hipóteses H2a, H2b, H2c, H2d suportadas apontam que *Data Lakes* promovem o uso de dados não estruturados, de *Advanced Analytics*, de *Outliers* e do método de pesquisa indutivo. Com o uso de *Data Lakes* os dados históricos e atuais ficam facilmente disponíveis para as análises (MITROVIC, 2017; WATSON, 2017); desta forma, entende-se que estas ferramentas serão mais utilizadas, na dimensão da cultura *data-driven*. Há também suporte empírico para afirmar que o uso de *Data Lakes* impacta indiretamente na identificação de SF. Associando com o uso de dados brutos, *Advanced Analytics* e computadores, o BDA pode revelar *insights* que anteriormente permaneciam invisíveis (OLBRICH, 2014).

Os resultados evidenciam que o uso de três práticas do BDA influencia positivamente a identificação de Sinais Fracos: dados não estruturados, *Advanced Analytics* e o método indutivo. Considera-se este achado é importante uma vez que a detecção de SF é o início do processo de Inteligência Antecipativa, é a base para o desenvolvimento de cenários futuros (JANISSEK-MUNIZ; FREITAS; LESCA, 2007). SF são as primeiras evidências de descontinuidades estratégicas e sintomas de mudanças (HOLOPAINEN; TOIVONEN, 2012). O passo seguinte do processo de inteligência, a criação de sentido (LESCA, 2003), é o elemento central necessário para construir o significado e gerar conhecimento para a tomada de decisão (FONSECA; BARRETO, 2011). Nesta etapa, quanto maior o volume de sinais relevantes, mais assertivos serão as interpretações futuras e os modelos criados. Os resultados mostram que estas práticas do BDA criam oportunidades para aumentar a quantidade de sinais percebidos, portanto, deduz-se, que com o uso delas, aumenta-se a chance de buscar significados futuros tangíveis e de tomar melhores decisões estratégicas.

Como exposto na Tabela 21, os três fatores com impacto na Identificação dos Sinais Fracos têm efeito semelhante: dados não estruturados, *Advanced Analytics* e o método indutivo. As hipóteses H3, H4 e H6 suportadas reforçam resultados apresentados no Artigo I, que mostraram que gestores identificam SF por meio destas práticas das ferramentas BDA. Conforme exposto nesta pesquisa, *insights* colhidos em grandes volumes de dados são possibilidades de SF, portanto SF podem ser percebidos por meio de ferramentas BDA. O uso de dados não estruturados, como textos coletados em mídias sociais (WELZ et al., 2012), base para análises comportamentais (EREVELLES; FUKAWA; SWAYNE, 2016), é uma das práticas que leva à descoberta de candidatos a SF. A análise desse tipo de dado, mesmo

sendo mais complexa (MCABEE; LANDIS; BURKE, 2017), gera informações para a tomada de decisão estratégica, inviáveis para as fontes de dados estruturados tradicionais (BHIMANI, 2015). A *internet*, em muitos casos, é uma relevante fonte de informações para as organizações. Portanto, o monitoramento do ambiente deve utilizar dados não-estruturados, que compõe a maioria dos dados coletados em *websites*, *blogs*, notícias, publicações em redes sociais (THORLEUCHTER; VAN DEN POEL, 2015).

O uso de técnicas avançadas de análise (*Advanced Analytics*) é igualmente importante para identificar um número maior de SF, já que é necessário empregar ferramentas adequadas para extrair informações destes tipos de dados (WELZ et al., 2012). Por exemplo, abordagens de mineração de texto são usadas para identificar evidências de fenômenos em grandes números de informações textuais (THORLEUCHTER; VAN DEN POEL, 2015). Dados não-estruturados e *Advanced Analytics* permitem a identificação de Sinais Fracos com propósitos econômicos e políticos (MÜHLROTH; GROTTKE, 2018; YOON, 2012), inovação (DUAN; CAO; EDWARDS, 2018; MILOVIDOV, 2018; VEUGELERS; BURY; VIAENE, 2010) e tecnologia (THORLEUCHTER; SCHEJA; VAN DEN POEL, 2014).

Para que haja maior espaço para descobertas, é preciso também adotar a abordagem indutiva de pesquisa (JEBB; PARRIGON; WOO, 2017), como indicado pela hipótese H6, também suportada pela amostra desta pesquisa. Com estes resultados, reafirma-se que, em razão da natureza dos SF – surpreendentes e não esperados – o método indutivo é apropriado e tem efeito positivo na sua identificação. Outro benefício apresentado por esta abordagem de análise é a redução do viés de confirmação, tendência de buscar e encontrar informações que reafirmam premissas anteriores (BHOLAT, 2015; KITCHIN, 2014). Indivíduos, em razão desse viés, não percebem sinais que não fazem parte de ideias prévias, ou do que esperam encontrar (SCHOEMAKER; DAY, 2009). Para Fonseca & Barreto (2011), há sempre a necessidade primária de que alguém entenda um sinal como relevante, para que ele possa ser considerado pela organização em seu curso de ação. Tendo em vista que, com o método indutivo, os candidatos a SF são apontados por algoritmos e podem ser analisados coletivamente, entende-se que este viés fica abreviado.

Contrariando as expectativas, a hipótese H5 não foi suportada, indicando que não há relação positiva significativa entre a análise de *Outliers* e a identificação de Sinais Fracos. Conjectura-se que a não confirmação deste caminho ocorra em razão de que, nas organizações da amostra, essa análise esteja sendo realizada para outros fins. Técnicas

analíticas para detecção de *outliers* são realizadas principalmente na fase de preparação dos dados, com o objetivo de remover “ruídos” (*data cleaning*) na geração e coleta (GE; BANGUI; BUHNOVA, 2018; LI; ZHANG, 2018; YOON; KIM, 2012). Há também outras aplicações operacionais e táticas baseadas na detecção de *outliers*, como, por exemplo, a detecção de fraudes ou de intrusos (ARUNACHALAM; KUMAR, 2018; BREUKER et al., 2016; TABESH; MOUSAVIDIN; HASANI, 2019). Se esta possibilidade for ratificada, seria um exemplo da baixa maturidade organizacional para a cultura *data-driven*. Neste caso, as empresas não estariam apropriando-se da análise de *outliers* para obter *insights* estratégicos. Este resultado carece de um número maior estudos, contudo, pode ser explicado também pelo número e composição da amostra desta pesquisa. Para Abbasi, Sarker e Chiang (2016), análises centradas apenas na probabilidade, que reduzem a relevância dos *outliers*, não bastam para gerenciar adequadamente os riscos relacionados a eventos disruptivos.

### 3.7 Considerações Finais

Este estudo foi motivado diante da transformação que ocorre na área da Inteligência Antecipativa, devido ao volume de informações digitais socioeconômicas disponíveis. Procurou-se entender melhor os potenciais usos do BDA como fonte de informações para esta disciplina. A pesquisa foi conduzida com o objetivo principal de medir o impacto do emprego do BDA, mediante algumas de suas práticas, nos mecanismos de identificação de Sinais Fracos. Para atender essa problemática e cumprir o objetivo, desenvolveu-se um modelo conceitual e realizou-se uma pesquisa quantitativa do tipo *survey*, a fim de suportá-lo.

Entende-se que o estudo cumpriu a finalidade proposta. Os dados empíricos suportam que práticas do BDA tem efeito positivo na identificação de SF. Foram encontrados três fatores com impacto na Identificação dos Sinais Fracos: dados não-estruturados, *Advanced Analytics* e o uso do método de pesquisa Indutivo. Também foram suportados os antecedentes *Data-Driven Culture* e *Data Lake*, que completam o modelo e indiretamente influenciam na identificação de SF.

No âmbito acadêmico, como contribuição teórica, cita-se o desenvolvimento e o suporte empírico ao modelo de mensuração, estruturado para responder a questão de pesquisa. Trata-se de um modelo inicial que relaciona construtos das áreas de *Analytics* e Inteligência Antecipativa. As **evidências suportando que a *Data-Driven Culture* e as práticas do BDA – dados não estruturados, *Advanced Analytics*, método Indutivo, *Data Lakes* – e a identificação de Sinais Fracos estão positivamente relacionados formam a**

**principal contribuição teórica deste estudo.** Reafirma-se, portanto, as conclusões do Artigo I, obtidas por abordagem de natureza qualitativa, sugerindo uma atualização dos métodos de identificação de SF para comportar as análises em grandes volumes de dados como uma forma de monitorar o ambiente. Ressalta-se a conceitualização e validação dos construtos ilustrados no modelo, a discussão dos dados empíricos e suas convergências com a literatura estudada. Ainda no que tange as contribuições teóricas, esta pesquisa ratifica ponderações de Constantiou e Kallinikos (2015), Duan et al. (2018), Kitchin e Lauriault (2015), J. Kim e Lee (2017) e S. Kim et al. (2013), no sentido de que o BDA é empregado como meio de entender o mercado, devendo assim ser incorporado aos processos de monitoramento do ambiente.

Em relação às implicações gerenciais, apresentou-se fundamentos para afirmar que investimentos realizados em BDA retornam vantagem competitiva e valor estratégico. Esta afirmação concorda com estudos anteriores (BARTON; COURT, 2012; CAESARIUS; HOHENTHAL, 2018; JIANG; GALLUPE, 2015; LAVALLE et al., 2011; SALLEH, 2016) e é reforçada devido aos achados deste estudo. A geração de *insights data-driven* e a identificação de um maior número de Sinais Fracos proporcionará, às organizações, um melhor conhecimento sobre seus negócios e a antecipação de eventos futuros, alavancando oportunidades e prevendo possíveis ameaças. Para obter estes retornos estratégicos, considera-se necessário o investimento: na promoção da **cultura da decisão guiada por dados**; na implementação de **Data Lakes** e seus dados brutos históricos; no emprego da **análise em dados não estruturados**, do **Advanced Analytics** e do **método indutivo**.

Entende-se como limitação desta pesquisa, primeiramente, a análise ter sido realizada com 123 respondentes de organizações que são, em sua maioria, de grande porte e da área de serviços. O estudo, ademais, foi conduzido entre praticantes brasileiros, que em sua maioria possuem pós-graduação e mais de dez anos de experiência profissional, conforme perfil apresentado na Tabela 13. Diferenças culturais podem mudar as percepções a respeito dos temas deste estudo, portanto, há possibilidade de vieses que cria oportunidades de estudos futuros mais abrangentes, em um maior e mais variado escopo de amostra. Outro limite desta pesquisa é a probabilidade de que a identificação de Sinais Fracos seja impactada por outros fatores não considerados. Observa-se, ainda referindo-se a limites, que o modelo foi testado usando dados transversais, portanto, outros estudos poderiam repetir a coleta e análise, investigando sua estabilidade. Por último, destaca-se o fato de que nesta pesquisa não foi desenvolvido o construto *Big Data Analytics*, partindo-se do pressuposto, conforme questionário aplicado, que as organizações dos participantes da *survey* já utilizam grandes

volumes de dados. Analisou-se a relação de algumas de suas práticas, conceituadas no referencial teórico e baseadas nas conclusões do Artigo I, como sendo as que têm maior influência na identificação de Sinais Fracos.

Sugere-se, também, novos estudos para investigar a relação entre os construtos análise de *Outliers* e identificação de Sinais Fracos, que não apontou efeito significativo, ao contrário do que foi deduzido e apresentado na revisão teórica. Propõe-se, ainda, que outros modelos sejam explorados e testados, procurando compreender a se a cultura *data-driven* ou o método indutivo moderam as relações com a variável dependente identificação de Sinais Fracos.

Por fim, conclui-se que, com o volume e complexidade de informações disponíveis no ambiente, é relevante para os processos de inteligência que o BDA seja utilizado como meio de detecção Sinais Fracos. Como o objetivo de encontrar Sinais Fracos é a antecipação de fenômenos futuros, o mais cedo possível (ANSOFF, 1975; ROHRBECK; THOM; ARNOLD, 2015), um sistema de monitoramento do ambiente deve processar o maior número de dados, a fim de perceber as alterações mínimas ocorridas (MÜHLROTH; GROTTKE, 2018). Portanto, o BDA é a ferramenta apropriada para esta finalidade. Neste mesmo sentido, Mendonça et al. (2012) observam que é necessário ir além das capacidades dos analistas para identificar e empregar um maior número destes sinais no estabelecimento de estratégias. Erevelles et al. (2016), Yoon (2012) e Mühlroth e Grottke (2018) sugerem que o uso de BDA seja o caminho.

### **3.8 Referências**

As referências bibliográficas deste artigo encontram-se no Capítulo 5 - Referências.

#### 4 Considerações Finais

Por muitos anos, Aguilar (1967), Ansoff (1975), Choo (2001), Lesca (2003), Rohrbeck (2010), Shoemaker, Day e Snyder (2013), entre outros pesquisadores seminais, têm influenciado acadêmicos e praticantes, modelando visões de como o ambiente pode ser compreendido a fim de criar possíveis cenários futuros e antecipar ações. Neste trabalho revisitou-se estes conceitos e teorias, a respeito da necessidade de monitorar o ambiente em processos sistemáticos, encontrar SF e criar sentido, obtendo conhecimento do que pode acontecer.

O contexto da área de SI e de inteligência mudou muito nos últimos anos, principalmente devido ao volume de informações disponíveis (CONSTANTIOU; KALLINIKOS, 2015). De fato, a disponibilidade de dados em grandes volumes, com velocidade e em vários formatos, resultou em uma transformação na tomada de decisão empresarial, com possíveis vantagens competitivas associadas (GHASEMAGHAEI; EBRAHIMI; HASSANEIN, 2018; LAVALLE et al., 2011). Está claro que nas últimas décadas o advento de ferramentas tecnológicas como o BDA possibilitou o tratamento e a análise de uma quantidade de dados que antes era impensável. Tornou-se possível coletar informações econômicas e sociais de clientes, negócios e mercados, correlacionando-as para entender o contexto empresarial (BRYNJOLFSSON; GEVA; REICHMAN, 2016; GHASEMAGHAEI; EBRAHIMI; HASSANEIN, 2018; KITCHENS et al., 2018; MILLER, 2010). Gestores buscam explorar de forma mais eficaz esses dados coletados (CHEN; PRESTON; SWINK, 2015). Em 2014, de acordo com um estudo do Gartner, o BDA foi a maior prioridade dos CIOs (HOLSAPPLE; LEE-POST; PAKATH, 2014). Em muitas organizações ele está no topo da lista dos principais investimentos em TI (CHEN; CHIANG; STOREY, 2012; JIANG; GALLUPE, 2015). Por certo, em razão dos ambientes turbulentos, organizações investem em TI e BDA para ter mais agilidade ao agir conforme as mudanças de mercado (PARK; EL SAWY; FISS, 2017).

O uso de BDA para monitoramento do ambiente e identificação de Sinais Fracos é ainda um tema pouco abordado na literatura de *Big Data* e Inteligência Antecipativa. Há poucos estudos e eles referem-se a características funcionais e tecnológicas (MÜHLROTH; GROTTKE, 2018). Não há pesquisas que retratem visões gerais de como as tecnologias podem ser empregadas (ECKHOFF et al., 2014) ou que apontem o qual prática do BDA deve ser priorizada para majorar a identificação de Sinais Fracos. Este trabalho abordou esse importante *gap* na literatura de Sistemas de Informação. Em dois artigos contendo pesquisas

empíricas, apresentou-se evidências de que organizações identificam candidatos a SF no uso de ferramentas BDA e que práticas do BDA tem efeito positivo na identificação de Sinais Fracos. Na amostra colhida houve suporte à relação estatística entre práticas do BDA e a identificação de SF, mediante um modelo conceitual deduzido, envolvendo também o construto antecedente cultura *data-driven*.

Entende-se, pelos achados dos artigos, que a questão de pesquisa foi respondida satisfatoriamente, evidenciando e medindo o impacto de práticas do BDA na identificação de Sinais Fracos. O Artigo I ilustra quais as estratégias de busca de *insight* que facilitam a identificação de SF: a coleta e análise de dados brutos e não estruturados, o estudo de *outliers*, o uso da abordagem indutiva, o uso de técnicas de análise disponíveis no BDA e a promoção da cultura *data-driven*. No Artigo 2, o modelo conceitual suportado mediu o impacto direto de três fatores, práticas do BDA, na identificação dos Sinais Fracos: uso de dados não estruturados, *Advanced Analytics* e o método de pesquisa indutivo. Também foram suportados os efeitos indiretos do uso de *Data Lakes* e da cultura *data-driven*.

Atendeu-se, pelos resultados do Artigo I e pelo modelo estrutural suportado apresentado no Artigo II, o objetivo geral desta pesquisa, de analisar a relação entre as práticas do BDA e a identificação de Sinais Fracos. Ao medir estatisticamente, usando modelagem de equações estruturais, o efeito das práticas do BDA na identificação de Sinais Fracos, analisou-se as relações causais – correlacionais e de dependência – entre esses elementos (FREITAS et al., 2000; HAIR et al., 2009). As hipóteses do modelo foram suportadas, com exceção à análise de *Outliers*. Dois objetivos específicos foram cumpridos pelo Artigo I e outros dois pelo Artigo II, conforme planejado e ilustrado na Tabela 1: Resumo da Pesquisa. Na revisão teórica do Artigo I analisou-se a relação entre os Sinais Fracos e os tipos de informações que são coletadas nos processos de análise usando BDA, resumida na Tabela 3: Paralelo entre Sinais Fracos e *Big Data*. Entre os resultados da pesquisa empírica do Artigo I, demonstrou-se que as organizações percebem SF em análises realizadas usando BDA, atendendo o segundo objetivo específico.

Como resultado da pesquisa qualitativa do Artigo I e do referencial teórico do Artigo II, identificou-se variáveis relacionadas ao BDA que impactam na identificação de SF. O modelo proposto e suas hipóteses, mensurando como a identificação de SF é impactada por essas variáveis, foi suportado na pesquisa quantitativa realizada. Portanto, compreende-se que os objetivos gerais e específicos desta pesquisa foram cumpridos.

Este estudo foi conduzido por um plano metodológico, apresentado no desenho de pesquisa da Figura 1. No primeiro artigo, com método qualitativo de natureza exploratório, realizou-se dez entrevistas individuais em profundidade, usando um roteiro de perguntas semiestruturado, com profissionais da área de inteligência e experiência em BDA. No segundo artigo, a pesquisa quantitativa foi desenvolvida e analisada em *survey* com 123 respostas válidas.

Este trabalho revela contribuições teóricas significativas para a literatura de SI e de Inteligência Antecipativa, a respeito do uso do BDA com a finalidade de encontrar Sinais Fracos, já apresentadas nas considerações finais de cada artigo e resumidas na Tabela 22. Destacam-se aqui, a título de consolidação, as seguintes implicações:

1. Cita-se, inicialmente, o paralelo realizado a partir do referencial teórico entre os Sinais Fracos e as informações obtidas nos processos de análise usando BDA. Apresentou-se as semelhanças entre as naturezas destes dois tipos de informações.

2. **Dados empíricos suportam que práticas do BDA tem efeito positivo na identificação de Sinais Fracos.** No Artigo I foram apresentadas evidências de que candidatos a SF podem ser identificados a partir do BDA e, no Artigo II, suportou-se esta relação mediante práticas do BDA;

3. O desenvolvimento e suporte ao modelo de mensuração contendo hipóteses que relacionam práticas de BDA – **uso de dados não estruturados**, de *Advanced Analytics* e do **método de pesquisa Indutivo** – com efeito na **identificação de Sinais Fracos**. Também foram suportados os antecedentes *Data-Driven Culture* e *Data Lake*, que indiretamente também impactam na identificação de SF;

4. Sugere-se que o BDA seja incorporado como um mecanismo de monitoramento do ambiente, como um meio que viabiliza a busca de informações de mercado;

5. Contribui-se com a literatura de Inteligência Antecipativa, demonstrando a importância das variáveis que têm impacto significativo e distinto para a identificação de SF.

No âmbito gerencial, apresentou-se implicações relevantes para os processos que usam BDA para fins de Inteligência Antecipativa. Os resultados indicam que **o BDA precisa desempenhar um papel maior no monitoramento do ambiente, para majorar o número de *insights* e Sinais Fracos gerados**. Ser mais efetivo no processo de inteligência, permitindo encontrar um maior número de Sinais Fracos, oportuna mais informações para a tomada de decisão. A melhora de desempenho da tomada de decisão estratégica é o principal

motivo para os investimentos em BDA (GHASEMAGHAEI; EBRAHIMI; HASSANEIN, 2018). Resume-se as implicações práticas de ambos os artigos:

1. Evidencia-se que o BDA é um meio de monitoramento do mercado, portanto, viabiliza a identificação de Sinais Fracos capturados no ambiente e disponibilizados nos seus dados;
2. Os resultados indicam que, **para identificar um maior número de Sinais Fracos, considera-se necessário investimentos: no desenvolvimento da cultura da decisão guiada por dados (*Data-driven Culture*); na adoção de *Data Lakes* e seus dados brutos históricos; no emprego da análise em dados não estruturados, *Advanced Analytics* e o método indutivo**. Portanto, com o fim de antecipar fenômenos de mercado a partir de SF identificados em dados, dados empíricos destes estudos indicam que organizações devem implementar essas práticas nas suas áreas de inteligência.
3. Apresentou-se novos argumentos para afirmar que investimentos realizados em BDA retornam vantagens competitivas e valor estratégico.

Esta investigação possui limites já manifestados em cada artigo. Destaca-se, em ambas pesquisas, o número de organizações que participaram das amostras. No primeiro artigo foram consideradas na análise de conteúdo entrevistas de dez especialistas e, no segundo, há 123 respostas válidas analisadas na *survey*. Observa-se, ainda, que cada artigo possui perfis de empresas e respondentes que atendem satisfatoriamente os requisitos metodológicos, porém, seus resultados podem trazer vieses em razão de percepções divergentes em outras culturas.

Partindo dos resultados, limites e conclusões, apontam-se novas oportunidades para estudos futuros, nesta temática de compreender melhor como o BDA contribui para a Inteligência Antecipativa. Seria importante investigar qual a assertividade dos *insights* gerados pelos Sinais Fracos encontrados nos grandes volumes de dados. Eles realmente reduzem a incerteza e apoiam de forma importante a tomada de decisão estratégica? Os possíveis cenários desenvolvidos mediante as práticas do BDA, com suas antecipações de eventos futuros, realmente aumentam o discernimento dos gestores? Neste sentido, sugere-se pesquisas que avaliem o quanto os Sinais Fracos encontrados nos dados são significativos para a tomada de decisão estratégica. Da mesma forma, cita-se como oportunidade de estudos futuros a necessidade de compreender como o BDA pode ser usado para aumentar a aceitação dos fenômenos futuros percebidos, de modo que as decisões sejam acolhidas facilmente.

Tabela 22: Gaps, contribuições e estudos futuros

	Dissertação	Artigo I	Artigo II
Gaps de Pesquisa	<p>Há uma mudança contextual nos processos de tomada de decisão (FRISK; BANNISTER, 2017; GOES, 2013; MCAFEE; BRYNJOLFSSON, 2012) e de inteligência, gerada pelo advento do BDA (CONSTANTIOU; KALLINIKOS, 2015) ainda não tratada de forma extensiva na literatura (JIANG; GALLUPE, 2015).</p> <p>Na literatura não há estudos que demonstrem o que é preciso priorizar em uma implementação de BDA para que o monitoramento do ambiente e a identificação de Sinais Fracos sejam mais eficientes.</p> <p>Revisões de literatura não citam a temática relacionada BDA e Sinais Fracos. (GÜNTHER et al., 2017; LIANG; LIU, 2018).</p> <p>Trabalhos encontrados tratam de aspectos específicos sem uma visão abrangente (ECKHOFF et al., 2014; MÜHLROTH; GROTTKE, 2018).</p> <p>Abbasi, Sarker e Chiang (2016) e Sharma, Mithas e Kankanhalli (2014) apontam a necessidade de pesquisas que demonstrem o resultado do <i>Big Data</i> em relação a tomada de decisão.</p>	<p>Embora existam pesquisas endereçadas ao BDA (GÜNTHER et al., 2017), poucas estudos empíricos o relacionam aos Sinais Fracos (ECKHOFF et al., 2014).</p> <p>BDA ainda é considerado como um tópico de pesquisa emergente. A demanda de uma base teórica na literatura de SI existe e é importante, já que tópicos relevantes e teorias ainda são desconhecidos (POSPIECH; FELDEN, 2013)</p>	<p>O <i>Big Data</i> abriu novas possibilidades para o suporte à decisão (MAIAH et al., 2017), porém, a maior parte da literatura sobre o BDA se concentra em como aprimorar capacidades organizacionais táticas (GROVER et al., 2018), pouco abordando sua utilização como apoio à decisão estratégica (VAJHALA; STRANG; SUN, 2015).</p> <p>Há estudos sobre a detecção de Sinais Fracos por meio de <i>Big Data</i>, entretanto são pesquisas em características funcionais e tecnológicas, não há estudos teóricos propondo modelos com estas duas áreas de pesquisa (MÜHLROTH; GROTTKE, 2018).</p> <p>Não encontram-se estudos sobre detecção de Sinais Fracos por meio de SI que apresentem uma visão geral de como as tecnologias podem ser empregadas (ECKHOFF et al., 2014).</p>
Contribuições Teóricas	<p>Dados empíricos suportam que práticas do BDA tem efeito positivo na identificação de Sinais Fracos.</p> <p>A investigação ratifica ponderações de que o BDA é empregado como meio de entender o mercado, devendo assim ser incorporado como uma ferramenta aos processos de monitoramento do ambiente.</p> <p>O desenvolvimento e suporte ao modelo de mensuração contendo hipóteses que relacionam práticas de BDA e a identificação de Sinais Fracos. A pesquisa empírica deu suporte às seguintes variáveis com efeito na identificação de sinais fracos: dados não estruturados, <i>Advanced Analytics</i>, o uso do método de pesquisa Indutivo, a <i>Data-Driven Culture</i> e o uso de <i>Data Lakes</i>.</p> <p>Contribui-se com a literatura de Inteligência Antecipativa, demonstrando a importância das variáveis que têm impacto significativo e distinto para a identificação de SF.</p>	<p>Apresentadas evidências de que dados candidatos a SF podem ser identificados a partir do BDA.</p> <p>Sugere-se que o BDA seja incorporado como um mecanismo de monitoramento do ambiente.</p>	<p>Dados empíricos suportaram que práticas do BDA tem efeito positivo na identificação de Sinais Fracos: dados não estruturados, <i>Advanced Analytics</i> e o uso do método de pesquisa Indutivo. <i>Data-Driven Culture</i> e <i>Data Lake</i> indiretamente influenciam na identificação de SF.</p>
Implicações Gerenciais	<p>Os resultados indicam que o BDA precisa desempenhar um papel maior no monitoramento do ambiente, para majorar o número de <i>insights</i> e Sinais Fracos gerados.</p> <p>Apresentou-se novos argumentos para afirmar que investimentos realizados em BDA retornam vantagens</p>	<p>Evidencia-se que as organizações podem utilizar BDA como uma forma de monitorar o mercado, antecipando fenômenos, oportunidades e ameaças.</p> <p>Resultados revelam que a coleta e análise de dados brutos e não estruturados, o estudo de <i>outliers</i>, o uso da abordagem indutiva, o</p>	<p>A geração de <i>insights data-driven</i> e a identificação de um maior número de Sinais Fracos por meio do BDA proporciona um melhor conhecimento sobre seus negócios e a antecipação de eventos futuros.</p> <p>Para tal, considerou-se necessário: promover a cultura da decisão guiada por dados; implementar</p>

	<p>competitivas e valor estratégico para a organização.</p> <p>Para identificar um maior número de Sinais Fracos, considera-se necessário investimentos: no desenvolvimento da cultura da decisão guiada por dados (<i>Data-driven Culture</i>); na adoção de <i>Data Lakes</i> e seus dados brutos históricos; no emprego da análise em dados não estruturados, <i>Advanced Analytics</i> e o método indutivo.</p>	<p>uso de técnicas de análise disponíveis no BDA e a promoção da cultura <i>data-driven</i>, favorecem a identificação de um número maior de Sinais Fracos.</p>	<p><i>Data Lakes</i> e seus dados brutos históricos; empregar análise em dados não estruturados, <i>Advanced Analytics</i> e o método indutivo.</p>
Estudos Futuros	<p>Sugere-se pesquisas que avaliem o quanto os Sinais Fracos encontrados nos dados são significativos para a tomada de decisão estratégica. Eles realmente aumentam o discernimento dos gestores?</p> <p>Há a necessidade de compreender como o BDA pode ser usado para aumentar a aceitação dos fenômenos futuros percebidos, de modo que as decisões sejam acolhidas mais facilmente.</p>	<p>Propõe-se a definição e análise de um modelo contendo elementos de impacto do BDA no monitoramento do ambiente e na percepção de possíveis Sinais Fracos.</p>	<p>Sugere-se estudos mais abrangentes, em um maior e mais variado escopo de amostra.</p> <p>Sugere-se novos estudos para investigar a relação entre os construtos análise de <i>Outliers</i> e Identificação de Sinais Fracos, que não apontou efeito significativo.</p> <p>Propõe-se que outros modelos sejam explorados e testados, procurando compreender a se a cultura <i>data-driven</i> ou o método indutivo moderam as relações com a variável dependente Identificação de Sinais Fracos.</p>

Fonte: elaborado pelo autor

Na Tabela 22 resume-se os *gaps* da literatura que motivaram este trabalho, bem como as implicações e sugestões de estudos futuros. Com isso, entende-se que os resultados encontrados aprofundaram e ampliaram o entendimento acerca da identificação de Sinais Fracos por meio do BDA, relacionando teoricamente estas duas áreas, a partir da literatura e de duas pesquisas empíricas.

Finaliza-se esse relatório de investigação para dissertação de mestrado reafirmando que, neste novo contexto de possibilidade de gerar conhecimento a partir de enormes quantidades de dados, ferramentas de tecnologia de informação são cada vez mais necessárias para os processos de inteligência organizacional. Mais especificamente, o uso de BDA e a cultura *Data-driven* viabilizam esse conhecimento, oferecendo Sinais Fracos que estimulam *insights* de cenários futuros e aumentam a capacidade de decidir estratégias. Conclui-se, que o BDA e suas práticas singulares – uso de dados não estruturados, de *Advanced Analytics* e do método de pesquisa Indutivo – são ferramentas relevantes, gerando competitividade e valor estratégico ao suportar decisões alinhadas ao ambiente dinâmico e incerto atual.

## 5 Referências

- ABBAS, A.; ZHANG, L.; KHAN, S. U. A literature review on the state-of-the-art in patent analysis. **World Patent Information**, [s. l.], v. 37, p. 3–13, 2014.
- ABBASI, A.; SARKER, S.; CHIANG, R. Big Data Research in Information Systems: Toward an Inclusive Research Agenda. **Journal of the Association for Information Systems**, [s. l.], v. 17, n. 2, p. I–XXXII, 2016.
- AGUILAR, F. J. **Scanning the business environment**. New York: Macmillan, 1967.
- ANSOFF, H. I. Managing Strategic Surprise by Response to Weak Signals. **California Management Review**, [s. l.], v. 18, n. 2, p. 21–33, 1975.
- ANSOFF, H. I. et al. **Implanting strategic management**. 3rd editio ed. [s.l: s.n.]. Revisão literatura artigo2.
- ARUNACHALAM, D.; KUMAR, N. Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. **Expert Systems with Applications**, [s. l.], v. 111, p. 11–34, 2018.
- ARUNACHALAM, D.; KUMAR, N.; KAWALEK, J. P. Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. **Transportation Research Part E: Logistics and Transportation Review**, [s. l.], v. 114, p. 416–436, 2018.
- BABBIE, E. **Métodos de pesquisas de Survey**. Belo Horizonte: UFMG, 1999.
- BARDIN, L. **Análise de Conteúdo**. São Paulo, SP: Edições 70, 2016.
- BARTON, D.; COURT, D. Making Advanced Analytics Work For You. **Harvard Business Review**, [s. l.], n. October, p. 78, 2012.
- BAWDEN, D.; ROBINSON, L. The dark side of information: Overload, anxiety and other paradoxes and pathologies. **Journal of Information Science**, [s. l.], v. 35, n. 2, p. 180–191, 2009.
- BEDENIK, N. O. et al. Early warning systems - Empirical evidence. **Trziste**, [s. l.], v. 24, n. 2, p. 201–218, 2012.
- BHIMANI, A. Exploring big data's strategic consequences. **Journal of Information Technology**, [s. l.], v. 30, n. 1, p. 66–69, 2015.
- BHOLAT, D. Big Data and central banks. **Big Data & Society**, [s. l.], v. 2, n. 1, p.

205395171557946, 2015.

- BORGES, N. M.; JANISSEK-MUNIZ, R. Individual environmental scanning as a barrier to collective processes in organizations. **Revista de Gestão**, [s. l.], v. 25, n. 3, p. 321–335, 2018.
- BORTOLI, L.; JANISSEK-MUNIZ, R.; BORGES, N. Criação de Conhecimento Organizacional e Análise de Sinais Fracos: Convergências. **CONF-IRM 2015 Proceedings**, [s. l.], n. May, 2015.
- BOSE, R. Advanced analytics: opportunities and challenges. **Industrial Management & Data Systems**, [s. l.], v. 109, n. 2, p. 155–172, 2009.
- BOUDREAU, M.-C.; GEFEN, D.; STRAUB, D. W. Validation in Information Systems Research: A State-of-the-Art Assessment. **MIS Quarterly**, [s. l.], v. 25, n. 1, p. 1, 2001.
- BREUKER, D. et al. Big Data & Analytics in Networked Business Comprehensible Predictive Models for Business Processes 1. **MIS Quarterly**, [s. l.], v. 40, n. 4, p. 1009–1034, 2016.
- BRYMAN, A. **Research Methods and Organization Studies**. London and New York: Routledge, 1989.
- BRYNJOLFSSON, E.; GEVA, T.; REICHMAN, S. Crowd-Squared: Amplifying the Predictive Power of Large-Scale Crowd-Based Data. **MIS Quarterly**, [s. l.], v. 40, n. 4, p. 1–41, 2016.
- BUMBLAUSKAS, D. et al. Big data analytics: transforming data to action. **Business Process Management Journal**, [s. l.], v. 23, n. 3, p. 703–720, 2017.
- CAESARIUS, L. M.; HOHENTHAL, J. Searching for big data: How incumbents explore a possible adoption of big data technologies. **Scandinavian Journal of Management**, [s. l.], v. 34, n. 2, p. 129–140, 2018.
- CAINELLI, A. de S. **Diagnóstico De Pré-Adoção Do Processo Estruturado De Inteligência Nas Organizações**. 2018. Universidade Federal do Rio Grande do Sul, [s. l.], 2018.
- CAINELLI, A.; JANISSEK-MUNIZ, R. **Proposta de ferramenta para diagnóstico de práticas antecedentes à adoção da atividade de Inteligência nas Organizações**. 2017. Universidade Federal do Rio Grande do Sul, [s. l.], 2017.
- CAINELLI, A.; JANISSEK-MUNIZ, R. Pre-Adoption Diagnosis of the Intelligence Process

- in Organizations: A Delphi Study with Intelligence Practitioners. **BAR - Brazilian Administration Review**, [s. l.], v. 16, n. 3, p. 1–28, 2019.
- CARAYANNIS, E. G. et al. An exploration of contemporary organizational artifacts and routines in a sustainable excellence context. **Journal of Knowledge Management**, [s. l.], v. 21, n. 1, p. 35–56, 2017.
- CATALDI, M.; CARO, L. Di; SCHIFANELLA, C. Personalized emerging topic detection based on a term aging model. **ACM Transactions on Intelligent Systems and Technology**, [s. l.], v. 5, n. 1, 2013.
- CEKULS, A. Leadership Values in Transformation of Organizational Culture To Implement Competitive Intelligence Management: the Trust Building Through Organizational Culture. **European Integration Studies**, [s. l.], v. 0, n. 9, 2015.
- CHEN, C. et al. The promises of big data and small data for travel behavior (aka human mobility) analysis. **Transportation Research Part C: Emerging Technologies**, [s. l.], v. 68, p. 285–299, 2016.
- CHEN, D. Q.; PRESTON, D. S.; SWINK, M. How the use of big data analytics affects value creation in supply chain management. **Journal of Management Information Systems**, [s. l.], v. 32, n. 4, p. 4–39, 2015.
- CHEN, H.; CHIANG, R. H. L.; STOREY, V. C. Business Intelligence and Analytics: From Big Data to Big Impact. **Management Information Systems Quarterly**, [s. l.], v. 36, n. 4, p. 1165–1188, 2012.
- CHEN, H. M. et al. How Lufthansa Capitalized on Big Data for Business Model Renovation. **MIS Quarterly Executive**, [s. l.], v. 1615, n. 14, p. 299–320, 2017.
- CHOO, C. W. **The art of scanning the environment**, Bulletin of the American Society for Information Science, 1999.
- CHOO, C. W. Environmental Scanning as Information Seeking and Organizational Learning. **Information Research**, [s. l.], v. 7, n. 1, p. 1–25, 2001.
- COHEN, J. **Statistical power analysis for the behavioral sciences**. Second Edition. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1988.
- CONSTANTIOU, I. D.; KALLINIKOS, J. New games, new rules: Big data and the changing context of strategy. **Journal of Information Technology**, [s. l.], v. 30, n. 1, p. 44–57, 2015.
- CORBETT, J.; WEBSTER, J. Organizational sensemaking and big data frames:

- Opportunity, control, and data limitation. **Proceedings of the Annual Hawaii International Conference on System Sciences**, [s. l.], v. 2015- March, p. 4772–4781, 2015.
- CÔRTE-REAL, N.; OLIVEIRA, T.; RUIVO, P. Assessing business value of Big Data Analytics in European firms. **Journal of Business Research**, [s. l.], v. 70, p. 379–390, 2017.
- DAVENPORT, T. H. Competing on analytics. **Harvard Business Review**, [s. l.], v. 22, n. 3, p. 139–145, 2006.
- DAVENPORT, T. H.; BARTH, P.; BEAN, R. How ‘Big Data’ is Different. **MIT Sloan Management Review**, [s. l.], v. 54, n. 1, 2012.
- DAVENPORT, T. H.; DYCHÉ, J. Big Data in Big Companies. **International Institute for Analytics**, [s. l.], n. May, p. 1–31, 2013.
- DAY, G. S. Closing the marketing capabilities gap. **The Journal of Marketing**, [s. l.], v. 75, n. 4, p. 183–195, 2011.
- DAY, G. S.; SCHOEMAKER, P. Are you a “vigilant leader”? **IEEE Engineering Management Review**, [s. l.], v. 37, n. 2, p. 102–102, 2009.
- DENZIN, N. K. **The SAGE Handbook of Qualitative Research**. Fifth Edit ed. Los Angeles / London / New Delhi / Singapore / Washington DC / Melbourne: SAGE Publications, Inc, 2018.
- DIEBOLD, F. X. The Origin(s) and Development of “Big Data”: The Phenomenon, the Term, and the Discipline. **University of Pennsylvania**, [s. l.], 2018.
- DREMEL, C. How AUDI AG Established Big Data Analytics in Its Digital Transformation. **MIS Quarterly Executive**, [s. l.], 2017.
- DUAN, Y.; CAO, G. An Analysis of the Impact of Business Analytics on Innovation. **Twenty-Third European Conference on Information Systems**, [s. l.], 2015.
- DUAN, Y.; CAO, G.; EDWARDS, J. S. Understanding the impact of business analytics on innovation. **European Journal of Operational Research**, [s. l.], v. 0, p. 1–14, 2018.
- DUTTA, D.; BOSE, I. Managing a big data project: The case of Ramco cements limited. **International Journal of Production Economics**, [s. l.], v. 165, p. 293–306, 2015.
- ECKHOFF, R. et al. Detecting Weak Signals with Technologies Overview of current technology-enhanced approaches for the detection of weak signals. **International Journal of Trends in Economics Management & Technology (IJTEMT)**, [s. l.], v.

- 3, n. 5, 2014.
- EREVELLES, S.; FUKAWA, N.; SWAYNE, L. Big Data consumer analytics and the transformation of marketing. **Journal of Business Research**, [s. l.], v. 69, n. 2, p. 897–904, 2016.
- FARAWAY, J. J.; AUGUSTIN, N. H. When small data beats big data. **Statistics and Probability Letters**, [s. l.], v. 136, p. 142–145, 2018.
- FONSECA, F.; BARRETO, L. F. B. P. de M. Processo Decisório e o Tratamento De Sinais Fracos. **Future Studies Research Journal: Trends and Strategies**, [s. l.], v. 03, n. 02, p. 34–58, 2011.
- FORNELL, C.; LARCKER, D. F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. **Journal of Marketing Research**, [s. l.], v. 18, n. 1, p. 39, 1981.
- FRASER, M. T. D.; GONDIM, S. M. G. Da fala do outro ao texto negociado: discussões sobre a entrevista na pesquisa qualitativa. **Paidéia (Ribeirão Preto)**, [s. l.], v. 14, n. 28, p. 139–152, 2004.
- FREITAS, H. et al. O método de pesquisa survey. **Revista de Administração da USP, RAUSP**, [s. l.], v. 35, n. 3, p. 105–112, 2000.
- FREITAS, H.; JANISSEK-MUNIZ, R. **Análise léxica e análise de conteúdo: técnicas complementares, sequenciais e recorrentes para exploração de dados qualitativos**. Porto Alegre: Sphinx: Editora Sagra Luzzatto, 2000.
- FRISK, J. E.; BANNISTER, F. Improving the Use of Analytics and Big Data By Changing the Decision-Making Culture-a Design Approach. **Management Decision**, [s. l.], n. just-accepted, p. 0, 2017.
- GASKELL, G. Entrevistas individuais e grupais. In: **BAUER, Martin; GASKELL, George. Pesquisa Qualitativa com Texto, Imagem e Som**. 7<sup>a</sup> ed. Petrópolis, RJ: Editora Vozes Ltda., 2002.
- GE, M.; BANGUI, H.; BUHNOVA, B. Big Data for Internet of Things: A Survey. **Future Generation Computer Systems**, [s. l.], 2018.
- GEORGE, G.; HAAS, M. R.; PENTLAND, A. BIG DATA AND MANAGEMENT. **Journal of Strategic Studies**, [s. l.], v. 37, n. 4, p. 477–481, 2014.
- GHASEMAGHAEI, M.; EBRAHIMI, S.; HASSANEIN, K. Data analytics competency for improving firm decision making performance. **Journal of Strategic Information**

- Systems**, [s. l.], v. 27, n. 1, p. 101–113, 2018.
- GHEORGHIU, R.; ANDREESCU, L.; CURAJ, A. A foresight toolkit for smart specialization and entrepreneurial discovery. **Futures**, [s. l.], v. 80, p. 33–44, 2015.
- GOES, P. B. **Editor'S Comments Information Systems Research and Behavioral EconomicsMIS Quarterly**. [s.l: s.n.].
- GOES, P. B. **EDITOR'S COMMENTS Big Data and IS ResearchMIS Quarterly**. [s.l: s.n.].
- GROVER, V. et al. Creating Strategic Business Value from Big Data Analytics: A Research Framework. **Journal of Management Information Systems**, [s. l.], v. 35, n. 2, p. 388–423, 2018.
- GÜNTHER, W. A. et al. Debating big data: A literature review on realizing value from big data. **Journal of Strategic Information Systems**, [s. l.], v. 26, n. 3, p. 191–209, 2017.
- HAIR, J. F. J. et al. **Análise multivariada de dados**. 6a Edição ed. [s.l: s.n.].
- HAIR, J. F. J. et al. An assessment of the use of partial least squares structural equation modeling in marketing research. **Journal of the Academy of Marketing Science**, [s. l.], v. 40, n. 3, p. 414–433, 2012.
- HAIR, J. F. J. et al. **A primer on partial least squares structural equation modeling (PLS-SEM)**. [s.l.] : SAGE Publications, Inc, 2016.
- HAIR, J. F. J.; RINGLE, C. M.; SARSTEDT, M. PLS-SEM: Indeed a silver bullet. **Journal of Marketing Theory and Practice**, [s. l.], v. 19, n. 2, p. 139–152, 2011.
- HAJI-KAZEMI, S.; ANDERSEN, B.; KLAKEGG, O. J. Barriers against effective responses to early warning signs in projects. **International Journal of Project Management**, [s. l.], v. 33, n. 5, p. 1068–1083, 2015.
- HENSELER, J.; RINGLE, C. M.; SARSTEDT, M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. **Journal of the Academy of Marketing Science**, [s. l.], v. 43, n. 1, p. 115–135, 2014.
- HEUPEL, T.; VON JUTERZENKA, M. From a Weak Signal to a Boiling Frog Syndrome Opportunities and Threats of Legal Boundaries in the Paper-based Packaging Supply Chain! **Procedia Economics and Finance**, [s. l.], v. 26, n. 15, p. 722–728, 2015.
- HILTUNEN, E. The future sign and its three dimensions. **Futures**, [s. l.], v. 40, n. 3, p. 247–260, 2008.
- HININGS, B.; GEGENHUBER, T.; GREENWOOD, R. Digital innovation and

- transformation: An institutional perspective. **Information and Organization**, [s. l.], v. 28, n. March, p. 52–61, 2018.
- HOLOPAINEN, M.; TOIVONEN, M. Weak signals: Ansoff today. **Futures**, [s. l.], v. 44, n. 3, p. 198–205, 2012.
- HOLSAPPLE, C.; LEE-POST, A.; PAKATH, R. A unified foundation for business analytics. **Decision Support Systems**, [s. l.], v. 64, p. 130–141, 2014.
- HOPPEN, N.; LAPOINTE, L.; MOREAU, E. Um guia para a avaliação de artigos de pesquisa em sistemas de informação. **Read**, [s. l.], v. 2, n. 2, p. 1–34, 1996.
- ILMOLA, L.; KUUSI, O. Filters of weak signals hinder foresight: Monitoring weak signals efficiently in corporate decision-making. **Futures**, [s. l.], v. 38, n. 8, p. 908–924, 2006.
- INTEZARI, A.; GRESSEL, S. Information and reformation in KM systems: big data and strategic decision-making. **Journal of Knowledge Management**, [s. l.], v. 21, n. 1, p. 71–91, 2017.
- JANISSEK-MUNIZ, R. **Veille anticipative stratégique en PMI: vers un nouvel usage des sites web pour provoquer des informations ‘terrain’ afin d’amorcer des innovations: concepts, instrumentation et validation**. 2004. Université Pierre Mendès, [s. l.], 2004.
- JANISSEK-MUNIZ, R. Fatores Críticos Em Projetos De Inteligência Estratégica Antecipativa E Coletiva. **Revista Inteligência Competitiva**, [s. l.], v. 6, p. 147–180, 2016.
- JANISSEK-MUNIZ, R.; BORGES, N. M.; BORTOLI, L. Gestão Dos Sinais Fracos No Contexto Brasileiro: Estado Da Arte. In: 8º CONGRESSO IFBAE 2015, Gramado, RS. **Anais...** Gramado, RS
- JANISSEK-MUNIZ, R.; FREITAS, H.; LESCA, H. A Inteligência Estratégica Antecipativa e Coletiva como apoio ao desenvolvimento da capacidade de adaptação das organizações. In: CONGRESSO INTERNACIONAL DE GESTÃO DE TECNOLOGIA E SISTEMAS DE INFORMAÇÃO (CONTECSI) 2007, São Paulo/SP. **Anais...** São Paulo/SP: CONTECSI, 2007.
- JANISSEK-MUNIZ, R.; LESCA, H.; FREITAS, H. Inteligência Estratégica Antecipativa e Coletiva para Tomada de Decisão. **Revista Organizações em Contexto-online**, [s. l.], v. 2, n. 4, p. 92–118, 2006.
- JANSSEN, M.; VAN DER VOORT, H.; WAHYUDI, A. Factors influencing big data decision-making quality. **Journal of Business Research**, [s. l.], v. 70, p. 338–345,

2017.

- JEBB, A. T.; PARRIGON, S.; WOO, S. E. Exploratory data analysis as a foundation of inductive research. **Human Resource Management Review**, [s. l.], v. 27, n. 2, p. 265–276, 2017.
- JEBLE, S.; KUMARI, S.; PATIL, Y. Role of Big Data in Decision Making. **Operations and Supply Chain Management**, [s. l.], v. 11, n. 1, p. 36–44, 2018.
- JIANG, J.; GALLUPE, R. B. Environmental Scanning and Business Insight Capability: The Role of Business Analytics and Knowledge Integration Full paper. In: TWENTY-FIRST AMERICAS CONFERENCE ON INFORMATION SYSTEMS 2015, **Anais...** [s.l: s.n.]
- JISSINK, T.; SCHWEITZER, F.; ROHRBECK, R. Forward-looking search during innovation projects: Under which conditions it impacts innovativeness. **Technovation**, [s. l.], v. 84–85, n. July 2018, p. 71–85, 2019.
- KAIVO-OJA, J. Weak signals analysis, knowledge management theory and systemic socio-cultural transitions. **Futures**, [s. l.], v. 44, n. 3, p. 206–217, 2012.
- KAYSER, V.; BLIND, K. Extending the knowledge base of foresight: The contribution of text mining. **Technological Forecasting and Social Change**, [s. l.], v. 116, p. 208–215, 2017.
- KELLER, J.; VON DER GRACHT, H. A. The influence of information and communication technology (ICT) on future foresight processes - Results from a Delphi survey. **Technological Forecasting and Social Change**, [s. l.], v. 85, p. 81–92, 2014.
- KIM, J.; LEE, C. Novelty-focused weak signal detection in futuristic data: Assessing the rarity and paradigm unrelatedness of signals. **Technological Forecasting and Social Change**, [s. l.], v. 120, n. June 2016, p. 59–76, 2017.
- KIM, J.; PARK, Y.; LEE, Y. A visual scanning of potential disruptive signals for technology roadmapping: investigating keyword cluster, intensity, and relationship in futuristic data. **Technology Analysis and Strategic Management**, [s. l.], v. 28, n. 10, p. 1225–1246, 2016.
- KIM, S. et al. NEST: A quantitative model for detecting emerging trends using a global monitoring expert network and bayesian network. **Futures**, [s. l.], v. 52, p. 59–73, 2013.
- KIRON, D.; FERGUSON, R. B.; PRENTICE, P. K. From value to vision: Reimagining the possible with data analytics. **MIT Sloan Management Review**, [s. l.], p. 54 (3), 1–19,

2013.

KIRON, D.; PRENTICE, P. K.; FERGUSON, R. B. Innovating With Analytics. **MIT Sloan Management Review**, [s. l.], 2012.

KIRON, D.; SHOCKLEY, R. Creating Business Value with Analytics. **MIT Sloan Management Review**, [s. l.], 2011.

KITCHENS, B. et al. Advanced Customer Analytics: Strategic Value Through Integration of Relationship-Oriented Big Data. **Journal of Management Information Systems**, [s. l.], v. 35, n. 2, p. 540–574, 2018.

KITCHIN, R. Big data and human geography: Opportunities, challenges and risks. **Dialogues in Human Geography**, [s. l.], v. 3, n. 3, p. 262–267, 2013.

KITCHIN, R. Big Data, new epistemologies and paradigm shifts. **Big Data & Society**, [s. l.], v. 1, n. 1, p. 205395171452848, 2014.

KITCHIN, R.; LAURIAULT, T. P. Small data in the era of big data. **GeoJournal**, [s. l.], v. 80, n. 4, p. 463–475, 2015.

KITCHIN, R.; MCARDLE, G. What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. **Big Data & Society**, [s. l.], v. 3, n. 1, p. 205395171663113, 2016.

KONONIUK, A.; SACIO-SZYMAŃSKA, A. **Assessing the maturity level of foresight in Polish companies—a regional perspective**, 2015.

KOSTIN, K. Foresight of the global digital trends. **Strategic Management**, [s. l.], v. 23, n. 2, p. 11–19, 2018.

KOUFTEROS, X. A. Testing a model of pull production: A paradigm for manufacturing research using structural equation modeling. **Journal of Operations Management**, [s. l.], v. 17, n. 4, p. 467–488, 1999.

KUOSA, T. Futures signals sense-making framework (FSSF): A start-up tool to analyse and categorise weak signals, wild cards, drivers, trends and other types of information. **Futures**, [s. l.], v. 42, n. 1, p. 42–48, 2010.

KUSHIRO, N.; MATSUDA, S.; TAKAHARA, K. Model oriented system design on big-data. **Procedia Computer Science**, [s. l.], v. 35, n. C, p. 961–968, 2014.

LAM, S. K. et al. Leveraging Frontline Employees' Small Data and Firm-Level Big Data in Frontline Management: An Absorptive Capacity Perspective. **Journal of Service Research**, [s. l.], v. 20, n. 1, p. 12–28, 2017.

- LANEY, D. 3D Data Management: Controlling Data Volume, Velocity, and Variety. **Application Delivery Strategies**, [s. l.], v. 949, n. February 2001, p. 4, 2001.
- LARSON, D.; CHANG, V. A review and future direction of agile, business intelligence, analytics and data science. **International Journal of Information Management**, [s. l.], v. 36, n. 5, p. 700–710, 2016.
- LAVALLE, S. et al. Big Data, Analytics and the Path From Insights to Value. **MIT Sloan Management Review**, [s. l.], v. 52, n. 2, p. 21–32, 2011.
- LEE, S. et al. Business planning based on technological capabilities: Patent analysis for technology-driven roadmapping. **Technological Forecasting and Social Change**, [s. l.], v. 76, n. 6, p. 769–786, 2009.
- LEE, Y. Identification of future signal based on the quantitative and qualitative text mining : a case study on ethical issues in artificial intelligence. **Quality & Quantity**, [s. l.], v. 52, n. 2, p. 653–667, 2018.
- LESCA, H. **Veille stratégique: la méthode L.E.SCanning** ®. EMS ed. [s.l: s.n.].
- LESCA, H.; JANISSEK-MUNIZ, R. **Inteligência Estratégica Antecipativa e Coletiva: o Método L.E.SCanning (r)**. Porto Alegre: Palotti, 2015.
- LESCA, N. et al. Drivers and barriers to pre-adoption of strategic scanning information systems in the context of sustainable supply chain. **Systèmes d'information & management**, [s. l.], v. 20, n. 3, p. 9, 2015.
- LI, Y.; ZHANG, Y. The Innovation Research of College Students' Academic Early-Warning Mechanism Under the Background of Big Data. **Proceedings of the Eleventh International Conference on Management Science and Engineering Management**, [s. l.], 2018.
- LIANG, T.-P.; LIU, Y.-H. Research Landscape of Business Intelligence and Big Data Analytics: A Bibliometrics Study. **Expert Systems with Applications**, [s. l.], v. 0, n. 128, p. 1–9, 2018.
- LINDSTROM, M. **Small data: Como poucas pistas indicam grandes tendências**. 1<sup>a</sup> ed. Rio de Janeiro: HarperCollins Brasil, 2016.
- LYCETT, M. “Datafication”: Making sense of (big) data in a complex world. **European Journal of Information Systems**, [s. l.], v. 22, n. 4, p. 381–386, 2013.
- LYYTINEN, K.; GROVER, V. Management Misinformation Systems: A Time to Revisit ? **Journal of the Association for Information Systems**, [s. l.], v. 18, n. 3, p. 1–44,

- 2017.
- MALHOTRA, N. K. **No Pesquisa de marketing: uma orientação aplicada**. Porto Alegre: Bookman, 2012.
- MAY, T. **Pesquisa social - questões, métodos e processos**. 3<sup>a</sup> ed. Porto Alegre, RS: Artmed, 2004.
- MCABEE, S. T.; LANDIS, R. S.; BURKE, M. I. Inductive reasoning: The promise of big data. **Human Resource Management Review**, [s. l.], v. 27, n. 2, p. 277–290, 2017.
- MCAFEE, A.; BRYNJOLFSSON, E. Big Data: The Management Revolution. **Harvard Business Review**, [s. l.], 2012.
- MEDEIROS, M. M. De. **O impacto da capacidade de inteligência analítica de negócios na tomada de decisões na era dos grandes dados**. 2018. Universidade do Vale do Rio dos Sinos - Unisinos, [s. l.], 2018.
- MEDEIROS, M. M. De; MAÇADA, A. C. G.; FREITAS JUNIOR, J. C. da S. O Efeito da Estratégia de Dados na Vantagem Competitiva. In: XXII SEMEAD 2019, São Paulo/SP. **Anais...** São Paulo/SP
- MENDONÇA, S. et al. Wild Cards, Weak Signals and Organizational Improvisation. **Futures**, [s. l.], n. 36, 2004.
- MENDONÇA, S.; CARDOSO, G.; CARAÇA, J. The strategic strength of weak signal analysis. **Futures**, [s. l.], v. 44, n. 3, p. 218–228, 2012.
- MIAH, S. J. et al. A Big Data Analytics Method for Tourist Behaviour Analysis. **Information and Management**, [s. l.], v. 54, n. 6, p. 771–785, 2017.
- MILLER, H. J. The data avalanche is here. Shouldn't we be digging? **Journal of Regional Science**, [s. l.], v. 50, n. 1, p. 181–201, 2010.
- MILOVIDOV, V. Hearing the sound of the wave: What impedes one's ability to foresee innovations? **Foresight and STI Governance**, [s. l.], v. 12, n. 1, p. 88–97, 2018.
- MITROVIC, S. Specifics of the integration of Business Intelligence and Big Data technologies in the processes of economic analysis. **Business Informatics**, [s. l.], v. 2017, n. 4, p. 40–46, 2017.
- MOREIRA, A. L. M. et al. A Clustering Method for Weak Signals to Support Anticipative Intelligence. **International Journal of Artificial Intelligence and Expert Systems**, [s. l.], v. 6, n. 1, p. 1–14, 2015.
- MÜHLROTH, C.; GROTTKE, M. A systematic literature review of mining weak signals

- and trends for corporate foresight. **Journal of Business Economics**, [s. l.], v. 88, n. 5, p. 643–687, 2018.
- OLBRICH, S. Madness of the Crowd - How Big Data Creates Emotional Markets and What Can Be Done To Control Behavioural Risk. **Ecis**, [s. l.], p. 0–11, 2014.
- PARK, Y.; EL SAWY, O. A.; FISS, P. The Role of Business Intelligence and Communication Technologies in Organizational Agility. **Journal of the Association for Information Systems**, [s. l.], v. 18, n. 9, p. 648–686, 2017.
- PETTY, N. J.; THOMSON, O. P.; STEW, G. Ready for a paradigm shift? Part 1: Introducing the philosophy of qualitative research. **Manual Therapy**, [s. l.], v. 17, n. 4, p. 267–274, 2012.
- PINSONNEAULT, A.; KRAEMER, K. L. Survey research methodology in management information systems: An assessment. **Journal of Management Information Systems**, [s. l.], v. 10, n. 2, p. 75–105, 1993.
- PLADS, S.; ROBERT, D.; STREET, F. Towards an Understanding of the Role of Business Intelligence Systems in. **ECIS Proceedings**, [s. l.], p. 1–12, 2013.
- POSPIECH, M.; FELDEN, C. A Descriptive Big Data Model Using Grounded Theory. **2013 IEEE 16th International Conference on Computational Science and Engineering**, [s. l.], p. 878–885, 2013.
- POZZEBON, M.; FREITAS, H. M. R. De. Pela aplicabilidade: com um maior rigor científico - dos estudos de caso em sistemas de informação. **Revista de Administração Contemporânea**, [s. l.], v. 2, n. 2, p. 143–170, 1998.
- RAGUSEO, E. Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. **International Journal of Information Management**, [s. l.], v. 38, n. 1, p. 187–195, 2018.
- RHISIART, M.; MILLER, R.; BROOKS, S. Learning to use the future: Developing foresight capabilities through scenario processes. **Technological Forecasting and Social Change**, [s. l.], v. 101, p. 124–133, 2015.
- RIOS, F. L. de C. et al. Inteligência competitiva, empresarial, estratégica ou de negócios? Um olhar a partir da administração de empresas. **FACEF PESQUISA**, [s. l.], v. 14, p. 225–238, 2011.
- ROHRBECK, R. **Corporate Foresight: Towards a Maturity Model for the Future Orientation of a Firm**. Berlin, Heidelberg.

- ROHRBECK, R.; BADE, M. Environmental scanning, futures research, strategic foresight and organizational future orientation: a review, integration, and future research directions. In: *ISPIM ANNUAL CONFERENCE 2012* 2012, **Anais...** [s.l: s.n.]
- ROHRBECK, R.; SCHWARZ, J. O. The value contribution of strategic foresight: Insights from an empirical study of large European companies. **Technological Forecasting and Social Change**, [s. l.], v. 80, n. 8, p. 1593–1606, 2013.
- ROHRBECK, R.; THOM, N.; ARNOLD, H. IT tools for foresight: The integrated insight and response system of Deutsche Telekom Innovation Laboratories. **Technological Forecasting and Social Change**, [s. l.], v. 97, p. 115–126, 2015.
- ROSSEL, P. Weak signals as a flexible framing space for enhanced management and decision-making. **Technology Analysis and Strategic Management**, [s. l.], v. 21, n. 3, p. 307–320, 2009.
- ROSSEL, P. Early detection, warnings, weak signals and seeds of change: A turbulent domain of futures studies. **Futures**, [s. l.], v. 44, n. 3, p. 229–239, 2012.
- SALLEH, K. A. Adoption of Big Data Solutions : A study on its security determinants using Sec-TOE Framework. **Conf-IRM 2016**, [s. l.], 2016.
- SAMMUT, G.; SARTAWI, M. Perspective-Taking and the Attribution of Ignorance. **Journal for the Theory of Social Behaviour**, [s. l.], v. 42, n. 2, p. 181–200, 2012.
- SAMPIERI, R. H.; COLLADO, C. F.; LUCIO, P. B. **Metodologia de Pesquisa**. 3a Ed ed. São Paulo/SP: Mc-Graw-Hill, 2006.
- SCHOEMAKER, P. J. H.; DAY, G. S. How to Make Sense of Weak Signals. **MIT Sloan Management Review**, [s. l.], v. 50, n. 3, p. 81–89, 2009.
- SCHOEMAKER, P. J. H.; DAY, G. S.; SNYDER, S. A. Integrating organizational networks, weak signals, strategic radars and scenario planning. **Technological Forecasting and Social Change**, [s. l.], v. 80, n. 4, p. 815–824, 2013.
- SCHULTE, C.; HOVORKA, D. Heuristics for Gaining Project Insights. **Australasian Conference on Information Systems**, [s. l.], p. 1–11, 2017.
- SHARMA, R. et al. Business analytics and competitive advantage: A review and a research agenda. **Frontiers in Artificial Intelligence and Applications**, [s. l.], v. 212, p. 187–198, 2010.
- SHARMA, R.; MITHAS, S.; KANKANHALLI, A. Transforming decision-making processes: A research agenda for understanding the impact of business analytics on

- organisations. **European Journal of Information Systems**, [s. l.], v. 23, n. 4, p. 433–441, 2014.
- SIMON, H. A. A Behavioral Model of Rational Choice. **The Quarterly Journal of Economics**, [s. l.], v. 69, n. 1, p. 99, 1955.
- SIVARAJAH, U. et al. Critical analysis of Big Data challenges and analytical methods. **Journal of Business Research**, [s. l.], v. 70, n. August, p. 263–286, 2017.
- SOON, K. W. K.; LEE, C. A.; BOURSIER, P. A study of the determinants affecting adoption of big data using integrated Technology Acceptance Model (TAM) and diffusion of innovation (DOI) in Malaysia. **International Journal of Applied Business and Economic Research**, [s. l.], v. 14, n. 1, p. 17–47, 2016.
- SOUZA, V. A. De. **Fatores Influenciadores da Cultura no Processo de Inteligência nas Organizações**. 2019. Universidade Federal do Rio Grande do Sul, [s. l.], 2019.
- SPRAGUE, R. H. A Framework for the Development of Decision Support Systems. **Source: MIS Quarterly**, [s. l.], v. 4, n. 4, p. 1–26, 1980.
- STARKS, H.; TRINIDAD, S. B. Choose your method: a comparison of phenomenology, discourse analysis, and grounded theory. **Qual. Health Res.**, [s. l.], v. 17, n. 10, p. 1372–1380, 2007.
- STEADMAN, I. Big data and the death of the theorist. **Wired**, [s. l.], 25 jan. 2013.
- TABESH, P.; MOUSAVIDIN, E.; HASANI, S. Implementing big data strategies: A managerial perspective. **Business Horizons**, [s. l.], v. 62, n. 3, p. 347–358, 2019.
- TAYLOR, L.; SCHROEDER, R.; MEYER, E. Emerging practices and perspectives on Big Data analysis in economics: Bigger and better or more of the same? **Big Data & Society**, [s. l.], v. 1, n. 2, 2014.
- TEKINER, F.; KEANE, J. A. Big data framework. In: PROCEEDINGS - 2013 IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN, AND CYBERNETICS, SMC 2013 2013, **Anais...** [s.l: s.n.]
- THORLEUCHTER, D.; SCHEJA, T.; VAN DEN POEL, D. Semantic weak signal tracing. **Expert Systems with Applications**, [s. l.], v. 41, n. 11, p. 5009–5016, 2014.
- THORLEUCHTER, D.; VAN DEN POEL, D. Idea mining for web-based weak signal detection. **Futures**, [s. l.], v. 66, p. 25–34, 2015.
- TIEFENBACHER, K.; OLBRICH, S. Increasing the value of big data projects - Investigation of industrial success stories. **Proceedings of the Annual Hawaii**

- International Conference on System Sciences**, [s. l.], v. 2015- March, p. 294–303, 2015.
- TRIVIÑOS, A. N. S. **Introdução à pesquisa social**. São Paulo/SP: Atlas, 1987.
- VAJJHALA, N. R.; STRANG, K. D.; SUN, Z. Statistical modeling and visualizing Open Big data using a terrorism case study. **Proceedings - 2015 International Conference on Future Internet of Things and Cloud, FiCloud 2015 and 2015 International Conference on Open and Big Data, OBD 2015**, [s. l.], p. 489–496, 2015.
- VASSAKIS, K.; PETRAKIS, E.; KOPANAKIS, I. Big Data Analytics: Applications, Prospects and Challenges. In: **Mobile Big Data**. [s.l: s.n.]. v. 10p. 346.
- VECCHIATO, R. Creating value through foresight: First mover advantages and strategic agility. **Technological Forecasting and Social Change**, [s. l.], v. 101, p. 25–36, 2015.
- VEUGELERS, M.; BURY, J.; VIAENE, S. Linking technology intelligence to open innovation. **Technological Forecasting and Social Change**, [s. l.], v. 77, n. 2, p. 335–343, 2010.
- VILLARS, R. L.; OLOFSON, C. W.; EASTWOOD, M. Big Data: What It Is and Why You Should Care. **International Data Corporation**, [s. l.], 2011.
- WALLER, M. A.; FAWCETT, S. E. Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. **Journal of Business Logistics**, [s. l.], v. 34, n. 2, p. 77–84, 2013.
- WARNKE, P.; SCHIRRMEISTER, E. Small seeds for grand challenges-Exploring disregarded seeds of change in a foresight process for RTI policy. **Futures**, [s. l.], v. 77, p. 1–10, 2016.
- WATSON, H. J. Preparing for the Cognitive Generation of Decision Support. **MIS Quarterly Executive**, [s. l.], v. 16, n. 3, p. 153–169, 2017.
- WELZ, K. et al. Weak signals detection : Criteria for social media monitoring tools. In: PROCEEDINGS OF THE 5TH ISPIM INNOVATION SYMPOSIUM: "STIMULATING INNOVATION: CHALLENGES FOR MANAGEMENT, SCIENCE & TECHNOLOGY 2012, **Anais...** [s.l: s.n.]
- WHITE, M. Digital workplaces: Vision and reality. **Business Information Review**, [s. l.], v. 29, n. 4, p. 205–214, 2012.
- YABLONSKY, S. A. Multidimensional Data-Driven Artificial Intelligence Innovation. **Technology Innovation Management Review**, [s. l.], v. 9, n. 12, p. 16–28, 2019.

- YOON, B.; PARK, Y. Development of new technology forecasting algorithm: Hybrid approach for morphology analysis and conjoint analysis of patent information. **IEEE Transactions on Engineering Management**, [s. l.], v. 54, n. 3, p. 588–599, 2007.
- YOON, J. Detecting weak signals for long-term business opportunities using text mining of Web news. **Expert Systems with Applications**, [s. l.], v. 39, n. 16, p. 12543–12550, 2012.
- YOON, J.; KIM, K. Detecting signals of new technological opportunities using semantic patent analysis and outlier detection. **Scientometrics**, [s. l.], v. 90, n. 2, p. 445–461, 2012.

## 6 Apêndice I – Tabela de categorias e códigos do Artigo I

Na Tabela 23 apresenta-se a lista de categorias e códigos gerados pela análise de conteúdo, realizada a partir da transcrição das gravações no software de QDA NVivo Plus. Gerou-se a tabela mediante a exportação dos nós disponibilizada pelo aplicativo. As colunas “Fontes” e “Ref.” indicam o número de fontes e o número de referências respectivas de cada código e categoria.

Tabela 23: Análise de conteúdo, tabela de categorias e códigos

Categorias e Códigos	Fontes	Ref.
P01.1 BDA coleta dados brutos para análise posterior	9	10
Coletam de forma bruta para análise posterior	7	8
Coletam em grande parte apenas dados transacionais	2	2
P02.2 BDA são coletados e analisados rapidamente	8	8
Dados NÃO são coletados e analisados rapidamente	4	4
Dados são coletados e analisados rapidamente	0	0
Dados são coletados e analisados rapidamente e permitem percepção de mudanças operacionais e táticas no ambiente	4	4
P03.3 Dados não estruturados	7	7
Estão coletando e analisando	1	1
Estão coletando, mas não analisando	2	2
Poucos dados não estruturados	2	2
Sim e possibilitam orientações operacionais e táticas em tempo real	2	2
P04.4 BDA <i>Insights</i> eventos futuros	9	11
BDA não está sendo utilizado para perceber eventos futuros	1	1
BDA oferece <i>insights</i> compreensão de eventos futuros	8	8
BDA pode considerar mais informações nas análises.	2	2
P04.5 Informações análises BDA fragmentadas	6	6
Resultados das análises de BDA apresentam informações incompletas, incertas, fragmentadas ou ambíguas	6	6
P04.6 Como as informações antecipativas são analisadas	7	8
Não há um processo definido de coleta, análise, julgamento e decisão	5	5
Possuem processos bem definidos para coleta, análise e tomada de decisão	2	2
Possui processo de inteligência sem BDA	1	1
P05.7 Filtros Ansoff	7	8
Dados facilitam o filtro de poder	2	2
Existem os 3 filtros ao utilizar BDA	3	3
Filtro de poder forte	1	1
Sistema facilita filtro vigilância	2	2
P06.8 Insights e <i>data-driven</i>	10	11
Confiança nos dados como barreira para <i>data-driven</i>	1	1
Decisões mais intuitivas do que <i>data-driven</i>	3	3
Empresa já é <i>data-driven</i> , mas há um caminho a percorrer	4	4
Há uma tendência por ser <i>data-driven</i>	2	2
Organização é <i>data-driven</i>	1	1
P07.9 <i>Outliers</i>	9	11
Consideram <i>Outliers</i> como importantes informações	3	3

Consideram <i>Outliers</i> como possíveis indicadores de fenômenos futuros	3	3
Não utilizam <i>Outliers</i> removendo-os das análises	4	4
Organizações não estão olhando para os <i>outliers</i>	1	1
<b>P08.10 Método de Pesquisa</b>	<b>6</b>	<b>6</b>
Não utilizam, mas é uma tendência	1	1
Pode utilizar com ressalvas	1	1
Utiliza em geral método dedutivo	2	2
Utiliza também método indutivo	2	2
<b>P09.11 BDA pode ser uma ferramenta de monitoramento do ambiente</b>	<b>8</b>	<b>9</b>
BDA ainda NÃO pode ser uma ferramenta de monitoramento do ambiente	1	1
BDA pode ser uma ferramenta de monitoramento do ambiente	7	8
<b>Tipos de Análise</b>	<b>10</b>	<b>28</b>
Análise de <i>Outliers</i>	2	2
Análise de sentimento	3	3
Classificação	3	4
Clustering	1	1
Correlações	2	2
Descritiva exploratória	5	6
Forecasting	1	2
Modelos preditivos	4	4
NLP	1	1
Reconhecimento de imagem	1	3
<b>Exemplos de resultados de análise</b>	<b>8</b>	<b>20</b>

Fonte: dados da pesquisa (2018)

## 7 Apêndice II – Telas iniciais do questionário online do Artigo II

Figura 12: Tela inicial com apresentação e termo de aceite

### Pesquisa científica em Big Data e Inteligência nas organizações

Olá, meu nome é Fabiano Chiapinotto Saffi, sou mestrando do Programa de Pós-Graduação em Administração da Escola de Administração da UFRGS (PPGA). A dissertação, orientada pela Prof. Dra. Raquel Janissek-Muniz, trata do uso do Big Data e suas ferramentas analíticas nos processos de inteligência das organizações.

Gostaríamos de lhe convidar a participar desta pesquisa científica, contamos com a sua experiência e conhecimento para responder um breve questionário de perguntas. O formulário demora entre 5 e 10 minutos para ser respondido. As opções devem ser marcadas conforme suas experiências profissionais, na organização onde trabalha atualmente. Por último perguntamos algumas informações pessoais e da sua organização, que servem somente para caracterização da amostra.

Os dados serão utilizados apenas para fins acadêmicos, de forma agrupada, preservando o sigilo individual de cada resposta. Caso queira receber os resultados finais da pesquisa, de maneira agregada, favor preencher o e-mail abaixo. Este campo não é obrigatório.

Muito obrigado!!

\* Required

Endereço de e-mail (opcional)

Your answer \_\_\_\_\_

Estou ciente que esta é uma pesquisa científica e autorizo o pesquisador a utilizar as informações para fins acadêmicos e de forma agregada, preservando o sigilo das informações individuais. \*

Sim, autorizo

Não quero participar

NextPage 1 of 4

Fonte: elaborado pelo autor

Figura 13: Ponto de triagem

### Pesquisa científica em Big Data e Inteligência nas organizações

\* Required

Sobre os dados na organização onde atuo \*

Minha organização possui grandes volumes de dados

Minha organização NÃO possui grandes volumes de dados

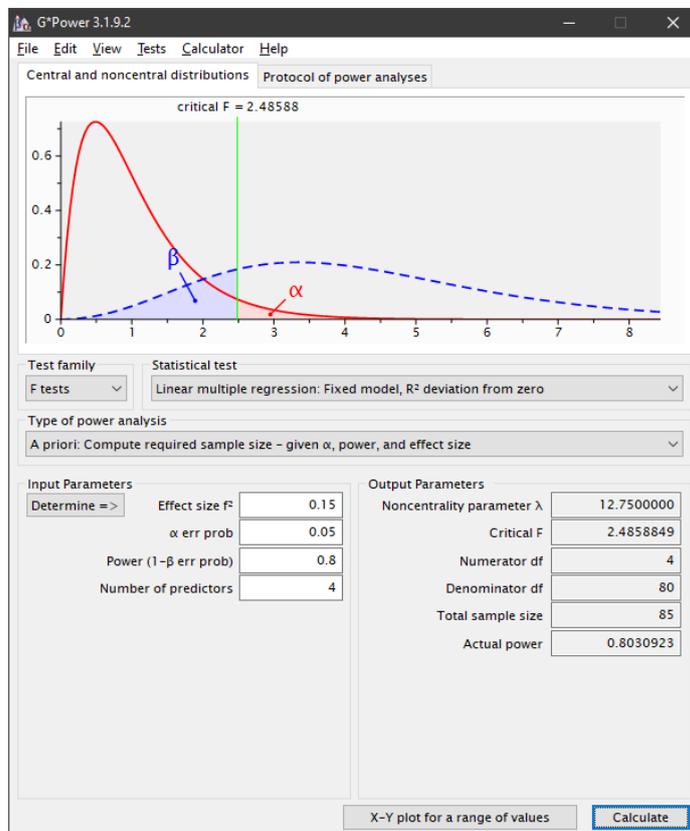
Back
NextPage 2 of 4

Fonte: elaborado pelo autor

## 8 Apêndice III – Estimativa do tamanho mínimo da amostra

A Figura 14 ilustra a tela parametrizada e o resultado do cálculo do tamanho mínimo da amostra.

Figura 14: Cálculo do tamanho da amostra mínima



Fonte: elaborado pelo autor