

FUNDAÇÃO GETULIO VARGAS  
ESCOLA DE ADMINISTRAÇÃO DE EMPRESAS DE SÃO PAULO

PEDRO HENRIQUE GOMES RUGGIERO

**INTELIGÊNCIA ANALÍTICA:  
COMPETÊNCIAS PARA ATUAÇÃO**

SÃO PAULO

2017

PEDRO HENRIQUE GOMES RUGGIERO

**INTELIGÊNCIA ANALÍTICA:  
COMPETÊNCIAS PARA ATUAÇÃO**

Dissertação apresentada à Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas, como requisito para obtenção do título de Mestre em Administração de Empresas

Campo de conhecimento: Administração, Análise e Tecnologia de Informação.

Orientador: Prof. Dr. Fernando de Souza Meirelles

SÃO PAULO

2017

Ruggiero, Pedro Henrique Gomes.

Inteligência analítica: competências para atuação / Pedro Henrique Gomes Ruggiero. - 2017.  
108 f.

Orientador: Fernando de Souza Meirelles

Dissertação (mestrado) - Escola de Administração de Empresas de São Paulo.

1. Gestão do conhecimento. 2. Aquisição de conhecimento (Sistemas especialistas). 3. Sistema de informação gerencial. 4. Desempenho. 5. Big data. 6. Sistemas de suporte de decisão. I. Meirelles, Fernando de Souza. II. Dissertação (mestrado) - Escola de Administração de Empresas de São Paulo. III. Título.

CDU 65.012.4

PEDRO HENRIQUE GOMES RUGGIERO

**INTELIGÊNCIA ANALÍTICA:  
COMPETÊNCIAS PARA ATUAÇÃO**

Dissertação apresentada à Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas, como requisito para obtenção do título de Mestre em Administração de Empresas

Campo de conhecimento: Administração, Análise e Tecnologia de Informação.

**Data de aprovação: 17/02/2017**

**Banca Examinadora:**

---

Prof. Dr. Fernando de Souza Meirelles  
FGV-EAESP

---

Prof. Dr. Alberto Luiz Albertin  
FGV-EAESP

---

Prof. Dr. Gustavo Hermínio Salati  
Marcondes de Moraes  
UNICAMP-Faculdade de Ciências Aplicadas

Dedico este trabalho a todos os seres humanos que procuram contribuir para a sociedade, seja através de estudo, seja através de ação.

Aos meus pais, Sra. Márcia Aparecida Gomes Ruggiero e Sr. Sérgio Ruggiero, que sempre foram meus exemplos de virtudes e inteligência.

## **AGRADECIMENTO**

Ao meu orientador Professor Dr. Fernando de Souza Meirelles, por simplificar os momentos mais difíceis e por contribuir com sua visão holística e com sua vasta experiência acadêmica. Grato pelos ensinamentos e por me ajudar a amadurecer.

Ao Professor Dr. Phokion Sotirios Georgiou, por me esclarecer, através de experiência e dedicação, como fazer uma análise de citações dentro do contexto científico.

Ao Professor Dr. Alberto Luiz Albertin, por me ensinar a ter uma visão organizada de todos os fenômenos e situações, algo que mudou minha capacidade profissional e científica.

À Professora Dra. Maria Alexandra Viegas Cortez da Cunha, por me ensinar a ter visão humana e me ajudar a conciliar desafios pessoais e acadêmicos.

À minha esposa Clarissa Conde, meu filho Henrique Conde Ruggiero, minha mãe Márcia Aparecida Gomes Ruggiero e meu pai Sérgio Ruggiero, por me darem todo o apoio e suporte necessários para cada um dos desafios que enfrentei até hoje.

Ao meu amigo da Fundação Getulio Vargas, Renato de Luca Pires, pela parceria e sinergia que desenvolvemos ao longo do tempo.

## RESUMO

Na era do *Big Data*, a área de Inteligência Analítica tem se tornado cada vez mais importante para comunidades acadêmicas e também para o mercado. A maioria dos pesquisadores acredita que é necessário formar e treinar profissionais cada vez mais completos, com um conjunto de competências em tecnologias, técnicas e negócio, mas há pouco entendimento e convergência sobre esse conjunto. O objetivo deste trabalho é apontar e descrever, segundo a literatura, um conjunto de competências para atuação em Inteligência Analítica, visto que há muitos estudos complementares e com diferentes respostas. Com foco nesse objetivo realizou-se uma análise de citações para levantamento de uma base de textos que foram escopo do trabalho; uma análise nesses textos, sob óptica da Teoria de Competências, para levantamento das competências; uma validação com especialistas, de forma a diminuir a subjetividade do uso da Teoria de Competências; e uma contagem de frequência das competências validadas em toda a base de textos, para obter material e descrever essas competências. Obteve-se um conjunto unificado de competências, entre conhecimentos, habilidades e comportamentos. Para pesquisadores, o resultado deste trabalho serve como uma base unificada de competências e suas descrições, o que poderia ser utilizado para realizar inúmeros estudos mais avançados e voltados para o uso de Inteligência Analítica. Para gerentes de negócio, o resultado deste trabalho poderia ser utilizado para seleção de profissionais para atuação em Inteligência Analítica, ou para desenvolver treinamentos internos, entre outros.

Palavras-chave: Inteligência Analítica, competências, profissional, atuação, *Big Data*.

## **ABSTRACT**

In the era of Big Data, the area of Business Intelligence and Analytics has become increasingly important for academic communities as well as for the market. Most researchers believe that it is necessary to train and train professionals with a set of skills in technologies, techniques and business, but there is little understanding and convergence on this set. The objective of this paper is to list and describe, according to the literature, a set of competencies to perform in Business Intelligence and Analytics, since there are many complementary studies and with different answers. Focusing on this objective, an analysis of citations was carried out to collect a base of texts that were the scope of the work; An analysis in these texts, from the perspective of the Theory of Competences, to find the competences; A validation with specialists, in order to diminish the subjectivity of the use of the Theory of Competences; And a frequency count of competencies validated throughout the text base, to obtain material and to describe those competencies. A unified set of skills was determined, between knowledge, skills and behavior. For researchers, the result of this work serves as a unified basis of competences and their descriptions, which could be used to carry out numerous more advanced studies and directed towards the use of Business Intelligence and Analytics. For business managers, the result of this work could be used to select professionals to perform in Business Intelligence and Analytics, or to develop internal training, among others.

**Keywords:** Business Intelligence Analytics, competences, professional, Big Data.



## LISTA DE ILUSTRAÇÕES

Figura 1: Principais passos para execução de uma análise de citações .....	21
Figura 2: Modelo holístico de Competências .....	25
Figura 3: Resumo dos Requisitos e Desafios de IA .....	36
Figura 4: Principais benefícios e as formas de mensuração encontradas na literatura.....	41
Figura 5: Arquitetura de IA na década de 90 .....	43
Figura 6: Típica arquitetura de IA .....	44
Figura 7: <i>Business Intelligence Framework</i> .....	46
Figura 8: Competências necessárias para profissionais atuarem com IA.....	51
Figura 9: Resumo das etapas do estudo.....	53
Figura 10: Representação visual dos grupos de competências.....	90

## LISTA DE QUADROS

Quadro 1: Conhecimentos para atuação em IA, segundo literatura .....	61
Quadro 2: Habilidades para atuação em IA, segundo literatura .....	62
Quadro 3: Comportamentos para atuação em IA, segundo literatura .....	63
Quadro 4: Validação (médias) por competência .....	66
Quadro 5: Frequência das competências na base de 206 textos .....	92
Quadro 6: Relação entre as competências e o fenômeno IA .....	95

## LISTA DE TABELAS

Tabela 1: Agenda de investimentos das empresas em 2016.....	18
Tabela 2: Textos que compõem o <i>Main Path</i> .....	57
Tabela 3: Textos entre os dez mais citados e que não fazem parte do <i>Main Path</i> .....	58
Tabela 4: Regra de conversão da escala Likert .....	65
Tabela 5: Tipos de textos na base de 208 textos.....	67
Tabela 6: <i>Journals</i> mais citados na base de 121 artigos, dentre os 208 textos.....	68
Tabela 7: Ano de publicação dos 208 textos .....	68

## LISTA DE ABREVIATURAS E SIGLAS

BA - *Business Analytics*

BI - *Business Intelligence*

BI&A - *Business Intelligence and Analytics*

CRM - *Customer Relationship Management*

DDS - *Decision Support Systems*

DOI - *Digital Object Identifier*

EAESP - *Escola de Administração de Empresas de São Paulo*

ERP - *Enterprise Resource Planning*

ETL - *Extract, Transform, Load*

FGV - *Fundação Getulio Vargas*

GIS - *Geographic Information Systems*

HDFS - *Hadoop Distributed File System*

IA - *Inteligência Analítica*

IS - *Information Systems*

NLM - *National Library of Medicine*

OLAP - *Online Analytical Processing*

OLTP - *Online Transaction Processing*

SQL - *Structured Query Language*

SVM - *Support Vector Machine*

TI - *Tecnologia da Informação*

XML - *eXtensible Markup Language*

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>16</b>
<b>2</b>	<b>PROBLEMA DE PESQUISA .....</b>	<b>17</b>
2.1	Objetivos.....	17
2.1.1	<i>Gerais .....</i>	<i>17</i>
2.1.2	<i>Específicos.....</i>	<i>17</i>
2.2	Justificativa.....	18
<b>3</b>	<b>ANÁLISES DE CITAÇÕES.....</b>	<b>20</b>
<b>4</b>	<b>TEORIA DE COMPETÊNCIAS .....</b>	<b>23</b>
<b>5</b>	<b>INTELIGÊNCIA ANALÍTICA .....</b>	<b>26</b>
5.1	Conceitos .....	26
5.2	Aplicações .....	31
5.3	Requisitos e Desafios .....	33
5.4	Benefícios e Desempenho .....	37
5.5	Arquitetura.....	41
5.6	Os profissionais que atuam com Inteligência Analítica .....	46
<b>6</b>	<b>MÉTODO.....</b>	<b>53</b>
6.1	Abordagem metodológica .....	55
6.2	Definição e obtenção da base de textos para revisão de literatura .....	55
6.3	Levantamento e composição dos Quadros de Competências .....	58
6.4	Validação dos Quadros de Competências .....	63
6.5	Método para descrição das Competências.....	67
<b>7</b>	<b>DESCRIÇÃO DAS COMPETÊNCIAS .....</b>	<b>70</b>
7.1	Conhecimentos .....	70
7.1.1	<i>Big Data.....</i>	<i>70</i>
7.1.2	<i>Cloud .....</i>	<i>71</i>
7.1.3	<i>Data Science.....</i>	<i>72</i>

7.1.4	<i>Decision Making</i> .....	72
7.1.5	<i>Decision Support Systems</i> .....	73
7.1.6	<i>Negócio</i> .....	74
7.1.7	<i>Real Time</i> .....	74
7.1.8	<i>Structured Data</i> .....	75
7.1.9	<i>Unstructured Data</i> .....	75
7.2	Habilidades analíticas .....	76
7.2.1	<i>Classification</i> .....	76
7.2.2	<i>Clustering</i> .....	77
7.2.3	<i>Forecasting</i> .....	77
7.2.4	<i>Modeling</i> .....	78
7.2.5	<i>Neural Network</i> .....	79
7.2.6	<i>Optimization</i> .....	79
7.2.7	<i>Predictive Analytics</i> .....	80
7.2.8	<i>Regression</i> .....	80
7.2.9	<i>Sentiment Analysis</i> .....	81
7.2.10	<i>Simulation</i> .....	82
7.2.11	<i>Statistical Analysis</i> .....	82
7.3	Habilidades computacionais .....	83
7.3.1	<i>Data Mining</i> .....	83
7.3.2	<i>Data Warehousing</i> .....	84
7.3.3	<i>ETL (Extract, Transform, Load)</i> .....	84
7.3.4	<i>Machine Learning</i> .....	85
7.3.5	<i>MapReduce</i> .....	86
7.3.6	<i>OLAP (Online Analytical Processing)</i> .....	86
7.3.7	<i>Programming</i> .....	87
7.3.8	<i>Web Mining</i> .....	88
7.4	Comportamentos.....	88
7.4.1	<i>Communicate</i> .....	88
7.4.2	<i>Creative</i> .....	89
<b>8</b>	<b>CONCLUSÕES</b> .....	<b>90</b>
	<b>REFERÊNCIAS</b> .....	<b>96</b>

<b>APÊNDICE 1: TEXTOS DA REVISÃO DE LITERATURA QUE CONTÉM AO MENOS UMA COMPETÊNCIA.....</b>	<b>107</b>
---	------------

## 1 INTRODUÇÃO

A área de Inteligência Analítica (IA) tem se tornado cada vez mais importante para comunidades acadêmicas e também para o mercado (Chen, Chiang, & Storey, 2012). Há estudos que sugerem relação entre a performance das grandes empresas e o uso de IA (Trkman, McCormack, Oliveira, & Ladeira, 2010). Pesquisas mostram que o maior limitante para Inteligência Analítica hoje é a falta de mão de obra qualificada, com as competências necessárias para aproveitar os potenciais benefícios (Russom, 2011).

Sobre os profissionais que atuam em Inteligência Analítica, existem duas linhas observadas na literatura. A primeira delas, defendida por um número menor de pesquisadores, esclarece que os sistemas de Inteligência Analítica devem evoluir o suficiente para serem utilizados por analistas de negócio. A segunda linha, defendida pela maioria dos pesquisadores, acredita que é necessário formar e treinar profissionais cada vez mais completos, com um conjunto de competências (Russom, 2011). Este trabalho está em sintonia com a segunda linha observada na literatura.

O objetivo deste trabalho é apontar e descrever, segundo a literatura, um conjunto de competências para atuação em Inteligência Analítica. Com foco nesse objetivo realizou-se uma análise de citações para levantamento de uma base de textos que foram escopo do trabalho; uma análise nesses textos, sob óptica da Teoria de Competências, para levantamento das competências; uma validação com especialistas, de forma a diminuir a subjetividade do uso da Teoria de Competências; e uma contagem de frequência das competências validadas em toda a base de textos, para obter material e descrever essas competências.

O trabalho está estruturado como se segue: a seção 2 PROBLEMA DE PESQUISA apresenta o problema de pesquisa, o objetivo do trabalho e a justificativa. As seções 3 ANÁLISES DE CITAÇÕES, 4 TEORIA DE COMPETÊNCIAS, e 5 INTELIGÊNCIA ANALÍTICA apresentam a revisão de literatura. A seção 6 MÉTODO descreve a abordagem metodológica e os procedimentos metodológicos utilizados nas fases do estudo. A seção 7 DESCRIÇÃO DAS COMPETÊNCIAS apresenta o resultado da execução do trabalho. A seção 8 CONCLUSÕES apresenta as conclusões e descobertas a partir da execução do trabalho, considerando-se o objetivo do trabalho.



## 2 PROBLEMA DE PESQUISA

A evolução do fenômeno Inteligência Analítica e a crescente atenção gerencial e acadêmica para o assunto criaram a necessidade de se entender melhor quais são as competências para atuação (Chen, Chiang, & Storey, 2012). A falta de uma resposta adequada deixa os praticantes e os gestores das empresas em dúvida sobre como proceder frente ao fenômeno, impossibilitando abordar problemas mais administrativos, como a seleção e retenção de talentos em Inteligência Analítica, ou a exploração dos potenciais benefícios, por exemplo. Em um estudo prévio da literatura, percebeu-se que as respostas obtidas em diferentes estudos se mostraram complementares (Chen, Chiang, & Storey, 2012; Davenport & Patil, 2012), o que indica que ainda não há uma resposta final para a questão.

Baseando-se nesse problema do conhecimento, destaca-se a importância de se entender melhor quais as competências para atuação em Inteligência Analítica. Sob fundamento dos fatos supracitados, essa dissertação pretende responder a seguinte pergunta de pesquisa: Segundo a literatura, quais são as competências para atuação em Inteligência Analítica?

### 2.1 Objetivos

#### 2.1.1 *Gerais*

Apontar e descrever, segundo a literatura, uma base unificada de competências para atuação em Inteligência Analítica.

#### 2.1.2 *Específicos*

- Definir, através de uma análise de citações, uma base adequada para revisão de literatura, onde estará a resposta para a pergunta de pesquisa;

- Analisar, sob óptica da Teoria de Competências, a base definida;
- Quantificar a ocorrência de competências e realizar validação com especialistas;
- Descrever as competências e como se relacionam com Inteligência Analítica.

## 2.2 Justificativa

Em 1986, um artigo publicado na *Sloan Management Review* destacou Inteligência Analítica (IA) como um fenômeno discreto, que já na década de 80 estava propagando-se pelos EUA, mas com pouca atenção acadêmica, até por ser um fenômeno que não foi gerado na academia (Gilad & Gilad, 1986).

Mais de 20 anos depois, em 2008, realizou-se uma revisão de literatura sobre IA nos 10 *top journals* da área *Information Systems*, e pode-se perceber que o assunto ganhou atenção acadêmica (Wixom, Watson, Reynolds, & Hoffer, 2008). Nesse estudo revelou-se que a maioria dos 167 artigos analisados foca na parte estratégica de IA, seguido da parte técnica de IA, implementação de IA, melhoria na tomada de decisão com o uso de IA e dos benefícios no uso de IA. Em 2008 já havia a percepção de que a maioria dos *papers* produzidos sobre IA apenas focavam nas experiências iniciais das empresas e não nas melhorias a longo prazo da adoção (Wixom, Watson, Reynolds, & Hoffer, 2008). Segundo o Gartner (2016), Inteligência Analítica está em primeiro lugar na agenda de investimentos das grandes empresas, dado obtido em uma *survey* com 951 empresas norte americanas e 2.944 empresas ao redor do mundo. A Tabela 1 mostra a porcentagem de respondentes que apontou cada área como o principal foco de investimento para 2016, o que indica que Inteligência Analítica está em primeiro lugar.

Tabela 1: Agenda de investimentos das empresas em 2016

Áreas de investimento	EUA (% de respondentes)	Global (% de respondentes)
BI/Inteligência Analítica	39	39
Infraestrutura e Data Center	28	27
Segurança	25	15
Cloud	24	25
ERP	19	21

Fonte: Traduzido e adaptado de Gartner (2016)

Chen, Chiang e Storey (2012) também destacam que, até 2012, a maioria dos estudos acadêmicos em IA foram publicados em congressos, não em revistas científicas, o que evidencia que o assunto ainda é novo e que os respectivos estudos são em maioria superficiais e sem peso teórico o suficiente para os *top journals* da área de IS.

Goes (2014) escreveu um texto editorial sobre o momento do *Big Data* (que pode ser considerado Inteligência Analítica em contextos mais amplos do que apenas nas empresas) e as oportunidades de estudo científico. O autor relata fatos e informações que reforçam que o fenômeno pode revolucionar praticamente todas as áreas de interesse dos seres humanos.

No contexto gerencial, no mundo todo há líderes seniores perguntando-se se estão obtendo todo o valor que poderiam acerca desse fenômeno (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). Não se pôde descartar que Inteligência Analítica impacta positivamente na performance das empresas (Mcafee & Brynjolfsson, 2012).

Frente a tanta atenção acadêmica e gerencial, surge a necessidade de estudar o fenômeno Inteligência Analítica. Por exemplo, sobre os profissionais que atuam diretamente na prática em IA, questiona-se se a melhor opção seria ter profissionais com mais competências, ou ter sistemas mais simples e que permitam usuários de negócio resolverem seus próprios problemas (Bose, 2009). Para abordar essa questão (e muitas outras) é necessário saber quais são as competências para atuação em IA. Para tentar resolver o problema, pesquisadores publicaram diferentes estudos, entretanto, percebeu-se que as respostas obtidas se mostraram complementares, o que indica que ainda não há uma resposta final para a questão.

Os resultados desta pesquisa podem servir de base para outros estudos mais complexos que envolvam os profissionais que atuam em Inteligência Analítica, inclusive na definição de práticas administrativas, além de poderem servir de base para futuros estudos acadêmicos.

### 3 ANÁLISES DE CITAÇÕES

O objetivo desta seção é indicar os principais conceitos e teorias utilizados como ferramentas para obtenção de material para estudo. Há informações sobre o conceito e como realizar uma análise de citações sobre determinado tema.

Uma análise de citações consiste em verificar todas as seções de referências de um conjunto de textos que caracteriza uma área de conhecimento para encontrar que texto cita o outro com o objetivo de construir uma rede que permita entender esta área. Segundo Garfield (1970), um dos maiores estudiosos em análises de citações, o objetivo de uma análise de citações é realizar medidas em determinada área científica, um processo geralmente lento. Com um número reduzido de textos (até 50 textos) é possível realizar este processo manualmente. Entretanto, o processo manual torna-se inadequado quando há mais textos a analisar, chegando a centenas ou milhares.

Na análise de citações há conceitos específicos para a representação de uma área de conhecimento. Um dos conceitos é o *Main Path*, que consiste em explicitar textos essenciais de uma rede (Georgiou, 2014). O *Main Path* analisa a conectividade entre textos e o tempo define a direção das conexões (Carley, Hummon, & Harty, 1993). Um texto que está conectado a muitos textos do passado e do futuro tende a fazer parte do *Main Path* (Lucio-Arias & Leydesdorff, 2008). O cálculo do *Main Path* baseia-se na contagem de conexões de quantos textos citaram determinado texto e quantos textos esse mesmo texto citou, para cada texto (Batagelj & Mrvar, 1998). O método mais usado para o cálculo do *Main Path* é o *Search Path Count* (SPC) (Jo, Jeung, Park, & Yoon, 2009; Liu, Lu, Lu, & Lin, 2013; Lu, Lin, Liu, & Yu, 2012), adotado neste trabalho.

Para Garfield (1979), uma análise de citações é uma medida genérica do nível de contribuição individual de um pesquisador, texto ou *journal* em determinado assunto, não respondendo perguntas como quem é mais importante, ou quem é melhor. Citações contêm importantes informações de como o conhecimento se dissemina (Liu, Lu, Lu, & Lin, 2013).

Uma alternativa à construção de uma rede e a realização da respectiva análise de citações seria utilizar o *Science Citation Index*, já existente e bem conhecido (Garfield, 1970), porém os

resultados seriam mais genéricos e menos conclusivos do que partir de uma rede específica do assunto.

Segundo alguns autores (Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011), quando se realiza uma análise de citações, há uma série de passos lógicos que envolvem análises e decisões até que se chegue em um resultado que possa ser interpretado. A Figura 1 foi elaborada com base nos passos lógicos definidos por esses autores. Os passos representados por hexágonos são os mais críticos nesse processo. A seguir, descrevem-se em mais detalhes os passos.

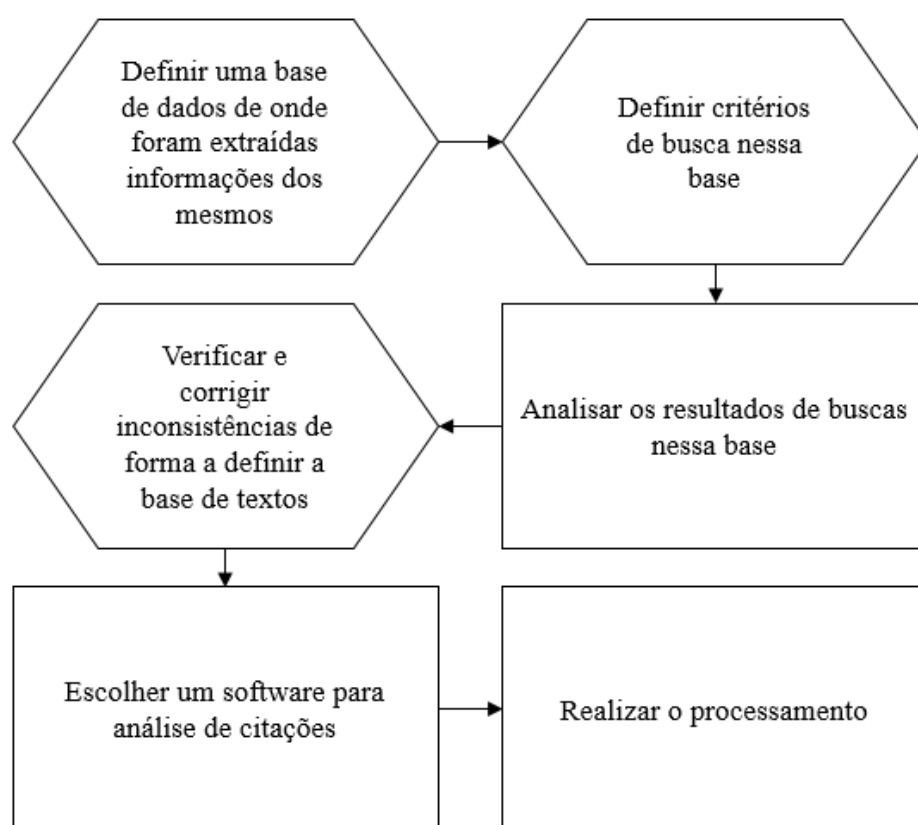


Figura 1: Principais passos para execução de uma análise de citações

Fonte: Elaboração própria, baseado em Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011

Algumas bases de dados são consideradas clássicas no contexto de análise de citações, pois contêm grande parte dos textos de interesse acadêmico (seja no padrão acadêmico, seja em padrões menos rigorosos) e permitem exportar informações. As principais são ISI *Web of Science*, Scopus, Google Scholar e NLM's MEDLINE (Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011).

Vários aspectos devem ser considerados no momento da escolha de qual base ou quais bases serão utilizadas, mas há dois que merecem destaque. O primeiro deles é que tipo de textos a base abrange. Nem todas as bases possuem todos os textos, por isso é apropriado buscar na literatura quais bases são mais adequadas a cada tema, de forma a obter um conjunto representativo de textos. Trabalhar com um conjunto incompleto de textos pode permitir conclusões que não correspondem à realidade da área. O segundo deles é que tipo de exportação a base permite: no momento de definir quais textos representam uma literatura, pode-se chegar a um número alto de textos que inviabilize o processo manual de verificação de citações.

Assim como a seleção incoerente da base, a seleção incoerente de critérios de busca pode excluir textos importantes para determinada área, invalidando análises e conclusões. Novamente, indica-se a busca em literatura para encontrar os melhores termos que definam determinada área.

O próximo passo é a realização de análises de consistência e correções, conhecido como pré-processamento de dados (Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011). Essa etapa consiste em detectar informações duplicadas ou incompletas, como por exemplo identificadores de textos duplicados, e pode demandar um grande esforço dependendo da base de dados de origem. O próximo passo é a seleção de *software* para análise de citações.

Há diversidade de *software* para realizar análises de citações. Alguns deles são (Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011): Bibexcel, CiteSpace II, Sci2Tool, VantagePoint, etc. Além dos programas específicos para análises de citações, existem outros mais matemáticos, utilizados em diferentes tipos de redes (não apenas de citações), por exemplo o *software* Pajek. A escolha do *software* deve ser orientada não somente pela facilidade de uso e pelas operações disponíveis, mas também pela capacidade de representação gráfica das redes (que permite que o ser humano entenda as relações de interesse). Utilizar um *software* matemático como o Pajek pode ser interessante, pois permite operações matemáticas avançadas e tratamentos mais específicos em uma rede (Batagelj & Mrvar, 1998). Isso garante algumas vantagens ao pesquisador, como ter maior imparcialidade ao realizar operações sobre a rede. Além disso, em 2013, o Pajek ganhou o prêmio da *International Network for Social Network Analysis* de melhor *software*, o que garante confiabilidade e justifica a sua escolha para a execução deste trabalho.

## 4 TEORIA DE COMPETÊNCIAS

Existem muitas definições na literatura sobre o que é competência. A definição que é adotada neste trabalho é competência como a posse de vários atributos, que incluem conhecimentos, habilidades relacionadas a resolução de problemas, análise, comunicação, reconhecimento de padrões, e atitudes de todos os tipos, mas sempre dependentes de contexto (Hager & Gonczi, 1996). Outra definição para o termo competência é o conjunto de características associadas à alta performance e motivação, através de uma interação entre o indivíduo e o ambiente (White, 1959). Uma terceira definição é competência como a habilidade de executar as atividades de um determinado trabalho, considerando os padrões e necessidades do empregador (MSC, 1986).

Poucos autores tentaram definir uma terminologia comum dentro do assunto competências. A maioria dos estudos assume pressupostos diferentes para construir modelos de competências, e quase não há diálogo (Le Deist & Winterton, 2005).

Nos EUA, White (1959) define competência como características pessoais associadas a boa performance e boa motivação, com foco na interação entre o indivíduo e o ambiente. Boyatzis (1982) criou um modelo que parte da definição de White e relaciona essas características com as funções gerenciais. Iniciativas como essa despertaram a atenção da AMA (*American Management Association*), que passou a pressionar a AACSB (*American Association on Colleges and Schools of Business*) para adoção da educação por competências (Albanese, 1989). Depois do modelo de Boyatzis (1982), muitos outros apareceram, sempre com o objetivo de conectar o desenvolvimento de recursos humanos e a estratégia organizacional. A maioria dos estudos nos EUA focam em competências funcionais e comportamentais (Aragon & Johnson, 2002).

No Reino Unido, o governo introduziu um modelo de competências baseado denominado VET (*vocational education and training*), que consistia em competências ocupacionais, que por sua vez são consideradas competências funcionais em determinados contextos (Mansfield & Mitchell, 1996). Esse modelo fracassou em uma série de situações, por ser considerado incompleto para alguns empregadores e por tornar o desenvolvimento de recursos humanos muito burocrático (Le Deist & Winterton, 2005). Por esse motivo, muitos outros modelos foram desenvolvidos nos anos seguintes (Reilly, 2003). Apesar de diferentes dos modelos dos EUA,

os modelos que gerados no Reino Unido também focam em competências funcionais e comportamentais (Winterton, Parker, Dodd, McCracken, & Henderson, 2000).

Na França, apenas na década de 90 as empresas adotaram um modelo de competências que considera as competências do indivíduo, diminuindo o foco na função e aumentando o foco no funcionário (Durand, 2000). Na Alemanha, em 1996, foi adotado um modelo de competências em atuação, no qual se mede competência pela qualidade do produto final da atuação (Straka, 2004). O que difere França e Alemanha dos modelos dos EUA e Reino Unido é o aparecimento de competências sociais (Le Deist & Winterton, 2005).

Considerando todo esse histórico de particularidades e pouca convergência entre o que é adotado em cada país influente no assunto competências, Le Deist e Winterton (2005) argumentam que o desenvolvimento de uma tipologia de competências é importante para integrar formação e treinamento, de forma a suprir as necessidades do mercado, e permitir a evolução individual do profissional. Esses autores explicam e argumentam sobre os pontos fortes e fracos de cada modelo, para então definirem o seu próprio modelo. O modelo holístico de competências definido, já traduzido para o português, encontra-se na Figura 2.

Segundo os autores, o conhecimento é capturado por competências cognitivas, habilidades são capturadas por competências funcionais, e comportamentos são capturados por competências sociais. A meta competência é a base que facilita a aquisição das outras competências (Le Deist & Winterton, 2005), ou seja, aprender a aprender. Entende-se, portanto, que competências são conjuntos compostos de conhecimento, habilidades e comportamentos.

Conhecimento é o entendimento do contexto de atuação de um indivíduo, que permite enxergar padrões e pensar sistematicamente (Boyatzis, 2006), sendo este o critério utilizado para mapear os conhecimentos para atuação em Inteligência Analítica, presentes na literatura estudada.

Existem muitos estudos que dividem as habilidades em habilidades funcionais e habilidades comportamentais (Le Deist & Winterton, 2005). O modelo adotado neste trabalho inclui as habilidades comportamentais no pilar de competência social, o que implica na definição de habilidade como competência funcional: *know-how*, ou o “saber fazer”, sendo este o critério utilizado para mapear as habilidades para atuação em Inteligência Analítica, presentes na literatura estudada. No contexto de Inteligência Analítica, são consideradas duas categorias de habilidades: as habilidades analíticas e as habilidades computacionais (Chen, Chiang, & Storey, 2012).



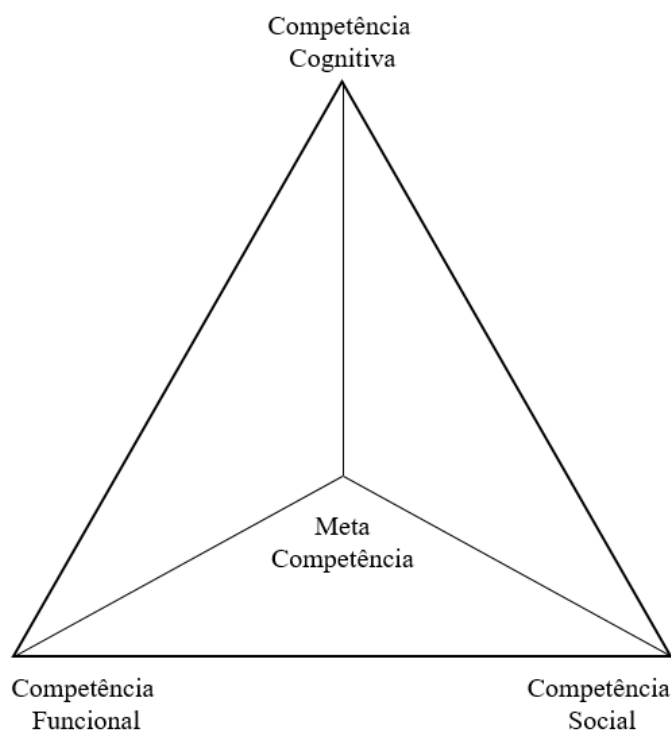


Figura 2: Modelo holístico de Competências  
Fonte: Traduzido de Le Deist e Winterton, 2005

Por fim, os comportamentos são definidos como um conjunto de características sociais que um indivíduo precisa para uma boa performance dentro de uma organização (Le Deist & Winterton, 2005). Nesse contexto, comportamento abrange a vontade e capacidade de experimentar e moldar relações, para identificar e entender interações sociais, de forma a manter a racionalidade, com o objetivo final de ser responsável e solidário (Straka, 2004), sendo este o critério utilizado para mapear os comportamentos para atuação em Inteligência Analítica, presentes na literatura estudada.

## 5 INTELIGÊNCIA ANALÍTICA

O objetivo desta seção é indicar os principais conceitos e teorias utilizados como referências sobre o tema Inteligência Analítica, na literatura também chamado de *Business Intelligence*, *Big Data*, *Business Analytics*, *Analytics*, entre outros. A seção consiste em uma revisão de literatura sobre o tema Inteligência Analítica, que abrange textos cuja obtenção foi detalhada na seção de 6 MÉTODO.

### 5.1 Conceitos

H. P. Luhn (1958) foi o primeiro pesquisador a falar sobre o termo *Business Intelligence* (BI). Naquele momento, BI era definido como um sistema automático (ainda em desenvolvimento) que sugeria possíveis ações reais para os usuários, partindo sempre de documentos. Luhn descreve de forma técnica o que tal sistema deve fazer, mas não são citados possíveis benefícios e nem para onde esperava-se que o BI se desenvolvesse.

Anos depois, em 1986, Gilad e Gilad (1986) definiram *Business Intelligence* como um processo no qual a entrada são os dados brutos e a saída é inteligência para tomada de decisão.

Outros autores confirmam o conceito de *Business Intelligence* como um processo, que utiliza informações operacionais e ferramentas analíticas para melhorar os processos de decisão das empresas (Negash, 2004). A maior diferença desde o surgimento do BI foi a evolução do contexto em que ele se insere, tornando-o mais poderoso. Por exemplo, houve evolução em tratamento de dados, em capacidades de *software* e *hardware*, etc. Em 2004, BI já estava amplamente presente nas empresas. Alguns autores também definem BI como um processo, mas também como um produto (Vedder, Vanecek, Guynes, & Cappel, 1999). Para eles, o processo é composto de métodos para tratar informação que as empresas podem usar para competir, e o produto é a informação que permite predição de comportamento de competidores, fornecedores, clientes, tecnologias, entre outros, com determinado grau de acerto.

De forma similar aos autores supracitados, Bose (2009) alega que BI pode ser definido sob duas perspectivas. A primeira perspectiva é gerencial e o BI é definido como uma forma de entregar as informações corretas para as pessoas corretas para que tomem decisões e melhorem a performance do negócio. A segunda perspectiva é técnica e o BI é definido como um conjunto de tecnologias para reunir, armazenar, analisar e prover acesso à dados como uma forma de melhorar a tomada de decisão (Bose, 2009).

Na década de 70 surgiram pacotes de *software* analíticos no mercado, como as planilhas, por exemplo; na década de 80 surgiram os primeiros sistemas de informação com foco executivo e os DDS (*Decision Support Systems*), que permitiam rápido acesso a informações internas, mas que ao longo dos anos provaram dar uma visão muito estreita para a tomada de decisão; na década de 90 surgiram os sistemas de BI, uma soma de *Data Warehouse*, ferramentas de ETL (*Extract, Transform, Load*), e *softwares* analíticos com capacidade OLAP (*Online Analytical Processing*) (Petrini & Pozzebon, 2009).

Originalmente o BI baseava-se em grandes quantidades de dados estruturados, geralmente na forma de *Data Warehouse* (Holsapple, Lee-Post, & Pakath, 2014). *Data Warehouse* pode ser definido como uma coleção de dados utilizada para fins de tomada de decisão (Baars & Kemper, 2008). Um *Data Warehouse* é constituído de dados de diversas fontes e operações de extração, transformação e carga (ETL) são necessárias. Para tratar os dados, surgiram novas tecnologias como o *Hadoop* e o MapReduce, com o objetivo final de melhorar ganhos de conhecimento, gerar *insights*, encontrar problemas e resolver problemas como forma de apoiar decisões (Holsapple, Lee-Post, & Pakath, 2014). As informações consideradas podem ser quantitativas e também qualitativas, uma vez que ambos os tipos interessam aos tomadores de decisão.

Meirelles (2016), em sua pesquisa anual de Administração e Uso da TI com mais de 2.500 empresas participantes, destaca *Business Intelligence* como um termo tão utilizado comercialmente, que há dificuldade em entender realmente que fenômeno ele descreve. Meirelles considera o termo Inteligência Analítica (IA) como substituto para o BI no contexto acadêmico (por não ser usado comercialmente). Este é o termo adotado ao longo deste trabalho.

Considerando-se os papéis dos sistemas de informação nos negócios ao longo das décadas, o termo inteligência surgiu tardiamente. Nas décadas de 50 e 60 os sistemas de informação tinham o papel de automatizar processamentos e organização de informações. A partir da década de 70 surgiu o papel de apoio à decisão, apesar de não trazer respostas específicas como observou-se

nas décadas seguintes. Da década de 80 em diante o papel dos sistemas de informação passou a ser estratégico para as empresas, permitindo o surgimento de negócios baseados nessas tecnologias, como o comércio eletrônico, por exemplo. Só a partir dos anos 2000 surgiu o termo inteligência para representar o papel dos sistemas de informação (O'brien & Marakas, 2013).

Em paralelo com a evolução do IA como tecnologia e como fenômeno, surgiu o termo *Big Data*. Luvizan, Meirelles e Diniz (2014) definem *Big Data* como o fenômeno do aumento da oferta e demanda de dados e de ferramentas para manejá-los, o que parece bastante relacionado com IA. Esses autores levantam dados de diferentes pesquisas para comprovar a atenção que este fenômeno tem em meio acadêmico e em meio executivo.

Uma definição aceita e utilizada de *Big Data* é o conjunto volume, velocidade, variedade, valor e veracidade (Demchenko, Grosso, de Laat, & Membrey, 2013). Essas dimensões podem ser utilizadas para descrever os requisitos e o que se deseja de tal fenômeno, o que, de certa forma, já demonstra no conceito quais são as expectativas comerciais. Sobre o volume de dados passível de utilização por técnicas de IA em 2011 a *internet* oferecia próximo a um zettabyte ( $2^{70}$  bytes) de tráfego de dados e que possivelmente antes do final do século chegara-se ao yottabyte, que equivale a 1000 zettabytes (Kiron & Shockley, 2011).

*Big Data* deixou de ser algo visto como um grande volume de dados, e hoje é visto como algo que agrega valor dependendo da quantidade de *insights* que podem potencialmente ser obtidos (George, Haas, & Pentland, 2014). Alguns autores definem cinco fontes de dados de *Big Data* (Tan, Blake, Saleh, & Dustdar, 2013): dados públicos (retidos pelo governo e empresas públicas), dados privados (retidos pelas empresas e gerados por transações internas e externas), dados de ambiente (gerados por sistemas ou dispositivos, mas sem valor central, precisam ser associados), dados comunitários (representam realidades sociais) e dados quantificados (contagem de características qualitativas).

Elizabeth D. Liddy, professora da School of Information Studies na Syracuse University, acredita que é mais importante discutir o que se pode aprender com *Big Data* do que permanecer na pergunta sobre o que é *Big Data* (Dumbill, Liddy, Stanton, Mueller, & Farnham, 2013).

Chen, Chiang e Storey (2012) adicionam o termo *Analytics* e definem *Big Data Analytics* como um termo para descrever conjuntos de dados e técnicas analíticas para aplicações muito grandes e complexas, que requerem soluções avançadas de armazenagem, gerenciamento, análise e ferramentas de visualização. Davenport (2006) definiu *Business Analytics* (BA) como o termo

adequado para *Big Data Analytics* no contexto *Business*. Para esse autor, BA é o componente analítico do BI, que surgiu pela crescente complexidade desse componente na presença do fenômeno *Big Data*. Essa mistura de termos e conceitos é uma das principais características da área, o que causa divergências entre pesquisadores, entre praticantes e entre pesquisadores e praticantes.

Chen, Chiang e Storey (2012), frente a essa mistura de termos, adotaram o termo BI&A (*Business Intelligence and Analytics*), que se refere ao conjunto de técnicas, tecnologias, sistemas, práticas, metodologias e aplicações para analisar dados críticos de negócio com objetivo de ajudar empresas a melhor entender o negócio, o mercado e realizar decisões em tempo adequado. Apesar da iniciativa de um termo que unifique os conceitos, a maioria dos autores atuais persiste na utilização do termo *Analytics*. Por isso, para esse trabalho, adotou-se o termo Inteligência Analítica, cuja utilização não é comercial. A diferença entre o clássico BI e o atual IA não é apenas composta de volume, mas de profundidade das informações e que tipos de perguntas pode-se responder (Debortoli, Müller, & vom Brocke, 2014).

*Analytics* é um termo genérico que significa aplicar várias técnicas analíticas em dados para responder perguntas ou resolver problemas, algo que facilita a realização de objetivos de negócio através do *report* de dados para análise de tendências, criação de modelos preditivos para previsão e otimização de processos para melhoria de performance (Robinson, Levis, & Bennett, 2010). Não é uma tecnologia, mas um grupo de ferramentas utilizadas em conjunto para obter informações, analisar informações e prever resultados das possíveis soluções (Bose, 2009). *Business Analytics* é grupo de abordagens, procedimentos organizacionais e ferramentas utilizadas em conjunto para obtenção de informação, análise desta, e predição de resultados de possíveis soluções para problemas (Trkman, McCormack, Oliveira, & Ladeira, 2010), que também pode ser entendido como um campo emergente, com potencial para aumentar o gerenciamento de performance, melhorar o entendimento da dinâmica da empresa e melhorar a qualidade da tomada de decisão (Schläfke, Silvi, & Möller, 2012). Outros autores definem *Business Analytics* como o fenômeno de obter valor a partir de dados, chegando a ser comparado com uma nova onda econômica, como foi o petróleo em sua época (Acito & Khatri, 2014).

Goes (2014), em um texto editorial na *MIS Quarterly*, definiu uma série de conceitos com objetivo de criar uma convergência e permitir o avanço de estudos acadêmicos na área de *Big Data* e *Analytics*. Para o autor, dados geram informação, que geram conhecimento, que geram

inteligência. *Analytics* concentra-se em conhecimento e inteligência. *Analytics* reúne visualização, exploração, explanação e predição. São técnicas de *Analytics*: estatística, econometria, *Machine Learning*, otimização e simulação.

As definições de *Business Analytics* podem ser organizadas em 6 grandes grupos (Watson, 2011): um movimento (mudança de *mind-set* para orientação a problemas, seja individual, seja grupo, seja organizacional); um conjunto de práticas e tecnologias (como usar BA, sendo esse o grupo mais comum); um processo de transformação (com o principal assunto sendo como transformar evidências em *insights*); um conjunto de habilidades (organizacionais e individuais); atividades específicas (uma visão simplista de BA como um processo); um paradigma de decisão (uma nova forma de decidir). O autor destaca perceber uma preocupação com a conexão entre BA e tomada de decisão através de evidências e entende que todos os grupos de definições se interconectam.

Alguns autores consideram que o termo *Big Data* é uma evolução do termo *Analytics*, mas que continua sendo um movimento para obter inteligência partindo-se de dados, com a diferença de que agora tem-se mais volume e variedade e precisa-se de maior velocidade (McAfee & Brynjolfsson, 2012). Na literatura ainda se vê o termo *Analytics* sendo amplamente usado e há poucos sinais da adoção de *Big Data* como uma evolução de *Analytics*.

Outros autores exploram a similaridade entre *Data Science* e *Analytics*. *Data Science* é a aplicação de métodos quantitativos e qualitativos para solucionar problemas relevantes e prever resultados, o que é muito similar ao conceito de *Analytics* (Waller & Fawcett, 2013a). Chen, Chiang e Storey (2012) consideram que *Business Intelligence* e *Analytics* é *Data Science* no contexto de *Business*. Na literatura não há muito material com o termo *Data Science* no contexto de *Business*, mas há bastante material com o termo *Data scientist*. *Data scientist* é um profissional de alto nível com o treinamento e a curiosidade necessárias para fazer descobertas no mundo do *Big Data* (Davenport & Patil, 2012). Justamente por ser um dos elementos chave em IA, existem muitos estudos sobre as competências para atuação do *Data scientist*.

## 5.2 Aplicações

Na base de artigos estudada, o primeiro exemplo de aplicação de Inteligência Analítica (IA) apareceu em 2002. IA é usado em *e-commerce*, vendas, *Marketing*, *supply chain* e detecção de fraude (Kohavi, Rothleder, & Simoudis, 2002), gestão de performance corporativa; otimização de relacionamento com clientes, monitoramento de atividades de negócio, suporte a decisão; como pacote para resolver um problema específico; criação de *reports* (Willen, 2002). Através de uma riqueza de informações, IA pode aprofundar o conhecimento em práticas de *Marketing*. Para muitas empresas esse conhecimento é essencial e pode ser obtido através de análises históricas de transações e de análises sobre utilização de *internet* (Li, Tao, Cheng, & Zhao, 2015).

Davenport (2006) exemplifica o uso de IA em diversas áreas das empresas. Por exemplo, pode-se usar IA em *Supply chain* para otimizar fluxos e diminuir estoques; no relacionamento com clientes para identificar oportunidades e potenciais; na precificação; no uso inteligente do capital humano interno; no ganho de qualidade a partir dos processos internos; de performance *Financeira*; e para pesquisa e desenvolvimento. Outros autores discorrem sobre a integração do uso de IA nas áreas, tem-se a maioria de empresas utilizando IA em várias unidades de negócio, mas não na totalidade do negócio (Anderson, Daniels, McDonald, & Edvalson, 2014).

IA pode ser usado na análise de mídias sociais para extrair opiniões de consumidores, através do uso de técnicas de análise de sentimento (Pang & Lee, 2008), ou para estudar opiniões de consumidores presentes em *tweets* (Twitter), *likes* (Facebook) e análises de produtos em outros sites (Waller & Fawcett, 2013a). Essa realidade tem causado o aparecimento e consolidação de muitas empresas que ofertem análises desses dados e oferece a possibilidade que as empresas clientes respondam mais rapidamente a padrões percebidos.

Em um estudo sobre o impacto de IA em *Supply chain* são levantados alguns exemplos de utilização: analisar dados para prever tendências de produtos e serviços em um determinado mercado; programar ferramentas para seleção de fornecedores; produção dos produtos ou partes corretas, não considerando apenas o estoque óbvio; na tomada de decisão da parte logística (Trkman, McCormack, Oliveira, & Ladeira, 2010).

Chen, Chiang e Storey (2012) apresentam uma série de aplicações de IA que extrapolam o uso nas empresas: no governo público e na política, com ganho de transparência e participação; na área de ciência e tecnologia, com amplo ganho de informações através de sensores e instrumentos (exemplo oceanografia, astrofísica, etc); na saúde, com ganho nas frentes de genoma e no bom uso dos dados gerados pelos pacientes; e na segurança pública, com o uso de dados de múltiplas fontes e o uso de sistemas de inteligência. O intenso uso de tecnologias de comunicação, como a *internet*, telefones móveis, sensores e redes sociais está gerando uma quantidade massiva de dados (Wang, White, & Chen, 2015). Com a *internet* como meio preferido para as pessoas se comunicarem, jogarem, comprarem, buscarem, etc, empresas como o Google, Wikipédia, Amazon, eBay, YouTube e outras possuem dados de como essas pessoas se relacionam e suas ações rotineiras. As empresas pioneiras no uso de IA foram o Google, Amazon e o Facebook, que são negócio orientados à *internet* (Debortoli, Müller, & vom Brocke, 2014).

O termo *Analytics* geralmente está associado a uma área contexto, por exemplo *Business Analytics*, *Health Analytics*, *Safe Analytics*, etc. O entendimento do contexto cria uma diferenciação do propósito do uso, o que clareia quem usa, quais habilidades são necessárias e as tecnologias envolvidas (Van Barneveld, Arnold, & Campbell, 2012). Os conceitos de *Analytics* e *Business Analytics* são frequentemente confundidos, mas um motivo para a adoção do termo Inteligência Analítica (IA) para esse trabalho.

Chung, Chen e Nunamaker Jr (2005) apresentam um estudo voltado para a aplicação de ferramentas de *Web Analytics* (uma forma de IA) a partir de um *framework* visual por eles proposto. Esse *framework* baseia-se em testes matemáticos e teve validação estatística durante o estudo. A partir de sistemas e tecnologias internas pode-se obter palavras chave para busca na *internet* através de sistemas de busca. Essa busca deve gerar um grande volume de dados, que deve passar por uma análise de ocorrência, que é um grafo que mostra a relação entre os resultados. Segundo o estudo, a composição e validade desse grafo constituem o maior desafio. Por fim, algoritmos de visualização disponibilizam os resultados em árvores, partindo de assuntos mais genéricos para assuntos mais específicos. Essa utilização de IA permite que usuários tenham mais chances de encontrar a informação que realmente precisam.

Dentro do grupo da maioria das empresas brasileiras, hoje há predominância nas soluções de IA da SAP, Oracle e TOTVS (Meirelles, 2016). Essas soluções podem ser consideradas sistemas gerenciais, de decisão ou mesmo executivos, dependendo da sua utilização. Como as



soluções dos grandes fornecedores pode ser usada de diferentes maneiras, há dificuldade em manter-se uma taxonomia que seja amplamente aceita pelo mercado. O mesmo se observa nos mercados fora do Brasil: existem muitos pacotes de IA utilizados de diferentes maneiras, assim como existem milhares de soluções específicas, que as próprias empresas estão desenvolvendo internamente.

### 5.3 Requisitos e Desafios

Na literatura há quatro aspectos sobre requisitos e desafios na adoção e manutenção de Inteligência Analítica (IA) que se destacam. Esses aspectos são: pessoas, tecnologias, custos e práticas gerenciais, que nos textos parece muito próximo ao conceito de cultura organizacional. Há estudos que abordam apenas um desses aspectos, mas a maioria consiste em misturas, com diferentes percepções de importância para os respectivos autores. Pela difícil divisão dos aspectos, optou-se por organizar essa seção em uma sequência cronológica, mas sem divisão pelos diferentes aspectos. Ao final da seção estão as boas práticas para o início da adoção e, na Figura 3, um resumo gráfico dos requisitos e desafio de IA.

Alguns dos desafios relacionados a IA são reduzir tempo de ciclo de trabalho entre analistas de negócio e analistas de dados; reduzir a necessidade de ser um *expert* para analisar dados; alinhar expectativas entre objetivos do negócio e o que os dados podem responder; selecionar e analisar os dados relevantes para cada desafio de negócio; distribuir resultados de análises de dados de uma forma menos técnica; integrar dados de diferentes fontes e demonstrar retorno na performance do negócio (Kohavi, Rothleder, & Simoudis, 2002). Mesmo com muitos avanços nas ferramentas de IA, analistas de negócio raramente são também especialistas em análise de dados. Para tomar decisões, os analistas de negócio dependem do trabalho de analistas de dados, que por sua vez demoram certo tempo para comunicar resultados de análises, que muitas vezes são questionáveis e geram novos ciclos de trabalho para todos os envolvidos. Outro desafio é garantir a segurança da informação nos diferentes níveis organizacionais (Negash, 2004; Boyd & Crawford, 2012).

Em uma linha parecida, outros autores destacam os custos e esforços necessários para manter os acessos corretos às novas informações disponíveis em determinado cenário de IA, a

dificuldade em manter a qualidade dos dados e o reuso eficiente de conhecimento gerado (Baars & Kemper, 2008). Os autores demonstram-se mais preocupados com o aspecto do uso correto das informações, pelas pessoas corretas, mas também falam sobre os custos: a adoção de IA tem custos relacionados a *hardware*, *software*, implementação (adaptação de soluções para a realidade da empresa), e de profissionais para executar e dar suporte ao IA e a infraestrutura de TI (Negash, 2004).

Bose (2009), em um estudo avançado sobre desafios de IA, reforça que a extração e integração dos dados são a base. Para esse autor, a extração é a capacidade de capturar informações implícitas nas bases de dados, definir um modo de exibição e interpretar para transformar em ação para um problema de negócio. Ainda para ele, quanto mais integração houver, maior a tendência de reconhecimento de padrões e relacionamentos, o que por sua vez tende a melhorar a tomada de decisão. Da base de artigos estudada, nenhum outro autor citou a palavra integração antes de Bose. Esse autor explica que para que o IA funcione bem do ponto de vista gerencial é necessário alinhamento entre a infraestrutura de TI e os objetivos do negócio, de forma a haver um fluxo contínuo de informações para apoio à tomada de decisão. Bose também discorre sobre o desafio de implementar IA em uma velocidade compatível com cultura e tamanho da alteração nos processos internos, definir regras de uso e acesso à informação, dominar as tecnologias necessárias e o compartilhamento de informações dentro da empresa.

O bom aproveitamento do IA depende de dois outros fatores que podem viabilizar ou inviabilizar a revolução dentro da empresa (Trkman, McCormack, Oliveira, & Ladeira, 2010). O primeiro fator é a orientação a processos. Se uma empresa possui processos integrados e já age baseando-se nisso, esta realidade tem um efeito multiplicador nas possibilidades oferecidas por IA. O segundo fator é o alinhamento entre TI e negócio (Bose, 2009; Trkman, McCormack, Oliveira, & Ladeira, 2010).

Uma dificuldade ao lidar com IA é entender do que isso se trata (Watson, 2011). Há uma grande quantidade de textos, livros, artigos, etc, que falam sobre o tema, tornando-o um "tema quente". Por outro lado, há muitos entendimentos errados, imprecisos ou incompletos, pois IA não é apenas uma palavra ou uma tendência. Nessa questão, pode-se considerar que essa revisão de literatura se faz necessária para melhor entendimento do fenômeno, antes de levantar as competências para atuação.

Através de uma *survey* com mais de 4.500 executivos respondentes, descobriu-se que os ganhos relacionados a IA não dependem apenas da obtenção das ferramentas adequadas e dos profissionais adequados, mas também há outras variáveis organizacionais importantes acerca da cultura a empresa, como aceitação do IA em todos os níveis gerenciais, uma mentalidade aberta a novas ideias e foco em decisões com informações (e menos intuitivas) (Kiron & Shockley, 2011; LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011).

A partir de 2012, observa-se mais textos sobre o aspecto humano dos desafios de adoção e manutenção de IA nas empresas. Davenport e Patil (2012), no artigo mais citado sobre o aspecto humano de IA, discutem a importância de profissionais qualificados para atuação em IA. Esses autores destacam que a falta de *Data Scientists* pode ser um dos maiores limitantes dessa nova onda, e que não há programas universitários para formação.

Poucos autores destacam o problema da qualidade dos dados utilizados para tomada de decisão. Muito se fala sobre tecnologias e algoritmos que fazem parte do processo cujo produto final é o apoio a decisão, mas pouco se fala sobre a qualidade básica dos dados, no que diz respeito à veracidade, integridade, etc, e a não consideração desse aspecto pode causar danos para organizações em diferentes aspectos, como relacionamento com clientes, fornecedores e até mesmo insatisfação dos funcionários (Baars & Kemper, 2008; Hazen, Boone, Ezell, & Jones-Farmer, 2014).

Em um estudo no estado de Utah, EUA, mapeou-se que as maiores dificuldades das empresas são: a ausência de uma solução de IA que sirva para todos os problemas organizacionais, problemas com inconsistência e baixa qualidade de dados, pouco conhecimento organizacional de IA, funcionários pouco qualificados em IA, falta de acesso aos dados corretos, funcionários com baixo entendimento do potencial de IA, gerentes muito ocupados para adotarem IA, alto custo de implantação e dificuldade em usar e customizar ferramentas de IA (Anderson, Daniels, McDonald, & Edvalson, 2014). Estudos como esse confirmam quantitativamente o que se definiu na literatura ao longo de muitos anos.

As organizações que têm consciência sobre o nível de adoção de IA que se encontram, têm mais chances em transformar desafios em oportunidades (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). Existem cinco recomendações para que IA gere valor para o negócio: começar pelos grandes desafios, que estão mais ligados à estratégia global da empresa e que têm mais atenção da alta gerência; começar pela pergunta a ser respondida e depois verificar

como os dados disponíveis podem respondê-la; não usar IA como substituto aos processos atuais, mas como complementar, com o objetivo de criar ramificações aos poucos na cultura da empresa; planejar que dados serão importantes no futuro para começar a criar um histórico hoje (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). Essa lista pode ser complementada com a necessidade de conscientizar os funcionários da empresa e em obter alinhamento (Barton & Court, 2012). Outros autores citam os mesmos pontos e ainda destacam que, para uma transição completa, também são necessários: liderança adequada, gestão de talentos considerando os novos conceitos e tecnologias, as novas tecnologias em si (adoção e bom uso) e a adequação da cultura organizacional (Mcafee & Brynjolfsson, 2012).

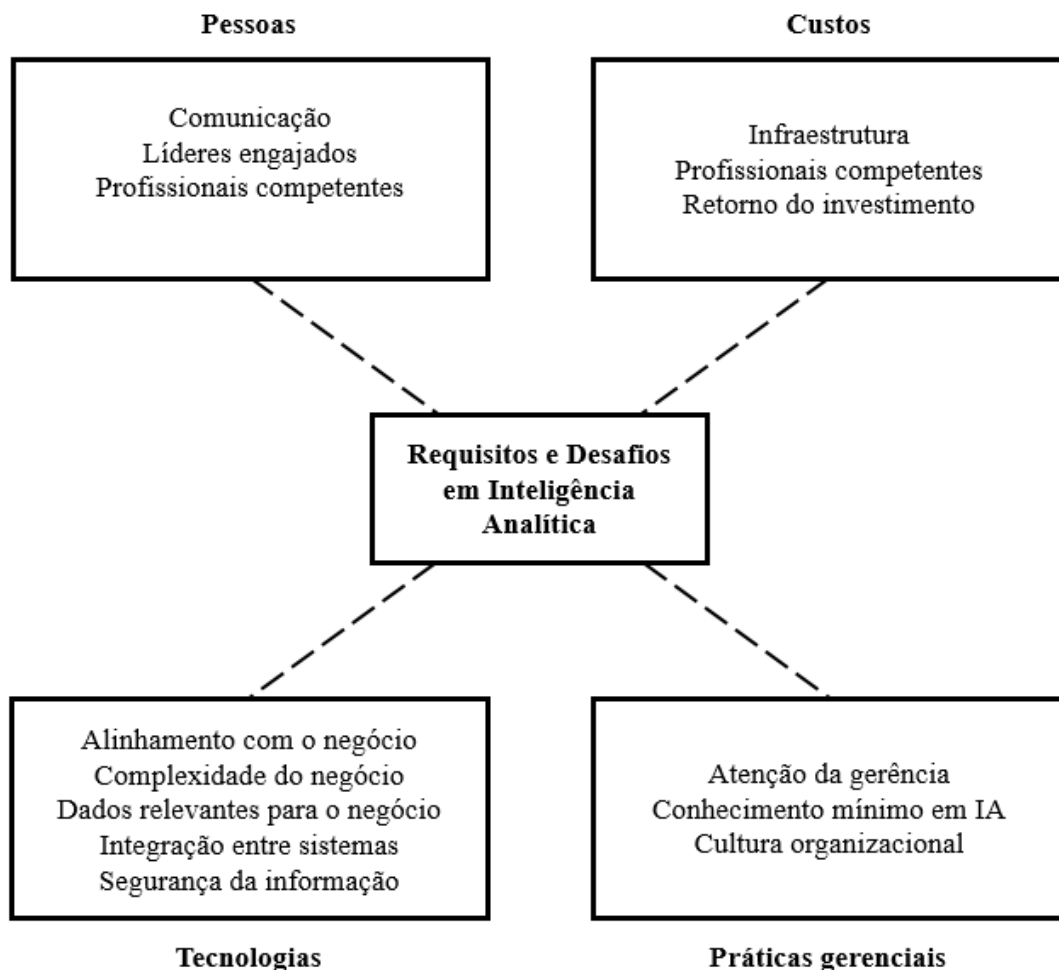


Figura 3: Resumo dos Requisitos e Desafios de IA  
Fonte: Elaboração própria

Em uma *survey* com mais de 4.500 executivos respondentes, revelou-se quantitativamente que há predominância de duas abordagens de adoção de IA nas empresas que estão obtendo sucesso nesse fenômeno: aquelas que iniciaram com uma adaptação na cultura gerencial, para suportar o uso efetivo de informações em todos os níveis da empresa; e aquelas que iniciaram com ganho de competências analíticas de forma mais técnica e tentam disponibilizar inteligência para os demais níveis da empresa (Kiron & Shockley, 2011). São abordagens bastante distintas, mas os elementos relacionados são os mesmos, a realidade não muda, apenas a priorização.

#### 5.4 Benefícios e Desempenho

Esta seção é dedicada a entender os potenciais benefícios que Inteligência Analítica (IA) pode oferecer a uma empresa. Sobre os benefícios, percebeu-se que não há muito material e que a maioria das indicações são em nível estratégico ou organizacional. Complementarmente, também há uma revisão da literatura sobre o desempenho de IA, ou seja, a mensuração do que se obteve dos potenciais benefícios. Sobre o desempenho, há convergência na literatura de que cada empresa tem um cenário bastante particular, o que dificulta a criação e utilização de ferramentas de mensuração inteligente. O mesmo se observa para a mensuração de benefícios resultantes de investimentos de TI como um todo. Sanchez & Albertin (2009) elaboraram um Modelo de Análise do uso de Tecnologia da Informação no nível da Firma, o qual pode ser utilizado para investigar benefícios obtidos *versus* benefícios esperados em projetos de TI. O modelo não foi testado para a adoção e uso de IA.

A maioria dos benefícios de IA são intangíveis antes de acontecerem e a maioria das empresas não consideram a diminuição de custos e de tempo de operação como benefícios de IA, embora geralmente existam. A expectativa geral é que o IA resulte, em longo prazo, em uma revolução para a empresa, mas geralmente não há garantias nem métricas (Negash, 2004).

Davenport (2006) relata que há um grande número de empresas que se beneficiam de IA como forma de obter vantagem competitiva, e que algumas delas foram criadas com base nesse fenômeno. Empresas tão imersas em IA não apenas conseguem saber o que os clientes querem, mas quanto aceitam pagar, e até quantos itens serão necessários ao longo de suas vidas. IA torna possível a descoberta de inter-relacionamentos ocultos, que permite uma gestão mais eficaz, e

o aprendizado de maneiras interessantes de armazenamento e distribuição de informações (Baars & Kemper, 2008).

Inteligência Analítica (IA) está no topo das agendas corporativas, pois podem transformar a maioria dos negócios, como não se via desde a década de 90, quando as organizações redesenharam os processos centrais. No entanto, a maioria dos líderes não sabe como proceder para aproveitar esses potenciais benefícios (Barton & Court, 2012). A tomada de decisão direcionada por dados é simplesmente melhor do que a decisão orientada a intuição, por isso IA tem o potencial de revolucionar a forma de gerenciar (Mcafee & Brynjolfsson, 2012). Um gerenciamento de performance mais analítico pode ser utilizado para identificar e provar relacionamentos quantitativos entre variáveis de contexto, entradas, processos, saídas e resultados (Schläfke, Silvi, & Möller, 2012).

Através de uma extensa revisão de literatura, autores resumizam tipos de valor criados para as empresas com IA: criação de transparência, possibilidade de experimentação para descobrir necessidades, melhoria de performance, segmentação de públicos para ações mais direcionadas, apoio a decisão com algoritmos automatizados e inovação em produtos, serviços e modelos de negócio (Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015).

A adoção de IA em *Cloud* permite que as empresas tenham menos preocupação com escalabilidade e escopo, já que esses dois direcionadores são terceirizados (Demirkan & Delen, 2013). Porém, ainda não se sabe como será estabelecida cobrança por esses serviços. Esse benefício em *Cloud* não é exclusivo do IA, mas é interessante citar que já existem iniciativas desse tipo. Em 2015 foi publicado um estudo sobre adoção de *cloud*. Este estudo revela que a adoção do fenômeno *cloud* não está associado a existência prévia de grandes infraestruturas de TI, o que indica que o fenômeno é adotado em empresas de todos os portes (Cappellozza & Moraes, 2015)

O desempenho em IA pode ser visto como forma de medir a realização dos potenciais benefícios. Em 2002 autores propuseram que as soluções de IA sempre estiveram ligadas à geração de valor para o negócio, onde o valor é medido em termos de interligar as necessidades do analista de negócio e a acessibilidade e usabilidade de sistemas de IA (Kohavi, Rothleder, & Simoudis, 2002). Os autores não definiram mais detalhes de como isso poderia ser feito, ou estudado. Quase dez anos depois, outros autores alegam que ainda é difícil medir o impacto de IA na performance das empresas, pois o grau de adoção em cada área possível ainda é muito

heterogêneo, dificultando conclusões (Trkman, McCormack, Oliveira, & Ladeira, 2010). Uma alternativa para mensuração é definir indicadores de tecnologia na percepção de TI e gestores de negócio. Indicadores permitem comparações dentro da empresa e fora da empresa, entre competidores (Moraes, 2015). Com mais informações históricas, potencialmente há mais chance de perceber a geração de valor de IA para o negócio.

O uso inteligente da grande quantidade de dados disponível para as empresas é algo que parece ter grande influência na performance. Empresas com alta performance tem maior probabilidade de estarem envolvidas no uso de dados para melhorar a operação e a tomada de decisão (Hopkins, Lavalle, & Balboni, 2010). Mais uma vez percebe-se ausência de detalhamento. Apesar das novas tecnologias que surgiram para IA, o desafio de como obter valor para o negócio permanece como uma questão central (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011).

Em um *survey* conduzido pelo MIT Sloan Management Review e pela IBM Institute for Business Value, com mais de 3000 executivos de negócio em mais de 100 países, revelou-se que empresas com alta performance usam IA cinco vezes mais que as de baixa performance (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). Nesta mesma pesquisa revelou-se que existe uma ampla crença de que IA resulte em valor para o negócio e 61% acredita que as respectivas organizações não conseguem usar grande parte dos dados disponíveis e que a inovação através do IA pode ser chave para tornar as empresas positivamente diferenciadas em termos de competição. Executivos seniores querem que as decisões de negócio sejam orientadas à dados, com simulações que permitam guiar ações em novos cenários. Para isso, existe o pressuposto de que o uso de IA esteja conectado com a estratégia da empresa em todos os níveis hierárquicos. Através de outro *survey* com mais de 4.500 executivos de 122 países, concluiu-se que empresas que usam IA estão ganhando vantagem competitiva, mas as abordagens e utilizações são variadas (Kiron & Shockley, 2011).

Dr. David Kreutter, vice-presidente de operações comerciais da empresa Pfizer, expõe que IA não pode ser encarada como um investimento comum, onde a área financeira verifica se há retorno financeiro no investimento (Kiron & Shockley, 2011). Para ele, o correto seria criar métricas para entender o valor criado por IA e, se estiver abaixo do esperado, planejar mudanças para novas tentativas através de experimentação, evidenciando que IA é algo sem volta para as grandes empresas. Os profissionais mais céticos têm dificuldade de acreditar que o uso de IA melhora a performance da empresa, quando não há dados e ferramentas para mostrar isso. A

verdade é que esse assunto ainda não tem sido abordado de forma rigorosa (McAfee & Brynjolfsson, 2012).

Em um estudo com 330 empresas norte-americanas, levantou-se que nem todas as empresas tomam decisões a partir de dados (McAfee & Brynjolfsson, 2012). A pesquisa também mostra que as que o fazem possuem melhor performance em termos financeiros e operacionais estatisticamente significantes e refletindo na valorização das ações das empresas. Esses resultados mostram uma transformação fundamental na economia e que quase nenhuma esfera de negócio permanecerá intacta ao fenômeno.

A eficácia de um profissional que atua com IA pode ser medida pela quantidade de oportunidades reais que eles descobrem (Waller & Fawcett, 2013a). Essa é uma ideia inovadora na literatura e pode ser o início de uma nova linha de estudos.

Em 2014, Acito e Khatri (2014) discutem uma questão levantada por Neely (1999) quase vinte anos antes: como medir a performance do negócio e que variáveis devem influenciar. Acito e Khatri destacam que o desenvolvimento do IA ao longo dos anos alterou a dinâmica da mensuração de resultados, e que ainda não há respostas que independem de contextos, confirmando explicitamente que não há uma resposta para essa questão.

Resumidamente, os principais benefícios e as formas de mensuração encontradas na literatura estão resumidos na Figura 4.



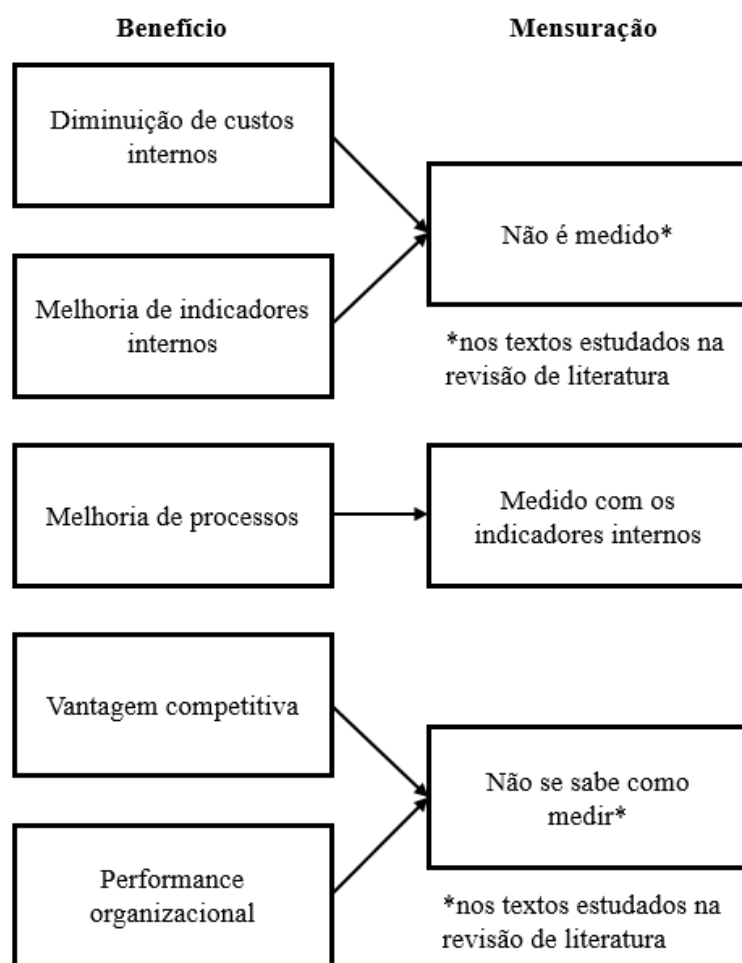


Figura 4: Principais benefícios e as formas de mensuração encontradas na literatura  
Fonte: Elaboração própria

## 5.5 Arquitetura

Esta seção visa relacionar conceitos, tecnologias e áreas da empresa envolvidos na adoção e manutenção de Inteligência Analítica (IA). Conforme seções anteriores, a grande variedade de termos comerciais e não comerciais (que muitas vezes representam um mesmo fenômeno) dificulta a criação de uma arquitetura única, que possa explicar o que é necessário para uso de IA. Ao comparar-se as dezenas de arquiteturas propostas na literatura, torna-se evidente que o fenômeno é complexo e pode ser visto de várias perspectivas, o que não invalida cada um dos estudos, mas torna-os complementares, como já afirmavam Bolman e Deal (1991), grandes estudiosos das perspectivas dos fenômenos organizacionais, mais de vinte anos atrás. Por

tratarem de diferentes perspectivas do fenômeno IA, muitas das arquiteturas propostas na literatura não compatibilizam entre si, pois não faz sentido representar diferentes aspectos em um mesmo esquema.

Em 2004, reportou-se que na década de 90, os sistemas de IA eram basicamente constituídos pelo que hoje entende-se como *Data Warehouse* (Golfarelli, Rizzi, & Cella, 2004). Na Figura 5 observa-se como era esse esquema: a estratégia da empresa deveria ser traduzida em indicadores, que eram usados para medir e equilibrar processos do nível operacional, tático e estratégico. Parte das informações das medidas dos indicadores eram usados nos próprios níveis da empresa para ajudar na tomada de decisão, mas o maior intuito era alimentar sistemas de informação, que abasteciam as chamadas *Data Warehouses*, que por fim eram consultadas para entendimento e criação de novas estratégias. É importante perceber que esse modelo oferecia informações *bottom-up*, uma vez que o nível estratégico não solicitava informações, mas apenas as recebia. A coleta desses dados nos diferentes níveis hierárquicos da empresa tornou-se tão complexa que surgiu a necessidade de automatizar partes das fases de coleta e análise de dados, resultando em uma complexidade denominada (*Business*) *Analytics*, que contempla o uso de informações externas e permite que o negócio demande inteligência a qualquer momento (Kohavi, Rothleder, & Simoudis, 2002), o que torna o esquema da Figura 5 desatualizado. Em 2008 já se considerava que decisões corretas e relevantes através de grandes volumes de dados (internos e externos) só seria possível através de IA (Sahay & Ranjan, 2008). Para Bose (2009) a IA passou por cinco fases evolucionárias: transformação e integração de dados de múltiplas fontes; relatórios em tempo real; informação entregue no momento certo para a pessoa certa; *Analytics* preditivo; e tomada de decisão para vantagem competitiva. Para suportar essa evolução, o BI passou de um foco operacional para um foco mais analítico, mas ainda suportando os níveis operacional, tático e estratégico.

BI consome informações de muitos outros sistemas de informação, como OLAP (*on-line Data processing*), CRM (*customer relationship management*), DSS (*Decision Support Systems*) e GIS (*geographic information systems*), entre outros (Negash, 2004). Conforme Figura 6, dados são coletados de fontes externas e internas, movimentados através de técnicas de ETL até servidores de *Data Warehouse*, onde são tratados por métodos matemáticos, estatísticos e computacionais para se encaixarem em banco de dados relacionais e poderem ser localizados através de técnicas de *MapReduce*, *Data Mining*, entre outras, para então serem exibidos em aplicações do tipo *front-end*, como planilhas, *softwares* de visualização, etc (Chaudhuri, Dayal, & Narasayya, 2011). Os autores destacam que a lógica pode parecer simples, mas que a falta

de qualidade e padrão nos dados cria inúmeros problemas, assim como cada tipo de tratamento pode levar a diferentes resultados.

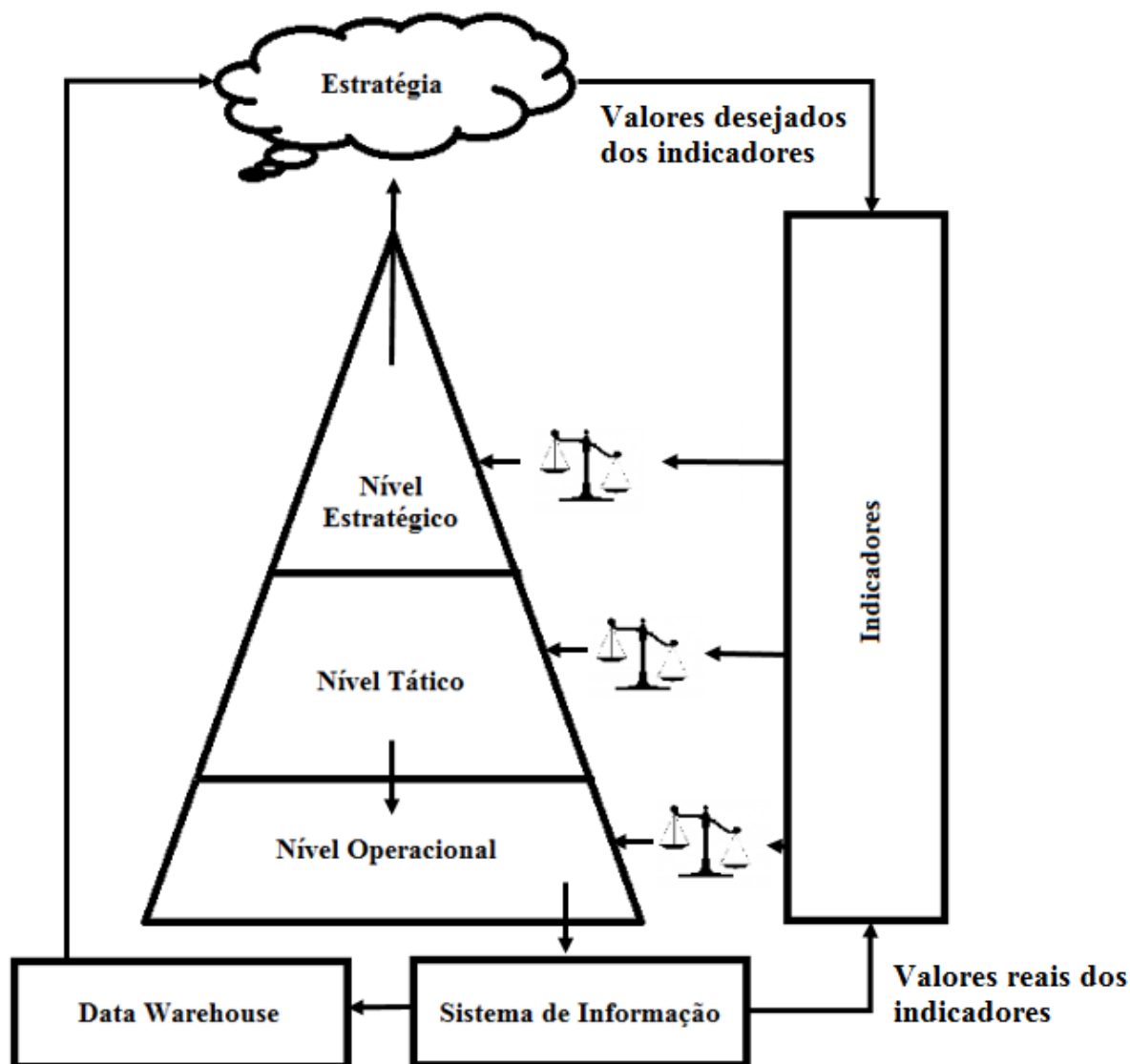


Figura 5: Arquitetura de IA na década de 90  
 Fonte: Traduzido de Golfarelli, Rizzi, & Cella, 2004

A possibilidade de coletar informações específicas, geolocalizadas, contextualizadas e altamente personalizadas através de dispositivos móveis, criou novas oportunidades para IA avançado e inovador (Chen, Chiang, & Storey, 2012). Para tratar essas informações são necessárias as técnicas *Mining*. *Data Mining* é a automatização da extração de informação preditiva não explícita (Bose, 2009). Chen, Chiang e Storey (2012) defendem que novos

algoritmos para otimização e *Machine Learning* contribuíram com o avanço no campo de *Data Mining*. Há duas modalidades de *Data Mining* que são importantes para as empresas: *Text Mining* que trata de obter padrões em resposta de *e-mails*, por exemplo; e *Web Mining* que trata de obter padrões em servidores *Web*. Considerar dados obtidos na *Web* não é tarefa fácil dentro do conceito de *Analytics*, pois não é apenas uma nova fonte de dados: há uma série de entendimentos e dificuldades particulares, conforme pode-se observar no estudo de Ning Zhong (Zhong, 2003): deve-se entender os aspectos semânticos, as redes de relacionamento social, a estrutura para aplicação de tecnologias *Mining*, entre outros.

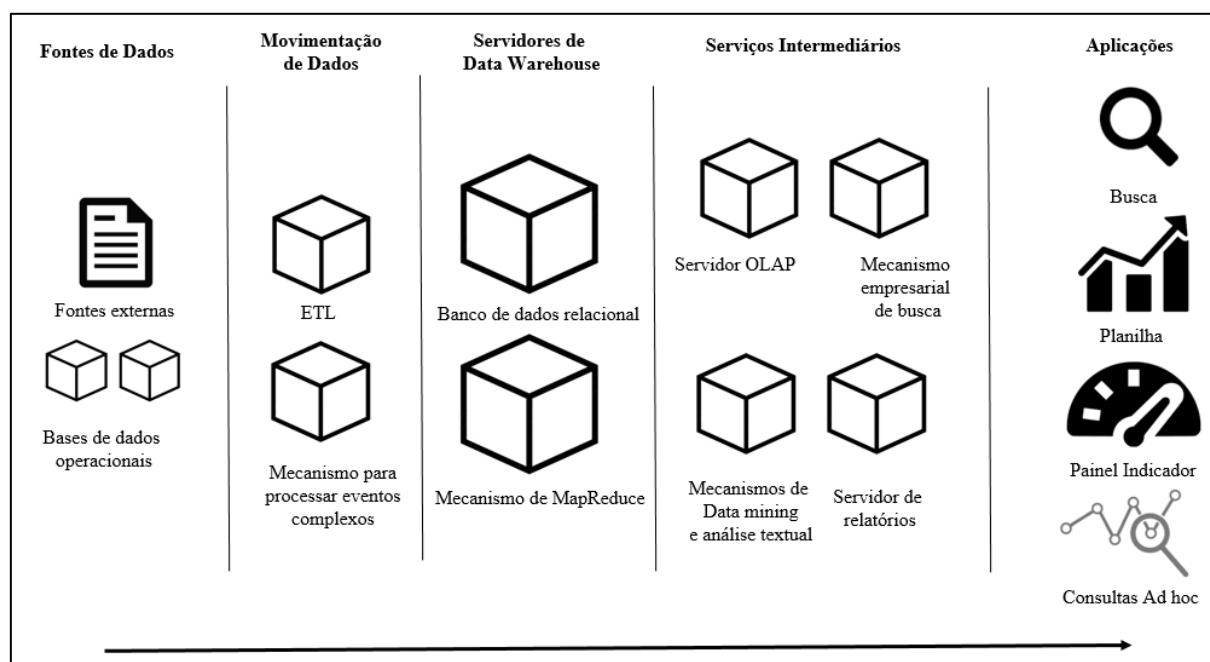


Figura 6: Típica arquitetura de IA

Fonte: Traduzido e adaptado de Chaudhuri, Dayal, & Narasayya, 2011

Inspirado no MapReduce, o *Hadoop* é um *framework* para processamento distribuído e análise massiva de dados. Os três maiores fornecedores de soluções (Oracle, IBM e Microsoft) adotaram o *Hadoop*, alguns poucos com infraestrutura em *Cloud* (Chen, Chiang, & Storey, 2012).

Existem dados estruturados e semiestruturados. Os dados estruturados são arquivos relacionais, onde pode-se organizar a informação no esquema clássico de colunas. Os dados semiestruturados são aqueles que tem algum tipo de padrão, mas que são constituídos de partes

que não podem ser automaticamente classificadas, como *e-mails*, gráficos, imagens, cartas, filmes, apresentações, pesquisas, planilhas de usuários finais, páginas de *internet* e documentos textuais (Negash, 2004). Em 2004 mais de 85% dos dados dentro de uma empresa eram semiestruturados. Em 2005 estimou-se que 90% das informações de uma empresa não eram registradas ou revisitadas, por exemplo cartas, *e-mails*, gravações de ligações, contratos, patentes, etc (Weiss, Indurkha, Zhang, & Damerau, 2005). Esse número provavelmente mudou após a disseminação de técnicas de *Text Mining*.

Baars e Kemper (2008) definem um *framework* que é também uma arquitetura IA composta de três camadas: camada de dados, camada lógica e camada de acesso, conforme Figura 7. A camada de dados é composta de todos os sistemas da empresa e de todos os dados por eles gerados ou armazenados. A camada lógica, mais abstrata, contém sistemas para análise, compilação e processamento dos dados da camada inferior e práticas de como distribuir conhecimento gerado. A camada de acesso contém portais de acesso para usuários consumirem os dados disponíveis pelas soluções de IA. Percebe-se que a principal diferença em relação ao modelo de Chaudhuri, Dayal e Narasayya (2011) é a preocupação com a distribuição de conhecimento.

De acordo com uma pesquisa Gartner há informação de que o mercado de *Data Warehouse* e de soluções de IA deveria alcançar US\$10.8 bilhões em 2011 (Moore, 2011). Empresas como Amazon Web Services, AT&T's Synaptic Hosting, AppNexus, GoGrid, Rackspace Cloud Hosting, HP/Yahoo/Intel Cloud Computing Testbed, IBM/ Google e MicroStrategy BI Cloud oferecem vários tipos de serviços em *cloud* para amenizar as dificuldades com armazenamento de dados que as empresas possam ter (Demirkan & Delen, 2013), entre eles o acesso a máquinas não configuradas que podem ser utilizadas para sistemas de apoio a decisão. Geralmente as soluções *cloud* abrangem um ou mais dos seguintes componentes: sistemas operacionais, *Data Warehouse*, *Analytics online* ou sistemas para usuário final. Arquitetura *cloud* permite um conveniente acesso remoto para um conjunto de recursos computacionais configuráveis que podem rapidamente ser aumentados com pouca interação com o provedor (Mell & Grance, 2009). Entre 2004 e 2010 foram publicados artigos sobre os riscos da adoção de soluções *cloud*, e após 2010 artigos sobre os benefícios de *cloud*, o que indica o amadurecimento do fenômeno, com um histórico de receio gerencial e uma evolução para uma visão mais otimista (Vieira & Meirelles, 2015).

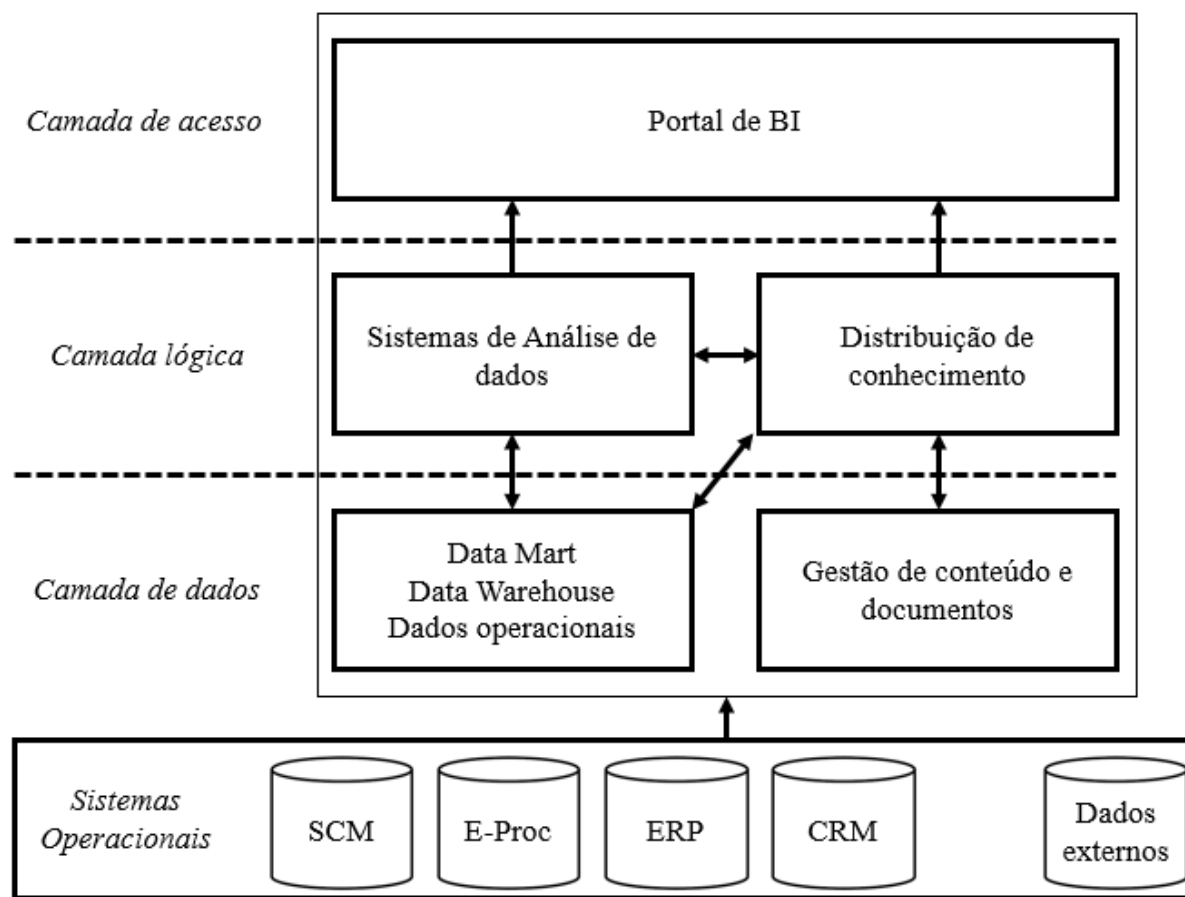


Figura 7: *Business Intelligence Framework*

Fonte: Traduzido e adaptado de Baars & Kemper, 2008

## 5.6 Os profissionais que atuam com Inteligência Analítica

Existem duas linhas observadas na literatura. A primeira delas, defendida por um número menor de pesquisadores, esclarece que os sistemas de Inteligência Analítica (IA) devem evoluir o suficiente para serem utilizados por analistas de negócio. A outra linha, defendida pela maioria dos pesquisadores, acredita que é necessário formar e treinar profissionais cada vez mais completos, com um conjunto de competências em tecnologias, técnicas e negócio. Não há convergência na literatura sobre onde tais profissionais devem ser inseridos nas empresas, e também não fica claro quanta autonomia devem ter. Diferentes estudos utilizam diferentes nomes para esses profissionais, como analista de dados, *Data scientist* (com destaque especial), etc. Questiona-se se a melhor opção seria ter profissionais com mais habilidades nas duas esferas (*Analytics* e negócio), ou ter sistemas mais simples e que permitam usuários de negócio

resolverem seus próprios problemas (Bose, 2009). Pesquisas mostram que o maior limitante para IA hoje é a falta de mão de obra qualificada, com as habilidades necessárias para aproveitar os potenciais benefícios (Russom, 2011).

Em 2004, Negash, defensor da primeira linha observada na literatura, reporta o desenvolvimento de soluções de BI para ser utilizado por todos (Negash, 2004). Após décadas do clássico IA, na qual apenas um indivíduo obtém e investiga um conjunto de dados de forma parcial e depois repassa as conclusões a quem vai tomar a decisão (independentemente do nível hierárquico), percebeu-se que talvez existam vantagens na adoção de sistemas de IA para massas, que permitam que os decisores introduzam perguntas e obtenham respostas com menor subjetividade e ao mesmo tempo com menor domínio de técnicas avançadas de estatística e programação. Em um estudo também em 2004, visualizou-se que a maioria dos usuários dos sistemas de IA não teriam tempo ou habilidades necessárias para operar sistemas OLAP, de forma a fazer-se necessária uma tradução sistêmica em relatórios e indicadores conhecidos (Golfarelli, Rizzi, & Cella, 2004).

Davenport, um defensor da segunda linha, acredita que para atuação em IA são necessários analistas capazes de conduzir experimentos quantitativos, com habilidades em modelagem e capacidade de aprender rapidamente a utilizar novas tecnologias, e também conhecimentos de áreas de negócio (Davenport, 2006). Para os funcionários que ainda não tem esse perfil, recomenda-se extensivo treinamento para que saibam que dados existem e que análises podem ser realizadas, quais são os problemas de falta de dados, duplicações e problemas de qualidade que podem ocorrer. Na mesma linha, define-se que IA é *Data Science* no contexto de *Business*, por isso chama-se os profissionais atuantes de *Data scientists* (Chen, Chiang, & Storey, 2012). Esses autores reportam que atualmente há óbvia falta de profissionais com o conhecimento necessário, que é considerado profundo. A formação de profissionais de IA deve ser interdisciplinar e deve abranger: habilidades com IA e TI, incluindo tópicos como o estudo estatístico e soluções de tecnologia para tratar dados estruturados e não estruturados; conhecimento apropriado do negócio, para garantir entendimento das estratégias da empresa e prover informações com maior qualidade; e a habilidade de comunicar descobertas para os especialistas de negócio da empresa (Chen, Chiang, & Storey, 2012).

Uma pesquisa realizada pela IBM em 2010 (IBM, 2010) sugere que decisores precisam adquirir novas habilidades (matemáticas, estatísticas, econômicas e de tecnologia) para aproveitarem os benefícios de IA. O McKinsey Global Institute (Manyika et al., 2011) prevê que até 2018 os

EUA terão demanda de 140mil a 190mil profissionais com habilidades de IA avançado e de 1,5 milhões de gestores capazes de analisar dados para tomada de decisão eficaz. Estes números mostram a importância de se entender mais sobre o conjunto de habilidades e conhecimentos necessários para que um profissional atue com BA de forma desenvolta, com geração de valor para o negócio. Através de uma grande *survey*, descobriu-se que a ordem de aquisição de habilidades de IA para casos de sucesso foram heterogêneas, dificultando a criação de um padrão de boas práticas (Kiron & Shockley, 2011).

Os participantes de IA dentro de uma empresa podem envolver-se em três tarefas: produzir e/ou gerar *insights* ou permitirem a criação de *insights* (Chandler, Hostmann, Rayner, & Herschel, 2011). Essas tarefas podem ser realizadas em diferentes papéis e áreas das empresas. Quem produz *insights* são analistas com foco na condução de análises, assim como quem consome *insights* deve tomar decisões. Os profissionais que permitem a criação de *insights* precisam de habilidades de tecnologia e de análise de dados, como: buscas difusas, estatística com inferências, estatística preditiva, simulação e otimização, predição e prescrição.

*Data scientists* precisam de autonomia para produzirem resultados, mas também precisam fazer uma ponte entre os interessados nos resultados e os dados disponíveis. Existe pouco consenso em que nível hierárquico um *Data scientist* deve ser encaixado na empresa, como extrair o máximo de valor do trabalho deles e como medir a performance dos indivíduos (Davenport & Patil, 2012). Hal Varian, o economista chefe da empresa Google, cita o trabalho do *Data scientist* como o mais *sexy* da década, similar ao que houve na década de 90 com os engenheiros de computação que eram procurados para trabalhar com algoritmos de busca, hoje já ensinados nas escolas de tecnologia.

Um *Data scientist* certamente irá se deparar com limitações técnicas e terá que contorná-las para continuar sua busca por soluções de negócio no universo dos dados (Davenport & Patil, 2012). Muitas vezes utilizam sua própria estratégia, de forma quase artística, que pode incluir um estilo mais acadêmico. A maioria dos *Data scientists* atuais são formados em Ciência da Computação, Matemática ou Economia, e depois adquirem conhecimentos em negócio. O professor Jeff Stanton da Syracuse University foi citado por Dumbill (Dumbill, Liddy, Stanton, Mueller, & Farnham, 2013) que diz que do ponto de vista educacional, pode-se ensinar testes estatísticos e pequenas habilidades de programação em menos de uma hora, mas não se pode ensinar com facilidade o conhecimento sobre o negócio. Assim, é isso que diferencia um



profissional de *Data Science*. O conhecimento no negócio é extremamente importante pois evita que o profissional aceite falsos positivos estatísticos (Waller & Fawcett, 2013a).

Em uma entrevista, Davenport e Patil (2012) obtiveram respostas de como um caçador de talentos de uma grande empresa seleciona profissionais para atuação em IA. Em primeiro lugar, esse caçador de talentos procura profissionais que saibam programar um protótipo exigido e mostrem desenvoltura com o uso de tecnologia. Depois, verifica se o profissional tem base em estatística, probabilidade e ciência da computação. Também procura profissionais com desenvoltura ao falar no setor de atuação do negócio e que tenham empatia por clientes. Esse conjunto de competências não foge muito do modelo de Chen, Chiang e Storey (2012). Muitas empresas estão tentando desenvolver essas habilidades internamente em vez de tentar encontrar profissionais disponíveis no mercado.

*Data scientists* de sucesso devem enxergar problemas do negócio na perspectiva de dados (Waller & Fawcett, 2013a). Desenvolver habilidades quantitativas requer muito tempo e dedicação, além da experiência, assim como desenvolver conhecimento do negócio. Por isso, em condições normais é muito difícil encontrar profissionais com tudo que se deseja encontrar em um *Data scientist*. Muitos métodos estatísticos aplicados são necessários, mas isso não significa que não há necessidade de conhecimento teórico em estatística. As seguintes habilidades estatísticas são consideradas as mais importantes para um *Data scientist*: estatística básica, como o conhecimento de métodos de estimação e amostragem; predição, como o entendimento da utilização de métodos preditivos; otimização, como o uso adequado de métodos de otimização; simulação discreta, como a capacidade de aplicar métodos desse tipo; probabilidade aplicada, como forma de estimar valores de variáveis aleatórias; e o uso de modelagem analítica, como o uso de métodos numéricos para estimar funções (Waller & Fawcett, 2013a). Destaca-se que todo esse conjunto de habilidades ainda deve ser confirmado com pesquisas quantitativas.

Para Jeff Stanton, professor da School of Information Studies na Syracuse University, o mais importante para um *Data scientist* é entender os conceitos e a maneira analítica de pensar (Dumbill, Liddy, Stanton, Mueller, & Farnham, 2013). Depois disso, um *Data scientist* pode se adaptar a outras ferramentas e linguagens, seja em contexto de negócios, seja em outros contextos. Um *Data scientist* precisa amar aprender e ter uma forte automotivação. O professor também revela um desejo de maior comunicação com os gerentes que contrariaram funcionários para atuarem com IA, pois para ele os acadêmicos entendem melhor quais são as habilidades e

conhecimentos necessários para que os funcionários tenham mais chances de obterem bons resultados nas empresas. Para o professor, um único *Data scientist* pode obter mais êxito do que uma grande equipe de analistas sem as habilidades adequadas.

Luvizan, Meirelles e Diniz (2014) levantaram a formação de capital humano como um dos principais temas de interesse para estudos futuros após o ano de 2013. De fato, encontra-se bastante material recente na literatura, com crescente aceitação da falta de profissionais com o conjunto adequado de competências para atuação. Ao mesmo tempo, há divergência sobre quais são essas competências, o que mostra o quanto a área ainda precisa amadurecer.

Paulo Goes, editor chefe da *MIS Quarterly*, destaca que em 2014 percebeu-se uma crescente demanda por *Data scientists* e uma crescente resposta das universidades, através de novos cursos e especializações que devem unir administração, estatística e programação (Goes, 2014).

Em um estudo sobre o uso de Inteligência Analítica (IA) no estado de Utah, EUA, identificou-se crescente interesse de alunos de sistemas da informação pela atuação com IA, o que reforça a necessidade de entender quais competências são necessárias (Anderson, Daniels, McDonald, & Edvalson, 2014). O estudo revelou que nas empresas de Utah as ferramentas de BI&A mais usadas são: planilhas, query e análise, *Web Analytics*, estatística geral, otimização, visualização de dados, previsão, *Data* e *Text Mining*, simulações, etc, confirmando as tecnologias encontradas na literatura. Sobre as áreas onde IA é usado, destacam-se vendas, finanças, *Marketing*, planejamento, relacionamento com clientes, *Supply chain*, entre outros. O conhecimento dessas ferramentas e áreas de uso oferecem pistas aos praticantes de IA sobre o que devem aprender em termos de competências.

Em um estudo estatístico com dados obtidos a partir de técnicas de *Text Mining* em anúncios de emprego (1357 de BI e 450 de *Big Data*), obtiveram-se competências necessárias para profissionais atuarem com IA. Conforme Figura 8, essas competências podem ser classificadas em duas grandes categorias: competências de negócio e competências de TI (Debortoli, Müller, & vom Brocke, 2014). Dentro dessas categorias, subdivide-se competências de negócio em competências de gerenciamento e de domínio, e as competências de TI em competências metodológicas, conceituais e de ferramentas específicas. Enquanto BI baseia-se em habilidades relacionadas ao uso de plataformas comerciais (dos fornecedores SAP, Microsoft, IBM, etc), IA baseia-se em engenharia de *software*, habilidades estatísticas e ferramentas de código-aberto (Debortoli, Müller, & vom Brocke, 2014). Os autores argumentam que na literatura há vasto

material que permite entender as competências necessárias para IA em nível organizacional, mas há pouco material sobre o nível individual.

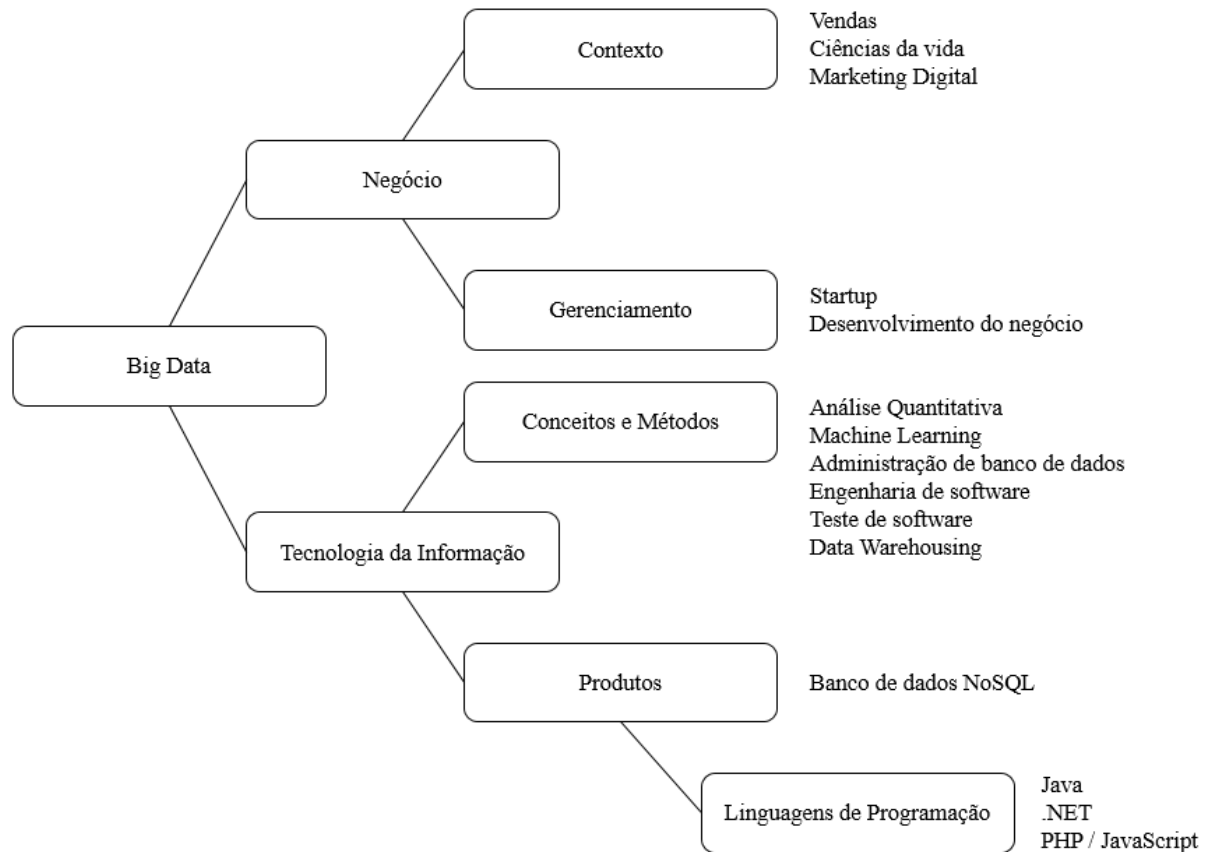


Figura 8: Competências necessárias para profissionais atuarem com IA

Fonte: Traduzido de Debortoli, Müller, & vom Brocke, 2014

Tradicionalmente, equipes de IA estão localizados como consultorias internas, centros de excelência, ou nos departamentos de TI, e atendem gerentes e executivos com relatórios específicos. As iniciativas de IA, por outro lado, estão em nível mais experimental e podem ser encontrados em diferentes áreas e níveis da empresa (Debortoli, Müller, & vom Brocke, 2014).

Através de uma série de entrevistas com profissionais praticantes de IA em grandes empresas no estado de Maharashtra, Índia, obtiveram-se algumas respostas genéricas sobre conhecimentos de habilidades necessárias para atuação com IA, mas contribuem para um primeiro entendimento (Dubey & Gunasekaran, 2015). As habilidades são divididas em dois conjuntos: *hard skills*, que consiste em estatística, predição, otimização, finanças quantitativas, estatística multivariada, tomada de decisão, *Marketing*, métodos de pesquisa e finanças; e os

*soft skills*, que consiste em liderança, trabalho em equipe, saber ouvir, facilidade em aprender, atitude positiva, boa comunicação, relacionamentos interpessoais, paciência e paixão. Uma boa comunicação entre áreas e funcionários é essencial para obtenção de sinergia, o que aumenta as chances de atender as expectativas dos interessados (Moraes, Meirelles, & Cappellozza, 2016). Uma comunicação ruim entre áreas permite a proposta, aprovação e execução de projetos de tecnologia, mas sem uma estrutura adequada, raramente as expectativas dos envolvidos são atingidas (Albertin & Albertin, 2012).

## 6 MÉTODO

Nesta seção estão descritos a abordagem metodológica e os procedimentos metodológicos utilizados nas fases do estudo. Na Figura 9 observa-se um resumo visual de cada etapa até o resultado final.

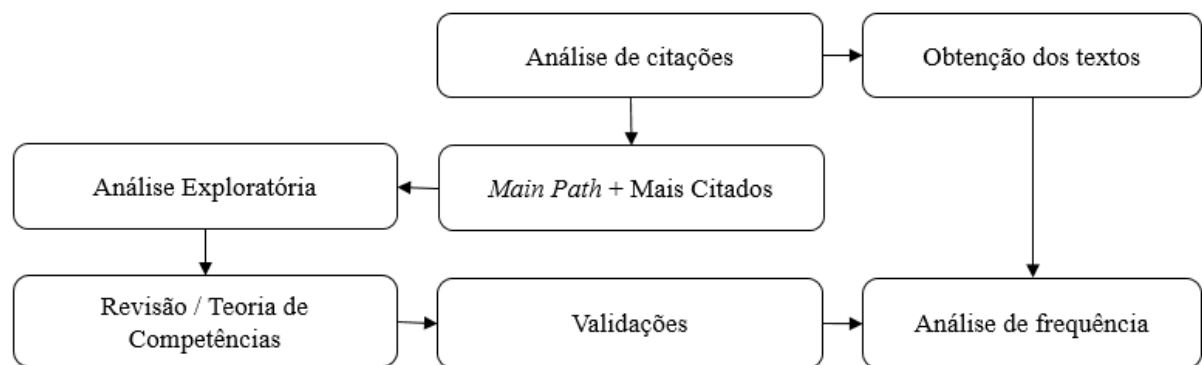


Figura 9: Resumo das etapas do estudo  
Fonte: Elaboração Própria

A etapa "Análise de citações" consistiu em definir critérios, obter uma base de informações sobre textos que representassem a literatura de Inteligência Analítica e realizar a análise de citações. Utilizando-se os critérios estabelecidos, foram encontrados mais de mil textos, dos quais 444 que citam ou são citados ao menos uma vez, e 208 são citados ao menos uma vez.

A etapa "*Main Path* + Mais Citados" consistiu em realizar o cálculo do *Main Path* (o conceito é explicado nesta seção) e em verificar quais textos são os mais citados. O cálculo do *Main Path* resultou um conjunto de textos que foi considerado para a revisão de literatura, juntamente com os 10 mais citados (alguns já faziam parte do *Main Path*).

Realizada em seguida, a etapa "Análise Exploratória" consistiu na leitura exploratória dos textos supracitados (*Main Path* + Mais Citados). Esse procedimento permitiu que fossem encontrados mais textos considerados pelo pesquisador como importantes, totalizando 62 textos.

Os 62 textos foram revisados utilizando-se a Teoria de Competências, etapa essa que foi denominada "Revisão / Teoria de Competências". Foram levantadas competências de três tipos: conhecimentos, habilidades e comportamentos para atuação em Inteligência Analítica.

Como todo o trabalho foi realizado por critérios e procedimentos definidos pelo pesquisador, realizou-se uma etapa denominada "Validações", na qual definiu-se seis especialistas no assunto Inteligência Analítica e apresentou-se as competências levantadas. O resultado foi a eliminação de algumas dessas competências, bem como o reforço da validade das remanescentes.

206 dos 208 textos que foram citados ao menos uma vez na análise de citações foram obtidos na etapa "Obtenção dos textos". Por fim, a etapa "Análise de Frequência" consistiu na contagem de frequência das competências validadas nos 208 textos. Essas contagens foram utilizadas para entendimento de cada competência, até que fosse obtida saturação.

Detalhes sobre a “Análise de citações”, “*Main Path* + Mais Citados” e “Análise Exploratória” podem ser encontrados na seção 6.2 Definição e obtenção da base de textos para revisão de literatura.

Detalhes sobre a “Revisão / Teoria de Competências” podem ser encontradas na seção 6.3 Levantamento e composição dos Quadros de Competências.

Detalhes sobre a “Validações” podem ser encontradas na seção 6.4 Validação dos Quadros de Competências.

Detalhes sobre a “Obtenção dos textos” e “Análise de Frequência” podem ser encontradas na seção 6.5 Descrição das Competências.

## 6.1 Abordagem metodológica

A abordagem metodológica adotada no trabalho contém elementos quantitativos e elementos qualitativos, o que caracteriza uma abordagem mista. Considerando o objetivo de levantar as competências para atuação em Inteligência Analítica (IA), uma análise quantitativa permite conclusões sobre convergências e tendências acerca deste fenômeno. Complementarmente, uma análise qualitativa permite um entendimento mais profundo sobre cada aspecto do fenômeno.

## 6.2 Definição e obtenção da base de textos para revisão de literatura

Para entender os relacionamentos dos textos na base estudada realizou-se uma análise de citações. Ao estudar quais textos foram mais influentes ao longo do tempo e o porquê, tem-se uma boa estimativa de quais devem ser lidos para uma imersão mais informada na área de IA. A utilização de um método matemático permitiu definir uma base de textos de forma mais certa e livre de intuições.

Para este trabalho optou-se pelo uso da base ISI *Web of Science*. Segundo alguns autores (Liu, Lu, Lu, & Lin, 2013) essa base de dados é adequada por incluir mais de 10.000 *journals* e materiais de mais de 120.000 congressos. Importante parte do conhecimento sobre Inteligência Analítica está sendo gerado e discutido em congressos (Chen, Chiang, & Storey, 2012). Adicionalmente, verificou-se a possibilidade de exportar informações importantes para a verificação automatizada de citações (através de programação), como a lista de referências de um texto, por exemplo.

O artigo “*Business Intelligence and Analytics: from Big Data to Big Impact*” (Chen, Chiang, & Storey, 2012), publicado em 2012 no periódico *MIS Quarterly*, foi escolhido como referência para as palavras-chave da busca. Havia a percepção prévia de que era um artigo importante, muito citado, e durante a realização deste trabalho confirmou-se ser também o mais citado da rede construída. Nesse trabalho, os autores fizeram uma busca em diversas bases, utilizando como critério os termos “*Business Intelligence*”, “*Business Analytics*” ou “*Big Data*” (através

dos campos título, *abstract* e *keywords*) e obtiveram cerca de 3610 registros. Assim, para refinar a pesquisa, aqui adicionou-se o termo “*Analytics*” aos termos anteriores, em textos publicados até 31 de março de 2016, sem restrição inicial de tempo. Foram considerados todos os registros de textos encontrados (não apenas artigos científicos). Com essa segunda estratégia obteve-se 1561 registros, dos quais 1234 estão na Principal Coleção *Web of Science* (contém todas as informações necessárias para automatizar parte da construção da rede), que foi o conjunto considerado para as próximas etapas.

Para realizar a análise, foram consideradas as seguintes informações exportadas do ISI *Web of Science*: título, local de publicação, lista de autores, lista de referências, DOI (*digital object identifier*) e ano de publicação. Realizou-se uma análise dessas informações e verificou-se que os 1234 continham título, local de publicação, lista de autores e ano de publicação. Por outro lado, 799 continham DOI (identificador único na base) e 1155 continham lista de referências. A ausência da lista de referências indica que o texto não é acadêmico.

Sempre que um texto da Principal Coleção *Web of Science* possui DOI, ele aparece nas listas de citação de outros textos com esse identificador, sendo esse o campo utilizado para cruzamentos. Porém, para os casos em que o DOI não existe, o padrão das listas de citação é PRIMEIRO\_AUTOR, ANO\_DE\_PUBLICAÇÃO, sendo esta informação a segunda maneira de realizar cruzamentos. Através do MS Excel cruzou-se cada um dos 1234 textos com os outros, totalizando 1.522.756 comparações.

De todas as comparações, verificou-se quase 700 ocorrências de citação, ou seja, quase 700 relacionamentos “um citou o outro”. Algumas ocorrências eram inválidas por duplicações nas citações de textos identificados pela mesma dupla PRIMEIRO\_AUTOR, ANO\_DE\_PUBLICAÇÃO. Após esta análise foram mantidas 634 ocorrências válidas.

A seguir, verificou-se que 444 dos 1234 textos foram responsáveis pelas 634 ocorrências de citação, sendo que apenas esses foram considerados para construção da rede de citações. Os demais não citam algum outro e nem são citados. Dos 444 textos que citam ou são citados ao menos uma vez, 208 são citados ao menos uma vez, sendo essa a base de textos considerada em um segundo momento nesse estudo.

De acordo com as vantagens em se utilizar um *software* matemático, descritas na seção 3 ANÁLISES DE CITAÇÕES, para este trabalho optou-se pelo uso do *software* Pajek. O *input* foi construído considerando-se os 444 textos e os 634 relacionamentos de citação que foram



encontrados. O processamento no *software* consistiu em dois grupos de operações: cálculo do *Main Path* (12 textos na Tabela 2) e análise dos textos mais citados (4 textos na Tabela 3, não repetidos em relação ao *Main Path*), o que resultou em um conjunto de 16 textos considerados como essenciais para a revisão de literatura deste trabalho. O cálculo do *Main Path* foi realizado no *software* Pajek com o método *Search Path Count* (SPC), através do comando *Net/Citation Weights/Search Path Count* (SPC). Os textos mais citados foram obtidos através do comando *Net/Partitions/Degree/Output*, que realiza a contagem de quantas vezes um texto foi citado.

Tabela 2: Textos que compõem o *Main Path*

#	Informações sobre o texto
1	Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). <i>Emerging Trends in Business Analytics</i> . Communications of the ACM, 45(8), 45-48.
2	Jourdan, Z., Rainer, R. K., & Marshall, T. E. (2008). <i>Business Intelligence: An Analysis of the Literature 1</i> . Information Systems Management, 25(2), 121-131.
3	Bose, R. (2009). <i>Advanced Analytics: opportunities and challenges</i> . Industrial Management & Data Systems, 109(2), 155-172.
4	Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. (2010). <i>The impact of Business Analytics on supply chain performance</i> . Decision Support Systems, 49(3), 318-327.
5	Chae, B., & Olson, D. L. (2013). <i>Business Analytics for Supply chain: A Dynamic - Capabilities Framework</i> . International Journal of Information Technology & Decision Making, 12(1), 9-26.
6	Davenport, T. H., & Patil, D. J. (2012). <i>Data scientist: The Sexiest Job of the 21st Century</i> . Harvard Business Review, 90, 70-76.
7	McAfee, A., & Brynjolfsson, E. (2012). <i>Big Data. The Management Revolution</i> . Harvard Business Review, 90(10), 61-67.
8	Chen, H., Chiang, R. H., & Storey, V. C. (2012). <i>Business Intelligence and Analytics: From Big Data to Big Impact</i> . MIS Quarterly, 36(4), 1165-1188.
9	Waller, M. A., & Fawcett, S. E. (2013). <i>Data Science, Predictive Analytics, and Big Data: a Revolution That Will Transform Supply chain Design and Management</i> . Journal of Business Logistics, 34(2), 77-84.

10	Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). <i>Data quality for Data Science, Predictive Analytics, and Big Data in Supply Chain Management: An introduction to the problem and suggestions for research and applications</i> . International Journal of Production Economics, 154, 72-80.
11	Chae, B. K. (2015). <i>Insights from hashtag #supplychain and Twitter Analytics: Considering Twitter and Twitter Data for supply chain practice and research</i> . International Journal of Production Economics, 165, 247-259.
12	Wang, X., White, L., & Chen, X. (2015). <i>Big Data research for the knowledge economy: past, present, and future</i> . Industrial Management & Data Systems, 115(9).

Fonte: Elaboração Própria

Tabela 3: Textos entre os dez mais citados e que não fazem parte do *Main Path*

#	Informações sobre o texto
1	LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). <i>Big Data, Analytics and the Path from Insights to Value</i> . MIT Sloan Management Review, 52(2), 21.
2	Boyd, D., & Crawford, K. (2012). <i>Critical Questions for Big Data</i> . Information, Communication and Society, 15(5), 662-679.
3	Demirkan, H., & Delen, D. (2013). <i>Leveraging the capabilities of service-oriented Decision Support Systems: Putting Analytics and Big Data in cloud</i> . Decision Support Systems, 55(1), 412-421.
4	Barton, D., & Court, D. (2012). <i>Making Advanced Analytics Work For You</i> . Harvard Business Review, 90(10), 78-83.

Fonte: Elaboração Própria

Os 16 textos foram lidos e revisados e esse processo resultou em uma adição de mais 46 textos considerados importantes, o que totalizou 62 textos, lidos e revisados durante três meses. Os 62 textos foram encontrados através da ferramenta de busca Google Scholar, no ambiente da FGV EAESP, o que garantiu os acessos necessários para a obtenção de todos eles.

### 6.3 Levantamento e composição dos Quadros de Competências

O levantamento foi baseado na divisão proposta pelo modelo de tipologias de competências (Le Deist & Winterton, 2005), que é composto de três grupos: competências cognitivas

(conhecimentos), competências funcionais (habilidades) e competências sociais (comportamentos). Esses três grupos são suportados por um quarto grupo, denominado meta competências, que por independe de contexto, foge do escopo deste trabalho, que visa entender as competências do profissional que atua no contexto de Inteligência Analítica. Assim, dividiu-se as Competências encontradas em conhecimentos, habilidades e comportamentos. Os termos que representam as Competências não foram traduzidos, pois a tradução poderia causar perda de fidelidade, pela especificidade dos termos.

Conhecimento é o entendimento do contexto de atuação de um indivíduo, que permite enxergar padrões e pensar sistematicamente (Boyatzis, 2006), sendo este o critério utilizado para mapear os conhecimentos para atuação em Inteligência Analítica, presentes na literatura estudada.

Existem muitos estudos que dividem as habilidades em habilidades funcionais e habilidades comportamentais (Le Deist & Winterton, 2005). O modelo adotado neste trabalho inclui as habilidades comportamentais no pilar de competência social, o que implica na definição de habilidade como competência funcional: *know-how*, ou o “saber fazer”, sendo este o critério utilizado para mapear as habilidades para atuação em Inteligência Analítica, presentes na literatura estudada. No contexto de Inteligência Analítica, são considerados duas categorias de habilidades: as habilidades analíticas e as habilidades computacionais (Chen, Chiang, & Storey, 2012). As habilidades encontradas foram divididas nessas categorias antes da descrição das competências.

Por fim, os comportamentos são definidos como um conjunto de características sociais que um indivíduo precisa para uma boa performance dentro de uma organização (Le Deist & Winterton, 2005). Nesse contexto, comportamento abrange a vontade e capacidade de experimentar e moldar relações, para identificar e entender interações sociais, de forma a manter a racionalidade, com o objetivo final de ser responsável e solidário (Straka, 2004), sendo este o critério utilizado para mapear os comportamentos para atuação em Inteligência Analítica, presentes na literatura estudada.

Os critérios supra descritos foram utilizados para analisar os textos escopo da revisão de literatura e compor os três Quadros de competências: conhecimentos, habilidades e comportamentos.

Dos 62 textos considerados na revisão de literatura, 38 textos contêm pelo menos uma competência. No Apêndice 1 pode-se observar detalhes desses textos, bem como os respectivos identificadores numéricos para identificação nos Quadros dessa seção.

<b>Conhecimentos</b>	<b>Referências encontradas</b>
<i>Accounting</i>	[5], [6], [14], [26], [33], [35], [37]
<i>Big Data</i>	[1], [2], [3], [4], [6], [7], [11], [12], [13], [14], [16], [17], [18], [21], [25], [26], [27], [28], [29], [34], [37]
<i>Cloud</i>	[1], [4], [12], [13], [14], [16], [17], [18], [21], [25], [29], [34], [36], [37]
<i>CRM</i>	[1], [7], [8], [10], [20], [21], [32]
<i>Data Science</i>	[1], [4], [6], [13], [26], [27], [37]
<i>Decision Making</i>	[1], [2], [3], [5], [6], [9], [10], [12], [13], [14], [15], [16], [18], [19], [24], [25], [26], [29], [30], [32], [33], [35], [36], [37]
<i>Decision Support Systems</i>	[1], [5], [12], [15], [16], [23], [33], [37]
<i>Economics</i>	[1], [5], [6], [10], [13], [14], [15], [16], [17], [18], [26], [30], [34], [35], [37]
<i>ERP</i>	[10], [13], [15], [20], [23], [32], [35]
<i>Finance</i>	[1], [2], [9], [16], [19], [22], [25], [26], [27], [29], [31], [33], [37]
<i>Fraud Detection</i>	[1], [34]
<i>Healthcare</i>	[1], [12], [14], [18], [25], [27], [28], [29], [32], [34], [36], [37]
<i>Human Resources</i>	[14], [19], [24], [33], [35]
<i>Logistics</i>	[1], [5], [6], [8], [10], [13], [15], [17], [18], [19], [26], [37]
<i>Marketing</i>	[1], [2], [5], [6], [7], [8], [9], [10], [11], [13], [14], [15], [16], [17], [18], [19], [20], [24], [25], [26], [29], [31], [32], [34], [35], [37], [38]
<i>Operations Management</i>	[15], [26], [37]
<i>Real Time</i>	[2], [3], [4], [5], [9], [12], [13], [15], [16], [18], [20], [21], [26], [28], [29], [30], [31], [32], [33], [34], [36], [37]
<i>Sales</i>	[2], [3], [5], [6], [7], [8], [9], [15], [17], [18], [19], [20], [21], [24], [25], [28], [31], [33], [34], [37]

<i>Structured Data</i>	[1], [18], [20], [21], [32], [34]
<i>Supply Chain Management</i>	[5], [6], [13], [15], [17], [37]
<i>Unstructured Data</i>	[4], [7], [9], [12], [16], [18], [20], [32], [34], [37]

Quadro 1: Conhecimentos para atuação em IA, segundo literatura  
Fonte: Elaboração Própria

Conforme Quadro 1, foram encontrados termos que representam conhecimentos para atuação em Inteligência Analítica, segundo a literatura e sob óptica da Teoria de Competências. Os termos *Big Data*, *Cloud*, *CRM*, *Data Science*, *Decision Making*, *Decision Support Systems*, *ERP*, *Real Time*, *Structured Data* e *Unstructured Data* representam conceitos que os profissionais devem dominar, como um pré-requisito para entenderem o contexto de atuação em Inteligência Analítica. Os termos *Accounting*, *Economics*, *Finance*, *Fraud Detection*, *Healthcare*, *Human Resources*, *Logistics*, *Marketing*, *Operations Management*, *Sales* e *Supply Chain Management* representam possíveis contextos de aplicação de Inteligência Analítica, o que sugere que o profissional não necessita todos os conhecimentos, mas apenas os necessários para entendimento do contexto de atuação. Esses últimos termos foram agrupados em uma competência denominada Negócio.

<b>Habilidades</b>	<b>Referências encontradas</b>
<i>Classification</i>	[1], [10], [18], [20], [25], [32], [38]
<i>Clustering</i>	[1], [2], [9], [15], [17], [20], [23]
<i>Data Mining</i>	[1], [6], [7], [8], [9], [9], [10], [12], [14], [15], [16], [17], [20], [21], [25], [32], [33], [34], [35], [36], [38]
<i>Data Visualization</i>	[2], [19], [34]
<i>Data Warehousing</i>	[1], [9], [10], [15], [16], [20], [21], [22], [25], [30], [33], [34]
ETL	[1], [8], [9], [12], [15], [20], [21], [25], [30]
<i>Forecasting</i>	[2], [3], [5], [6], [8], [9], [10], [12], [14], [15], [17], [19], [26], [34], [35]
Hive	[25]
<i>Machine Learning</i>	[1], [10], [17], [25], [28]

MapReduce	[1], [21], [25], [34]
<i>Mathematical Optimization</i>	[2], [15]
<i>Modeling</i>	[1], [2], [5], [6], [7], [9], [10], [12], [13], [15], [16], [20], [22], [24], [25], [29], [30], [31], [34], [35], [35], [36], [37]
<i>Network Analysis</i>	[1], [17]
<i>Neural Network</i>	[1], [9], [10], [15]
NoSQL	[12], [25]
OLAP	[1], [9], [15], [16], [20], [21], [25], [30], [32], [33], [34]
<i>Optimization</i>	[1], [2], [4], [5], [6], [7], [8], [9], [10], [12], [14], [15], [19], [21], [22], [23], [24], [26], [29], [34]
<i>Parallel DBMS</i>	[1], [21], [21]
<i>Partitioning</i>	[21], [23]
<i>Predictive (Analytics)</i>	[1], [2], [5], [6], [7], [9], [12], [13], [14], [15], [16], [19], [21], [22], [24], [25], [26], [29], [31], [34], [35], [36], [37]
<i>Prescriptive (Analytics)</i>	[14], [15], [16]
<i>Process Mining</i>	[1], [5]
<i>Programming</i>	[1], [4], [10], [15], [17], [25]
<i>Project Management</i>	[10], [25]
<i>Regression</i>	[1], [2], [9], [10], [35], [37]
<i>Search Engines</i>	[1], [21], [23]
<i>Sentiment Analysis</i>	[1], [17], [34], [37]
<i>Simulation</i>	[5], [6], [9], [10], [12], [14], [15], [16], [22], [29], [37]
<i>Social Media Analytics</i>	[1], [19], [37]
<i>Statistical Analysis</i>	[1], [9], [34], [36]
<i>Text Mining</i>	[1], [9], [12], [17], [19], [20], [25], [34], [38]
<i>Web Mining</i>	[1], [9], [23], [38]

Quadro 2: Habilidades para atuação em IA, segundo literatura

Fonte: Elaboração Própria

Conforme Quadro 2, foram encontrados termos que representam habilidades para atuação em Inteligência Analítica, segundo a literatura e sob óptica da Teoria de Competências. Essas habilidades podem ser divididas em duas categorias: as habilidades analíticas, mais ligadas ao uso correto de técnicas; e as habilidades computacionais, mais ligadas ao uso correto de

tecnologias (Chen, Chiang, & Storey, 2012). As habilidades encontradas foram divididas nessas categorias antes da descrição das competências.

<b>Comportamentos</b>	<b>Referências encontradas</b>
<i>Communicate</i>	[4], [25], [27], [35]
<i>Creative</i>	[4], [7], [36]
<i>Leadership</i>	[3], [5], [6], [15], [17], [18], [24], [26], [29], [31], [33]
<i>Passion</i>	[26]
<i>Patient</i>	[1]
<i>Problem Solving</i>	[16]
<i>Self-Motivated</i>	[27]

Quadro 3: Comportamentos para atuação em IA, segundo literatura

Fonte: Elaboração Própria

Conforme Quadro 3, foram encontrados termos que representam comportamentos para atuação em Inteligência Analítica, segundo a literatura e sob óptica da Teoria de Competências. Observa-se uma frequência menor dos termos que representam comportamentos, em relação as frequências dos termos de conhecimentos e habilidades. Dentro da amostra de textos considerada, isso mostra que existe menos preocupação em relação aos requisitos comportamentais para atuação em Inteligência Analítica. Outro ponto é que os comportamentos encontrados são desejáveis em praticamente todos os cargos dentro de uma empresa, segundo o senso comum.

#### 6.4 Validação dos Quadros de Competências

Apesar de todo o trabalho basear-se em literatura, métodos específicos e teorias, considerou-se necessário realizar-se uma validação com especialistas no assunto de Inteligência Analítica. Como o assunto é ainda pouco discutido no cenário brasileiro, as grandes empresas costumam

proteger informações sobre como utilizam Inteligência Analítica, o que dificulta uma validação quantitativa de larga escala.

Assim, optou-se por consultar um grupo seletivo de acadêmicos e de praticantes, escolhidos por terem conhecimentos em Inteligência Analítica e também por conveniência, ou seja, por serem acessíveis ao autor deste trabalho.

Foram selecionados três professores doutores da FGV EAESP com pesquisas publicadas em Inteligência Analítica, e três especialistas, que já atuaram ou atuam diretamente com Inteligência Analítica, totalizando seis especialistas.

Os Quadros de competências foram apresentados para o grupo de seis especialistas em Inteligência Analítica e houve uma validação prévia.

Os especialistas verificaram a quantidade de vezes que cada competência apareceu na base de textos considerada no estudo, e sugeriram que todas as competências que aparecessem três vezes ou menos fossem eliminadas, por serem demasiadamente específicas, ou por contribuírem pouco para o objetivo do estudo. Através dessa regra não foram eliminados conhecimentos, mas foram eliminadas 12 habilidades. Se a mesma regra fosse seguida para comportamentos, seriam eliminados 5 do conjunto de 7 comportamentos. Por isso, decidiu-se eliminar os comportamentos que apareceram somente uma vez.

Com os critérios acima, eliminou-se as seguintes habilidades: *Data Visualization*, *Hive*, *Mathematical Optimization*, *Network Analysis*, *NoSQL*, *Parallel DBMS*, *Partitioning*, *Prescriptive (Analytics)*, *Process Mining*, *Project Management*, *Search Engines* e *Social Media Analytics*; e os seguintes comportamentos: *Passion*, *Patient*, *Problem Solving* e *Self Motivated*.

A segunda parte da validação consistiu em uma avaliação quantitativa de cada competência, considerando-se a escala Likert (Likert, 1932). Essa escala é frequentemente utilizada para medição de qualidade de percepção, geralmente em cinco ou sete níveis, e podendo ser o nível do meio neutro (Allen & Seaman, 2007). Para este estudo optou-se pela utilização de cinco níveis: Muito importante, Importante, Neutro, Sem importância e Totalmente sem importância.

Para cada competência foi calculada uma média simples entre as respostas dos seis especialistas, sendo a regra de conversão utilizada para transformar em números conforme Tabela 4.



Tabela 4: Regra de conversão da escala Likert

Valor na escala Likert	Valor numérico considerado
Muito importante	5
Importante	4
Neutro	3
Sem importância	2
Totalmente sem importância	1

Fonte: Elaboração Própria

As competências que foram avaliadas como “Totalmente sem importância” (valor 1) ou como “Sem importância” (valor 2) por pelo menos um dos especialistas, tiveram suas médias analisadas e, nos casos em que a média fosse inferior a 4,0, a respectiva competência foi desconsiderada nas fases posteriores do trabalho. A validação desse conjunto foi realizada através de escala Likert e pode ser observada no Quadro 4.

Competência	E1	E2	E3	E4	E5	E6	MÉDIA
Conhecimento - <i>Big Data</i>	5	4	5	4	5	5	4,67
Conhecimento - <i>Cloud</i>	4	3	4	4	4	4	3,83
Conhecimento - CRM	3	4	5	4	4	<u>2</u>	3,67
Conhecimento - <i>Data Science</i>	3	5	5	3	5	5	4,33
Conhecimento - <i>Decision Making</i>	4	3	5	4	5	4	4,17
Conhecimento - <i>Decision Support Systems</i>	4	4	5	4	3	4	4,00
Conhecimento - ERP	4	<u>2</u>	5	4	3	<u>1</u>	3,17
Conhecimento - Negócio (conhecer o contexto da empresa em que vai atuar)	4	5	5	5	5	4	4,67
Conhecimento - <i>Real Time</i>	3	3	4	4	4	3	3,50
Conhecimento - <i>Structured Data</i>	3	3	5	3	5	5	4,00
Conhecimento - <i>Unstructured Data</i>	3	3	5	3	5	4	3,83
Habilidade - <i>Classification</i>	4	5	5	4	5	4	4,50
Habilidade - <i>Clustering</i>	4	5	5	4	5	4	4,50

Habilidade - <i>Data Mining</i>	5	5	5	4	5	5	4,83
Habilidade - <i>Data Warehousing</i>	5	4	5	5	4	3	4,33
Habilidade - ETL ( <i>Extract, Transform, Load</i> )	3	5	4	4	5	5	4,33
Habilidade - <i>Forecasting</i>	5	4	4	4	5	4	4,33
Habilidade - <i>Machine Learning</i>	3	5	4	4	5	4	4,17
Habilidade - MapReduce	3	4	5	3	4	3	3,67
Habilidade - <i>Modeling</i>	5	5	5	4	4	5	4,67
Habilidade - <i>Neural Network</i>	3	5	4	4	5	4	4,17
Habilidade - OLAP ( <i>Online Analytical Processing</i> )	4	3	4	5	4	3	3,83
Habilidade - <i>Optimization</i>	5	5	4	3	5	4	4,33
Habilidade - <i>Predictive Analytics</i>	3	5	5	4	5	5	4,50
Habilidade - <i>Programming</i>	4	3	4	3	5	5	4,00
Habilidade - <i>Regression</i>	5	5	5	4	5	5	4,83
Habilidade - <i>Sentiment Analysis</i>	3	5	4	3	5	3	3,83
Habilidade - <i>Simulation</i>	5	3	5	4	5	4	4,33
Habilidade - <i>Statistical Analysis</i>	5	5	5	4	5	5	4,83
Habilidade - <i>Text Mining</i>	4	<u>2</u>	5	4	5	3	3,83
Habilidade - <i>Web Mining</i>	4	3	5	4	4	4	4,00
Comportamento - <i>Communicate</i>	4	4	5	4	5	4	4,33
Comportamento - <i>Creative</i>	3	5	4	5	5	5	4,50
Comportamento - <i>Leadership</i>	3	<u>2</u>	5	4	4	4	3,67

Quadro 4: Validação (médias) por competência

Fonte: Elaboração Própria

Conforme Quadro 4, observa-se que 24 competências foram consideradas pelos especialistas, na média, importantes ou muito importantes (acima de 4,0). Outras 10 competências tiveram valores inferiores a 4,0 (mas sempre superiores a 3,0), sendo que 4 tiveram ao menos uma resposta “Totalmente sem importância” (valor 1) ou “Sem importância” (valor 2), e foram,

consequentemente, desconsideradas nas fases posteriores do estudo: CRM, ERP, *Text Mining* e *Leadership*.

## 6.5 Método para descrição das Competências

Para descrever as competências considerou-se uma base de textos numericamente bastante superior à base de textos utilizada para levantamento das competências.

Enquanto o levantamento de competências baseou-se em um método manual da observação de competências através de uma teoria e em uma base de 62 textos, para encontrar material descritivo para essas competências considerou-se um método semi automatizado e 208 textos, que são todos aqueles com pelo menos uma citação da base original de 444 textos, descrita na seção 6.2 Definição e obtenção da base de textos para revisão de literatura.

Essa base de 208 textos abrange diferentes tipos de textos, conforme Tabela 5.

Tabela 5: Tipos de textos na base de 208 textos

<b>Tipo de Texto</b>	<b>Contagem</b>	<b>Porcentagem</b>
Artigo	121	58%
Texto de Congresso	65	31%
Editorial	13	6%
<i>Review</i>	9	5%

Fonte: Elaboração Própria

Dos 121 textos que são artigos, 31 estão concentrados em apenas 10 *journals* mais citados, que incluem alguns dos *top journals* da área de *Information Systems*, o que destaca a relevância da base considerada no estudo. A distribuição dos dez *journals* mais citados nessa base estão na Tabela 6.

Tabela 6: *Journals* mais citados na base de 121 artigos, dentre os 208 textos

Nome do <i>journal</i>	Contagem de textos na base	Soma de citações	Porcentagem
MIS QUARTERLY	1	77	22%
MIT SLOAN MANAGEMENT REVIEW	3	61	18%
HARVARD BUSINESS REVIEW	5	60	18%
DECISION SUPPORT SYSTEMS	8	50	15%
INFORMATION COMMUNICATION & SOCIETY	1	24	7%
COMMUNICATIONS OF THE ACM	2	20	6%
JOURNAL OF BUSINESS LOGISTICS	1	15	4%
HEALTH AFFAIRS	5	13	4%
INDUSTRIAL MANAGEMENT & DATA SYSTEMS	3	12	3%
INFORMATION SYSTEMS MANAGEMENT	2	12	3%

Fonte: Elaboração Própria

Por fim, como mais um argumento sobre a qualidade da base de 208 textos, destaca-se a atualidade da base de informações, através de uma análise do ano de publicação dos textos, conforme Tabela 7.

Tabela 7: Ano de publicação dos 208 textos

Ano de Publicação	Contagem de textos	Porcentagem
2002	1	~0%
2003	1	~0%
2004	1	1%
2007	2	1%
2008	4	2%
2009	4	2%
2010	9	4%
2011	10	5%
2012	20	10%
2013	53	25%
2014	66	32%
2015	33	16%

2016	4	2%
------	---	----

Fonte: Elaboração Própria

Percebe-se uma concentração entre os anos 2010 e 2015, o que indica que a base é atualizada. Além disso, estão presentes alguns textos mais antigos (entre 2002 e 2009), o que indica que a base contém textos seminais, assim como alguns textos mais novos, em 2016.

Os 208 textos foram encontrados através da ferramenta de busca Google Scholar, no ambiente da FGV EAESP, o que garantiu os acessos necessários para a obtenção de 206 todos eles. Alguns foram solicitados diretamente aos autores, através da rede social *Research Gate*. Apenas dois textos não puderam ser obtidos, sendo ambos textos de congresso e com poucas citações (menor relevância para o estudo).

Os 206 textos foram submetidos, um a um, na ferramenta do *site Textalyser*, que oferece diversas funcionalidades estatísticas de um texto, inclusive o cálculo de frequência de termos compostos de 1, 2, 3, 4 e 5 palavras. Cada saída da ferramenta foi copiada para MS Excel e tratada para permitir rastreabilidade da frequência de cada termo em cada texto. Como resultado final, obteve-se uma base que permite a verificação da frequência absoluta de cada termo, bem como em quais textos e com que frequência aparecem.

As competências que passaram pela fase de validação foram verificadas nessa base e foram considerados os textos em que apareceram com maior frequência, até ter-se a percepção de saturação do entendimento. A decisão de considerar em primeiro lugar os textos que os termos aparecem com maior frequência tem o intuito de separar os textos mais focados em determinado assunto. A estratégia de trabalhar com saturação está sujeita à percepção do pesquisador, até pela ausência de um valor parâmetro na literatura, dada a especificidade do método utilizado.

As competências do tipo habilidade foram divididas em duas categorias. As habilidades analíticas: *Classification, Clustering, Forecasting, Modeling, Neural Network, Optimization, Predictive Analytics, Regression, Sentiment Analysis, Simulation* e *Statistical Analysis*; e as habilidades computacionais: *Data Mining, Data Warehousing, ETL (Extract, Transform, Load), Machine Learning, MapReduce, OLAP (Online Analytical Processing), Programming* e *Web Mining*.

## 7 DESCRIÇÃO DAS COMPETÊNCIAS

### 7.1 Conhecimentos

#### 7.1.1 *Big Data*

*Big Data* não é um conceito fácil de explicar. Existem muitos estudos que verificam diferentes aspectos desse fenômeno, por exemplo aspectos técnicos, aspectos culturais, aspectos estruturais, requisitos, vantagens, expectativas, etc. Definir *Big Data*, por si só, exigiria um estudo mais específico. Assim, essa seção visa caracterizar o fenômeno e relacioná-lo à Inteligência Analítica.

*Big Data* pode ser entendido como o fenômeno que se refere a um volume de dados que é difícil de tratar através de tecnologias clássicas de banco de dados (Hashem, Yaqoob, Anuar, Mokhtar, Gani, & Khan, 2015). Outra definição amplamente adotada é caracterizar o fenômeno sob alguns aspectos, os conhecidos “V”s, que podem ser 3, 4 ou 6, dependendo dos estudos. Por exemplo, um dos estudos que adota os 4 “V”s os definem como: volume, variedade, velocidade e valor. O volume é uma característica sobre a quantidade de dados a tratar; a variedade é uma característica sobre os tipos de dados encontrados, comumente classificados em estruturados, semi estruturados e não estruturados; a velocidade é uma característica acerca da necessidade de rápida execução e resposta, dado os problemas de volume e variedade; por fim, o valor é uma característica sobre a utilidade dos dados, que devem gerar resultados alinhados as expectativas de quem utiliza os dados (Gantz & Reinsel, 2011).

Considerando-se os 4 “V”s, o conceito de *Big Data* aproxima-se do contexto de Inteligência Analítica. A análise de dados que se enquadram no conceito de *Big Data* são frequentemente chamadas de *Big Data Analytics*, e há autores que garantem ser o mesmo que Inteligência Analítica (Chen, Chiang, & Storey, 2012), principalmente no que tange as tecnologias utilizadas para lidar com dados com tais características. *Big Data* não tem valor para tomada de decisão, a menos que seja realizada uma análise dos dados (Tien, 2013) para ganho de inteligência. Esse é o elo que une *Big Data* e Inteligência Analítica.

*Big Data* é algo visto com grande potencial para o contexto *Business*, sendo cada vez mais crítico o estabelecimento de práticas, técnicas e a criação de ferramentas adequadas para explorar esse potencial (Zheng, Zhu, & Lyu, 2013). Considera-se, inclusive, que *Big Data* é uma nova forma de capital no mercado atual (Mayer-Schönberger & Cukier, 2013). Trata-se de algo que se movimenta com uma aura de expectativa e com o potencial poder de romper paradigmas atuais em diversas áreas de atuação humana.

### 7.1.2 *Cloud*

Assim como *Big Data*, *Cloud* é um fenômeno que, por si só, exigiria um estudo mais específico. Assim, essa seção visa caracterizar o fenômeno e relacioná-lo à Inteligência Analítica.

*Cloud* é uma importante tecnologia para computação complexa e de grande escala, pois elimina a necessidade de manutenção de *hardware* e *software*, que podem ser utilizados apenas sob demanda, com mais tolerância a falhas e com escalabilidade adequada (Hashem, Yaqoob, Anuar, Mokhtar, Gani, & Khan, 2015). Esse conjunto de vantagens oferece mais flexibilidade para as empresas e há grande variedade de implementações e modelos de cobrança, por exemplo o pagamento somente pelo que for usado (Assunção, Calheiros, Bianchi, Netto, & Buyya, 2015).

Recentemente, o fenômeno *Cloud* é um dos mais promissores e discutidos tópicos da área de tecnologia da informação por oferecer um ambiente adequado para aplicações de Inteligência Analítica (Sandhu & Sood, 2015). Todos os processamentos, que podem envolver NoSQL, MapReduce e Hadoop, podem ser realizados em ambiente *Cloud*, sendo essa a relação entre Inteligência Analítica e esse fenômeno (Hashem, Yaqoob, Anuar, Mokhtar, Gani, & Khan, 2015).

Basicamente, *Cloud* consiste em grandes *Data Centers* em diferentes localizações e que podem ser acessados de diferentes maneiras. Além do simples armazenamento de dados, muitos serviços relacionados à informação podem ser oferecidos. Tudo que pode ser realizado com dados em um ambiente terceirizado e externo à empresa pode ser considerado *Cloud* (Sandhu & Sood, 2015).

### 7.1.3 *Data Science*

*Data Science* pode ser entendido como a aplicação de métodos quantitativos e qualitativos para resolver problemas relevantes e prever ocorrências (Waller & Fawcett, 2013a). Esse conhecimento aparece junto a muitas habilidades mapeadas neste trabalho, como *Data Mining*, *Predictive Analytics* e *Machine Learning*, o que indica que *Data Science* é um conhecimento interligado com a prática de Inteligência Analítica. Nessa linha, Chen, Chiang, & Storey (2012) definem Inteligência Analítica como *Data Science* no contexto *Business*.

Outros autores consideram *Data Science* como a integração de *Data Warehousing*, *Data Mining* e *Data Visualization* (Fernández, del Río, López, Bawakid, del Jesus, Benítez, & Herrera, 2014), o que sugere que conhecer *Data Science* significa dominar e entender uma série de habilidades. Ainda para esses autores, *Data Science* é qualquer tipo de análise feita com objetivos de conhecimento ou negócio.

Na literatura há certa confusão entre os termos *Data Science* e *Predictive Analytics*. Considera-se que *Predictive Analytics* seja um conjunto de ferramentas dentro do contexto de *Data Science* (Hazen, Boone, Ezell, & Jones-Farmer, 2014). *Data Science* é particularmente útil frente aos atuais desafios de *Big Data*.

### 7.1.4 *Decision Making*

Esse conhecimento define o foco com que a Inteligência Analítica deve ser operacionalizada. Dados por si só não tem valor para tomada de decisão (*Decision Making*), a menos que seja realizada uma análise dos dados (Tien, 2013) para ganho de inteligência, com foco na tomada de decisão. Há senso comum sobre a Inteligência Analítica servir como suporte a tomada de decisão, sendo este o tema de estudos na última década (Holsapple, Lee-Post, & Pakath, 2014), inclusive em áreas críticas, como a militar, por exemplo (Mohamed & Al-Jaroodi, 2014).

Antes da massiva adoção de sistemas de informação a tomada de decisão gerencial era realizada através de uma mistura de intuição e fatos (Sharma, Mithas, & Kankanhalli, 2014). Com a evolução dos sistemas e dos métodos de análise, as decisões são cada vez mais baseadas em



informações. Na era do *Big Data*, ampliam-se as expectativas em obter-se respostas úteis para perguntas de negócio, e a consequente tomada de decisão com foco em obtenção de vantagem e criação de valor (Bose, 2009; Arnott & Pervan, 2014). Existem até alguns casos de tomada de decisão automática, mas que trazem comprovadas consequências e certas perdas (Newell & Marabelli, 2015).

A tomada de decisão pode ser realizada em diferentes níveis do negócio, e pode ser classificada em sistêmica, estratégica, tática ou operacional. Com exceção de decisões básicas no nível operacional, todas as decisões têm como requisito uma análise prévia da situação, através de dados (Tien, 2012). Há estudos que relacionam o fenômeno de *Big Data* com uma possível radical (ou até disruptiva) mudança na tomada de decisão (Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015).

#### 7.1.5 *Decision Support Systems*

*Decision Support Systems* (DSS) são sistemas de informação criados com o intuito de ajudar na tomada de decisão. Esses sistemas existem há mais de 30 anos e passaram por grandes evoluções conceituais e de complexidade, tornando-os base para as novas tecnologias de Inteligência Analítica (Abrahams, Jiao, Wang, & Fan, 2012; Holsapple, Lee-Post, & Pakath, 2014). Desde as implementações originais, os DSS melhoram a eficiência e eficácia das empresas, juntamente com outros tipos de sistema de informação, como o ERP e o CRM (Chang, Kauffman, & Kwon, 2014).

A maioria das empresas possui algum tipo de DSS, que geralmente tratam uma mistura de dados transacionais (internos) e dados externos (Delen & Demirkan, 2013). A tendência é que esses sistemas continuem existindo, mas sejam desenvolvidos para um uso mais ágil, orientado à serviço e em ambiente *cloud* (Demirkan & Delen, 2013).

O suporte à decisão é o quarto assunto mais estudado dentro da literatura de Inteligência Analítica (Chen, Chiang, & Storey, 2012), o que evidencia a importância gerencial que esses sistemas têm.

### 7.1.6 *Negócio*

Os termos *Accounting* (Elbashir, Collier, & Sutton, 2011; Bhimani & Willcocks, 2014), *Economics* (Chang, Kauffman, & Kwon, 2014), *Finance* (Bhimani & Willcocks, 2014), *Fraud Detection* (Kambatla, Kollias, Kumar, & Grama, 2014), *Healthcare* (Berger & Doban, 2014), *Human Resources* (Varshney & Mojsilović, 2011), *Logistics* (Waller & Fawcett, 2013a), *Marketing* (Rust & Huang, 2014), *Operations Management* (Huang & Rust, 2013), *Sales* (Ko, Maciejewski, Jang, & Ebert, 2012) e *Supply Chain Management* (de Oliveira, McCormack, & Trkman, 2012) representam possíveis contextos de aplicação de Inteligência Analítica, o que sugere que o profissional não necessita conhecer todos esses contextos, mas precisa conhecer em detalhe aquele que atua ou vai atuar.

O professor Jeff Stanton, especialista em *Data Science* da Universidade de Siracusa, explica que pode ensinar testes estatísticos e programação em pouco tempo, mas que o domínio do negócio não é algo fácil de se ensinar. Na mesma linha, a especialista Shelly Farnham, da Microsoft, explica que procura profissionais de Inteligência Analítica que estejam altamente engajados no contexto de atuação (Dumbill, Liddy, Stanton, Mueller, & Farnham, 2013). Opiniões parecidas são encontradas em um outro estudo, em que se sugere que a rotação de trabalhos entre negócio e tecnologia pode ser uma possível solução para essa dificuldade (Wixom, Watson, & Werner, 2011).

### 7.1.7 *Real Time*

Se é desejável que as empresas possam tomar decisões em tempo real (*Real Time*), então é necessário planejar aplicações de Inteligência Analítica que considerem atualizações em tempo real como uma de suas variáveis. Basicamente, a ideia é que cada pequena mudança nos dados analisados seja refletida imediatamente no resultado final de uma análise de dados (Mohamed & Al-Jaroodi, 2014). A ideia é simples, mas a execução é extremamente complexa (Li, Özsu, Chen, & Ooi, 2014).

*Real Time* está interligado com *Big Data* no “V” correspondente à velocidade, ou seja, não faz sentido obter inteligência em um momento posterior à necessidade da empresa (Mohamed &

Al-Jaroodi, 2014). *Real Time* é um conceito simples e de fácil entendimento. Sua aplicação, por outro lado, é um enorme desafio computacional, por tratar-se de uma variável que depende de ambientes, infraestruturas, configurações, etc.

De forma resumida, Inteligência Analítica em tempo real significa a leitura de várias fontes de dados, uma análise para consolidação em dados abstratos e uma forma de modelagem para repassar ao usuário, tudo isso da forma mais otimizada possível e obedecendo à uma forte restrição temporal, e nunca haverá uma solução totalmente em tempo real, ou seja, com zero de atraso entre a pergunta e a resposta (Tien, 2012).

#### 7.1.8 *Structured Data*

O conceito de *Structured Data* é simples para profissionais da área de computação. Trata-se de um conceito que representa a organização para armazenamento e recuperação de dados em diferentes formatos, como por exemplo, os bancos de dados relacionais. Dados desse tipo são fáceis no aspecto de entrada nos sistemas, recuperação, armazenamento e análise, e são amplamente utilizados nos sistemas transacionais, presentes em praticamente todas as empresas. (Hashem, Yaqoob, Anuar, Mokhtar, Gani, & Khan, 2015).

Uma tecnologia criada para tratamento e manipulação de dados é o SQL (Lomotey & Deters, 2014), que significa *Structured Query Language*, ou seja, uma linguagem criada para padronizar operações em dados estruturados. Técnicas mais avançadas de Inteligência Analítica também precisam contemplar dados estruturados. São exemplos de dados estruturados: informações demográficas, informações transacionais e informações de cadastro (Gruhl, Nagarajan, Pieper, Robson, & Sheth, 2010).

#### 7.1.9 *Unstructured Data*

O conhecimento *Unstructured Data* (em português dados não estruturados), aparece bem mais na literatura de Inteligência Analítica do que *Structured Data* (em português dados estruturados): os dados não estruturados são mais recentes e estão mais relacionados com a

realidade de Inteligência Analítica, podendo ser considerados parte do terceiro “V” - Variedade (Lomotey & Deters, 2014). De forma muito básica, o maior desafio é encontrar e tratar dados diferentes, com pouco padrão e em diferentes locais de armazenamento. Dados não estruturados são geralmente desprovidos de um modelo de dados e não se encaixam bem em bancos de dados relacionais (Bakshi, 2012). São exemplos de dados não estruturados: vídeos, filmes, fotos, mapas, *e-mails*, Tweets, etc, que são normalmente gerados por humanos, não por sistemas (Khan, Yaqoob, Hashem, Inayat, Mahmoud Ali, Alam, & Gani, 2014).

Analogamente aos bancos de dados relacionais e a linguagem SQL criada para tratamento de dados estruturados, foram criados bancos de dados não relacionais (por exemplo, o NoSQL), que possui uma estrutura menos rígida e permite adicionar e remover elementos sem ter que atualizar Tabelas inteiras (Bakshi, 2012). Existem técnicas e ferramentas para análise de dados não estruturados, como por exemplo o MapReduce.

Atualmente, massiva quantidade de dados não estruturados são gerados em redes sociais (Gandomi & Haider, 2015). Por mais que existam alguns padrões para obtenção desses dados, sua análise permanece um grande desafio. Grande parte do que é chamado *Big Data*, que já discutiu-se ser a base para Inteligência Analítica, são dados não estruturados. Logo, pode-se aferir que grande parte da Inteligência Analítica consiste em analisar dados não estruturados.

## 7.2 Habilidades analíticas

### 7.2.1 *Classification*

*Classification* é tipo de método, ou algoritmo, que visa classificar determinado dado em grupos pré-definidos. A principal medida de performance para *Classification* é a precisão com que acerta cada nova classificação, geralmente calculado através da divisão do número de acertos pelo número total de testes (Ma, Chen, & Wei, 2013). Um exemplo muito conhecido e utilizado são as SVM - *Support Vector Machine* (Zhai, Ong, & Tsang, 2014). Outros exemplos são: regressão logística, regressão logística com stepwise, random forests e regressão logística condicional (Zhu, Lou, Zhou, Ballester, Kong, & Parikh, 2015).

Para alguns autores, *Classification* é parte do *Data Mining*, pois para encontrar padrões em bases de dados, é necessário criar regras e então classificar as informações, de forma a entender se atendem a determinados critérios (Fernández, del Río, López, Bawakid, del Jesus, Benítez, & Herrera, 2014). Para Chen, Chiang e Storey (2012), *Classification* é uma tecnologia fundamental para Inteligência Analítica, por ser uma de suas fases (Khan, Yaqoob, Hashem, Inayat, Mahmoud Ali, Alam, & Gani, 2014). A definição de qual método será utilizado para *Classification* depende muito da amostra disponível para criação das regras de classificação (Varshney & Mojsilović, 2011).

### 7.2.2 *Clustering*

*Clustering* é o processo de agrupamento de dados por algum tipo definido de similaridade, um processo posterior ao processo de *Classification*, com objetivo de diminuir os erros de predição, realizados sobre esses dados (Stefanovic, Stefanovic, & Radenkovic, 2008). *Clustering* facilita, por exemplo, na predição baseada em dados de clientes, pois os agrupa e organiza (Simmhan & Noor, 2013), outro exemplo é o uso de *Clustering* para dividir dados de vendas e predizer padrões com maior confiabilidade (Stefanovic, Stefanovic, & Radenkovic, 2008).

Uma dificuldade comum em *Clustering* é a dificuldade em se obter exemplos para treinamento nas categorias, principalmente quando as categorias são dinâmicas e mudam com a necessidade do negócio (Hu, Singh, & Mojsilovic, 2008).

O processo de *Clustering* envolve classificações baseadas em probabilidade. Até recentemente, não havia pesquisas sobre como classificar casos duvidosos, que não se encaixam perfeitamente nas categorias definidas (Pei, 2013). Não foram identificados textos na literatura que apresentam soluções para classificações duvidosas.

### 7.2.3 *Forecasting*

*Forecasting* adiciona a dimensão temporal ao *Predictive Analytics* (Simmhan & Noor, 2013). É uma análise com objetivo voltado para previsão do futuro, enquanto *Predictive Analytics*

pode ajudar a entender o passado (Waller & Fawcett, 2013a). *Forecasting* pode ser confundido com planejamento baseado em informações históricas, mas é melhor definido como algo necessário para tomada de decisão.

Alguns autores alertam que *Forecasting* é algo que traz muitas expectativas, mas que os decisores devem ser conservadores e por isso não devem confiar cegamente nos resultados, uma vez que a validade de métodos e resultados sempre é questionável e muitas vezes com utilização quase única (Armstrong, Green, & Graefe, 2015). Geralmente utiliza-se dados próximos temporalmente próximos ao acontecimento que se deseja prever (Chou & Telaga, 2014).

As utilizações mais comuns para *Forecasting* são para clima, consumo de energia e previsão de demanda (Simmhan & Noor, 2013).

#### 7.2.4 Modeling

Dentre os textos estudados, não há um texto que ofereça diretamente o conceito de *Modeling*, mas há textos que explicam sua utilização e finalidade. Trata-se do algoritmo que representa a maneira que uma organização coleta, atualiza e armazena as informações. Para Tien (2013), *Modeling* é uma fase posterior a coleta de informações, que indica como as informações serão consideradas e organizadas até a tomada de decisão.

Outros autores explicam que *Modeling* se refere a capacidade de produzir modelos para predição (Xu, Duan, & Whinston, 2014), tanto que em alguns estudos encontra-se o termo *Predictive Modeling* (Chen, Chiang, & Storey, 2012). Um modelo pode ser criado de forma manual, ou parcialmente automatizada, através de *Machine Learning* (Shin, Woo, & Rachuri, 2014).

Há um estudo que sugere que em cenários muito complexos ou em que se tem pouco conhecimento das variáveis, o *Modeling* torna-se pouco prática e de difícil uso. Nesse caso, indica-se o uso de simulações (*Simulation*) para obter relações válidas em uma determinada base de dados (Li & Kauffman, 2012).

### 7.2.5 *Neural Network*

*Neural Networks* são algoritmos que fazem parte de *Machine Learning* (Chae & Olson, 2013; Berger & Doban, 2014). São complexos e simulam o funcionamento do cérebro humano na resolução de problemas, com funções simulando cada neurônio conectado (Chou & Telaga, 2014). Este tipo de algoritmo costuma ser utilizado em situações em que se precisa de aprendizado com relacionamentos complexos (Shin, Woo, & Rachuri, 2014). *Neural Networks* envolve entradas, saídas e uma camada oculta (Chen & Yang, 2014).

Na base de textos estudada não há material que permita um melhor entendimento sobre o conceito e utilização *Neural Network*, o que restringe também o entendimento dessa competência frente aos desafios de Inteligência Analítica.

### 7.2.6 *Optimization*

*Optimization* consiste em alterar a maneira como algo é realizado (algoritmos, configurações e estratégias de busca e processamento), de forma a obter melhores resultados, por exemplo diminuir a quantidade de dados resultante de determinada operação de busca (Pedersen, Pedersen, & Riis, 2013).

Algumas técnicas de *Optimization* são: organizar dados não estruturados em colunas, replicar Tabelas de meta dados mais utilizadas e realizar iterações sobre blocos, ao invés de sobre linhas de Tabelas. Ações mais simples incluem incluir processamentos paralelos, reordenar operações e agrupar operações pequenas em uma única tarefa. *Optimization* pode ser usado em diversas dimensões de Inteligência Analítica, como nos programas de MapReduce, por exemplo (Doulkeridis & Nørnvåg, 2014).

Por geralmente envolver aspectos lógicos e matemáticos, *Optimization* quase sempre envolve o uso de métodos matemáticos para definir as melhores opções para cada situação. Há casos em que as alterações são muito pequenas, mas trazem grandes resultados (Slavakis, Kim, Mateos, & Giannakis, 2014).

### 7.2.7 *Predictive Analytics*

*Predictive Analytics* é uma parte de *Data Science*, que se diferencia da estatística por ser, além de quantitativa, também qualitativa. Além de poder ser utilizada para previsões futuras, é possível prever o que teria acontecido no passado, em diferentes condições (Waller & Fawcett, 2013a). Autores destacam que *Predictive Analytics* é considerado algo pouco apoiado em teorias, que busca respostas, mas sem explicações detalhadas. Mesmo assim, este é um assunto cada vez mais discutido nas empresas e na academia (Waller & Fawcett, 2013b).

O principal objetivo de *Predictive Analytics* é capturar padrões em bases de dados, bem como capturar relacionamentos entre variáveis. Pode-se dividir as técnicas preditivas em dois grupos: aquelas que procuram padrões em dados históricos e tentam extrapolar esses padrões para o futuro; e aquelas que tentam capturar interdependência entre variáveis e usar para tentar fazer previsões. Alguns cientistas definem que *Predictive Analytics* é a nova estatística, que é mais adequada para a era do *Big Data* e desconsidera problemas de amostragem (Gandomi & Haider, 2015).

Existe expectativa revolucionária, por parte das áreas de negócio, nos resultados positivos subsequentes a adoção de *Predictive Analytics*. Por exemplo, para a área de saúde, existem autores que definem *Predictive Analytics* como o uso de algoritmos eletrônicos para prever eventos futuros em tempo real (Cohen, Amarasingham, Shah, Xie, & Lo, 2014). Na prática, é muito difícil cumprir a promessa embutida nessa definição.

### 7.2.8 *Regression*

Regressão (*Regression*) é uma técnica estatística que busca correlações entre variáveis (Khan, Yaqoob, Hashem, Inayat, Mahmoud Ali, Alam, & Gani, 2014), mas no contexto de Inteligência Analítica, é um algoritmo englobado pelo conceito de *Data Mining*, utilizado especialmente para obtenção de respostas para questões baseadas em dados históricos (Varshney & Mojsilović, 2011). Existem vários tipos de regressão, com diferentes níveis de complexidade e com diferentes aplicações (Jin, Wu, Vidyanti, Di Capua, & Wu, 2015). Por exemplo, regressão



logística é um algoritmo utilizado para classificação, e é considerado um algoritmo de *Data Mining* (Fernández, del Río, López, Bawakid, del Jesus, Benítez, & Herrera, 2014).

Na base de textos não há muito material que permita conceituar regressão, mas é um termo que aparece em dezenas de textos da base estudada, o que destaca o quão comum é dentro do contexto de Inteligência Analítica.

#### 7.2.9 *Sentiment Analysis*

Técnicas de análise de *Sentiment Analysis* são adotadas em estudos em mídias sociais, o que tem se tornado frequente, pois as empresas têm realizado tais estudos para obter *insights* e consequente vantagem competitiva (He, Wu, Yan, Akula, & Shen, 2015). Essas técnicas podem ser definidas como procedimentos de *Text Mining*, que permitem descobertas sobre elementos textuais que contenham emoções humanas (Abrahams, Jiao, Wang, & Fan, 2012), como opiniões, sentimentos, emoções e outros tipos de subjetividade (He, Wu, Yan, Akula, & Shen, 2015). Outro nome para *Sentiment Analysis* é *Opinion Mining*, e essas técnicas geralmente são aplicadas para que as empresas entendam a opinião de seus consumidores (Gandomi & Haider, 2015).

Basicamente, inicia-se por uma análise de palavras-chave e são atribuídos valores emocionais, por exemplo, positivo ou negativo. Depois, cada ocorrência é estudada em seu contexto (frase) e são atribuídos valores emocionais nesse conjunto maior. Geralmente procura-se os extremos positivos e negativos, através de *Machine Learning*, ou seja, é necessário treinar a solução para reconhecer os extremos (He, Wu, Yan, Akula, & Shen, 2015).

Uma mídia social muito estudada com o uso de *Sentiment Analysis* é o Twitter, que é uma rede social que reúne opiniões e informações textuais. Um exemplo de *software* utilizado para esse tipo de análise é o SentiStrenght (Chae, 2015).

### 7.2.10 *Simulation*

*Simulation* envolve a criação artificial de dados, através da simulação de um processo (Jourdan, Rainer, & Marshall, 2008). Para problemas que envolvem variáveis de difícil previsão, como variáveis influenciadas por aspectos temporais sem padrão, o uso de *Optimization* não é indicado, pois é difícil estabelecer uma fórmula matemática otimizada. Nesses casos, indica-se o uso de *Simulation* (Li & Kauffman, 2012).

Com o crescente aumento de performance computacional, *Simulation* tem evoluído para um patamar superior de fidelidade e complexidade, tornando-se particularmente útil para contextos com muitas variáveis, climatologia, por exemplo. No contexto de climatologia, *Simulation* é usado para verificar o impacto de alteração em variáveis, para tentar identificar mudanças abruptas no sistema (Steed, Ricciuto, Shipman, Smith, Thornton, Wang, & Williams, 2013). No contexto *Business*, uma variável importante é a qualidade de serviço (do inglês *Quality of Service*, QoS), que reflete a performance de sistemas (Chen, Li, & Wang, 2015).

### 7.2.11 *Statistical Analysis*

No texto mais citado na base estudada Chen, Chiang e Storey (2012) definem análise estatística (*Statistical Analysis*) como a aplicação de técnicas avançadas de estatística, como regressão, análise de fator, *Clustering* e análise discriminante. O limite entre *Data Mining* e análise estatística não é claro e entende-se que os dois andam juntos, a ponto de ser definido o termo *Analytics*, como união dos dois (Chen, Chiang, & Storey, 2012; Khan, Yaqoob, Hashem, Inayat, Mahmoud Ali, Alam, & Gani, 2014).

Análise estatística baseia-se em teoria estatística, que consiste em modelar incertezas de acordo com a teoria de probabilidades. Em uma definição mais simples, análise estatística é algo que serve para descrever e inferir *Big Data* (Khan, Yaqoob, Hashem, Inayat, Mahmoud Ali, Alam, & Gani, 2014).

Para alguns autores, análise estatísticas são particularmente úteis para definir o modelo que se utilizará para realizar *Data Mining* (Afshari & Peng, 2015), e deve-se olhar para variância da base de dados, visto que o olhar clássico para *p-value* pouco vale em uma base grande de dados

(Xiang, Schwartz, Gerdes, & Uysal, 2015). Muitos *softwares* podem ser utilizados para essas análises, como por exemplo o D3, SPSS, Matlab, Mathematica, Microsoft Excel, Stata e Weka (Louridas & Ebert, 2013).

### 7.3 Habilidades computacionais

#### 7.3.1 *Data Mining*

*Data Mining* é um conjunto de algoritmos utilizados para navegar, de forma automatizada, em um conjunto de dados, identificar padrões, planejar regras e fazer previsões para o futuro (Fuchs, Höpken, & Lexhagen, 2014). A identificação de padrões é muito importante para tomada de decisão, por revelar oportunidades de melhoria (Bose, 2009). O processo básico de *Data Mining* consiste em uma análise humana dos dados para definir um modelo, treinar o modelo com uma amostra de dados, testar o modelo e realizar previsões através desse modelo, um processo geralmente iterativo (Stefanovic, Stefanovic, & Radenkovic, 2008).

Inteligência Analítica inclui a elaboração de relatórios, visualizações e *Data Mining*, sendo que o objetivo é prover *insights* para usuários de negócio. Deve-se ter cuidado para não confundir correlações com causalidades (Kohavi, Mason, Parekh, & Zheng, 2004). São tipos de *Data Mining*: métodos estatísticos, raciocínio baseado em casos, redes neurais, árvores de decisão, algoritmos genéticos e conjuntos difusos (Bose, 2009).

*Data Mining* torna-se especialmente complexo quando são incluídos dados não estruturados (Lomotey & Deters, 2014). Existem muitos *softwares* comerciais que foram criados com o intuito de permitir *Data Mining* no contexto de Inteligência Analítica, como por exemplo o Clementine, Oracle *Data Mining*, SAS Text Miner, etc (Bose, 2009).

*Data Mining* é uma tecnologia complexa que é composta de diversas outras, como redes neurais, *Text Mining*, *Web Mining*, séries temporais, regressão, etc. Muitos dos termos que são considerados dentro do conceito *Data Mining*, apareceram neste trabalho como habilidades, foram validados e há seções subsequentes para explica-los.

### 7.3.2 *Data Warehousing*

*Data Warehousing* é conceituado como um processo de gerenciamento e recuperação de dados, que serve de suporte para soluções de Inteligência Analítica, e que inclui tecnologias de ETL (*Extract, Transform, Load*), funções para inserir e remover dados de sistemas transacionais (ERP, CRM, etc), funções para limpeza de dados, no sentido de remover automaticamente erros e valores inesperados e funções para completar informações (Bose, 2009; Chae & Olson, 2013).

Apesar das tecnologias ETL terem evoluído bastante, o *Data Warehousing* não é, e nunca poderá ser em tempo real (Li, Özsu, Chen, & Ooi, 2014). Por isso, é importante que se entenda como funciona e de que forma priorizar as configurações, conforme necessidade.

O principal objetivo de *Data Warehousing* é integrar dados de diferentes sistemas. Esse procedimento costuma revelar problemas de qualidade dos dados, como redundância, por exemplo, o que exige ações de limpeza e padronização, também abrangidas pelo *Data Warehousing* (Chae & Olson, 2013). Consiste em dados propriamente ditos e meta dados, que suportam o uso de tecnologias analíticas (Fernández, del Río, López, Bawakid, del Jesus, Benítez, & Herrera, 2014).

### 7.3.3 *ETL (Extract, Transform, Load)*

ETL é uma sigla que significa *Extract, Transform, Load*, que são as fases para busca, consolidação e correção de dados com finalidade de Inteligência Analítica. Estima-se que 70% do esforço da Inteligência Analítica resida em ETL. É um conceito anterior ao ano 2000, mas que tem evoluído em conjunto com as técnicas avançadas de Inteligência Analítica (Agarwal, Shroff, & Malhotra, 2013; Mahmood & Afzal, 2013; Demirkan & Delen, 2013). Para Chen, Chiang e Storey (2012), o papel de ETL é garantir a integração dos dados para que se possa trabalhar com eles, tornando ETL uma tecnologia base para Inteligência Analítica.

Operações de ETL podem ser aplicadas em dois níveis: em um nível genérico, sobre uma base de dados inteira, ou sobre uma parte específica da base de dados, que pode ser posteriormente processada por técnicas mais avançadas de Inteligência Analítica. Um exemplo de *software* de ETL com tais funcionalidades é o Pentaho (Mahmood & Afzal, 2013).

Através de ETL, dados presentes em sistemas transacionais devem ser periodicamente exportados para sistemas OLAP (Li, Özsu, Chen, & Ooi, 2014). A definição de OLAP encontra-se em uma seção posterior neste trabalho.

#### 7.3.4 *Machine Learning*

*Machine Learning* é um ramo da área de Inteligência Artificial que estuda sistemas capazes de aprender e criar regras a partir de dados (Berger & Doban, 2014). A evolução de *Machine Learning* é um dos pilares que sustenta os constantes avanços em Inteligência Analítica (Fernández, del Río, López, Bawakid, del Jesus, Benítez, & Herrera, 2014). São técnicas utilizadas para aprendizado automático para melhorar regras e construir modelos tolerantes a falhas, com objetivo de aplicar em dados históricos e novos dados em tempo real (Li, Parikh, He, Qian, Li, Fang, & Hampapur, 2014). A principal diferença entre *Data Mining* e *Machine Learning* é justamente o aprendizado com *Data Mining* focar em propriedades desconhecidas de dados, enquanto *Machine Learning* trabalha com dados ajustáveis, porém conhecidos (Berger & Doban, 2014).

Operações de *Machine Learning* podem ser tão complexas a ponto de se tornarem de difícil entendimento para seres humanos. Para contornar essa dificuldade e melhorar a manutenção para uso em Inteligência Analítica, geralmente extrai-se regras localizadas das complexas saídas dos algoritmos, ou utiliza-se técnicas de visualização para extrair regras fáceis de entender e implementar (Li, Parikh, He, Qian, Li, Fang, & Hampapur, 2014).

Existem dois tipos de algoritmos de *Machine Learning*: aqueles supervisionados, ou seja, com intermédio humano e com uma variável estudada pré-definida; e aqueles não supervisionados. Os algoritmos supervisionados ainda se dividem nos algoritmos de classificação (*Classification*), para dados qualitativos, e de regressão (*Regression*), para dados quantitativos.

### 7.3.5 MapReduce

O MapReduce é uma ferramenta para processamento distribuído e paralelo de massivas quantidades de dados, sem a necessidade de escrever código para isso. As principais qualidades do MapReduce são a escalabilidade, tolerância a falhas, facilidade em programa-lo e flexibilidade (Doulkeridis & Nørnvåg, 2014).

MapReduce e Hadoop então inseridos nos desafios de ETL (Agarwal, Shroff, & Malhotra, 2013). Hadoop e Google MapReduce são implementações do MapReduce. MapReduce consiste em funções de Map, nas quais os usuários entram com as regras de busca ou execução a serem executadas em cada um dos computadores participantes; e em funções de Reduce, nas quais o programa retorna, de forma organizada, os resultados em cada um dos computadores participantes (Shang, Jiang, Hemmati, Adams, Hassan, & Martin, 2013; Fernández, del Río, López, Bawakid, del Jesus, Benítez, & Herrera, 2014).

Assim como este estudo consistiu na análise de frequência de palavras em 208 textos, o mesmo já foi realizado para milhares de textos, ou para livros inteiros, com ajuda do MapReduce (Mohammed, Far, & Naugler, 2014).

### 7.3.6 OLAP (*Online Analytical Processing*)

Sistemas OLAP (*Online Analytical Processing*) permitem a análise de grandes quantidades de dados de bancos de dados de sistemas transacionais. Também é possível incluir dados externos através de interfaces XML (*eXtensible Markup Language*) (Pedersen, Pedersen, & Riis, 2013). A ideia do OLAP é prover análises descritivas para criar relatórios online (Fuchs, Höpken, & Lexhagen, 2014).

É comum que sistemas OLTP (*Online Transaction Processing*) sejam usados para inserir, atualizar e deletar dados transacionais, que por sua vez são periodicamente exportados para sistemas OLAP através de ferramentas de ETL. Para permitir que análises sejam feitas de forma online (tentando-se chegar próximo ao *Real Time*), pode-se utilizar, por exemplo, o MapReduce (Li, Özsu, Chen, & Ooi, 2014).

O que difere OLAP de sistemas de banco de dados comuns é que a informação é organizada de forma multidimensional, conforme necessidades de análise (Chae & Olson, 2013). As vantagens de tal organização são: classificação automática, filtrar informação de forma visual e performance superior (Pedersen, Pedersen, & Riis, 2013).

### 7.3.7 *Programming*

*Programming* é uma competência que trata do domínio de diferentes linguagens e ferramentas de programação, com objetivo de conseguir utilizar as atuais tecnologias relacionadas à Inteligência Analítica.

Frente aos desafios relacionados aos requisitos da era do *Big Data*, surgiram ferramentas de programação (por exemplo, o MapReduce) para permitir processamentos em paralelo, com a utilização de muitos computadores comuns, ao invés de um supercomputador (Mohammed, Far, & Naugler, 2014). Tais ferramentas permitem a execução de tarefas programadas em linguagens particulares. São exemplos de linguagem de busca: Hive (Facebook), Pig (Yahoo), Jaql (IBM), Dremel (Google), Scope (Microsoft) (Mohammed, Far, & Naugler, 2014).

Algumas tecnologias de primeira geração utilizadas para análise de dados e análises estatísticas, que são encontradas na literatura são SAS e R. Essas tecnologias evoluíram, de forma a acompanhar o Hadoop/MapReduce. A segunda geração é constituída do Mahout, Pentaho e Cascading, e a terceira geração do Spark, GraphLab, Pregel e Giraph (Mohammed, Far, & Naugler, 2014).

Destaca-se que nenhuma linguagem de programação ou ferramenta atende às necessidades para solução de todos os problemas relacionados à Inteligência Analítica. Dado o número de tecnologias e a velocidade com que evoluem, muitos praticantes escolhem utilizar as tecnologias mais difundidas, mesmo que não se adequem perfeitamente ao problema que se tenta resolver (Lin, 2013).

### 7.3.8 *Web Mining*

Utilizado geralmente por empresas com foco em consumidores que utilizam *Web*, *Web Mining* é um ramo dentro de *Data Mining* cujo foco é procurar padrões em dados de servidores *Web*. Existem três categorias de *Web Mining*: estudos de conteúdo *Web*, estudos de estrutura *Web* e estudos de uso de *Web* (comportamento humano na *Web*) (Bose, 2009), como cliques, por exemplo (Chen, Chiang, & Storey, 2012).

Um dos maiores problemas relacionados a *Web Mining* é algo que foge do aspecto técnico e entra no aspecto de privacidade: os dados *Web* analisados e os resultados obtidos respeitam a privacidade dos usuários que os geraram? (Brown, Famili, Paass, Smith-Miles, Thomas, Weber, & Maldonado, 2011).

## 7.4 Comportamentos

### 7.4.1 *Communicate*

A habilidade de comunicação de empresa para cliente, empresa para empresa, e cliente para cliente torna-se cada vez mais fácil e necessária na era do *Big Data*, a ponto de se utilizar o termo comunicação ubíqua (Lemon & Huang, 2011; Rust & Huang, 2014).

A principal função de uma comunicação adequada no contexto de Inteligência Analítica é explicitar os valores básicos da empresa, o propósito e a direção da organização, de forma que essas variáveis sejam consideradas no momento de decidir estratégias de IA (Elbashir, Collier, & Sutton, 2011). Outra dimensão importante da comunicação está no sentido de reportar aprendizados e implicações sobre análises de dados para os demais interessados na empresa, em uma linguagem que todos entendam. Tal comunicação pode ser realizada verbalmente, ou através de *Data Visualization* (Davenport & Patil, 2012).

Conforme senso comum e também conforme Davenport (2014), profissionais com habilidades analíticas não costumam ser conhecidas por serem bons comunicadores. Mesmo assim, esse é um ponto crucial para obtenção de bons resultados em Inteligência Analítica.



#### 7.4.2 *Creative*

Há pouco material disponível sobre essa competência comportamental para atuação com Inteligência Analítica. Apenas alguns dos estudos analisados citam essa competência de forma explícita. A criatividade em ambiente profissional pode resultar em vantagem competitiva, e empresas sem profissionais com essa competência podem ter dificuldade em obter resultados com Inteligência Analítica (Erevelles, Fukawa, & Swayne, 2016).

Além da utilização da criatividade necessária para decidir quais tecnologias e ferramentas utilizar e como, a obtenção de Inteligência Analítica também requer criatividade para *Data Visualization* (Davenport & Patil, 2012). Também há material que sugere que por mais que o uso de Inteligência Analítica seja planejado, os resultados são positivamente afetados quando há criatividade e curiosidade do profissional envolvido (Sharma, Mithas, & Kankanhalli, 2014). Por fim, acredita-se que a criatividade também seja importante no momento de decidir quais bases de dados são interessantes para obtenção de Inteligência Analítica (Barton & Court, 2012).

## 8 CONCLUSÕES

O objetivo geral de apontar e descrever, segundo a literatura, uma base unificada de competências para atuação em Inteligência Analítica foi cumprido. Foram apontadas e descritas 30 competências, sendo 9 conhecimentos, 19 habilidades (11 habilidades analíticas e 8 habilidades computacionais) e 2 comportamentos. A Figura 10 é uma representação visual dos três tipos de competências.

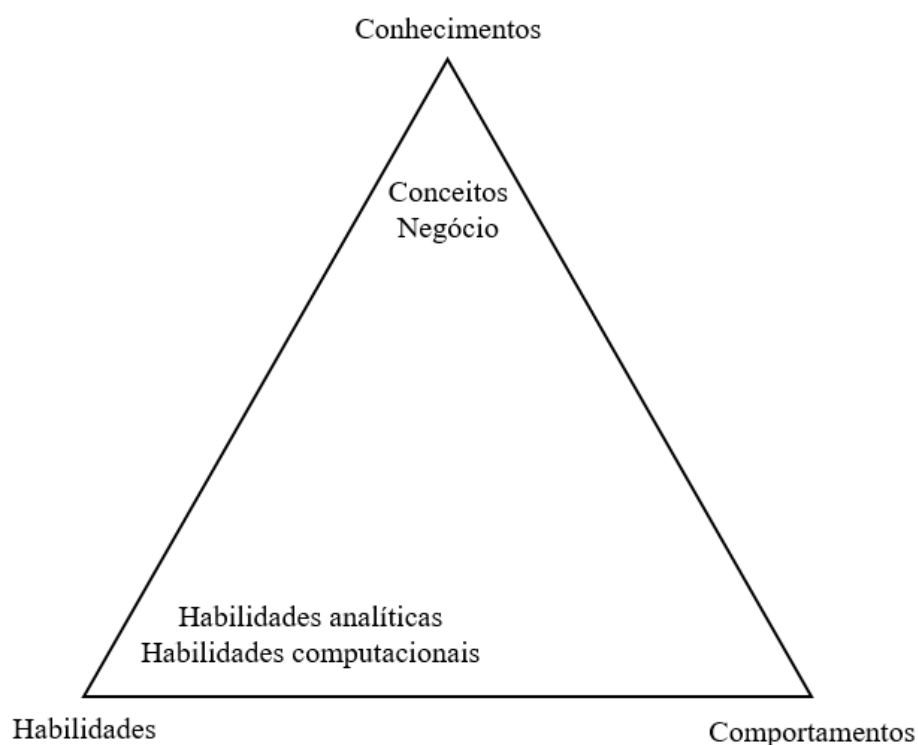


Figura 10: Representação visual dos grupos de competências

Fonte: Traduzido e adaptado de Le Deist e Winterton, 2005, no contexto de IA

As competências apontadas nesse trabalho formam um conjunto mais completo que os observados nos textos escopo deste estudo, o que demonstra que o resultado do trabalho pode colaborar com pesquisadores e gerentes de negócio que precisem conhecer e entender o conjunto de competências para atuação em Inteligência Analítica. Para pesquisadores, o resultado deste trabalho serve como uma base unificada de competências e suas descrições, o que poderia ser utilizado para realizar inúmeros estudos mais avançados e voltados para o uso

de Inteligência Analítica. Para gerentes de negócio, o resultado deste trabalho poderia ser utilizado para seleção de profissionais para atuação em Inteligência Analítica, ou para desenvolver treinamentos internos, entre outros.

Os resultados de cada fase do trabalho foram obtidos através de métodos científicos e baseados em teorias e métodos. Mesmo assim, pode haver subjetividade, por exemplo, ao mapear competências na literatura, por isso optou-se por realizar uma validação com especialistas, que resultou na eliminação de poucas competências. Como o resultado da utilização inicial dos métodos foi quase totalmente validado pelos especialistas, há um reforço sobre a acurácia e validade desses métodos e utilizações.

Pelo resultado obtido, percebe-se que na literatura há muita informação sobre conhecimentos e habilidades, mas pouco material sobre comportamentos. Isso pode significar que as competências comportamentais são genéricas em relação à outras funções na empresa, ou que o assunto ainda não foi estudado o suficiente para explorar tais competências. Sobre as competências de conhecimentos, observou-se que há certa convergência, ou seja, a maioria aparece em muitos textos e de forma repetida. O mesmo não foi observado para competências de habilidades, que formam um subconjunto expressivamente maior que o de conhecimentos, e que contém competências mais pulverizadas na literatura, com pouca convergência.

A frequência de cada competência na literatura está representada no Quadro 5. A coluna “# de textos” representa o número de textos em que a competência foi encontrada. A coluna “# relativo de textos” representa a porcentagem de quantos textos a competência foi encontrada em relação ao total de 206 textos analisados. A coluna “Média Validação” foi importada do Quadro 4 para comparação entre a média da validação dos especialistas e a frequência de determinada competência na literatura. O Quadro 5 está ordenado em ordem decrescente em relação a frequência de cada competência na literatura.

A competência Negócio, por ser composta de um total de 11 termos agrupados (*Accounting, Economics, Finance, Fraud Detection, Healthcare, Human Resources, Logistics, Marketing, Operations Management, Sales* e *Supply Chain Management*), não foi considerado na elaboração do Quadro 5.

Competência	# de textos	# relativo de textos	Média Validação
Conhecimento - <i>Big Data</i>	149	72,33%	4,67
Conhecimento - <i>Real Time</i>	90	43,69%	3,5
Habilidade analítica - <i>Modeling</i>	85	41,26%	4,67
Habilidade analítica - <i>Optimization</i>	82	39,81%	4,33
Conhecimento - <i>Cloud</i>	77	37,38%	3,83
Conhecimento - <i>Decision Making</i>	73	35,44%	4,17
Habilidade computacional - <i>Data Mining</i>	73	35,44%	4,83
Habilidade computacional - <i>Programming</i>	61	29,61%	4
Habilidade computacional - MapReduce	48	23,30%	3,67
Habilidade analítica - <i>Clustering</i>	45	21,84%	4,5
Habilidade computacional - <i>Machine Learning</i>	44	21,36%	4,17
Habilidade analítica - <i>Regression</i>	43	20,87%	4,83
Habilidade analítica - <i>Simulation</i>	42	20,39%	4,33
Comportamento - <i>Communicate</i>	40	19,42%	4,33
Habilidade analítica - <i>Forecasting</i>	38	18,45%	4,33
Conhecimento - <i>Unstructured Data</i>	32	15,53%	3,83
Habilidade analítica - <i>Predictive Analytics</i>	30	14,56%	4,5
Habilidade analítica - <i>Classification</i>	28	13,59%	4,5
Habilidade computacional - OLAP ( <i>Online Analytical Processing</i> )	24	11,65%	3,83
Conhecimento - <i>Structured Data</i>	22	10,68%	4
Habilidade computacional - ETL ( <i>Extract, Transform, Load</i> )	18	8,74%	4,33
Conhecimento - <i>Decision Support Systems</i>	16	7,77%	4
Habilidade computacional - <i>Data Warehousing</i>	16	7,77%	4,33
Conhecimento - <i>Data Science</i>	12	5,83%	4,33
Habilidade analítica - <i>Statistical Analysis</i>	11	5,34%	4,83
Habilidade analítica - <i>Neural Network</i>	10	4,85%	4,17
Habilidade analítica - <i>Sentiment Analysis</i>	9	4,37%	3,83
Comportamento - <i>Creative</i>	8	3,88%	4,5
Habilidade computacional - <i>Web Mining</i>	3	1,46%	4

Quadro 5: Frequência das competências na base de 206 textos

Fonte: Elaboração Própria

Conforme Quadro 5, as 10 competências com maior frequência na literatura são: *Big Data*, *Real Time*, *Modeling*, *Optimization*, *Cloud*, *Decision Making*, *Data Mining*, *Programming*, MapReduce e *Clustering*. Esta lista é composta por 4 competências do tipo conhecimento, 3 competências do tipo habilidade analítica e 3 competências do tipo habilidade computacional, o que indica que não há predominância de um tipo de competência na literatura. Para reforçar

essa não predominância de um tipo de competência, observa-se que das 10 competências com menor frequência na literatura (*Structured Data*, ETL, *Decision Support Systems*, *Data Warehousing*, *Data Science*, *Statistical Analysis*, *Neural Network*, *Sentiment Analysis*, *Creative* e *Web Mining*), 3 são conhecimentos, 3 são habilidades analíticas, 3 são habilidades computacionais e 1 é comportamento.

Os comportamentos *Communicate* e *Creative* aparecem sem grande destaque. Isso significa que os pesquisadores estudam mais os conhecimentos e habilidades, em relação aos comportamentos.

Por fim, destaca-se que *Big Data* é a competência com maior frequência. Isso demonstra que praticamente toda a literatura de Inteligência Analítica tange esse fenômeno, que é considerado por muitos como o contexto sobre o qual Inteligência Analítica faz sentido atualmente.

Para completar a discussão, optou-se pela criação do Quadro 6, no qual são discutidas as competências levantadas, validadas e descritas neste trabalho. A discussão de cada competência consiste em tentar explicar, conforme descrições na literatura, como cada uma delas se relaciona com o fenômeno Inteligência Analítica. O Quadro 6 foi ordenado por tipo de competência e depois pelo número relativo de textos, ou seja, a frequência nos 206 textos analisados. À competência negócio, por ser composta de outros termos mais específicos, atribuiu-se “N/A” (não aplicável) na coluna “# relativo de textos”.

Competência	# relativo de textos	Discussão
Conhecimento - <i>Big Data</i>	72,33%	<i>Big Data</i> é o cerne da Inteligência Analítica (George, Haas, & Pentland, 2014). Tanto que a expressão <i>Big Data Analytics</i> pode ser vista como equivalente para Inteligência Analítica (Chen, Chiang e Storey, 2012)
Conhecimento - <i>Real Time</i>	43,69%	<i>Real Time</i> é um requisito desejável para as soluções de Inteligência Analítica, pois consiste em entregar informações imediatas para a pessoa correta (Bose, 2009)
Conhecimento - <i>Cloud</i>	37,38%	<i>Cloud</i> é uma opção mais escalável e inteligente (Demirkan & Delen, 2013) para empresas que desejam investir em Inteligência Analítica, por reduzir muitos dos custos para processamento de massivo de dados e reduzir alguns riscos: a tolerância à falha e redundância (Hashem, Yaqoob, Anuar, Mokhtar, Gani, & Khan, 2015), por exemplo
Conhecimento - <i>Decision Making</i>	35,44%	<i>Decision Making</i> é o foco com que se operacionaliza Inteligência Analítica (Bose, 2009)
Conhecimento - <i>Unstructured Data</i>	15,53%	<i>Unstructured Data</i> precisa ser entendido pois é o núcleo do problema das análises de Inteligência Analítica. Praticamente tudo que é produzido pelo ser humano é dado não estruturado, ou semiestruturado (Negash, 2004)
Conhecimento - <i>Structured Data</i>	10,68%	<i>Structured Data</i> é amplamente encontrado nas empresas e os profissionais que forem atuar com Inteligência Analítica se depararão com grandes volumes de dados desse tipo (Holsapple, Lee-Post, & Pakath, 2014)

Conhecimento - <i>Decision Support Systems</i>	7,77%	Os DSS são utilizados por usuários de negócio e por gestores para tomada de decisão (Chang, Kauffman, & Kwon, 2014)
Conhecimento - <i>Data Science</i>	5,83%	<i>Data Science</i> é um conhecimento necessário por ser um conceito abrangente e que contempla várias técnicas para análise de dados (Waller & Fawcett, 2013a), o que está intrinsecamente ligado a Inteligência Analítica
Conhecimento - Negócio	N/A	Sem entender do negócio, o profissional de Inteligência Analítica não entende os problemas que necessita resolver (Dumbill, Liddy, Stanton, Mueller, & Farnham, 2013). Assim, sugere-se que esses profissionais precisam ser desenvolvidos no contexto de atuação
Habilidade analítica - <i>Modeling</i>	41,26%	<i>Modeling</i> define a maneira com que se faz a transformação da informação em inteligência dentro da empresa (Tien, 2013)
Habilidade analítica - <i>Optimization</i>	39,81%	<i>Optimization</i> consiste em alterar a maneira como algo é realizado, de forma a obter melhores resultados (Pedersen, Pedersen, & Riis, 2013). Por isso, é algo muito necessário em Inteligência Analítica, visto que qualquer ganho de performance na maneira de fazer algo pode trazer grandes resultados nos algoritmos de Inteligência Analítica
Habilidade analítica - <i>Clustering</i>	21,84%	<i>Clustering</i> consiste em efetivamente reorganizar logicamente dados, de forma a aumentar muito a performance e qualidade das técnicas de Inteligência Analítica (Stefanovic, Stefanovic, & Radenkovic, 2008)
Habilidade analítica - <i>Regression</i>	20,87%	Não há muito material na literatura que permita conceituar regressão, mas é um termo que aparece em dezenas de textos da base estudada, o que destaca o quão comum é dentro do contexto de Inteligência Analítica. A ocorrência dessa competência e sua relação com Inteligência Analítica não pôde ser explicada a partir dos textos escopo
Habilidade analítica - <i>Simulation</i>	20,39%	<i>Simulation</i> é uma competência necessária para Inteligência Analítica pois se faz necessário em certas situações nas quais é mais indicado gerar dados a partir de regras do que utilizar dados históricos (Li & Kauffman, 2012)
Habilidade analítica - <i>Forecasting</i>	18,45%	<i>Forecasting</i> é muito importante pois descreve um dos maiores objetivos de Inteligência Analítica: a previsão de eventos ou valores em um determinado momento do futuro, geralmente para prever um evento único (Simmhan & Noor, 2013)
Habilidade analítica - <i>Predictive Analytics</i>	14,56%	Considera-se que <i>Predictive Analytics</i> seja um conjunto de ferramentas dentro do contexto de <i>Data Science</i> (Hazen, Boone, Ezell, & Jones-Farmer, 2014). <i>Data Science</i> é particularmente útil frente aos atuais desafios de Inteligência Analítica
Habilidade analítica - <i>Classification</i>	13,59%	<i>Classification</i> consiste em discriminar dados em categorias, algo fundamental para Inteligência Analítica (Fernández, del Río, López, Bawakid, del Jesus, Benítez, & Herrera, 2014)
Habilidade analítica - <i>Statistical Analysis</i>	5,34%	Análises estatísticas são importantes para um entendimento inicial da base de dados e para o desenvolvimento de regras para técnicas de <i>Data Mining</i> , que são o fundamento para a obtenção de informação para Inteligência Analítica (Waller & Fawcett, 2013a)
Habilidade analítica - <i>Neural Network</i>	4,85%	<i>Neural Network</i> parece ser uma técnica avançada que permite um uso diferenciado de IA, mas é pouco presente na base de textos estudada, o que sugere que seu uso ainda é complexo e, talvez por esse motivo, pouco difundido
Habilidade analítica - <i>Sentiment Analysis</i>	4,37%	<i>Sentiment Analysis</i> é uma técnica mais específica, para mídias sociais, o que facilita seu domínio e utilização (Abrahams, Jiao, Wang, & Fan, 2012). É bem mais restrito do que <i>Predictive Analysis</i> , a ponto de haver uma série de <i>softwares</i> prontos, com funcionalidades bem definidas
Habilidade computacional - <i>Data Mining</i>	35,44%	<i>Data Mining</i> é um conjunto de algoritmos utilizados para navegar, de forma automatizada, em um conjunto de dados, identificar padrões, planejar regras e fazer previsões para o futuro (Fuchs, Höpken, & Lexhagen, 2014) e é utilizado para obter <i>insights</i> dentro do contexto de Inteligência Analítica
Habilidade computacional - <i>Programming</i>	29,61%	<i>Programming</i> é uma competência básica para Inteligência Analítica, pois trata da habilidade de moldar determinada técnica ao problema que se deseja resolver

Habilidade computacional - MapReduce	23,30%	MapReduce é uma tecnologia muito discutida na literatura de Inteligência Analítica avançada, por tratar de uma tecnologia que permite a análise de imensas quantidades de dados, utilizando-se computadores comuns (Mohammed, Far, & Naugler, 2014), o que democratiza o acesso a Inteligência Analítica mais atual
Habilidade computacional - Machine Learning	21,36%	Machine Learning é importante para Inteligência Analítica pois permite o aprendizado automático (e infactível por humanos) de regras a partir de bases de dados (Berger & Doban, 2014)
Habilidade computacional - OLAP (Online Analytical Processing)	11,65%	OLAP é um dos termos mais difundidos no contexto de Inteligência Analítica, bem como um dos mais antigos. Não permite análises muito avançadas, mas oferece boas possibilidades de visualização quase em tempo real de informações transacionais (Fuchs, Höpken, & Lexhagen, 2014), o que pode ajudar na tomada de decisão de forma menos intuitiva
Habilidade computacional - ETL (Extract, Transform, Load)	8,74%	ETL é a base para gerenciamento de dados que serão utilizados para Inteligência Analítica (Chaudhuri, Dayal, & Narasayya, 2011)
Habilidade computacional - Data Warehousing	7,77%	Data Warehousing e ETL garantem a base inicial para aplicação de técnicas mais avançadas de Inteligência Analítica. Na literatura há convergência na opinião que 70% dos esforços gastos são nessas fases (Agarwal, Shroff, & Malhotra, 2013; Mahmood & Afzal, 2013; Demirkan & Delen, 2013)
Habilidade computacional - Web Mining	1,46%	Na base de textos estudada há pouco muito material descrevendo esta competência, o que poderia indicar pouca relação direta entre Inteligência Analítica, exceto em situações em que as empresas possuem plataformas online para clientes. Os especialistas validaram esta competência como uma das mais importantes
Comportamento - Communicate	19,42%	O uso de tecnologias e técnicas de Inteligência Analítica requer uma capacidade técnica que geralmente é pouco observada em pessoas reconhecidas como boas comunicadoras (Davenport, 2014)
Comportamento - Creative	3,88%	Visto a diversidade de tecnologias e ferramentas disponíveis para Inteligência Analítica, saber qual utilizar e como utilizar parece ser uma mistura de conhecimento, habilidades e criatividade

Quadro 6: Relação entre as competências e o fenômeno IA

Fonte: Elaboração Própria

A média das médias das validações de especialistas das 10 competências com maior frequência na literatura é 4,22 e a média das médias das validações de especialistas das 10 competências com menor frequência na literatura é 4,23. Esses números indicam que os especialistas não consideram as competências mais discutidas na literatura como as mais importantes. Como estudo futuro, uma sugestão é a validação em maior escala dessas competências. A utilização de uma *survey* para centenas de respondentes, desde que qualificados no assunto Inteligência Analítica, poderia oferecer novas percepções e entendimentos sobre a importância de cada competência, seja individualmente, seja agrupada com outras competências. Estudos qualitativos, com entrevistas e análises de conteúdo, poderiam também complementar o trabalho, com novos relacionamentos entre as competências e o fenômeno Inteligência Analítica.

## REFERÊNCIAS

- Abrahams, A. S., Jiao, J., Wang, G. A., & Fan, W. (2012). *Vehicle defect discovery from Social Media*. *Decision Support Systems*, 54(1), 87-97.
- Acito, F., & Khatri, V. (2014). *Business Analytics: Why now and what next?*. *Business Horizons*, 57(5), 565-570.
- Afshari, H., & Peng, Q. (2015). *Modeling and quantifying uncertainty in the product design phase for effects of user preference changes*. *Industrial Management & Data Systems*, 115(9), 1637-1665.
- Agarwal, P., Shroff, G., & Malhotra, P. (2013, June). *Approximate incremental big-Data harmonization*. In 2013 IEEE International Congress on Big Data (pp. 118-125). IEEE.
- Albanese, R. (1989). *Competency-based management education*. *Journal of Management Development*, 8(2), 66-76.
- Albertin, A. L., & Albertin, R. M. D. M. (2012). *Dimensions of the use of information technology: A diagnostics and analysis instrument*. *Revista de Administração Pública*, 46(1), 125-151.
- Allen, I. E., & Seaman, C. A. (2007). *Likert scales and Data analyses*. *Quality progress*, 40(7), 64.
- Anderson, J. E., Daniels, J., McDonald, D., & Edvalson, R. (2014). *The Current State of Business Intelligence and Analytics in Utah*. *Optimization*, 6, 8.
- Aragon, S. R., & Johnson, S. D. (2002). *Emerging roles and competencies for training in e-learning environments*. *Advances in developing human resources*, 4(4), 424-439.
- Armstrong, J. S., Green, K. C., & Graefe, A. (2015). *Golden rule of Forecasting: Be conservative*. *Journal of Business Research*, 68(8), 1717-1731.
- Arnott, D., & Pervan, G. (2014). *A critical Analysis of Decision Support Systems research revisited: the rise of design science*. *Journal of Information Technology*, 29(4), 269-293.
- Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A., & Buyya, R. (2015). *Big Data computing and clouds: Trends and future directions*. *Journal of Parallel and Distributed Computing*, 79, 3-15.
- Baars, H., & Kemper, H. G. (2008). *Management support with structured and Unstructured Data—an integrated Business intelligence framework*. *Information Systems Management*, 25(2), 132-148.
- Bakshi, K. (2012, March). *Considerations for Big Data: Architecture and approach*. In *Aerospace Conference, 2012 IEEE* (pp. 1-7). IEEE.



- Barton, D., & Court, D. (2012). *Making Advanced Analytics Work For You*. Harvard Business Review, 90(10), 78-83.
- Batagelj, V., & Mrvar, A. (1998). *Pajek - Program For Large Network Analysis*. Connections, 21(2), 47-57.
- Berger, M. L., & Doban, V. (2014). *Big data, advanced analytics and the future of comparative effectiveness research*. Journal of comparative effectiveness research, 3(2), 167-176.
- Bhimani, A., & Willcocks, L. (2014). *Digitisation, 'Big Data' and the transformation of Accounting information*. Accounting and Business Research, 44(4), 469-490.
- Bolman, L. G., & Deal, T. E. (1991). *Reframing organizations*. San Francisco: Jossey-Bass Publishers.
- Bose, R. (2009). *Advanced Analytics: opportunities and challenges*. Industrial Management & Data Systems, 109(2), 155-172.
- Boyatzis, R. E. (2006). *Using tipping points of emotional intelligence and cognitive competencies to predict financial performance of leaders*. Psicothema, 18(1), 124-131.
- Boyd, D., & Crawford, K. (2012). *Critical Questions for Big Data*. Information, Communication and Society, 15(5), 662-679.
- Brown, D. E., Famili, F., Paass, G., Smith-Miles, K., Thomas, L. C., Weber, R., & Maldonado, S. (2011). *Future trends in Business Analytics and Optimization*. Intelligent Data Analysis, 15(6), 1001-1017.
- Cappelozza, A., & de Moraes, G. H. S. M. (2015). *A influência da infraestrutura de tecnologia da informação sobre a mobilidade computacional de usuários e a computação em nuvem*. REVISTA DE TECNOLOGIA APLICADA, 2(3).
- Carley, K. M., Hummon, N. P., & Harty, M. (1993). *Scientific influence: An analysis of the main path structure*. Journal of Conflict Resolution. Knowledge, 14(4), 417-447.
- Chae, B. K. (2015). *Insights from hashtag #supplychain and Twitter Analytics: Considering Twitter and Twitter Data for supply chain practice and research*. International Journal of Production Economics, 165, 247-259.
- Chae, B., & Olson, D. L. (2013). *Business Analytics for Supply chain: A Dynamic - Capabilities Framework*. International Journal of Information Technology & Decision Making, 12(1), 9-26.
- Chandler, N., Hostmann, B., Rayner, N., & Herschel, G. (2011). *Gartner's Business Analytics Framework*. Gartner Report G, 219420.
- Chang, R. M., Kauffman, R. J., & Kwon, Y. (2014). *Understanding the paradigm shift to computational social science in the presence of Big Data*. Decision Support Systems, 63, 67-80.
- Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). *An overview of Business intelligence technology*. Communications of the ACM, 54(8), 88-98.

- Chen, H., Chiang, R. H., & Storey, V. C. (2012). *Business Intelligence and Analytics: From Big Data to Big Impact*. MIS Quarterly, 36(4), 1165-1188.
- Chen, K., Li, X., & Wang, H. (2015). *On the model design of integrated intelligent Big Data Analytics systems*. Industrial Management & Data Systems, 115(9), 1666-1682.
- Chen, Y., & Yang, H. (2014, August). *Heterogeneous postsurgical Data Analytics for Predictive Modeling of mortality risks in intensive care units*. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 4310-4314). IEEE.
- Chou, J. S., & Telaga, A. S. (2014). *Real-time detection of anomalous power consumption*. Renewable and Sustainable Energy Reviews, 33, 400-411.
- Chuah, M. H., & Wong, K. (2011). *A review of Business intelligence and its maturity models*. African journal of Business management, 5(9), 3424-2428.
- Chung, W., Chen, H., & Nunamaker Jr, J. F. (2005). *A visual framework for knowledge discovery on the Web: An empirical study of Business intelligence exploration*. Journal of Management Information Systems, 21(4), 57-84.
- Cobo, M. J., López - Herrera, A. G., Herrera - Viedma, E., & Herrera, F. (2011). *Science Mapping Software Tools: Review, Analysis, and Cooperative Study Among Tools*. Journal of the American Society for Information Science and Technology, 62(7), 1382-1402.
- Cohen, I. G., Amarasingham, R., Shah, A., Xie, B., & Lo, B. (2014). *The legal and ethical concerns that arise from using complex Predictive Analytics in Health care*. Health affairs, 33(7), 1139-1147.
- Davenport, T. H. (2006). *Competing on Analytics*. Harvard Business Review, 84(1), 98.
- Davenport, T. H. (2014). *Keep up with your quants*. Harvard Business Review, 91(7/8).
- Davenport, T. H., & Patil, D. J. (2012). *Data scientist: The Sexiest Job of the 21st Century*. Harvard Business Review, 90, 70-76.
- de Oliveira, M. P. V., McCormack, K., & Trkman, P. (2012). *Business Analytics in supply chains—The contingent effect of Business process maturity*. Expert Systems with Applications, 39(5), 5488-5498.
- Debortoli, S., Müller, O., & vom Brocke, J. (2014). *Comparing Business Intelligence and Big Data Skills*. Business & Information Systems Engineering, 6(5), 289-300.
- Delen, D., & Demirkan, H. (2013). *Data, information and Analytics as services*. Decision Support Systems, 55(1), 359-363.
- Demchenko, Y., Grosso, P., de Laat, C., & Membrey, P. (2013). *Addressing Big Data issues in scientific Data infrastructure*. In Collaboration Technologies and Systems (CTS), 2013 International Conference on (pp. 48-55). IEEE.
- Demirkan, H., & Delen, D. (2013). *Leveraging the capabilities of service-oriented Decision Support Systems: Putting Analytics and Big Data in cloud*. Decision Support Systems, 55(1), 412-421.

- Doulkeridis, C., & Nørnvåg, K. (2014). *A survey of large-scale analytical query processing in MapReduce*. The VLDB Journal, 23(3), 355-380.
- Dubey, R., & Gunasekaran, A. (2015). *Education and training for successful career in Big Data and Business Analytics*. Industrial and Commercial Training, 47(4), 174-181.
- Dumbill, E., Liddy, E. D., Stanton, J., Mueller, K., & Farnham, S. (2013). *Educating the next generation of Data scientists*. Big Data, 1(1), 21-27.
- Durand, J. P. (2000) *Les enjeux de la logique compétences, Gérer et Comprendre*, 62, pp. 16 – 24.
- Elbashir, M. Z., Collier, P. A., & Sutton, S. G. (2011). *The role of organizational absorptive capacity in strategic use of Business intelligence to support integrated management control systems*. The Accounting Review, 86(1), 155-184.
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). *Big Data consumer Analytics and the transformation of Marketing*. Journal of Business Research, 69(2), 897-904.
- Fernández, A., del Río, S., López, V., Bawakid, A., del Jesus, M. J., Benítez, J. M., & Herrera, F. (2014). *Big Data with Cloud Computing: an insight on the computing environment, MapReduce, and Programming frameworks*. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 4(5), 380-409.
- Fuchs, M., Höpken, W., & Lexhagen, M. (2014). *Big Data Analytics for knowledge generation in tourism destinations—A case from Sweden*. Journal of Destination Marketing & Management, 3(4), 198-209.
- Gandomi, A., & Haider, M. (2015). *Beyond the hype: Big Data concepts, methods, and Analytics*. International Journal of Information Management, 35(2), 137-144.
- Gantz, J., & Reinsel, D. (2011). *Extracting value from chaos*. IDC iVIEW, 1142, 1-12.
- Garfield, E. (1970). *Citation Indexing for Studying Science*. Nature, 227, 669-671.
- Garfield, E. (1979). *Is Citation Analysis a Legitimate Evaluation Tool?*. Scientometrics, 1(4), 359-375.
- Gartner (2016). *2016 CIO Agenda: A U.S. Perspective*.
- George, G., Haas, M. R., & Pentland, A. (2014). *Big Data and management*. Academy of Management Journal, 57(2), 321-326.
- Georgiou, I. (2014). *Seeing the Forest for the Trees: An Atlas of the Politics–Administration Dichotomy*. Public Administration Review, 74(2), 156-175.
- Gilad, T., & Gilad, B. (1986). *SMR forum: Business intelligence-the quiet revolution*. Sloan Management Review (1986-1998), 27(4), 53.
- Goes, P. B. (2014). *Editor's Comments: Big Data and IS Research*. MIS Quarterly, 38(3), iii-viii.

- Golfarelli, M., Rizzi, S., & Cella, I. (2004). *Beyond Data Warehousing: what's next in Business intelligence?*. In Proceedings of the 7th ACM international workshop on Data Warehousing and OLAP (pp. 1-6).
- Gruhl, D., Nagarajan, M., Pieper, J., Robson, C., & Sheth, A. (2010). *Multimodal social intelligence in a real-time dashboard system*. The VLDB Journal—The International Journal on Very Large Data Bases, 19(6), 825-848.
- Hager, P., & Gonczi, A. (1996). *What is competence?*. Medical teacher, 18(1), 15-18.
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). *The rise of “Big Data” on cloud computing: Review and open research issues*. Information Systems, 47, 98-115.
- Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). *Data quality for Data Science, Predictive Analytics, and Big Data in Supply Chain Management: An introduction to the problem and suggestions for research and applications*. International Journal of Production Economics, 154, 72-80.
- He, W., Wu, H., Yan, G., Akula, V., & Shen, J. (2015). *A novel Social Media competitive Analytics framework with sentiment benchmarks*. Information & Management, 52(7), 801-812.
- Holsapple, C., Lee-Post, A., & Pakath, R. (2014). *A unified foundation for Business Analytics*. Decision Support Systems, 64, 130-141.
- Hopkins, M., Lavallo, S., & Balboni, F. (2010). *19 insights: A first look at the new intelligent enterprise survey*. MIT Sloan Management Review 52(1).
- Hu, J., Singh, M., & Mojsilovic, A. (2008, December). *Categorization using semi-supervised Clustering*. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on* (pp. 1-4). IEEE.
- Huang, M. H., & Rust, R. T. (2013). *IT-related service a multidisciplinary perspective*. Journal of Service Research, 16(3), 251-258.
- Hummon, N. P., & Dereian, P. (1989). *Connectivity in a Citation Network: The Development of DNA Theory*. Social Networks, 11(1), 39-63.
- IBM (2010), “*Global CEO study 2010*”. Disponível em: [www.ibm.com](http://www.ibm.com) (Acesso em Março de 2016).
- Jin, H., Wu, S., Vidyanti, I., Di Capua, P., & Wu, B. (2015). *Predicting Depression among Patients with Diabetes Using Longitudinal Data*. Methods of information in medicine, 54(6), 553-559.
- Jo, S. J., Jeung, C. W., Park, S., & Yoon, H. J. (2009). *Who is citing whom: Citation network analysis among HRD publications from 1990 to 2007*. Human Resource Development Quarterly, 20(4), 503-537.
- Jourdan, Z., Rainer, R. K., & Marshall, T. E. (2008). *Business Intelligence: An Analysis of the Literature 1*. Information Systems Management, 25(2), 121-131.

- Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). *Trends in Big Data Analytics*. Journal of Parallel and Distributed Computing, 74(7), 2561-2573.
- Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., & Gani, A. (2014). *Big Data: survey, technologies, opportunities, and challenges*. The Scientific World Journal, 2014.
- Kiron, D., & Shockley, R. (2011). *Creating Business value with Analytics*. MIT Sloan Management Review, 53(1), 57.
- Ko, S., Maciejewski, R., Jang, Y., & Ebert, D. S. (2012, June). *Marketanalyzer: an interactive visual Analytics system for analyzing competitive advantage using point of sale Data*. In Computer Graphics Forum (Vol. 31, No. 3pt3, pp. 1245-1254).
- Kohavi, R., Mason, L., Parekh, R., & Zheng, Z. (2004). *Lessons and challenges from mining retail e-commerce Data*. Machine Learning, 57(1-2), 83-113.
- Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). *Emerging Trends in Business Analytics*. Communications of the ACM, 45(8), 45-48.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). *Big Data, Analytics and the Path from Insights to Value*. MIT Sloan Management Review, 52(2), 21.
- Le Deist, F. D., & Winterton, J. (2005). *What is competence?*. Human resource development international, 8(1), 27-46.
- Lemon, K. N., & Huang, M. H. (2011). *IT-Related Service: A Multidisciplinary Perspective*. Journal of Service Research, 14, 251.
- Li, F., Özsu, M. T., Chen, G., & Ooi, B. C. (2014, March). *R-Store: a scalable distributed system for supporting real-time Analytics*. In 2014 IEEE 30th International Conference on Data Engineering (pp. 40-51). IEEE.
- Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., & Hampapur, A. (2014). *Improving rail network velocity: A Machine Learning approach to Predictive maintenance*. Transportation Research Part C: Emerging Technologies, 45, 17-26.
- Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). *Big Data in product lifecycle management*. The International Journal of Advanced Manufacturing Technology, 81(1-4), 667-684.
- Li, T., & Kauffman, R. J. (2012). *Adaptive learning in service operations*. Decision Support Systems, 53(2), 306-319.
- Likert, R. (1932). *A technique for the measurement of attitudes*. Archives of psychology.
- Lin, J. (2013). *Mapreduce is good enough? if all you have is a hammer, throw away everything that's not a nail!*. Big Data, 1(1), 28-37.
- Liu, J. S., Lu, L. Y., Lu, W. M., & Lin, B. J. (2013). *Data envelopment Analysis 1978–2010: A citation-based literature survey*. Omega, 41(1), 3-15.

- Lomotey, R. K., & Deters, R. (2014, April). *Towards knowledge discovery in Big Data*. In Service Oriented System Engineering (SOSE), 2014 IEEE 8th International Symposium on (pp. 181-191). IEEE.
- Louridas, P., & Ebert, C. (2013). *Embedded Analytics and Statistics for Big Data*. IEEE software, 30(6).
- Lu, L. Y., Lin, B. J., Liu, J. S., & Yu, C. Y. (2012). *Ethics in nanotechnology: What's being done? What's missing?*. Journal of business ethics, 109(4), 583-598.
- Lucio-Arias, D., & Leydesdorff, L. (2008). *Main - path analysis and path - dependent transitions in HistCite<sup>TM</sup> - based historiograms*. Journal of the American Society for Information Science and Technology, 59(12), 1948-1962.
- Luhn, H. P. (1958). *A Business Intelligence System*. IBM journal of Research and Development, 2(4), 314-319.
- Luvizan, S., Meirelles, F. S., & Diniz, E. H. (2014). *Big Data: Publication Evolution and Research Opportunities*. 11th CONTECSI-International Conference on Information Systems and Technology Management. São Paulo, 30.
- Ma, Y., Chen, G., & Wei, Q. (2013, June). *A novel fuzzy associative classifier based on information gain and rule-covering*. In IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013 Joint (pp. 490-495). IEEE.
- Mahmood, T., & Afzal, U. (2013, December). *Security Analytics: Big Data Analytics for cybersecurity: A review of trends, techniques and tools*. In Information assurance (ncia), 2013 2nd national conference on (pp. 129-134). IEEE.
- Manpower Services Commission. (1986). SASU Note 16: *Guidance on Designing Modules for Accreditation*. mimeo (Sheffield: Standards and Assessment Support Unit).
- Mansfield, B., & Mitchell, L. (1996). *Towards a competent workforce*. Gower Publishing.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big Data: The next frontier for innovation, competition, and productivity*.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- McAfee, A., & Brynjolfsson, E. (2012). *Big Data. The Management Revolution*. Harvard Business Review, 90(10), 61-67.
- Meirelles, F. S. (2016). *Pesquisa Anual – Administração de Recursos de Informática. GVcia – Centro de Tecnologia de Informação Aplicada da FGV-EAESP, São Paulo, 2016*.
- Mell, P., & Grance, T. (2009). *The NIST definition of cloud computing*. Disponível em, <http://www.nist.gov/itl/cloud/upload/cloud-def-v15.pdf>. Acesos em on Mai 16.
- Mohamed, N., & Al-Jaroodi, J. (2014, July). *Real-time Big Data Analytics: Applications and challenges*. In HPCS (pp. 305-310).

- Mohammed, E. A., Far, B. H., & Naugler, C. (2014). *Applications of the MapReduce Programming framework to clinical Big Data Analysis: current landscape and future trends*. *BioData Mining*, 7(1), 1.
- Moore, S. (2011). *Gartner forecasts global Business intelligence market to grow 9.7 percent in 2011*. Gartner Research, 2.
- Moraes, G. H. S. M. (2015). *Indicators of technology in the perception of it and Business managers—a comparative Analysis*. *Race: revista de administração, contabilidade e economia*, 14(3), 863-888.
- Moraes, G. H. S. M., Meirelles, F.S., & Cappellozza, A. (2016). *Information technology and e-government in Brazil: A single case study in the State of São Paulo*. *Revista ESPACIOS*| Vol. 37 (Nº 05) Ano 2016.
- Neely, A. (1999). *The performance measurement revolution: Why now and what next?* *International Journal of Operations and Production Management*, 19(2), 205—228.
- Negash, S. (2004). *Business intelligence*. The communications of the Association for Information Systems, 13(1), 54.
- Newell, S., & Marabelli, M. (2015). *Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'*. *The Journal of Strategic Information Systems*, 24(1), 3-14.
- O'brien, J. A., & Marakas, G. M. (2013). *Administração de Sistemas de Informação*. Porto Alegre: AMGH. Capítulo 11 - Desafios éticos e de segurança.
- Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis*. *Foundations and trends in information retrieval*, 2(1-2), 1-135.
- Pedersen, T. B., Pedersen, D., & Riis, K. (2013). *On-demand multidimensional Data integration: toward a semantic foundation for cloud intelligence*. *The Journal of Supercomputing*, 65(1), 217-257.
- Pei, J. (2013, October). *Some new progress in analyzing and mining uncertain and probabilistic Data for Big Data Analytics*. In *International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing* (pp. 38-45).
- Petrini, M., & Pozzebon, M. (2009). *Managing sustainability with the support of Business intelligence: Integrating socio-environmental indicators and organisational context*. *The Journal of Strategic Information Systems*, 18(4), 178-191.
- Robinson, A., Levis, J., & Bennett, G. (2010). *INFORMS to officially join Analytics movement*, *OR/MS Today* 37(5).
- Russom, P. (2011). *Big Data Analytics*. *TDWI Best Practices Report*, Fourth Quarter, 1-35.
- Rust, R. T., & Huang, M. H. (2014). *The service revolution and the transformation of Marketing science*. *Marketing Science*, 33(2), 206-221.

- Sahay, B. S., & Ranjan, J. (2008). *Real Time Business intelligence in supply chain Analytics*. Information Management & Computer Security, 16(1), 28-48.
- Sanchez, O. P., & Albertin, A. L. (2009). *A racionalidade limitada das decisões de investimento em tecnologia da informação*. RAE-Revista de Administração de Empresas, 49(1), 86-106.
- Sandhu, R., & Sood, S. K. (2015). *Scheduling of Big Data applications on distributed cloud based on QoS parameters*. Cluster Computing, 18(2), 817-828.
- Schläfke, M., Silvi, R., & Möller, K. (2012). *A framework for Business Analytics in performance management*. International Journal of Productivity and Performance Management, 62(1), 110-122.
- Shang, W., Jiang, Z. M., Hemmati, H., Adams, B., Hassan, A. E., & Martin, P. (2013, May). *Assisting developers of Big Data Analytics applications when deploying on hadoop clouds*. In Proceedings of the 2013 International Conference on Software Engineering (pp. 402-411). IEEE Press.
- Sharma, R., Mithas, S., & Kankanhalli, A. (2014). *Transforming decision-making processes: a research agenda for understanding the impact of Business Analytics on organisations*. European Journal of Information Systems, 23(4), 433-441.
- Shin, S. J., Woo, J., & Rachuri, S. (2014). *Predictive Analytics model for power consumption in manufacturing*. Procedia CIRP, 15, 153-158.
- Simmhan, Y., & Noor, M. U. (2013, October). *Scalable prediction of energy consumption using incremental time series Clustering*. In Big Data, 2013 IEEE International Conference on (pp. 29-36).
- Slavakis, K., Kim, S. J., Mateos, G., & Giannakis, G. B. (2014). *Stochastic approximation vis-a-vis online learning for Big Data Analytics*. IEEE Signal Processing Magazine, 31(6), 124-129.
- Steed, C. A., Ricciuto, D. M., Shipman, G., Smith, B., Thornton, P. E., Wang, D., & Williams, D. N. (2013). *Big Data visual Analytics for exploratory earth system Simulation Analysis*. Computers & Geosciences, 61, 71-82.
- Stefanovic, N., Stefanovic, D., & Radenkovic, B. (2008). *Application of Data Mining for supply chain inventory Forecasting*. In Applications and Innovations in Intelligent Systems XV (pp. 175-188). Springer London.
- Straka, G. A. (2004). *Measurement and evaluation of competence*. The foundations of evaluation and impact research. Luxembourg, 263-311.
- Tan, W., Blake, M. B., Saleh, I., & Dustdar, S. (2013). *Social-network-sourced big data analytics*. IEEE Internet Computing, 17(5), 62-69.
- Tien, J. M. (2012). *The next industrial revolution: Integrated services and goods*. Journal of Systems Science and Systems Engineering, 21(3), 257-296.
- Tien, J. M. (2013). *Big Data: Unleashing information*. Journal of Systems Science and Systems Engineering, 22(2), 127-151.



- Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. (2010). *The impact of Business Analytics on supply chain performance*. *Decision Support Systems*, 49(3), 318-327.
- Van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). *Analytics in higher education: Establishing a common language*. EDUCAUSE learning initiative, 1, 1-11.
- Varshney, K. R., & Mojsilović, A. (2011). *Business Analytics based on financial time series*. *IEEE Signal Processing Magazine*, 28(5), 83-93.
- Vedder, R. G., Vanecek, M. T., Guynes, C. S., & Cappel, J. J. (1999). *CEO and CIO Perspectives on Competitive Intelligence*. *Communications of the ACM*, 42(8), 108-116.
- Vieira, C. S., & Meirelles, F.S. (2015). *Computação em Nuvem—Análise bibliométrica da produção científica sobre os fatores que influenciam as empresas no seu uso*. *Revista Eletrônica Gestão e Serviços*, 6(2), 1215-1230.
- Waller, M. A., & Fawcett, S. E. (2013a). *Data Science, Predictive Analytics, and Big Data: a Revolution That Will Transform Supply chain Design and Management*. *Journal of Business Logistics*, 34(2), 77-84.
- Waller, M. A., & Fawcett, S. E. (2013b). *Click here for a Data scientist: Big Data, Predictive Analytics, and theory development in the era of a maker movement supply chain*. *Journal of Business Logistics*, 34(4), 249-252.
- Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). *How 'Big Data' can make big impact: Findings from a systematic review and a longitudinal case study*. *International Journal of Production Economics*, 165, 234-246.
- Wang, X., White, L., & Chen, X. (2015). *Big Data research for the knowledge economy: past, present, and future*. *Industrial Management & Data Systems*, 115(9).
- Watson, H. J. (2011). *Business Analytics insight: hype or here to stay*. *Business Intelligence Journal*, 16(1), 4-8.
- Weiss, S.M., Indurkha, N., Zhang, T., & Damerau, F.J. (2005), *Text Mining: Predictive Methods for Analyzing Unstructured Information*, Springer, New York, NY.
- White, R. W. (1959). *Motivation reconsidered: the concept of competence*. *Psychological review*, 66(5), 297.
- Willen, C. (2002) "Airborne Opportunities", *Intelligent Enterprise*, (5)2, pp. 11-12.
- Winterton, J., Parker, M., Dodd, M., McCracken, M. G., & Henderson, I. (2000). *Future skill needs of managers*.
- Wixom, B. H., Watson, H. J., & Werner, T. (2011). *Developing an Enterprise Business Intelligence Capability: The Norfolk Southern Journey*. *MIS Quarterly Executive*, 10(2).
- Wixom, B. H., Watson, H. J., Reynolds, A. M., & Hoffer, J. A. (2008). *Continental airlines continues to soar with Business intelligence*. *Information Systems Management*, 25(2), 102-112.

Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). *What can Big Data and text Analytics tell us about hotel guest experience and satisfaction?*. International Journal of Hospitality Management, 44, 120-130.

Xu, L., Duan, J. A., & Whinston, A. (2014). *Path to purchase: A mutually exciting point process model for online advertising and conversion*. Management Science, 60(6), 1392-1412.

Zhai, Y., Ong, Y. S., & Tsang, I. W. (2014). *The Emerging "Big Dimensionality"*. IEEE Computational Intelligence Magazine, 9(3), 14-26.

Zheng, Z., Zhu, J., & Lyu, M. R. (2013, June). *Service-generated Big Data and Big Data-as-a-service: an overview*. In 2013 IEEE international congress on Big Data (pp. 403-410). IEEE.

Zhong, N. (2003). *Toward Web intelligence* (pp. 1-14). Springer Berlin Heidelberg.

Zhu, K., Lou, Z., Zhou, J., Ballester, N., Kong, N., & Parikh, P. (2015). *Predicting 30-day Hospital Readmission with Publicly Available Administrative Database*. Methods of information in medicine, 54(6), 560-567.

## APÊNDICE 1: TEXTOS DA REVISÃO DE LITERATURA QUE CONTÉM AO MENOS UMA COMPETÊNCIA

#	Informações sobre o texto
[1]	Chen, H., Chiang, R. H., & Storey, V. C. (2012). <i>Business Intelligence and Analytics: From Big Data to Big Impact</i> . MIS Quarterly, 36(4), 1165-1188.
[2]	LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). <i>Big Data, Analytics and the Path from Insights to Value</i> . MIT Sloan Management Review, 52(2), 21.
[3]	McAfee, A., & Brynjolfsson, E. (2012). <i>Big Data. The Management Revolution</i> . Harvard Business Review, 90(10), 61-67.
[4]	Davenport, T. H., & Patil, D. J. (2012). <i>Data scientist: The Sexiest Job of the 21st Century</i> . Harvard Business Review, 90, 70-76.
[5]	Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. (2010). <i>The impact of Business Analytics on supply chain performance</i> . Decision Support Systems, 49(3), 318-327.
[6]	Waller, M. A., & Fawcett, S. E. (2013). <i>Data Science, Predictive Analytics, and Big Data: a Revolution That Will Transform Supply chain Design and Management</i> . Journal of Business Logistics, 34(2), 77-84.
[7]	Barton, D., & Court, D. (2012). <i>Making Advanced Analytics Work For You</i> . Harvard Business Review, 90(10), 78-83.
[8]	Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). <i>Emerging Trends in Business Analytics</i> . Communications of the ACM, 45(8), 45-48.
[9]	Bose, R. (2009). <i>Advanced Analytics: opportunities and challenges</i> . Industrial Management & Data Systems, 109(2), 155-172.
[10]	Jourdan, Z., Rainer, R. K., & Marshall, T. E. (2008). <i>Business Intelligence: An Analysis of the Literature 1</i> . Information Systems Management, 25(2), 121-131.
[11]	Boyd, D., & Crawford, K. (2012). <i>Critical Questions for Big Data</i> . Information, Communication and Society, 15(5), 662-679.
[12]	Demirkan, H., & Delen, D. (2013). <i>Leveraging the capabilities of service-oriented Decision Support Systems: Putting Analytics and Big Data in cloud</i> . Decision Support Systems, 55(1), 412-421.
[13]	Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). <i>Data quality for Data Science, Predictive Analytics, and Big Data in Supply Chain Management: An introduction to the</i>

	<i>problem and suggestions for research and applications</i> . International Journal of Production Economics, 154, 72-80.
[14]	Acito, F., & Khatri, V. (2014). <i>Business Analytics: Why now and what next?</i> . Business Horizons, 57(5), 565-570.
[15]	Chae, B., & Olson, D. L. (2013). <i>Business Analytics for Supply chain: A Dynamic - Capabilities Framework</i> . International Journal of Information Technology & Decision Making, 12(1), 9-26.
[16]	Holsapple, C., Lee-Post, A., & Pakath, R. (2014). <i>A unified foundation for Business Analytics</i> . Decision Support Systems, 64, 130-141.
[17]	Chae, B. K. (2015). <i>Insights from hashtag #supplychain and Twitter Analytics: Considering Twitter and Twitter Data for supply chain practice and research</i> . International Journal of Production Economics, 165, 247-259.
[18]	Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). <i>How 'Big Data' can make big impact: Findings from a systematic review and a longitudinal case study</i> . International Journal of Production Economics, 165, 234-246.
[19]	Anderson, J. E., Daniels, J., McDonald, D., & Edvalson, R. (2014). <i>The Current State of Business Intelligence and Analytics in Utah</i> . Optimization, 6, 8.
[20]	Baars, H., & Kemper, H. G. (2008). <i>Management support with structured and Unstructured Data—an integrated Business intelligence framework</i> . Information Systems Management, 25(2), 132-148.
[21]	Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). <i>An overview of Business intelligence technology</i> . Communications of the ACM, 54(8), 88-98.
[22]	Chuah, M. H., & Wong, K. (2011). <i>A review of Business intelligence and its maturity models</i> . African journal of Business management, 5(9), 3424-2428.
[23]	Chung, W., Chen, H., & Nunamaker Jr, J. F. (2005). <i>A visual framework for knowledge discovery on the Web: An empirical study of Business intelligence exploration</i> . Journal of Management Information Systems, 21(4), 57-84.
[24]	Davenport, T. H. (2006). <i>Competing on Analytics</i> . Harvard Business Review, 84(1), 98.
[25]	Debortoli, S., Müller, O., & vom Brocke, J. (2014). <i>Comparing Business Intelligence and Big Data Skills</i> . Business & Information Systems Engineering, 6(5), 289-300.
[26]	Dubey, R., & Gunasekaran, A. (2015). <i>Education and training for successful career in Big Data and Business Analytics</i> . Industrial and Commercial Training, 47(4), 174-181.

[27]	Dumbill, E., Liddy, E. D., Stanton, J., Mueller, K., & Farnham, S. (2013). <i>Educating the next generation of Data scientists</i> . Big Data, 1(1), 21-27.
[28]	George, G., Haas, M. R., & Pentland, A. (2014). <i>Big Data and management</i> . Academy of Management Journal, 57(2), 321-326.
[29]	Goes, P. B. (2014). <i>Editor's Comments: Big Data and IS Research</i> . MIS Quarterly, 38(3), iii-viii.
[30]	Golfarelli, M., Rizzi, S., & Cella, I. (2004). <i>Beyond Data Warehousing: what's next in Business intelligence?</i> . In Proceedings of the 7th ACM international workshop on Data Warehousing and OLAP (pp. 1-6). ACM.
[31]	Kiron, D., & Shockley, R. (2011). <i>Creating Business value with Analytics</i> . MIT Sloan Management Review, 53(1), 57.
[32]	Negash, S. (2004). <i>Business intelligence</i> . The communications of the Association for Information Systems, 13(1), 54.
[33]	Petrini, M., & Pozzebon, M. (2009). <i>Managing sustainability with the support of Business intelligence: Integrating socio-environmental indicators and organisational context</i> . The Journal of Strategic Information Systems, 18(4), 178-191.
[34]	Russom, P. (2011). <i>Big Data Analytics</i> . TDWI Best Practices Report, Fourth Quarter, 1-35.
[35]	Schläfke, M., Silvi, R., & Möller, K. (2012). <i>A framework for Business Analytics in performance management</i> . International Journal of Productivity and Performance Management, 62(1), 110-122.
[36]	Van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). <i>Analytics in higher education: Establishing a common language</i> . EDUCAUSE learning initiative, 1, 1-11.
[37]	Wang, X., White, L., & Chen, X. (2015). <i>Big Data research for the knowledge economy: past, present, and future</i> . Industrial Management & Data Systems, 115(9).
[38]	Zhong, N. (2003). <i>Toward Web intelligence</i> (pp. 1-14). Springer Berlin Heidelberg.

Fonte: Elaboração Própria