

**INTELIGÊNCIA DE NEGÓCIOS OU CIÊNCIA DE DADOS? O QUE DADOS  
BIBLIOGRÁFICOS INICIALMENTE NOS DIZEM?**

**Márcio Noveli,**

Universidade Estadual de Maringá,  
[mnoveli2@uem.br](mailto:mnoveli2@uem.br)

**Paula Mitiko Heller,**

Universidade Estadual de Maringá,  
[paula.mitiko@gmail.com](mailto:paula.mitiko@gmail.com)

**Cristiani Passolongo Noveli,**

Universidade Estadual de Maringá,  
[mnoveli2@uem.br](mailto:mnoveli2@uem.br)

**RESUMO**

A constante evolução tecnológica tem permitido que se gerem dados com maior volume, variedade e velocidade. Sob a denominação de *big data* e dataficação, esse contexto pode gerar dúvidas acerca da relação entre essas áreas, uma vez que, por um lado, continuará o modelo tradicional de *business intelligence* e que se adequará à essa nova realidade, ou se imporá o novo mundo que se estabelece com a chamada *data science* e suas novas e complementares formas de análise de dados. De forma a lançar luz sobre a relação entre essas duas áreas, no âmbito acadêmico, esse artigo objetivou identificar se há relação na produção acadêmica da área de *business intelligence* e *data science*, considerando características de artigos das respectivas áreas. Nesse sentido, buscou analisar, com base em uma pesquisa bibliográfica que considerou artigos de periódicos publicados na base de dados Scopus, se há alguma forma de sobreposição das duas áreas, considerando características bibliográficas das suas publicações, especificamente seus: autores, palavras-chave e citações. Ao final, com base nos dados coletados, foi realizada uma análise baseada em estatística descritiva, a partir da qual foi possível constatar uma tênue sobreposição entre os trabalhos das áreas estudadas e apontar questões para estudos futuros.

**Palavras-chave:** Inteligência de Negócios. Ciência de Dados. Bibliometria.

## 1 INTRODUÇÃO

A quantidade de dados disponíveis para análise tem sido cada vez mais alta. Isso tem ocorrido em função de pelo menos duas dimensões, uma tecnológica e outra humana. Na dimensão tecnológica, dispositivos como *smartphones* e sensores, tem permitido a coleta em tempo real de um alto volume de dados; e infra-estruturas de rede e telecomunicações tem evoluído e permitido sua transmissão em alta velocidade e volume. A isso se associa a cada vez maior capacidade de armazenamento desses dados. Essa recente realidade tecnológica, em que o conjunto de dados disponíveis para análise tem alto volume, alta variedade e alta velocidade, tem sido denominada *big data*. (DHAR 2013).

Do outro lado, na dimensão humana, as pessoas tem se apropriado dessas tecnologias disponíveis e gerado, voluntariamente, por meio de fotos, vídeos e textos, por exemplo, dados em alto volume e alta variedade. Esse fenômeno, que envolve as interações humanas serem mediadas e realizadas por meio de tecnologias de informação e comunicação, tem recebido o nome de dataficação. (VAN DIJCK 2014).

Nesse contexto de dataficação e *big data* as organizações deixam de ter que lidar apenas com dados gerados internamente em função de seus processos de negócio, de suas transações com clientes, fornecedores e governo, ou com outras unidades organizacionais, e passam a ter que acompanhar aqueles dados gerados externamente.

Contudo, não é novidade para as organizações, lidar com dados, e nesse âmbito, o termo que vem sendo utilizado tradicionalmente é *Business Intelligence* (BI) ou Inteligência de Negócios. Entretanto, atualmente tem surgido outro termo, denominado *Data Science* (DS) ou ciência de dados, que lida com a mesma situação de análise de dados.

Na medida em que dois termos aparentemente tratam da mesma situação, fica uma dúvida, sobre em qual dessas iniciativas, as organizações devem considerar: a primeira, a segunda, ou ambas?

Ademais, tem havido uma publicidade exagerada em torno do termo ciência de dados, que trouxe discussões sobre a novidade deste campo; se é vinho velho em novas garrafas, sendo apenas uma maneira de chamar a atenção para a já conhecida inteligência de negócios. Além disso, como um campo recente, a ciência de dados não é clara (CARTER; SHOLLER, 2016).

Nesse sentido, o objetivo desse artigo foi identificar se há relação na produção acadêmica da área de *business intelligence e data science*, considerando as características de

artigos das respectivas áreas.

Para isso realizou-se uma pesquisa bibliográfica, que buscou caracterizar, por meio de análise bibliométrica, utilizando estatística descritiva, se há alguma forma de sobreposição das duas áreas, BI e DS, considerando características bibliográficas das suas publicações, especificamente seus: autores, palavras-chave e citações.

Com base na análise dos dados coletados foi possível evidenciar que elementos bibliográficos apontam para uma tênue sobreposição das áreas, sugerindo que estudos futuros permitam descrever essa sobreposição e uma possível conversão das áreas.

## 2 REVISÃO DA LITERATURA

### 2.1 DATAFICAÇÃO E BIG DATA

A dataficação tem permeado a realidade social, na medida em que as relações sociais tem se dado por meio das tecnologias da informação e comunicação. Isso fica claro com o advento da *Web 2.0*, que permitiu que as pessoas não apenas pudessem acessar informações na Internet, ou por meio de tecnologias de telecomunicação e rede, mas que pudessem usar esses mesmos meios para colocar suas próprias informações online.

Saindo de uma perspectiva mais passiva e partindo para uma perspectiva de interação online. Essa possibilidade de interagir online com essas tecnologias e por meio delas é denominado *Web 2.0*, em contraste com a *Web 1.0*, na qual o usuário somente tinha acesso a conteúdo disponível online, sem poder colocar seus próprios conteúdos. (LAUDON; LAUDON, 2009). De acordo com Van Dijck (2014, p. 198):

com o advento da *Web 2.0* e sua proliferação de sites de redes sociais, muitos aspectos da vida social que nunca haviam sido quantificados antes, foram codificados - amizades, interesses, conversas casuais, buscas de informação, expressões de gostos, respostas emocionais e assim por diante. À medida que as empresas de tecnologia começaram a se especializar em um ou vários aspectos da comunicação on-line, elas convenceram muitas pessoas a mover partes de suas interações sociais para ambientes da web.

A relação da dataficação com o *big data* é que o último abrange o conjunto tecnológico que permite lidar com o volume de dados gerados pela dataficação. Nesse sentido, o *big data* é caracterizado pela “[...] capacidade de renderizar dados em muitos aspectos do mundo que

nunca foram quantificados antes [...] As palavras são tratadas como dados quando os computadores minam séculos de livros. Até amizades e ‘curtidas’ são arquivadas via Facebook.” (CUKIER; MAYER-SCHOENBERGER, 2013, p. 29).

Nessa relação explícita, a tecnologia que nos permite interagir de novas formas e por novos meios são as mesmas tecnologias que codificam e armazenam nossas relações, transações, localização, entre diversos outros aspectos de nossa vida.

Na medida em que isso é possível e vem acontecendo, as organizações têm que se haver com uma nova possibilidade, a de explorar um conjunto de dados, em quantidade, velocidade e variedade, até então não existente. Para isso, aparentemente existem dois caminhos: a utilização de BI ou de DS.

### 2.2 INTELIGÊNCIA DE NEGÓCIOS (BI)

O termo inteligência de negócio não é recente, sendo remetido ao trabalho de Luhn (1958), segundo o qual, a inteligência seria a “capacidade de apreender as inter-relações dos fatos apresentados de modo a orientar a ação em direção a um objetivo desejado” (LUHN 1958, p. 314) e negócio diria respeito a “uma coleção de atividades realizadas com qualquer finalidade, seja ciência, tecnologia, comércio, indústria, lei, governo, defesa, etc”. (LUHN 1958, p. 314).

Atualmente, considera-se a inteligência de negócios um termo guarda-chuva, usado para designar um conjunto de elementos que envolvem análise de dados, tecnologias, sistemas e processos. (MOHAMADINA et al. 2012).

Enquanto sistema, Elbashir, Collier e Davern (2008, P. 138), por exemplo, entendem a inteligência de negócios, como “uma importante classe de sistemas para análise de dados e relatórios que fornecem aos gerentes em vários níveis da organização informações oportunas, relevantes e fáceis de usar, que lhes permitem tomar melhores decisões (Hannula e Pirttimaki, 2003).”

Enquanto processo, Wieder e Ossimitz (2015, p. 1164), definem inteligência de negócios como um “processo analítico, tecnologicamente apoiado, que reúne e transforma dados fragmentados de empresas e mercados em informação ou conhecimento sobre objetivos, oportunidades e posições de uma organização.”

E, por fim, enquanto tecnologias e análise de dados, para Mishra et al. (2016, p. 86),

inteligência de negócio é: “um conceito de aplicação de um conjunto de tecnologias para converter dados em informação significativa. [...] principalmente [para] análise de dados e coleta de conhecimento e aplicá-los a vários métodos diferentes”. Corroborando nesse sentido, Anusha e Krishnan (2012, p. 1) afirmam que:

inteligência de negócios pode ser definida como um conjunto de modelos matemáticos e metodologias de análise que sistematicamente exploram os dados disponíveis para recuperar informações e conhecimentos úteis no apoio a processos complexos de tomada de decisão.

Considerando os elementos apresentados, entende-se que a inteligência de negócios envolve todos esses elementos, na medida que os relaciona para seu funcionamento. Assim, a inteligência de negócio utiliza um conjunto de tecnologias, para desenvolver sistemas, que permitem, por meio de um processo, que envolve modelos de análise, analisar dados nas organizações, que auxiliam, por exemplo, na tomada de decisões.

A inteligência de negócio evolui, de acordo com Panian (2009), passando especificamente por três fases: (1) a capacidade de coletar, integrar e armazenar dados; (2) a capacidade de analisar e relatar os dados armazenados e a (3) a capacidade de agir sobre a inteligência gerada a partir do ambiente.

Enfatizando as duas primeiras fases, a fase 1 diz respeito ao trato com os dados internos da empresa, e a fase 2, com as ferramentas de análise. O conjunto tecnológico utilizado nessas fases envolvia, na fase 1, ferramentas de coleta, integração e armazenamento como o ETL (Extração, Transformação e Carga de dados), os *data warehouses* e os *datamarts*, armazéns de dados utilizados para armazenar e ter acesso a dados integrados. E na fase 2, ferramentas para resumos matemáticos denominadas OLAP (Processamento Analítico Online) e *datamining*, para mineração de dados com base, principalmente, em modelos estatísticos. (ELBASHIR; COLLIER; DAVERN 2008).

Atualmente, com o advento da internet e da web 2.0, que é a Internet como a vemos hoje, interativa, a quantidade de dados gerados aumentou exponencialmente. Além disso, as tecnologias de armazenamento e análise de dados também evoluíram. Ao “novo” contexto, em que temos um conjunto de dados com maior variedade, volume e velocidade, do que no passado, tem se dado o nome de big data.

Nesse novo contexto, as tecnologias e as ferramentas de análise de inteligência de negócios tradicionais não mais se adequam. E a área tem buscado essa adequação. Assim, o *big*

*data*, com sua velocidade, volume e variedade de dados mudaram como os projetos de BI eram tradicionalmente abordados. (LARSON; CHANG, 2016). Atualmente, conforme afirmam Nagar et al. (2016, p. 3586):

[...] ferramentas de inteligência de negócios podem ser categorizadas em generalizadas ou ferramentas que funcionam em conjuntos de dados normais, de tamanho menor, em formato estruturado ou semi-estruturado; e ferramentas específicas de *big data* que utilizam *big data*, ou seja, dados grandes o suficiente, mas sem tamanho explícito em qualquer formato estrutural, seja estruturado, semi-estruturado ou não estruturado.

Contudo, e apesar da adequação buscada pela inteligência de negócio, com o surgimento do *big data*, também surgiu outra área, denominada ciência de dados, pronta para explorar o novo conjunto de dados que se apresentam, bem como o novo conjunto tecnológico em termos de tecnologias de coleta, integração e armazenamento, bem como novos modelos de análise.

## 2.3 CIÊNCIA DE DADOS (DS)

Relativamente à inteligência de negócios, o termo ciência de dados é mais recente. Em 2001, Cleveland (2001, p. 21) propunha a ciência de dados como “um plano de ação para ampliar as áreas técnicas de estatística concentrando-se no analista de dados.” E apesar de não se aprofundar aqui neste trabalho especificamente sobre esse plano de ação, aparentemente ele foi posto em prática. E hoje em dia, a ciência de dados é uma realidade. Nesse sentido, de acordo com Watson (2014, p. 1248):

atualmente, muitas organizações estão coletando, armazenando e analisando grandes quantidades de dados. Esses dados são comumente chamados de “*big data*” devido ao seu volume, à velocidade com que chegam e à variedade de formas necessárias. O *big data* está criando uma nova geração de gerenciamento de dados de suporte à decisão. As empresas estão reconhecendo o valor potencial desses dados e estão implementando tecnologias, pessoas e processos para aproveitar as oportunidades.

Uma definição comum em artigos que tratam do tema remete a Dhar (2013, p. 64), segundo a qual ciência de dados é “o estudo da extração generalizada de conhecimento a partir de dados.”

Ampliando essa definição Provost e Fawcett (2013, p. 52) afirmam que a ciência de

dados pode ser definida como "um conjunto de princípios fundamentais que apóiam e orientam a extração de princípios de informação e conhecimento a partir dos dados".

Fazendo um paralelo com o que foi definido acerca dos sistemas de inteligência de negócios, observa-se em Provost e Fawcett (2013) que a ciência de dados também envolve tecnologias, sistemas, processos e análise. Contudo, aparentemente, o conteúdo de alguns desses elementos é diferente, principalmente no que tange especificamente ao conteúdo tecnológico.

Na dimensão tecnológica a inteligência de negócios era baseada principalmente em mineração de dados e análise estatística (CHEN; CHIANG; STOREY, 2012). E apesar de, conforme descrevem Barlas, Lanning e Heavey (2015), a ciência de dados apresentar matemática, modelos de probabilidade e reconhecimento de padrões, também envolve aprendizado de máquina, que não é uma tecnologia comum à inteligência de negócios.

Além disso, diferentemente da inteligência de negócio, que geralmente lida com um conjunto de dados relativamente menor, a ciência de dados “é a ciência da extração de ‘conhecimento acionável’, geralmente de ‘*Big Data*’, ou seja, muitas vezes grandes volumes de dados não estruturados ou estruturados gerados por sistemas, pessoas, sensores ou dispositivos, ou traços pessoais, sociais e digitais. de informação das pessoas.” (DAS et al., 2015, p. 2072).

Portanto, pode-se afirmar, com base no que foi apresentado, que, sobre a evolução das áreas de inteligência de negócios e de ciência de dados, cada uma delas começou em diferentes pontos no tempo. Enquanto a inteligência de negócios é mais antiga e remete à década de 1950, a ciência de dados é relativamente mais recente. Além disso, outro elemento que as diferenciaria, a princípio, seriam as tecnologias utilizadas em cada uma delas.

Contudo, toda a publicidade exagerada que vem ocorrendo em torno do termo ciência de dados trouxe discussões sobre a novidade deste campo, se é vinho velho em novas garrafas, sendo apenas uma maneira de chamar a atenção para a já conhecida inteligência de negócios. Além disso, como um campo recente, a ciência de dados não é clara (CARTER; SHOLLER, 2016).

Para auxiliar a esclarecer esse ponto, a bibliometria possui um arcabouço teórico que fornece conceitos e ferramentas, que, uma vez utilizados na análise de dados empíricos, auxiliam no entendimento de uma área.

Nesse sentido, uma análise pesquisa bibliométrica foi proposta, conforme discutido na seção de metodologia, a seguir.



### **3METODOLOGIA**

Considerando o objetivo geral proposto, este estudo foi de natureza descritiva, buscando descrever características associadas aos artigos que compuseram a amostra do estudo. (COOPER; SCHINDLER, 2003). Para isso, o estudo também foi de natureza quantitativa, pois buscou a quantificação dos dados e a aplicação de análise estatística (MALHOTRA, 2006). Quanto às estratégias de pesquisa, realizou-se uma pesquisa bibliográfica, seguindo a perspectiva descritiva, uma vez que se buscou estudar determinadas características da literatura científica. (NARIN, 1976 apud URBIZAGASTEGUI, 2007).

Os dados coletados a partir de artigos acadêmicos em periódicos que compuseram a amostra do estudo, composta por 4179 artigos encontrados na base Scopus com o uso da palavra-chave “business intelligence” e 1297 artigos encontrados na mesma base com o uso da palavra-chave “data science”.

Para análise desses artigos utilizou-se de técnicas de análise bibliométricas, especificamente, a análise de citações. As citações podem ser consideradas uma ferramenta para indicar as estruturas e tendências científicas; por meio de acesso ao material citado, indica-se o uso do citado pelo citante e permite-se a quantificação de artigos, periódicos e autores, por exemplo, e a aceitação de alguns desses, por parte da comunidade que os cita, considerando a quantidade de vezes que foram citados. (URBIZAGASTEGUI, 2007). “A citação é primeiramente uma medida de impacto, então a maior habilidade da análise de citação é encontrar documentos, autores e revistas que são os mais influentes para responder diferentes questões de pesquisa.” (ZUPIC, ČATER, 2015, p. 438).

Isso porque, de acordo com Narin (1994, apud URBIZAGASTEGUI, 2007), a contagem de artigos proporciona indicadores válidos da atividade de pesquisa e desenvolvimento no assunto de diversas áreas e das instituições que as originam, o número de vezes com que esses artigos são citados em subseqüentes artigos, proporcionam indicadores válidos do impacto ou importância do artigo citado e as citações de artigos a artigos proporcionam indicadores de ligações de conhecimento entre suas áreas ou assuntos.

Para realizar a análise dessas contagens foram utilizadas técnicas de estatística descritiva, considerando seu objetivo de representar a informação de um conjunto de dados, de forma concisa e compreensível, através da elaboração de tabelas e gráficos. (MARCONI;



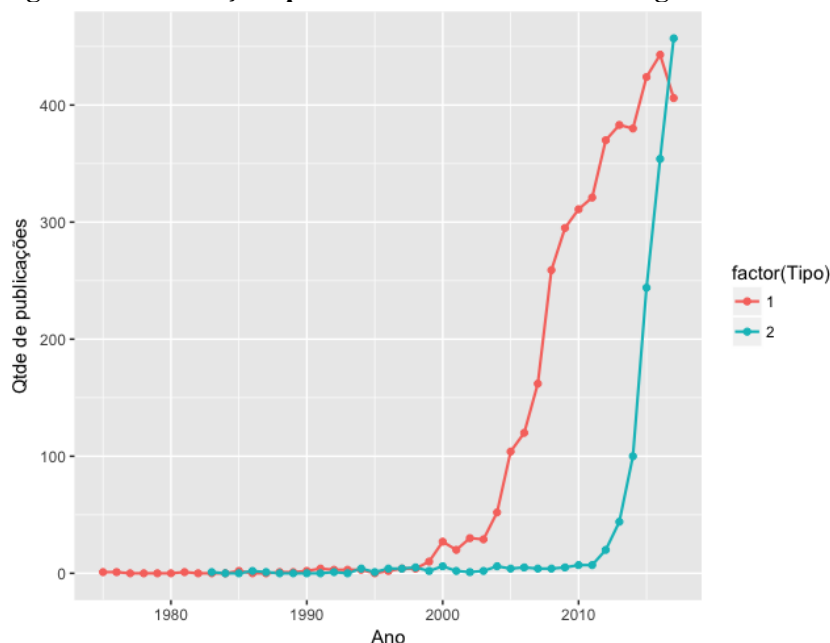
LAKATOS, 1996).

## 4 APRESENTAÇÃO E ANÁLISE DOS DADOS

### 4.1 NÚMERO DE PUBLICAÇÕES AO LONGO DOS ANOS.

De acordo com os dados coletados sobre as publicações na base Scopus relativos à Business Intelligence, conforme apresentado na Figura 1, as primeiras publicações aconteceram nos anos de 1975, 1976, 1981 e 1985, não sendo apontando o artigo de Luhn (1958) na base. E a partir do ano de 1988 até 1998 foram publicados artigos sobre o tópico, apesar de poucos por ano. Porém, de 1999 até 2016, o volume de publicações anuais cresceu de forma constante, atingindo seu ápice com 443 publicações em um ano. Apesar da redução da quantidade de publicações no ano de 2017, não é possível afirmar que há uma tendência de queda na quantidade de publicações anuais.

**Figura 1 - Publicações por ano sobre Business Intelligence e Data Science**



Fonte: Dados da Pesquisa

A primeira publicação sobre Data Science se deu aproximadamente 25 anos de Business Intelligence, sendo em 1983, antes do artigo de Cleveland (2001), mencionado na revisão da literatura. Similar ao primeiro, o tema teve poucas publicações ao longo de uma

década, sendo elas em 1983, 1986, 1987 e 1992. Este só voltou a ser discutido em 1994, e deste ponto até 2011 a quantidade de publicações anuais foi baixa, porém havendo pelo menos uma por ano. A partir de 2012 a quantidade de publicações sobre o tema apresentou um crescimento acentuado, conforme apresentado no Figura 1.

Por fim, é visível na Figura 1 que, apesar de mais recentemente, as discussões sobre o tema de Data Science aumentaram de forma mais intensa que o tema Business Intelligence. Comparativamente, a quantidade de publicações sobre BI cresceu de forma mais linear, relativamente a DS, que teve um crescimento mais exponencial. Esse comportamento pode estar atrelado à maior popularidade recente da DS, como apontado na revisão da literatura.

#### 4.2 AUTORES QUE SE DESTACARAM EM BI E DS

Ao comparar as listas individuais dos dez autores mais citados em artigos ambas as áreas, BI e DS, apresentadas na Tabela 1, é possível identificar o autor J. Han como autor comum à base de conhecimento das áreas, sendo o quinto mais citado em ambos. Além disso, ao comparar a lista de autores mais produtivos com a de autores mais citados, é possível identificar o autor U. Dayal em ambas as listas de Business Intelligence, e os autores S. R. Kalidindi, Y. Wang, C. K. Leung e J. Wang presentes em ambas as listas sobre Data Science.

Além disso, é importante observar que o nono autor mais citado em Data Science, Y. Liu, também está presente como terceiro autor mais produtivo em Business Intelligence, o que é um indicativo de comunicação entre as áreas.

**Tabela 1 – Autores mais citados**

Business Inteligence		Data Science		Ambas as áreas	
Autores	Citações	Autores	Citações	Autores	Citações
1° WATSON, HJ	629	KALIDINDI, SR	140	WATSON, HJ	630
2° CHEN, H	512	WANG, Y	116	CHEN, H	564
3° KIMBALL, R	468	MEZEY, PG	112	HAN, J	474
4° WIXOM, BH	440	LEUNG, CK	94	KIMBALL, R	474
5° HAN, J	394	HAN, J	82	WIXOM, BH	440
6° DAVENPORT, TH	384	KUMAR, V	81	DAVENPORT, TH	436

7º	DAYAL, U	348	TIBSHIRANI, R	78	DAYAL, U	351
8º	INMON, WH	310	WANG J	71	INMON, WH	314
9º	RIZZI, S	306	LIU, Y	69	RIZZI, S	306
10º	ROSS, M	279	DEAN, J	66	ROSS, M	283

Fonte: Dados da Pesquisa

Ao analisar as duas áreas em uma única base de dados, a lista dos dez autores mais citados (Tabela 1) apresenta que, com exceção de B. H. Wixom e S. Rizzi, os autores mais citados de Business Intelligence são, mesmo que pouco, citados em algumas publicações sobre Data Science. Isso adiciona aos indícios de que alguns artigos dessa área podem abordar pontos em comum com o Business Intelligence.

#### 4.4 ARTIGOS MAIS CITADOS CONSIDERANDO CADA UMA DAS ÁREAS

As Tabelas 2 e 3, assim como a Tabela 1, apresenta apenas destaques, nesse caso, enquanto a Tabela 1 apresentava os autores mais citados, as Tabelas 2 e 3 apresentam os artigos mais citados.

**Tabela 2 – Artigos mais citados em publicações sobre Business Intelligence**

	<b>Autores</b>	<b>Artigos</b>	<b>Citações</b>
1º	CHEN, H, CHIANG, RHL, STOREY, VC	Business intelligence and analytics: from big data to big impact (2012)	93
2º	YEOH, W, KORONIOS, A	A critical success factors for business intelligence systems (2010)	47
3º	ELBASHIR, MZ, COLLIER, PA, DA.VERN, MJ	Measuring the effects of business intelligence systems: the relationship between business process and organizational performance (2008)	42
4º	NEGASH, S	Business intelligence (2004)	40
5º	CHAUDHURI, S, DAYAL, U, NARASAYYA, V	An overview of business intelligence technology (2011)	38
6º	WIXOM, BH, WATSON, HJ	An empirical investigation of the factors affecting data warehousing success (2001)	37

7º	WATSON, HJ, WIXOM, BH	The current state of business intelligence (2007)	35
8º	HEVNER, AR, MARCH, ST, PARK, J, RAM, S	Design science in information systems research (2004)	31
9º	LUHN, HP	A business intelligence system (1958)	28
10º	DELONE, WH, MCLEAN, ER	The Delone and Mclean model of information systems success: a ten-year update (2003)	23

Fonte: Dados da Pesquisa

Ao comparar os artigos mais citados em ambas as áreas, como apresentado nas Tabelas 2 e 3, foi possível identificar a presença do artigo “Business Intelligence and Analytics: From Big Data to Big Impact” de Chen H., Chiang RHL e Storey VC, como o mais citado em Business Intelligence e o segundo mais citado em Data Science.

**Tabela 3 – Artigos mais citados em publicações sobre Data Science**

	<b>Autores</b>	<b>Artigos</b>	<b>Citações</b>
1º	DEAN, J, GHEMAWAT, S	Mapreduce: simplified data processing on large clusters (2008)	13
2º	CHEN, H, CHIANG, RH, STOREY, VC	Business intelligence and analytics: from big data to big impact (2012)	13
3º	DHAR, V	Data science and prediction (2013)	12
4º	PROVOST, F, FAWCETT, T	Data science and its relationship to big data and data-driven decision making (2013)	09
5º	BREIMAN, L	Random forests (2001)	08
6º	BLEI, DM, NG, AY, JORDAN, MI	Latent dirichlet allocation (2003)	06
7º	CLEVELAND, WS	Data science: an action plan for expanding the technical areas of the field of statistics (2001)	06
8º	SCHUTT	Educating the next generation of data scientists	06
9º	BREIMAN, L	Random forests (2001)	05
10º	DHAR, V	Data science and prediction (2013)	05

Fonte: Dados da Pesquisa

Essa é mais uma evidência de que, apesar das áreas não possuírem muitos artigos de citações em comum, há a presença dessa intersecção. Com isso é possível a pressuposição sobre

uma possível sobreposição dessas áreas.

#### 4.3 COMPARAÇÃO DAS PALAVRAS-CHAVES DAS ÁREAS DE BI E DS

Analisando as palavras chaves mais frequentes nas publicações relativas a ambas as áreas, é possível identificar a presença comum de termos como “*Big Data*” e “*Data Mining*”. Dessa forma é possível ver que a intersecção principal entre essas suas áreas é a utilização de tecnologias de mineração de dados relacionados a grandes volumes de dados. As principais palavras chave estão apresentadas na Figura 2.

Figura 2 – Wordcloud das palavras chaves dos artigos de Business Intelligence e Data Science



Fonte: Dados da Pesquisa

É importante observar que os termos “*Business Intelligence*”, “*Data Science*” e “*Machine Learning*” tiveram que ser abreviados na Figura 2 com “BI”, “DS” e “ML”, respectivamente, para que este coubesse na imagem dada a sua maior presença entre as palavras chave utilizadas e o comprimento do termo.

Destaca-se, corroborando o que foi apresentado na revisão da literatura, que aprendizado de máquina (*machine learning*), aparece nas palavras-chave associadas a DS mas não a BI, dado que, conforme tratado na revisão da literatura, enquanto conjunto tecnológico, é mais comum à primeira que à segunda.

Ademais, observa-se que, o conjunto tecnológico do BI, conforme apresentado na revisão da literatura, releva tecnologias como *data warehouse*, que lida com um conjunto de dados internos e estruturados. Enquanto o DS apresenta o *hadoop*, voltado para banco de dados mais ligados a dados semi-estruturados e não estruturados. Contudo, em ambos há destaque ao termo *big data*, o que pode apontar para o que foi tratado na revisão da literatura, acerca da adequação do BI em relação à nova realidade trazido pelo *big data*, bem como, do DS enquanto ferramenta que surgiu movida por essa realidade.

### 5 CONSIDERAÇÕES FINAIS

A evolução das tecnologias de informação e comunicação tem trazido à tona recentes fenômenos, como a dataficação e o *big data*, o que tem causado uma necessidade de lidar com dados em volume, variedade e velocidade antes não existentes.

Para lidar com essa demanda, as organizações aparentemente vem adequando ferramentas tradicionais de análise de dados já conhecidas, como business intelligence. Mas também, tem lançado mão de um aparato também aparentemente novo de análise denominado *data science*.

Com base na pesquisa bibliográfica, de natureza quantitativa e descritiva, foi possível observar que em diversas dimensões ambas as áreas se sobrepõem. Artigos são citados em ambas as áreas, autores publicam e são citados em ambas as áreas, palavras-chaves são comuns). Isso levou ao alcance do objetivo do estudo, de identificar uma tênue relação na produção acadêmica da área de *business intelligence e data science*, considerando as características de artigos das respectivas áreas.

Esses achados levam a algumas questões que precisam ser aprofundadas. Os autores de uma área estão migrando para a outra? Uma área vai deixar de existir em detrimento da outra ou elas vão se fundir? As duas coexistirão futuramente?

Essas questões não puderam ser exploradas nessa pesquisa, que contou com limitações, como os próprios termos utilizados para levantamento de artigos, que são insuficientes para mapear toda a área. E a base, Scopus, que não apresenta todos os trabalhos da área, não incluindo livros.

Além disso, para se responder àquelas questões, se faz necessário que, em pesquisa futuras, as teorias, princípios e técnicas de bibliometria sejam aprofundados, para dar suporte a

melhores e mais abrangentes respostas e técnicas mais rigorosas de coleta e análise de dados.

Ainda, é necessário ir além dos termos “business intelligence” e “data science” e encontrar termos associados, que possam aumentar a base de dados da pesquisa bibliométrica e permitir um melhor mapeamento das áreas. Bem como, considerar o *corpus* de artigo como um todo, e não apenas os 10 mais citados.

Contudo, entende-se que este trabalho contribuiu como um passo inicial de forma a permitir delinear um tracejado que permita separar as áreas e entender a área cinzenta, composta por ambas, auxiliando na evolução teórica de estudos dentro desta temática.

## REFERÊNCIAS

ANUSHA, R.; KRISHNAN, N. A Conceptual framework to organize large volume of data for business intelligence. **2012 Ieee International Conference on Computational Intelligence and Computing Research (Iccic)**, p. 752–755, 2012.

BARLAS, P.; LANNING, I.; HEAVEY, C. A survey of open source data science tools. **International Journal of Intelligent Computing and Cybernetics**, v. 8, n: 3, p.232-261, 2015.

CARTER, D.; SHOLLER, D. Data science on the ground: Hype, criticism, and everyday work. **Journal of the Association for Information Science and Technology**, v. 67, n. 10, p. 2309–2319, 2016.

CHEN, H.; CHIANG, R. H. L.; STOREY, V. C. Business intelligence and analytics: From big data to big impact. **MIS Quarterly: Management Information Systems**, v. 36, n. 4, p. 1165–1188, 2012.

CLEVELAND, W. S. Data science: an action plan for the field of statistics. **International Statistical Review**, v. 69, n. 1, p. 21–26, 2001.

COOPER, D. R.; SCHINDLER, P. S. **Métodos de pesquisa em administração**. Porto Alegre: Bookman, 2003.

CUKIER, K.; MAYER-SCHOENBERGER, V. The rise of big data. **Foreign Affairs**, v. 93, n. 3, p. 27-40, may-june, 2013.

DAS, M. et al. Towards methods for systematic research on big data. **Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015**, p. 2072–2081, 2015.

DHAR, V. Data science and prediction. **Communications of the ACM**, v. 56, n. 12, p. 64–73, 1 dez. 2013.

ELBASHIR, M. Z.; COLLIER, P. A.; DAVERN, M. J. Measuring the effects of business intelligence systems: the relationship between business process and organizational



performance. **International Journal of Accounting Information Systems**, v. 9, n. 3, p. 135–153, 2008.

LARSON, D.; CHANG, V. A review and future direction of agile, business intelligence, analytics and data science. **International Journal of Information Management**, v. 36, n. 5, p. 700–710, 2016.

LAUDON, K. C.; LAUDON, J. P. **Sistemas de Informações Gerenciais**. 9. Ed. - São Paulo: Pearson Prentice Hall, 2009.

LUHN, H. P. A business intelligence system. **IBM Journal of Research and Development**, v. 2, n. 4, p. 314–319, 1958.

MALHOTRA, Naresh K. **Pesquisa de marketing: uma orientação aplicada**. – Porto Alegre: Brookman, 2006.

MARCONI, M. A; LAKATOS, E. M. **Técnicas de pesquisa: planejamento e execução de pesquisas, amostragens e técnicas de pesquisas, elaboração e interpretação de dados**. 3.ed. São Paulo: Atlas, 1996.

MISHRA, B. K.; HAZRA, D.; TARANNUM, K.; KUMAR, M. Business intelligence using data mining techniques and business analytics. In. International Conference on System Modeling & Advancement in Research Trends. 5. 25th\_27th November, 2016 College of Computing Sciences & Information Technology, Teerthanker Mahaveer University, Moradabad, India... Proceedings of the SMART -2016, IEEE Conference ID: 39669. 2016.

MOHAMADINA, A. A. et al. Business intelligence: concepts, issues and current systems. **2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT)**, p. 234–237, 2012.

NAGAR, P. et al. **Comparison of generalized and big data business intelligence tools**. (H. M.N., Ed.) Proceedings of the 10th INDIACom; 2016 3rd International Conference on Computing for Sustainable Global Development, INDIACom 2016. **Anais...**Institute of Electrical and Electronics Engineers Inc., 2016. Disponível em:

<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84997261208&partnerID=40&md5=0f4c1b4d11a138b26e2a5a993053da01>

PANIAN, Z. Expected progress in the field of business intelligence 2 development path of business intelligence. **Proceedings of the 8th WSEAS Int. Conf. on ARTIFICIAL INTELLIGENCE, KNOWLEDGE ENGINEERING & DATA BASES (AIKED '09)**, p. 170–175, 2009.

PROVOST, F.; FAWCETT, T. **Data science for business: what you need to know about data mining and data-analytic thinking**. USA, O'reilly, 2013.

URBIZAGASTEGUI, R. A. A Bibliometria: história, legitimação e estrutura. **Para entender a ciência da informação**, p. 185–217, 2007.

## II Encontro Internacional de Gestão, Desenvolvimento e Inovação

20 a 23 de novembro de 2018 - Naviraí - MS



VAN DIJCK, J. Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. **Surveillance and Society**, v. 12, n. 2, p. 197–208, 2014.

WATSON, H. J. Tutorial: big data analytics : concepts , technologies , and applications. **Communications of the Association for Information Systems**, p. 1247-1268, apr., 2014.

WIEDER, B.; OSSIMITZ, M. L. The impact of business intelligence on the quality of decision making - a mediation model. **Procedia Computer Science**, v. 64, p. 1163–1171, 2015.

ZUPIC, I.; ČATER, T. Bibliometric methods in management and organization. **Organizational Research Methods**, v. 18, n. 3, p. 429–472, 2015.