

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**REGIS OSCAR LAND**

***BIG DATA* E SUAS TECNOLOGIAS EM ORGANIZAÇÕES ORIENTADAS  
POR DADOS: UMA REVISÃO SISTEMÁTICA DE LITERATURA**

**PONTA GROSSA**

**2022**

**REGIS OSCAR LAND**

***BIG DATA E SUAS TECNOLOGIAS EM ORGANIZAÇÕES ORIENTADAS POR DADOS: UMA REVISÃO SISTEMÁTICA DE LITERATURA***

***BIG DATA AND IT'S TECHNOLOGIES IN DATA-DRIVEN ORGANIZATIONS: A SYSTEMATIC LITERATURE REVIEW***

Trabalho de Conclusão de Curso apresentado como requisito à obtenção do título de Bacharel em Engenharia de Produção, do Departamento de Engenharia de Produção, da Universidade Tecnológica Federal do Paraná.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Claudia Tania Picinin

Coorientadora: Prof<sup>a</sup>. Me. Juliana Moletta

**PONTA GROSSA**

**2022**



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**REGIS OSCAR LAND**

***BIG DATA E SUAS TECNOLOGIAS EM ORGANIZAÇÕES ORIENTADAS  
POR DADOS: UMA REVISÃO SISTEMÁTICA DE LITERATURA***

Trabalho de Conclusão de Curso apresentado como requisito à obtenção do título de Bacharel em Engenharia de Produção, do Departamento de Engenharia de Produção, da Universidade Tecnológica Federal do Paraná.

Data de aprovação: 22 de junho de 2022

---

Claudia Tania Picinin  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Eliane Fernandes Pietrovski  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Gustavo Tadra Waldmann  
Mestrado  
Universidade Tecnológica Federal do Paraná

**PONTA GROSSA**

**2022**

## Resumo

*Big Data*, termo que é utilizado para classificar dados que possuem características significantes de volume, velocidade e variedade, está cada vez mais em evidência, não apenas em indústrias orientadas a dados, mas também na área de pesquisa científica. É possível identificar potencial inexplorado em pesquisas desta área. Este trabalho tem o objetivo de buscar as principais ferramentas, tecnologias e impactos da utilização de *Big Data* em empresas com uma cultura orientada por dados almejando produzir contribuições nos aspectos sociais, acadêmicos e organizacionais. Foi realizada uma Revisão Sistemática de Literatura, primeiro realizando uma Análise Bibliométrica de um portfólio amplo de artigos, utilizando os softwares *RStudio* e o *Nvivo*. Posteriormente, foi feita uma Análise de Conteúdo com um portfólio de artigos gerado por meio da *Methodi Ordinatio*. Foi demonstrado um aumento no número de publicações com este tema nos últimos anos e identificada a relevância de produções de instituições e autores Chineses e Americanos. As principais tecnologias e ferramentas identificadas foram *Internet of Things*, Inteligência Artificial, Realidade Virtual, *Digital Twins* e Mineração de Dados, gerando impactos como redução de desperdícios, diminuição de custos, automação de processos, aperfeiçoamento da experiência de usuários, melhora nos 3Vs, entre outros. Foi possível apresentar uma visão geral de ferramentas e tecnologias em destaque na literatura recente do tema, assim como seus impactos com o objetivo de melhor preparar as organizações orientadas por dados.

Palavras-chave: *Big Data*; Ferramentas; Revisão bibliográfica; Impactos.

## **ABSTRACT**

Big Data, a term that is used to classify data that have significant characteristics of Volume, Velocity and Variety, is increasingly in evidence, not only in data-oriented industries, but also in the scientific research area. It is possible to identify unexplored potential in research in this topic. This work aims to identify the main tools, technologies and impacts of the use of Big Data in companies with a data-driven culture aiming to produce contributions in the social, academic and organizational aspects. A Systematic Literature Review was carried out, first performing a Bibliometric Analysis through the Methodi Ordinatio methodology of a broad portfolio of articles, using RStudio and Nvivo software. An increase in the number of publications on this topic in recent years has been demonstrated and the relevance of productions by Chinese and American institutions and authors has been shown. The main technologies and tools identified were Internet of Things, Artificial Intelligence, Virtual Reality, Digital Twins and Data Mining, generating impacts such as waste reduction, cost reduction, process automation, improvement of user experience, improvement in 3Vs, among others. It was possible to present an overview of tools and technologies highlighted in recent literature on the subject, as well as their impacts to better prepare data-driven organizations.

Keywords: Big Data; Tools; Systematic Literature Review; Impacts.

## LISTA DE FIGURAS

|   |    |
|---|----|
| Figura 1 - Evolução da geração de dados .....   | 12 |
| Figura 2 - Três Vs do Big Data .....  | 19 |
| Figura 3 - Dados, informação e Business Intelligence .....  | 24 |
| Figura 4 - Quadrante de Softwares de Business Intelligence .....                                      | 26 |
| Figura 5 - Etapas do AM .....   | 27 |
| Figura 6 - Subáreas da Inteligência Artificial .....  | 28 |
| Figura 7 - Exemplos de dados, informação e conhecimento .....   | 29 |
| Figura 8 - Síntese de Conceitos e Ferramentas .....   | 32 |
| Figura 9 - Delineamento da Pesquisa .....   | 35 |
| Figura 10 - Formação dos portfólios de artigos .....  | 38 |
| Figura 11 - Clusters de palavras .....  | 44 |
| Figura 12 - Evolução temática das palavras-chave .....  | 45 |
| Figura 13 - Relação entre autores e coautores .....   | 48 |
| Figura 14 - Mapa Temático da pesquisa Big data analytics and enterprises: a bibliometric review ..... | 49 |
| Figura 15 - Nuvem de Palavras .....   | 69 |

## LISTA DE GRÁFICOS

|   |    |
|---|----|
| Gráfico 1 - Como 1000 executivos relatam o uso de Big Data em tomada de decisão ..... | 13 |
| Gráfico 2 - Evolução da IoT ao passar dos anos .....                                  | 15 |
| Gráfico 3 - Pesquisa de <i>Big Data</i> em Google Trends 2022 .....                   | 15 |
| Gráfico 4 - Publicações com Temática <i>Big Data</i> por Ano .....                    | 16 |
| Gráfico 5 - Número de publicações por ano .....                                       | 41 |
| Gráfico 6 - Publicações por periódico .....   | 42 |
| Gráfico 7 - Periódicos mais citados .....   | 42 |
| Gráfico 8 - Instituições com maior quantidade de produções .....                      | 43 |
| Gráfico 9 - Produção por países .....   | 44 |
| Gráfico 10 - Principais autores em número de publicações.....                         | 47 |
| Gráfico 11 - Principais Ferramentas e Tecnologias de <i>Big Data</i> .....            | 63 |

## LISTA DE QUADROS

|   |    |
|---|----|
| Quadro 1 - Tipos de serviço da Computação em Nuvem .....  | 22 |
| Quadro 2 – Benefícios da utilização da nuvem .....  | 22 |
| Quadro 3 - Classificação da pesquisa .....  | 34 |
| Quadro 4 - Fases para elaboração da Methodi Ordinatio .....   | 36 |
| Quadro 5 - Parâmetros para Análise Bibliométrica .....  | 39 |
| Quadro 6 - Parâmetros para Análise Conteúdo .....   | 39 |
| Quadro 7- Palavras-chave e trechos de artigos .....   | 46 |
| Quadro 8 - Metodologia e Objetivos de Artigos Segmentados por Utilização e seus Impactos.....                             | 51 |
| Quadro 9 - Resultados de Artigos Segmentados por Utilização do Big Data e seus Impactos.....                              | 53 |
| Quadro 10 - Metodologia e Objetivos de Artigos Segmentados por Descrição, Uso ou Criação de Ferramentas de Big Data ..... | 56 |
| Quadro 11 - Resultados de Artigos Segmentados por Descrição, Uso ou Criação de Ferramentas de Big Data .....              | 57 |
| Quadro 12 - Metodologia e Objetivos de Artigos Segmentados por Tecnologias, Processos e Métodos de Uso do Big Data .....  | 59 |
| Quadro 13 - Resultados de Artigos Segmentados por Tecnologias, Processos e Métodos de Uso do Big Data .....               | 61 |
| Quadro 14 - Explicação sobre <i>Big Data</i> .....  | 64 |
| Quadro 15 - Ferramentas, tecnologias e impactos do Big Data .....   | 70 |

## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 1 - Pesquisa realizada no Scopus .....                         | 37 |
| Tabela 2 - Palavras-chave mais frequentes .....                       | 45 |
| Tabela 3 - Autores com maior número de publicações.....               | 47 |
| Tabela 4 - Publicações de Big Data ao longo dos anos .....            | 48 |
| Tabela 5 - Principais autores na pesquisa de Assandre e Martins ..... | 50 |

## LISTA DE ABREVIATURAS E SIGLAS

|      |                                      |
|------|--------------------------------------|
| AM   | Aprendizado de Máquina               |
| BI   | <i>Business Intelligence</i>         |
| DT   | <i>Digital Twins</i>                 |
| IA   | Inteligência Artificial              |
| Ilot | <i>Industrial Internet of Things</i> |
| IoT  | <i>Internet of Things</i>            |
| RA   | Realidade Aumentada                  |

## SUMÁRIO

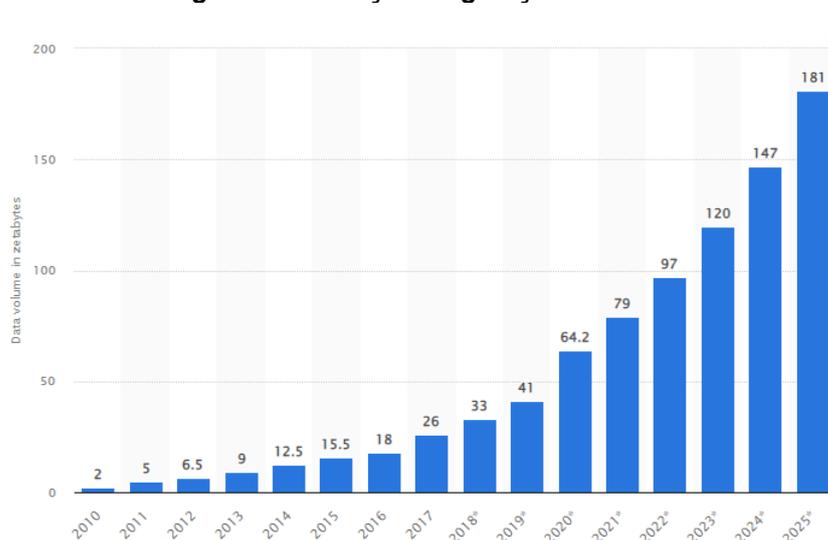
|   |           |
|---|-----------|
| <b>1 INTRODUÇÃO</b> .....   | <b>12</b> |
| 1.2 OBJETIVO GERAL .....  | 13        |
| 1.3 OBJETIVOS ESPECÍFICOS .....   | 14        |
| 1.4 JUSTIFICATIVA .....   | 14        |
| <b>2 REFERENCIAL TEÓRICO</b> .....                                      | <b>17</b> |
| 2.1 <i>BIG DATA</i> .....   | 17        |
| 2.2 INTERNET DAS COISAS .....   | 19        |
| 2.3 COMPUTAÇÃO EM NUVEM .....   | 21        |
| 2.4 BUSINESS INTELLIGENCE .....   | 23        |
| 2.4.1 PRINCIPAIS FERRAMENTAS DE <i>BUSINESS INTELLIGENCE</i> .....      | 25        |
| 2.5 APRENDIZADO DE MÁQUINA .....  | 26        |
| 2.5.1 Aprendizagem Profunda .....                                       | 28        |
| 2.5.2 Redes Neurais .....   | 28        |
| 2.6 <i>MINERAÇÃO DE DADOS</i> .....                                     | 29        |
| 2.7 <i>SOFTWARES DE PROCESSAMENTO DE BIG DATA</i> .....                 | 30        |
| 2.7.1 <i>HADOOP</i> .....   | 30        |
| 2.7.2 <i>SPARK</i> .....  | 31        |
| 2.8 SÍNTESE DE TECNOLOGIAS E FERRAMENTAS .....                          | 32        |
| <b>3 METODOLOGIA</b> .....  | <b>34</b> |
| 3.1 CLASSIFICAÇÃO DA PESQUISA .....                                     | 34        |
| 3.2 DELINEAMENTO DA PESQUISA .....                                      | 35        |
| 3.2.1 FORMAÇÃO DO PORTFÓLIO DE ARTIGOS .....                            | 36        |
| 3.3 ANÁLISE DE RESULTADOS .....   | 38        |
| 3.3.1 ANÁLISE BIBLIOMÉTRICA .....                                       | 38        |
| 3.3.2 ANÁLISE DE CONTEÚDO .....   | 39        |
| <b>4 RESULTADOS E DISCUSSÕES</b> .....                                  | <b>41</b> |
| 4.1 ANÁLISE BIBLIOMÉTRICA .....   | 41        |
| 4.1.1 NÚMERO DE PUBLICAÇÕES POR ANO .....                               | 41        |
| 4.1.2 PERIÓDICOS .....  | 42        |
| 4.1.3 INSTITUIÇÕES E DADOS POR LOCALIZAÇÃO .....                        | 43        |
| 4.1.4 TEMAS E PALAVRAS-CHAVE .....                                      | 44        |
| 4.1.5 AUTORES E COAUTORES .....   | 46        |
| 4.1.6 DISCUSSÃO DA ANÁLISE BIBLIOMÉTRICA .....                          | 48        |
| 4.2 ANÁLISE DE CONTEÚDO .....   | 51        |
| 4.2.1 METODOLOGIA, OBJETIVOS E RESULTADOS .....                         | 51        |
| 4.2.1.1 UTILIZAÇÃO E IMPACTOS DO USO DE <i>BIG DATA</i> .....           | 51        |
| 4.2.1.2 DESCRIÇÃO, USO OU CRIAÇÃO DE FERRAMENTAS DE <i>BIG DATA</i> ... | 56        |
| 4.2.1.3 TECNOLOGIAS, PROCESSOS E MÉTODOS DE USO DO <i>BIG DATA</i> .... | 59        |
| 4.2.2 FERRAMENTAS E TECNOLOGIAS .....                                   | 63        |
| 4.2.3 EXPLANAÇÃO SOBRE <i>BIG DATA</i> .....                            | 64        |
| 4.2.4 TEMÁTICA .....  | 69        |

|   |           |
|---|-----------|
| 4.3 FERRAMENTAS, TECNOLOGIAS E IMPACTOS DO BIG DATA .....   | 70        |
| 4.4 LACUNAS E SUGESTÕES PARA TRABALHOS FUTUROS .....        | 71        |
| <b>5 CONCLUSÃO .....</b>                                    | <b>72</b> |
| <b>REFERÊNCIAS .....</b>                                    | <b>73</b> |
| <b>APÊNDICE A – PORTFÓLIO DE ARTIGOS SELECIONADOS .....</b> | <b>84</b> |
| <b>APÊNDICE B – NÚMERO DE PUBLICAÇÕES POR PAÍS .....</b>    | <b>87</b> |

## 1 INTRODUÇÃO

De acordo com Gartner (2021), são produzidos cerca de 40 trilhões de gigabytes de dados diariamente, originados de postagens em redes sociais, fotos, vídeos, registros de transações comerciais, *Global Position System* (GPS), sensores e outros pontos de coleta. Entre estes pontos, existem mais de 729 milhões de tuítes, 7 bilhões de vídeos vistos no Youtube, 6 bilhões de buscas no Google, 4 bilhões de interações no Facebook, 2 bilhões de postagens no Instagram e 215 bilhões de e-mails enviados diariamente (GEYSEY, 2022). É possível observar a evolução da geração de volume de dados criados, capturados, copiados e consumidos em todo o mundo, e suas estimativas até 2025 na Figura 1:

**Figura 1 - Evolução da geração de dados**

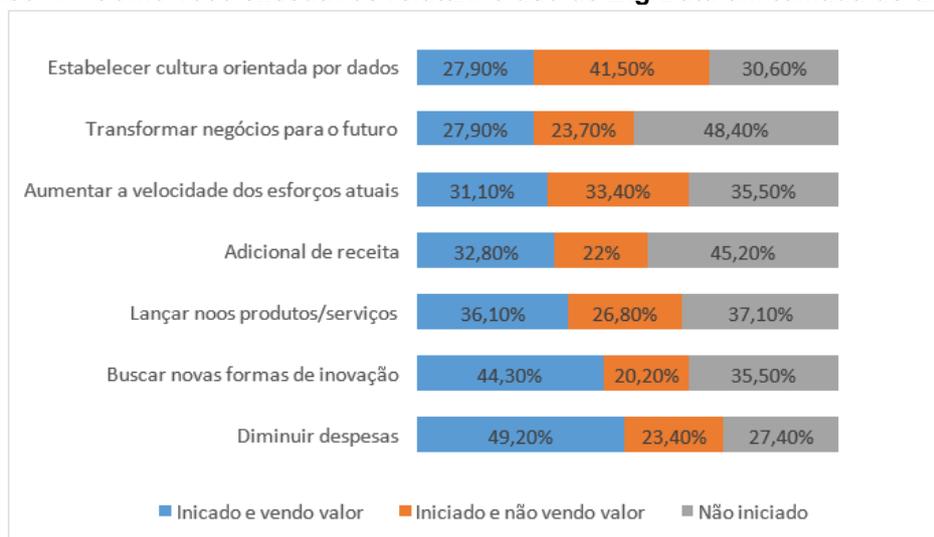


Fonte: Statista (2021)

Esta grande quantidade de dados (Volume), junto com Frequência e tipos de fontes de dados (Variedade) caracterizam o *Big Data*. Ao organizar estes dados, é possível produzir visualizações de padrões, relações e correlações e com estas informações é possível criar modelos descritivos, e diagnósticos preditivos utilizados como embasamento para a tomada de decisão de profissionais em cargos de liderança (MOHAMED *et. al.*, 2019).

O potencial da utilização destes dados é mostrado em resultados, de acordo com Bean (2017). O Gráfico 1 mostra os motivos por trás das iniciativas de *Big Data* que estão em andamento em algumas empresas pesquisadas, sendo a redução de despesas a medida mais vantajosa, com quase metade de todos os executivos indicando que diminuíram as despesas como resultado direto de seus investimentos em *Big Data*:

**Gráfico 1 - Como 1000 executivos relatam o uso de Big Data em tomada de decisão**



Fonte: Adaptado de Bean (2017)

Desta forma, é possível identificar que existe um potencial significativo na aplicação de ferramentas e tecnologias de *Big Data* na otimização de processos e tomadas de decisão de empresas.

A minoria dos dados produzidos é de fato analisada. Isto acontece devido a velocidade do desenvolvimento de tecnologias de coleta, organização e armazenamento de dados se deu muito mais rápida que a de ferramentas de análise de dados (LV *et. al.*, 2021). Partindo desta ideia, foi definido o seguinte tema de pesquisa: Quais são as tecnologias, ferramentas e os impactos do *Big Data* em empresas orientadas por dados?

## 1.2 Objetivo Geral

Identificar as principais ferramentas, tecnologias e impactos do *Big Data* em organizações orientadas por dados.

### 1.3 Objetivos Específicos

Para ser possível o cumprimento do objetivo geral, este foi segmentado em três objetivos específicos, como são descritos a seguir:

1. Realizar um estudo bibliométrico, identificando principais autores, artigos, países de origem de pesquisa e palavras-chave relacionados ao tema de pesquisa nos últimos três anos na base de dados Scopus;
2. Efetuar uma análise de conteúdo, explorando objetivos, resultados, ferramentas e temas abordados sobre *Big Data* em um portfólio selecionado de artigos;
3. Identificar as principais tendências de utilização de tecnologias e ferramentas do *Big Data* e seus impactos em organizações orientadas por dados.

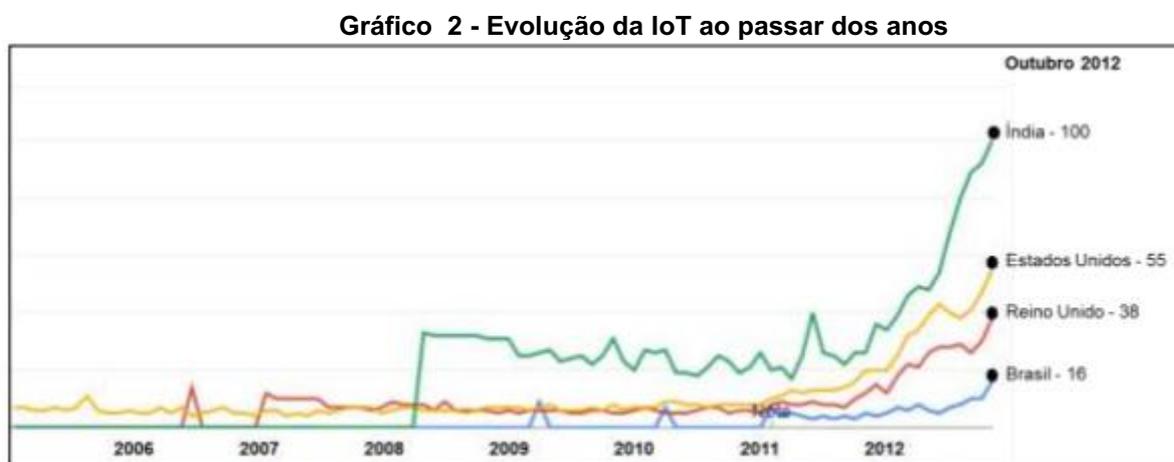
### 1.4 Justificativa

O presente trabalho possui o intuito de contribuir em três aspectos:

- a) Organizacional: apresentando as principais tecnologias e ferramentas de *Big Data* e possíveis impactos gerados, permitindo que sejam replicados ou aplicados de forma que gerem valor para as organizações;
- b) Social: incentivando o desenvolvimento tecnológico, para que possa servir como agente de transformação social. Artigos como o de Alsunaidi *et. al.* (2021) e Baig *et. al.* (2020) mostram a aplicação de tecnologia de *Big Data* nas áreas de educação e saúde pública, por exemplo;
- c) Acadêmica: permitir reconhecer o estado recente das produções científicas voltadas para a área de *Big Data* e auxiliar na identificação de como futuros trabalhos podem fazer a pesquisa desta área evoluir.

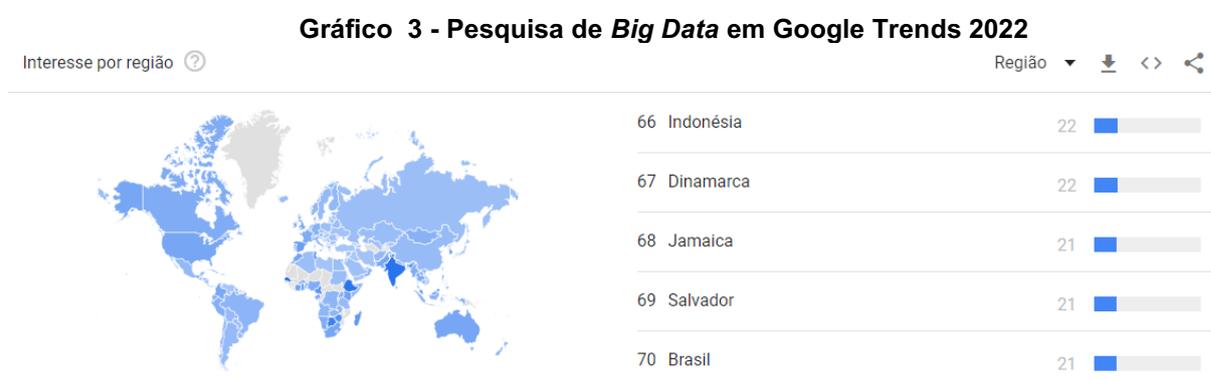
Estes três aspectos estão diretamente relacionados aos objetivos de desenvolvimento sustentável das ações unidas, mais especificamente ao nono objetivo: “Construir infraestruturas resilientes, promover a industrialização inclusiva e sustentável e fomentar a inovação” (NAÇÕES UNIDAS, 2016).

Segundo Sodré (2016), por mais que o termo *Big Data* tivesse uma alta repercussão no começo do século 21, fazendo uma consulta no *Google Trends* em 2012 como mostrado no Gráfico 2, pode ser observado que este possuía uma baixa procura no Brasil comparado a outros países:



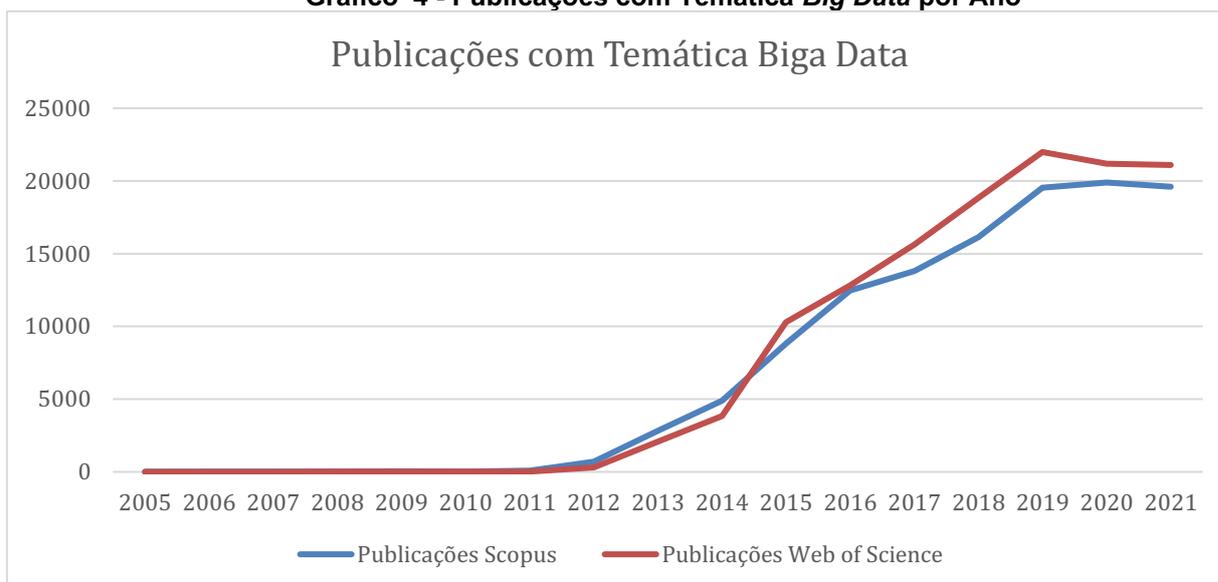
Fonte: Sodré (2016).

Realizando uma análise no *Google Trends*, em 2022, pode-se observar no Gráfico 3 que o Brasil está na septuagésima posição em quantia de pesquisas realizadas com o tema *Big Data*, podendo-se inferir que ainda há potencial desperdiçado no Brasil, principalmente em comparação com outros países:



Fonte: Google (2022).

É possível identificar que existe um aumento recente de pesquisas publicadas com o tema *Big Data*, conforme pode ser observado no Gráfico 4, onde “*Big Data*” foi pesquisado nas bases de dados *Scopus* e *Web of Science*:

**Gráfico 4 - Publicações com Temática *Big Data* por Ano**

Fonte: Dados de Pesquisa (2022)

Para Vianna e Dutra (2016), de forma geral, o tema *Big Data* ainda se encontra em fase inicial de discussão dentro da literatura. Levando em consideração as informações apresentadas nos gráficos 2, 3 e 4, como a produção de conteúdo nesta área aumentou na última década, pode ser inferido que ainda há tópicos a serem explorados dentro desta temática. Pode ser observado em pesquisas preliminares nas bases de dados presentes na pesquisa do Gráfico 4 que há uma lacuna de produções que apresentem uma revisão bibliográfica do tema e de uma síntese das ferramentas e impactos do *Big Data* na indústria, objetivos que o presente trabalho possui.

O trabalho possui em um primeiro momento a introdução, a fim de discutir o contexto do *Big Data*, a justificativa e objetivos do trabalho. Posteriormente, a secção de referencial teórica foi realizada com o objetivo de mostrar definições e conceitos importantes para a compreensão do estudo. Após a revisão teórica, a metodologia elabora como foi realizado o trabalho e a secção de resultados e discussões mostra como os objetivos foram atingidos. Por fim, a conclusão resume os conceitos e resultados apresentados no trabalho.

## 2 REFERENCIAL TEÓRICO

A seção a seguir apresenta conceitos importantes para o entendimento do tema deste Trabalho de Conclusão de Curso. Serão abordados conceitos e ferramentas que são relevantes na utilização do *Big Data* para os mais diversos fins.

### 2.1 *BIG DATA*

*Big Data*, dentro da Tecnologia da Informação, é o termo usado para se referir à grandes volumes e quantidades de dados que devem ser processados em alta velocidade a fim de possuírem uma maior perceptibilidade em situações de tomada de decisão (MACHADO, 2018). Neste sentido, é importante ressaltar que dado é um fato coletado, digitalmente ou analogicamente, e de alguma forma armazenado (AMARAL, 2016).

O conceito do *Big Data* passou a existir para denominar aplicações de Tecnologia da Informação que fazem uso de grandes volumes de dados nos seus mais diversos formatos que são captados, agrupados, transformados e analisados com técnicas estatísticas e computacionais, produzindo o conhecimento definido como “*Data Insight*” (NETO, 2021).

As informações (*Insights*), formadas a partir dos dados, derivam em um importante apoio na decisão orientada à negócios, seja na resolução de problemas ou criação de produtos. Neste segundo caso, na situação em que um novo produto é criado por meio de *Insights*, ele é denominado *Data-Driven Product* (Produto Orientado a Dados). Já em situações em que *Insights* transformam organizações de forma inovadora através do *Big Data*, estas empresas são chamadas de *Data-Driven Companies*, ou Companhias Orientada a Dados, em português (NETO, 2021).

Para Mohamed (2019), um grupo de dados pode ser considerado *Big Data*, caso eles possuam as seguintes características, também conhecidas como os “três Vs do *Big Data*”: volume, variedade e velocidade:

a) Volume: É estimado que, apenas em 2021, foram gerados 350 zettabytes de dados e grande parte destes dados podem ser analisadas por empresas (TERRA, 2021). São gerados *petabytes* de dados diariamente e esta quantidade dobra a cada

dezoito meses. O Volume do *Big Data* refere-se a toda quantidade de dados que são criados digitalmente (VRANOPOULOS; CLARKE; ATKINSON, 2022). É importante ressaltar o potencial que esta abundância de dados possui ao serem coletados e analisados (MACHADO, 2018).

b) Variedade: Variedade dentro do *Big Data* pode ser definida como as diferentes maneiras em que os dados podem se manifestar. Esta heterogeneidade se torna um dos principais desafios para os cientistas de dados. É estimado que as empresas passam de 70 a 80% do tempo total de seu trabalho tentando compreender e preparando os dados, para só então poder realizar análises e produzir *insights* (VRANOPOULOS; CLARKE; ATKINSON, 2022). Esta característica se refere aos tipos de formatos em que os dados são encontrados e processados, podendo ser estruturados, semiestruturados e não-estruturados. Dados não-estruturados são os mais comuns, pois são gerados por mídias sociais, câmeras de segurança, buscas em navegadores, entre outros. Estas fontes de dados, apesar de normalmente serem mais complexas e não serem apresentadas em formatos rígidos, como tabelas (dados estruturados), também acabam derivando em informações cruciais para profissionais em cargos de liderança embasarem a tomada de decisão dentro de organizações (MACHADO, 2018).

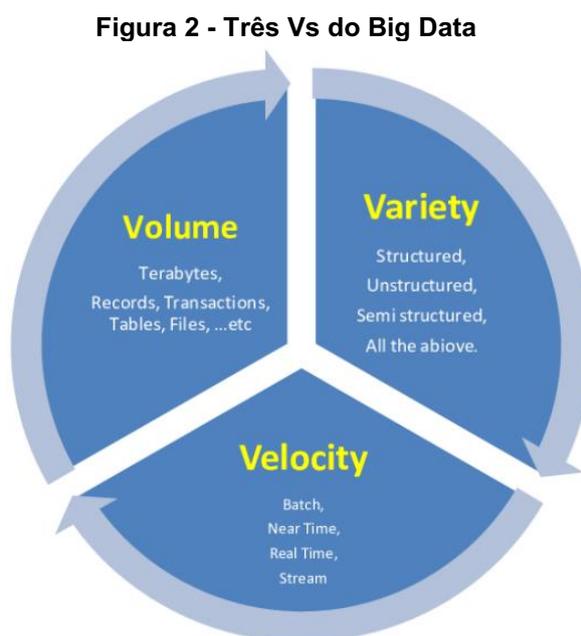
c) Velocidade: Para Vranopoulos, Clarke e Atkinson (2022) Velocidade é a rapidez com que os dados mudam. A Velocidade se refere à periodicidade constante em que novos dados são gerados e transmitidos entre pessoas, organizações e países, por meio de tecnologias da informação. Esta característica auxilia na rápida tomada de decisões, e como consequência faz com que eventos gerem impactos em questão de segundos (LOH, 2019).

Amaral (2016) também cita “Veracidade” e “Valor” como outros 2 “Vs” que caracterizam o *Big Data*, mas que não são considerados unânimes na literatura. Veracidade é um indicativo da integridade dos dados e a habilidade da organização em confiar nos mesmos a fim de tomar decisões cruciais. Vranopoulos, Clarke e Atkinson (2022) afirmam que Veracidade é a medida de homogeneidade e governança dos dados. Este fundamento do *Big Data* se torna relevante, já que não faz sentido embasar decisões em cenários e informações que não são reais e como consequência podem gerar decisões equivocadas e resultados negativos. A fim de

mitigar este problema, deve-se averiguar se o grau de confiança das fontes de dados está em níveis aceitáveis, o que varia de negócio para negócio e nos tipos de dados coletados (UNIVERSITY ALLIANCE, 2015).

Marquesone (2016) em seus estudos relata a necessidade de aprimorar os processos de extração, transformação e carregamento de dados com o objetivo de extrair o máximo da característica de Valor do *Big Data* para as estratégias de empresas. Para Vranopoulos, Clarke e Atkinson (2022), Valor está diretamente relacionado ao rendimento e vantagem competitiva adquiridos com o uso do *Big Data*

De forma geral, o *Big Data* pode ser sintetizado pela Figura 2:



Fonte: (AL-BARHAMTOSHY, 2014)

Ranjan e Foropon (2021) mencionam a crescente importância da Competitividade Inteligente, onde organizações monitoram os concorrentes para ter *insights* que entreguem resultados e inteligência de negócios. Neste aspecto, desenvolver e aproveitar as tecnologias relacionadas a *Big Data* possuem um papel fundamental

## 2.2 INTERNET DAS COISAS

A *Internet of Things* (IoT) ou Internet das Coisas representa uma rede sem fio que liga diversos objetos com equipamentos (como sensores) que podem transportar

dados estáticos ou dinâmicas gerados por eles. Nos últimos anos alguns fatores aumentaram consideravelmente o número de dados gerados, como a expansão de “coisas” conectadas, a transformação de várias atividades do método tradicional para o digital e o incremento da capacitação tecnológica para geração e transmissão de dados (LV *et. al.*, 2021).

IoT se refere ao relacionamento de objetos, humanos e animais por meio de dispositivos capazes de conexão e comunicação. Estes objetos, também chamados de “coisas” podem tomar uma série de ações diariamente sem a necessidade de comando de usuários (RIBEIRO; BRASILEIRO, 2022).

A IoT modifica como as pessoas vivem, trabalham e aprendem, pois permite que atividades básicas sejam automatizadas. Além de tornar trabalhos diários mais fáceis, a IoT também produz informações relevantes para o bem-estar público, e para empresas tomarem as melhores decisões voltadas para as mais diversas áreas (*marketing*, vendas, recrutamento, entre outras). A conexão virtual de dados, pessoas, processos e “coisas”, cria um mundo de novas oportunidades econômicas, entre as quais ao nível das *Smart Cities*, *Smart Homes*, e Industrial (IIoT) (LIBELIUM, 2020):

- a) *Smart Homes*: também chamadas de casas inteligentes, são casas com diversos dispositivos conectados que juntos providenciam conforto e segurança para seus moradores. Entre estes serviços podem ser citados lâmpadas inteligentes, caixas de som com inteligência artificial e ar-condicionado com sensor de temperatura (RIBEIRO, 2019).
- b) *Smart Cities*: O uso da IoT em cidades acarreta em benefícios aos cidadãos e ao governo, atuando em áreas como saúde, segurança e transporte. A IoT consegue, por exemplo, detectar a movimentação de pessoas pela cidade, dando um embasamento para a tomada de decisão no desenvolvimento de políticas públicas relacionadas à locomoção de pessoas e transporte público (BNDES 2018).
- c) Industrial Internet of things: A *Internet of Things Industrial* (IIoT), também conhecida como Indústria 4.0, representa dispositivos e sensores conectados digitalmente á aplicações industriais de manufatura e gerenciamento. A IIoT possibilita melhorar o grau de automação utilizando a computação em nuvem, e assim otimizando processos industriais (BOYES *et. al.*, 2018). Fábricas que

fazem uso da IoT são capazes de gerar dados que são compartilhados entre as próprias máquinas. O fluxo de informações que ocorre em tempo real, permite que as decisões sejam tomadas de forma mais rápida e assertiva. Entre vários benefícios do uso da IIoT, podem ser citados: menor custo em manutenções preventivas, automatização de processos manuais e aumento na velocidade de operações (SYNNEX WESTCON, 2020).

### 2.3 COMPUTAÇÃO EM NUVEM

A Computação em Nuvem se refere a capacidade de usar *softwares* sem necessidade de instalação prévia nos dispositivos locais, sendo a utilização feita de forma online (RIBEIRO; BRASILEIRO, 2022).

Segundo Belcastro (2017), a computação em nuvem é uma forma escalável e remota para armazenamento de recursos. Na sua implementação, a nuvem consiste em uma considerável quantidade de recursos localizados na internet que podem ser acessados sob demanda e de forma virtual.

A computação em nuvem é uma tecnologia fundamental para o pleno funcionamento da IoT, já que seria inviável gerar, processar e armazenar a quantidade significativa de dados dos milhões de dispositivos que estão conectados à internet produzindo informações. Um fato extremamente relevante da computação em nuvem, é que ela permite que recursos como *softwares* sejam utilizados sem necessariamente estarem instalados em *hardware*. Esta funcionalidade permite que um leque de serviços possa ser oferecido (GENG, 2017).

Computadores de alta performance, com processadores potentes capazes de rodar algoritmos complexos são comumente utilizados por analistas de dados para solucionar desafios de *Big Data*. Neste sentido, a nuvem é uma alternativa interessante devido ao seu custo-benefício, pois permite que os algoritmos sejam rodados através da internet, e muitas vezes, de forma mais rápida (Belcastro, 2017).

Os serviços de computação em nuvem são classificados de acordo com o recurso computacional apresentado. Os mesmos podem ser classificados em três tipos: *Software* como Serviço, *Plataforma* como Serviço e *Infraestrutura* como Serviço, apresentados no Quadro 1:

**Quadro 1 - Tipos de serviço da Computação em Nuvem**

| <b>Serviço</b>                     | <b>Descrição</b>  |
|------------------------------------|---|
| Software como Serviço (SaaS)       | Neste tipo de serviço, os programas são previamente instalados pelo fornecedor da nuvem.  |
| Plataforma como Serviço (PaaS)     | Este tipo de serviço consiste em permitir que o usuário instale suas próprias aplicações na nuvem, porém não é permitido que o usuário controle a infraestrutura e sistemas operacionais do sistema.                        |
| Infraestrutura como Serviço (IaaS) | Este serviço tem como objetivo proporcionar ao usuário acesso ao processamento, armazenamento, comunicação de rede, tornando o usuário capaz de ser administrador da infraestrutura digital da nuvem (porém não da física). |

Fonte: Belcastro *et. al.* (2017)

O tipo de segmentação apresentado no Quadro 1 é crucial, pois determina o grau de acesso que um usuário deve possuir para administrar os recursos.

Averiguar os riscos e benefícios da utilização da computação em nuvem é fundamental. Ao optar por aderir a este modelo tecnológico para implementar seus recursos operacionais ou aumentar o portfólio de serviços, os gestores de uma organização precisam identificar qual o custo-benefício de colocar em mãos de terceiros sua capacidade operacional, em termos tecnológicos. Como é apresentado no Quadro 2, podem ser observados alguns pontos positivos gerados para empresas ao utilizar a Computação em Nuvem. Pode-se observar de forma geral, como a nuvem provém não apenas mais segurança, como também menores despesas para as empresas e um acesso mais fácil dos usuários para utilizar os recursos disponibilizados pela organização (VIEIRA; MEIRELLES, 2015):

**Quadro 2 – Benefícios da utilização da nuvem**

| <b>Benefícios</b>              | <b>Características</b>   |
|--------------------------------|--|
| Redução dos Custos             | Recursos compartilhados e elasticidade conforme a demanda, otimiza os custos com TI.   |
| Crescimento do foco no negócio | Ao contratar a nuvem de uma empresa terceira, a empresa possui a possibilidade de focar mais nos negócios invés de alocar mão de obra qualificada em questões de TI. |

|  |   |
|--|---|
| Infraestrutura reutilizável              | A utilização de nuvem permite que empresas tenham um modelo de negócio em que a infraestrutura de tecnologia da informação é reutilizável.                                    |
| Resolução de problemas de forma Coletiva | A computação em nuvem permite que vários usuários possam trabalhar no mesmo projeto ao mesmo tempo, facilitando assim, o trabalho em equipe e compartilhamento de informação. |
| Escalabilidade                           | Expandir ou reduzir a alocação de recursos, baseado na demanda dos usuários   |
| Acessibilidade e Mobilidade              | Os usuários podem acessar remotamente informações da nuvem, por exemplo através de celulares, <i>tablets</i> e <i>notebooks</i> .   |
| Segurança                                | A infraestrutura é monitorada constantemente, podendo identificar irregularidades em tempo real.  |

Fonte: Adaptado de Vieira e Meirelles (2015).

Devido aos motivos apresentados no Quadro 2, atualmente existe uma tendência de claro aumento de empresas que utilizam serviços em nuvem. Existem empresas que são referências neste ramo. São elas: Amazon, Google, Hewlett-Packard Development Company, International Business Machines Corporation, Microsoft, Salesforce, JoyentCloud, Rackspace e VMware (BERRY, 2017).

## 2.4 BUSINESS INTELLIGENCE

*Business Intelligence* se refere à utilização de dados que auxiliam na mecanização de análises, tomada de decisão, formação de estratégia e previsões. O objetivo de utilizar o *Business Intelligence* é coletar, processar e analisar uma grande quantidade de dados e fazer com que isto agregue valor para as empresas (Yang *et. al.*, 2022).

O *Business Intelligence* é um termo que engloba tecnologias, ferramentas e metodologias que auxiliam na tomada de decisão através da utilização e manipulação de dados históricos, oferecendo valiosas informações aos gestores de organizações. Dentro do *Business Intelligence*, pode ser encontrado uma fonte importante de dados para geração de indicadores de desempenho, principalmente oriundos de *Enterprise Resource Planning* (ERPs), quando tratamos de grandes empresas (MACHADO, 2018). A Figura 3 resume o conceito de Business Intelligence:

**Figura 3 - Dados, informação e Business Intelligence**



Fonte: Know Solutions (2022)

Existe mais de uma definição para *Business Intelligence*, portanto é importante definir um escopo para impor limites aos diversos significados dados a este termo. *Business Intelligence*, ao ser traduzido de forma literal, significa “Inteligência do Negócio”, ou “Inteligência Empresarial”. A implementação do sistema de *Business Intelligence* ajuda na dinâmica de fluxo de informações valiosas aos usuários, diminuindo a quantidade de informações insignificantes. Em empresas do setor financeiro, o *Business Intelligence* permitiu uma melhor visualização e acompanhamento do fluxo de caixa, pagamentos e outras métricas importantes (COSER, 2020).

Para Sultan (2017), o *Business Intelligence* provê aos gestores do negócio, acesso aos dados e modelos que possibilitam a condução de análises que levam à tomada de decisão. O autor separa o *Business Intelligence* em cinco etapas sequenciais e dependentes uma da outra:

- a) Dados: a geração e estruturação dos dados;
- b) Informação: diz respeito à produção de informações utilizando os dados previamente armazenados;
- c) Conhecimento: refere-se à produção de *insights* para otimizar os processos organizacionais, a partir das informações que estão em mãos da empresa;
- d) Decisão: se refere ao processo de tomada de decisão baseado nos itens anteriormente descritos;
- e) Lucro: fase desejada de maturidade do BI, atingida ao tomar as melhores decisões e ao possuir um sistema de coleta a análise de dados eficaz.

O *Business Intelligence* é um aspecto fundamental do *Big Data*. Apenas organizações com um sistema sólido de *Business Intelligence* conseguem ter projetos de *Big Data* bem-sucedidos, já que os indicadores do *Business Intelligence* possuem uma importância significativa ao analisar o desempenho dos mesmos. O BI, com seus dados estruturados, junto com algoritmos de *Data Mining*, e outras novas tecnologias de processamentos em larga escala, armazenamento e memória, é capaz de fornecer análises de negócio extremamente valiosas a uma organização. Principalmente utilizando-se de processamento paralelo e distribuído sobre uma gama gigantesca de fontes de informação heterogêneas e das mais variadas origens (MACHADO, 2018).

Para Romero (2021), o *Business Intelligence* representa uma necessidade para empresas situadas em ambientes dinâmicos e complexos, principalmente os que envolvem a Indústria 4.0. Neste caso, o *Business Intelligence* é necessário para processar informações e tomar as decisões corretas.

De acordo com Machado (2018), o *Business Intelligence* é dependente de dados estruturados e, em muitos casos, retrata acontecimentos que já ocorreram, e para desenvolver bons estudos de caso de BI, cientistas de dados qualificados são necessários. Pelo BI, é possível criar gráficos e relatórios em formato de dashboard que refletem os indicadores da empresa, sendo desta forma, mais fácil para que profissionais em cargos de liderança embasem sua tomada de decisão. Porém, para ir além e buscar soluções e melhorias mais complexas para o negócio, são necessárias soluções com *Big Data*, onde seria possível produzir análises sobre dados não estruturados, e assim realizar análises preditivas e prescritivas (ou seja, trazendo informações não só sobre o passado, mas também definindo que ações tomar em cenários futuros e menos previsíveis).

#### 2.4.1 Principais Ferramentas de *Business Intelligence*

O Gartner Group (2019) realiza frequentemente uma análise de mercado dos softwares de *Business Intelligence*, relatando o posicionamento de cada empresa fabricante e suas plataformas. O posicionamento é definido por uma matriz chamada de o “quadrante mágico”. Nesta matriz, o eixo “x” é a abrangência de visão do fabricante e o eixo “y” a habilidade em executar o que se propõe. Podemos observar o resultado na Figura 4:

Figura 4 - Quadrante de Softwares de Business Intelligence



Fonte: Gartner Group, 2019.

Podemos inferir pelo gráfico que a Microsoft e seu programa, Power BI, está dentro do quadrante de empresas líderes, e é o principal *software* para se trabalhar com *Business Intelligence*, provavelmente devido ao seu ótimo custo-benefício.

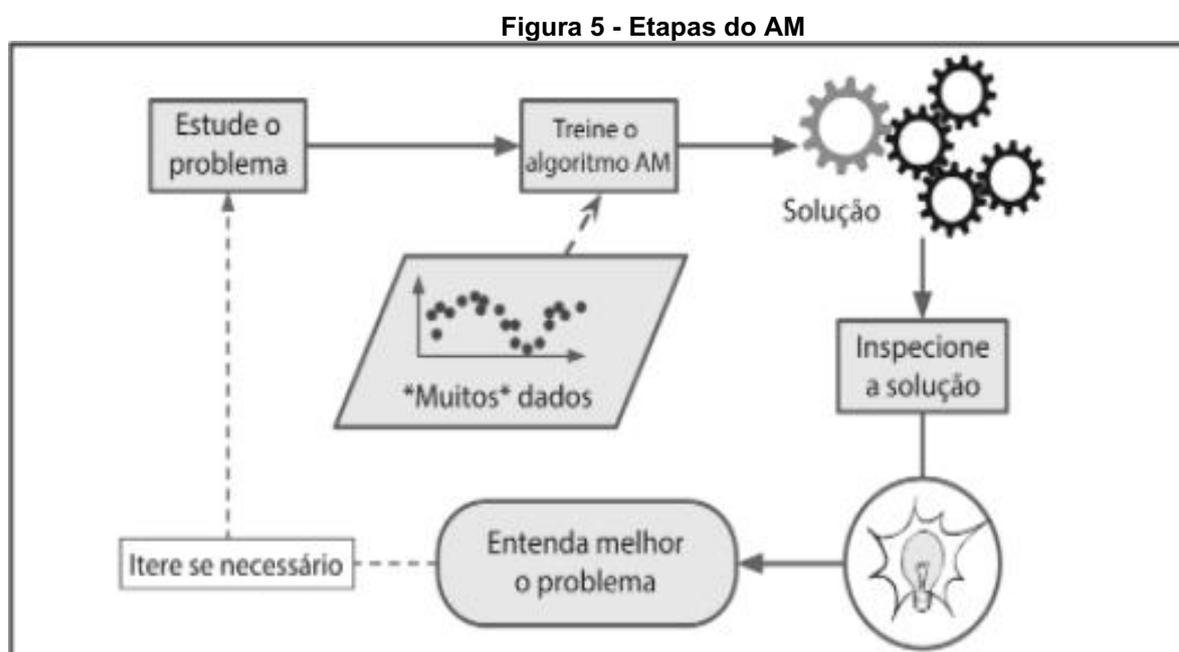
## 2.5 APRENDIZADO DE MÁQUINA

*Machine Learning* ou Aprendizado de Máquina (AM) é o ramo da Inteligência Artificial (IA) em que os computadores “aprendem” utilizando dados. Diz-se que uma máquina aprendeu algo, caso a máquina consiga desempenhar uma tarefa melhor após ela realizar uma leitura de dados, chamados de conjuntos de treinamento (GUNDERSEN, 2021).

O AM é um campo importante da inteligência artificial que auxilia os computadores em produzir modelagens e realizar previsões baseado em experiências ou fatos passados. Esse aprendizado pode ser supervisionado, quando os conjuntos

de treinamento já possuem entradas e saídas previamente mapeadas, ou não-supervisionadas, quando não há mapeamento ou rótulos, e a saída é resultante em percepções do conjunto de treinamento (DOGAN; BIRANT, 2020).

A Figura 5 apresenta um esquema das etapas a serem seguidas no processo de AM:



Fonte: Géron (2019)

Como pode-se observar na Figura 5, em um primeiro momento é necessário estudar a fundo o problema em questão e tentar identificar suas causas, padrões e formas de evitar ou solucionar o mesmo. Segundo, deve-se detectar um algoritmo para detectar padrões baseado em conjuntos de treinamento. E por fim, o programa deve ser testado para definir se o resultado do AM foi satisfatório ou não, fazendo modificações caso necessário (GÉRON, 2019).

Segundo Duana *et. al.* (2019), desde os anos 1950, a IA teve seus pontos altos, “Primaveras da IA” e pontos de baixa evolução, “Invernos da IA”. Com o avanço de tecnologias de *Big Data*, a IA tem se tornado uma ferramenta importante na transformação tecnológica de empresas de referência. Para Panetta (2018) Inteligência Artificial deveria ser a principal estratégia a ser utilizada por empresas de referência para alavancar sua vantagem competitiva, e acredita que ao menos até 2025 será o principal foco destas organizações.

### 2.5.1 Aprendizagem Profunda

Os modelos de AM não possuem a capacidade de fazer a extração de dados, por isso os conjuntos de treinamento são fornecidos de forma manual. Para que isto não precise ser feito de forma manual, pode ser utilizada a aprendizagem por representação, que é fornecer uma quantidade significativa de exemplos para que o modelo consiga automaticamente averiguar padrões. Neste caso, tem-se a Aprendizagem Profunda (SILVA, 2020).

A Figura 6 apresenta um diagrama relacionando Inteligência Artificial (*Artificial Intelligence*), AM (*Machine Learning*) e Aprendizagem Profunda (*Deep Learning*):

**Figura 6 - Subáreas da Inteligência Artificial**



Fonte: Lavagnoli (2019)

Como é observado na Figura 6, o AM é uma área dentro da Inteligência Artificial e a Aprendizagem Profunda é uma sub-área do AM.

### 2.5.2 Redes Neurais

Redes Neurais, ou mais precisamente Redes Neurais Artificiais possuem a intenção de copiar o funcionamento do cérebro humano, e estão na vanguarda de

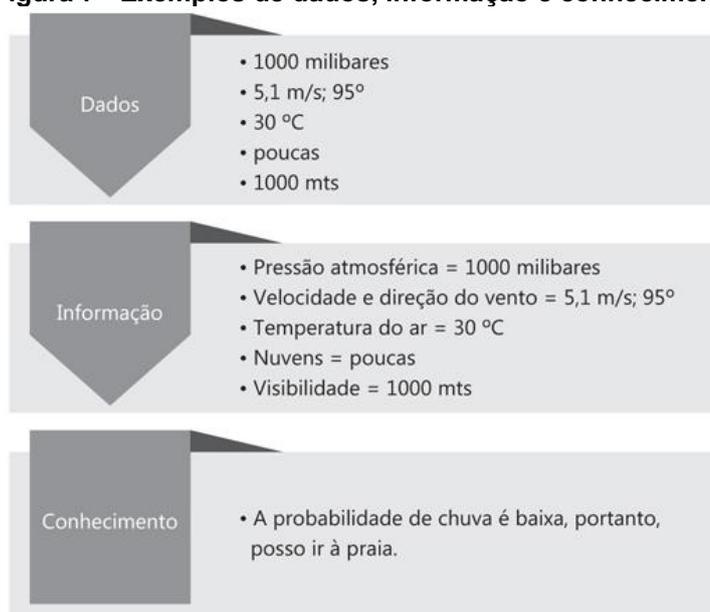
como a Inteligência Artificial funciona atualmente (DUANA; EDWARDSB; DWIVEDIC, 2019).

Redes Neurais Artificiais são modelos computacionais que consistem de unidades de processamento simples (neurônios, de acordo com a analogia com o corpo humano) que funcionam em paralelo e estão arranjados em camadas interconectadas. Essas conexões neurais consistem de uma camada de entrada e uma camada de saída; quando mais camadas são criadas, elas são chamadas de redes profundas. A Rede Neural aprende a desempenhar tarefas através do treinamento, realizando conexões entre os nós/neurônios que após o treinamento são capazes de realizar novas entradas de atividades (CICHY; KAISER, 2019).

## 2.6 MINERAÇÃO DE DADOS

Segundo Dogan (2020), a mineração de dados é o processo digital em que computadores capturam dados, principalmente a partir da internet. A Figura 7 mostra exemplos e diferenças entre dados, informações e conhecimento:

**Figura 7 - Exemplos de dados, informação e conhecimento**



Fonte: De Castro e Ferrari (2016)

Como apresentado na Figura 7, dados são símbolos, normalmente não estruturados e sem significado. Já a informação traz sentido aos dados, dando

utilidade aos mesmos. O conhecimento, ou *insight*, é o valor tirado do conjunto de informações que permite uma tomada de decisão (DE CASTRO; FERRARI, 2016).

O nome *Data Mining* ou mineração de dados vem de uma analogia feita com a exploração de minas a fim de descobrir pedras preciosas. Este conceito se refere ao processo de exploração de uma base de dados utilizando algoritmos para obter conhecimento (CASTRO; FERRARI, 2016).

## 2.7 SOFTWARES DE PROCESSAMENTO DE *BIG DATA*

A fim de extrair informações valiosas oriundas do *Big Data*, novas arquiteturas de sistemas de processamento de grandes volumes de dados têm sido desenvolvidas para superar a alta velocidade e complexidade destes elementos gerados (BELCASTRO et al., 2022). A seção a seguir descreve algumas das principais ferramentas relacionados ao processamento de *Big Data*:

### 2.7.1 *Hadoop*

O *Hadoop* é um programa projetado para explorar dados de forma paralela por meio de *MapReduce*. Os dados são alimentados por lotes e divididos em subgrupos que são processados por diferentes máquinas (BELCASTRO et al., 2022).

*Hadoop* é o resultado de um programa *open-source* originado de um trabalho em conjunto de empresas de referência como Yahoo!, Facebook, Cloudera, Microsoft, Intel, Uber, entre outros. O *Hadoop* foi resultante de dois projetos: um original do Google para *MapReduce* e o segundo do Google File System (GFS), um sistema de arquivos distribuído da própria empresa. Sua função é de agrupar e processar grandes quantidades de dados não-estruturados de forma escalável e distribuída, através de um conjunto de ferramentas, aplicações e frameworks Java. O programa em si foi produzido para trabalhar com *petabytes* e *exabytes* de dados distribuídos em nós em paralelo. Atualmente, o *Hadoop* é administrado pela empresa Apache Software Foundation, onde diversos profissionais otimizam constantemente a sua tecnologia (SYED, 2013).

Um exemplo de aplicação do processamento e armazenamento através do *Hadoop* foi citado por White (2015), na situação em que o jornal New York Times pode

processar 4 *terabytes* de arquivos digitalizados de jornal físico em *Portable Document Format* (PDF) em uma nuvem. O processamento teve duração de pouco menos de 24 horas utilizando 100 máquinas, e este projeto não teria ocorrido sem a tecnologia de programação paralela do *Hadoop*.

Seguindo as diretrizes de *Apache Hadoop* (2016), as principais características da tecnologia *Hadoop* são as seguintes:

- a) As aplicações são codificadas em linguagem de alto nível (Java);
- b) A política de operação reduz ao mínimo a comunicação entre nós;
- c) Os dados, são alocados em blocos de 64MB ou 128MB;
- d) As *map tasks* trabalham com poucos dados, normalmente um bloco por tarefa;
- e) As *map tasks* são executadas nos nós onde estão os blocos de dados;
- f) Na ocasião de ocorrer uma falha em um nó, o mesmo será detectado pelo nó mestre e o trabalho do nó com falha é repassado para outro nó.
- g) Após a reinicialização do nó que apresentou falha, ele será automaticamente adicionado ao sistema para a realização de tarefas;
- h) A reinicialização de uma tarefa não apresentará consequência aos outros nós;
- i) Caso um nó apresente muita lentidão, o mestre poderá executar outra instância da mesma tarefa.

O *Hadoop* requer mais linhas de código do que programas parecidos como *Spark*, *Hive* ou *Pig*. Porém, é muito mais eficiente por ser altamente customizável (BELCASTRO et al., 2022).

### 2.7.2 *Spark*

O *Spark* é comumente utilizado para desenvolver aplicações em memória, como consultas iterativas e processamento em lotes. Muitas bibliotecas de programação são feitas com base em *Spark*, o tornando um sistema flexível para uma gama de aplicações, como o *Spark SQL* (BELCASTRO et al., 2022).

O *Apache Spark* é um programa de computação em cluster (agrupamento de dados) para realizar processamento de uma grande quantidade de dados. Este

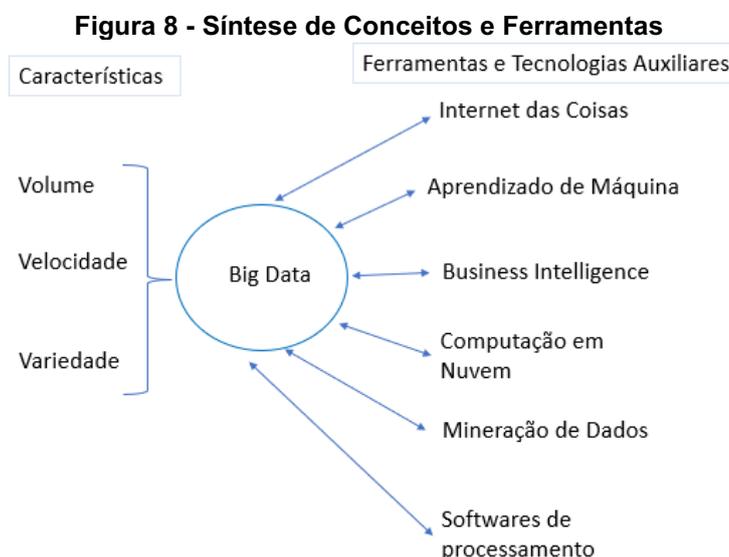
programa possui seu próprio mecanismo de agrupamento de dados, onde não é utilizado o *MapReduce* como em outros programas de processamento de dados. Contudo, o *MapReduce* pode ser usado utilizando o YARN e HDFS (WHITE, 2015).

O *Spark* possui como diferencial uma melhor usabilidade do que outras aplicações de *Big Data*, já que seus APIs podem executar muitas tarefas de processamento de dados usando linguagens como Scala, Java e Python. Também está presente um módulo de AM (MLib) processamento em grafos (GraphX) e podem ser realizadas consultas pela linguagem SQL (Spark SQL) (APACHE, 2016)

O *Spark* possui a característica de guardar grandes grupos de dados em sua memória no ato de processamento. Isto faz com que o *Spark* possa ter uma capacidade maior de suprir o fluxo de *Jobs* do *MapReduce*. Esta capacidade permite ao *Spark* superar o fluxo de trabalho do *MapReduce*, quando os dados são carregados do disco rígido. Este tipo de aplicação é interessante principalmente em modelos de algoritmos iterativos (quando uma função é acionada a um grupo de dados diversas vezes até que um resultado desejado seja obtido) e análises iterativas. (WHITE, 2015).

## 2.8 SÍNTESE DE TECNOLOGIAS E FERRAMENTAS

Tendo em vista os conceitos apresentados na seção de Referencial Teórico, a Figura 8 sintetiza as principais características, conceitos, ferramentas e tecnologias que auxiliam na implementação e utilização do *Big Data*:



Fonte: Autoria Própria (2022)

Como mencionado na seção 2.1, volume, velocidade e variedade são as três dimensões que caracterizam dados como *Big Data*. Apesar de não serem unânimes, veracidade e valor são outras duas qualidades que alguns autores atribuem ao *Big Data*.

O Referencial Teórico destacou alguns conceitos tecnológicos que complementam o *Big Data*, na sua coleta e utilização para geração de valor. A IoT mostra-se nos materiais pesquisados como uma fonte importante de geração e coleta de dados, por meio da integração de sensores e processamento de informações entre os mais diversos equipamentos e até mesmo seres vivos. Dentro da IoT, destaca-se a IIoT, também conhecida como Indústria 4.0.

O AM apresenta-se como um braço da Inteligência Artificial que através da simulação e detecção de padrões de um grande volume de dados pode gerar *insights* de situações que já ocorreram, e até mesmo prever por meio de modelos estatísticos acontecimentos futuros.

A computação em nuvem permite utilizar a internet como meio de substituir equipamentos físicos que muitas vezes limitam o armazenamento e processamento de *Big Data*. Existem diferentes tipos de *softwares* de processamento de *Big Data*, os quais se destacam *Hadoop* e *Spark* para a extração, transformação e carregamento de dados. O *Business Intelligence* é o conjunto de técnicas e tecnologias que permitem produção de *insights* dos dados coletados, desde modelos de algoritmos até a criação de relatórios ou *dashboards* para visualização de dados.

### 3 METODOLOGIA

Esta seção detalha a classificação da pesquisa, execução do trabalho, como a bibliografia foi definida e como ocorreu a análise do portfólio de artigos selecionado.

#### 3.1 CLASSIFICAÇÃO DA PESQUISA

Gil (2008) afirma que pesquisa é um mecanismo racional e sistemático que busca propiciar resposta à problemas apresentados. A pesquisa pode ser classificada considerando a abordagem, a natureza, os objetivos e procedimentos. Levando em consideração a pesquisa produzida neste trabalho, tem-se a classificação apresentada no Quadro 3:

**Quadro 3 - Classificação da pesquisa**

| Aspecto       | Classificação      | Conceito   | Justificativa   |
|---------------|--------------------|--|---|
| Abordagem     | Quali-Quantitativa | A pesquisa é feita fundamentada na observação de dados, com finalidade de compreender as informações obtidas após a aplicação da metodologia (GÜNTHER, 2006).  | Neste trabalho foram estudados aspectos quantitativos de artigos científicos por meio de métricas e qualitativos por meio da Análise de Conteúdo                        |
| Natureza      | Aplicada           | Fornecer uma base para aplicações reais (PRODANOV & FREITAS, 2009).  | O trabalho, entre seus objetivos, possui o intuito de permitir que organizações possam identificar e replicar usos práticos do <i>Big Data</i> descritos neste trabalho |
| Objetivos     | Descritiva         | O trabalho é considerado sistemático, pois tem a finalidade de descrever fenômenos e propor relações entre fatos sem interferir nos mesmos (PRODANOV & FREITAS, 2009).   | O trabalho expõe métricas e o conteúdo de temas e situações já definidas ou que já ocorreram  |
| Procedimentos | Bibliográfica      | A pesquisa é desenvolvida baseado em materiais já elaborados, composto principalmente por livros e artigos científicos, com finalidade de acercar o pesquisador ao material já desenvolvido sobre a pesquisa (PRODANOV & FREITAS, 2009). | O presente trabalho foi baseado em material científico já produzido por pesquisadores de diversas regiões, instituições e áreas de pesquisa.                            |

Fonte: Autoria própria (2022)

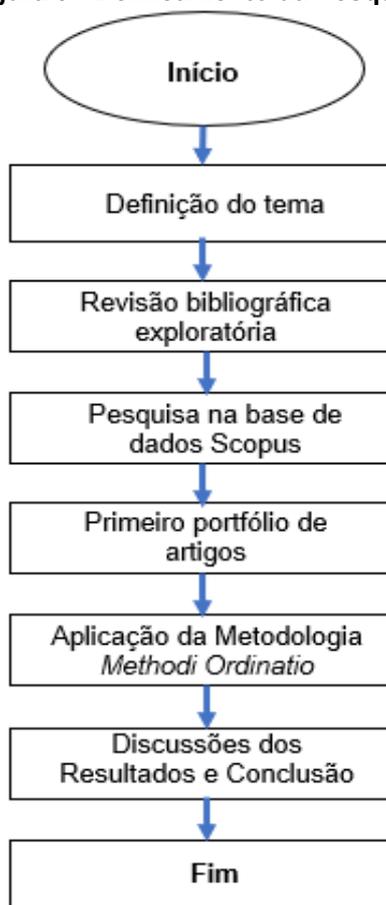
De acordo com Prodanov e Freitas (2009), planejamento é a etapa principal da pesquisa, pois devem ser previamente considerados todos os procedimentos para alcançar o objetivo final esperado, com relação ao problema proposto.

### 3.2 DELINEAMENTO DA PESQUISA

A revisão sistemática de literatura, utilizada nesta pesquisa, consiste em analisar diversas contribuições de pesquisa em um determinado tempo ao invés de analisar apenas um caso isolado (TRANFIELD *et. al.*, 2003).

Com o intuito de alcançar os objetivos estabelecidos no trabalho, o fluxograma da execução pesquisa é representado na Figura 9:

**Figura 9 - Delineamento da Pesquisa**



Fonte: Autoria Própria (2022)

Segundo Treinta (2011), a efetividade da pesquisa é dependente das ações do autor, desde a definição inicial das palavras-chave, filtragem e especialmente na avaliação dos resultados de pesquisa com o tema proposto.

### 3.2.1 Formação do portfólio de artigos

A *Methodi Ordinatio* foi utilizada como ferramenta para selecionar artigos científicos na formação de um portfólio bibliográfico (PAGANI; KOVALESKI; RESENDE, 2015). A metodologia leva em consideração três características:

- a) Número de citações: mostra o reconhecimento da comunidade científica pela obra em questão;
- b) Fator de Impacto: reflete a importância do *Journal* em que o artigo está presente, sendo maior quanto mais importante o *Journal*;
- c) Ano de publicação: indica o quão recente é o artigo. Pesquisas mais atuais tendem a proporcionar maiores novidades em suas pesquisas.

As etapas da *Methodi Ordinatio*, utilizadas neste trabalho bem como a sua descrição, estão apresentadas no Quadro 4:

**Quadro 4 - Fases para elaboração da Methodi Ordinatio**

|   | <b>Fase</b>   | <b>Descrição</b>   |
|---|---|--|
| 1 | Definir intenção de pesquisa  | Para este trabalho, <i>Big Data</i> , suas ferramentas e impactos em empresas orientadas por dados   |
| 2 | Pesquisa preliminar nas bases de dados                                    | Foram testadas palavras-chave baseado no tema proposto.  |
| 3 | Definição das palavras-chave, base de dados e delimitação temporal        | As palavras-chave e a base de artigos foram definidas considerando a disponibilidade e qualidade dos seus materiais científicos. Os resultados da consulta que foram considerados irrelevantes à pesquisa foram descartados. Foram considerados artigos dos últimos três anos.   |
| 4 | Pesquisa definitiva na base de dados                                      | A pesquisa das palavras-chave deste trabalho ocorreu no Scopus.  |
| 5 | Procedimentos de filtragem  | O material resultante da pesquisa definitiva foi filtrado para eliminar duplicatas, registros sem autor, título, ano e <i>journal</i> por meio do <i>software Mendeley</i> . Por meio do <i>Jabref</i> , são filtrados e separados os artigos com e sem ISSN. Por fim, o <i>software Gedit</i> filtra os artigos com "NoAuthor" (artigos sem autores). |
| 6 | Identificação do fator de impacto, ano de publicação e número de citações | Os artigos resultantes da filtragem inicial são inseridos na planilha <i>Finder</i> , onde as informações dos materiais serão ordenadas e padronizados para serem posteriormente utilizados na ferramenta <i>Rankin</i> . (Pagani <i>et. al.</i> , 2021)   |
| 7 | Cálculo do InOrdinatio  | O <i>Rankin</i> é uma planilha que automaticamente procura o fator de impacto das publicações, e no momento que o usuário preenche o número de citações o InOrdinatio é calculado baseado no meio onde o material foi publicado, ano de publicação do artigo e a soma da quantia de citações encontradas (Pagani <i>et. al.</i> , 2021).               |

|   |  |   |
|---|--|---|
| 8 | Localização, leitura e análise sistemática dos artigos | Os artigos foram localizados e foi realizada uma leitura sistemática dos materiais. |
|---|--|---|

Fonte: Adaptado de Pagani, Resende e Kovaleski (2015)

O cálculo do InOrdinatio relatado na etapa 7 é feito com a seguinte equação:

$$InOrdinatio = \left( \frac{\text{Fator de Impacto}}{1000} \right) + \alpha * [10 - (\text{Ano da Pesquisa} - \text{Ano da Publicação})] + \sum \text{Número de citações} \quad (1)$$

- O Fator de Impacto é dividido por 1000 para o deixar em padrão similar aos outros critérios;
- O fator de ponderação é definido pelo pesquisador, variando de 1 a 10 de acordo com a importância atribuída ao ano de publicação, quanto maior o fator, maior a importância;
- O número de citações é a quantidade absoluta de citações mostrada na base de dados.

A base de dados selecionada para a pesquisa foi a *Scopus*, por possuir uma característica multidisciplinar, por sua relevância e a disponibilidade dos seus artigos (PRANCKUTÉ, 2021). As palavras-chave, filtros da pesquisa e número de resultados estão relatados na Tabela 1:

**Tabela 1 - Pesquisa realizada no Scopus**

| Palavra-chave             | Pesquisa  | Número de artigos |
|---------------------------|---|-------------------|
| Big Data, tool*, industr* | TITLE-ABS-KEY ("big data" AND "tool*" AND "industry*" AND PUBYEAR > 2018 AND (LIMIT-TO (DOCTYPE, "ar")) | 207               |

Fonte: Autoria própria (2022)

A pesquisa inicial realizada na base de dados *Scopus*, no dia 27 de fevereiro de 2022 resultou em 207 artigos. O booleano "AND" foi utilizada para filtrar as palavras-chave "*Big Data*", tool\*, "industr\*" e "technolog\*" nos resumos ou palavras-chave. A pesquisa foi filtrada também por meio de publicação (*Journals*) e tipo de documento (artigos). Do resultado obtido, um artigo foi removido por fugir completamente do tema de pesquisa proposto. Os 206 artigos restantes foram utilizados para a realização da análise bibliométrica deste trabalho.

A *Methodi Ordinatio* foi aplicada nestes artigos e como critério de seleção para a Análise de Conteúdo, foram selecionados os que possuíram *InOrdinatio* acima de 93 (tomando como base a média aritmética dos *InOrdinatio* do portfólio em questão), eliminando assim 157 artigos e sendo considerado os 49 artigos presentes no Apêndice A. A Figura 10 apresenta um resumo da formação do portfólio de artigos:



Fonte: Dados de pesquisa (2022)

Ao ser finalizada a etapa de classificação dos artigos por meio do *InOrdinatio*, o pesquisador possui autonomia para selecionar quais artigos estarão em sua revisão de literatura, considerando que os trabalhos já estão classificados por ordem de relevância. A etapa de análise de resultados será descrita na próxima seção.

### 3.3 ANÁLISE DE RESULTADOS

Ao obter o portfólio de artigos, tem-se a fase de análise e discussão de resultados, divididos em Análise Bibliométrica e Análise de Conteúdo, que serão apresentadas a seguir.

#### 3.3.1 Análise Bibliométrica

Segundo Costa e Vanz (2012) a bibliometria serve como instrumento aos pesquisadores por ponderar quantitativamente características de materiais de produção científica, podendo assim haver um panorama geral da bibliografia. Por meio deste conceito, a Análise Bibliométrica deste trabalho apresenta os parâmetros presentes no Quadro 5:

**Quadro 5 - Parâmetros para Análise Bibliométrica**

| <b>Parâmetro</b>           | <b>Significado</b>   |
|----------------------------|--|
| Ano                        | Número de artigos publicados por ano                                       |
| Instituições e localização | Número de artigos por afiliações e países                                  |
| Periódicos                 | Principais periódicos em termos de contribuição ao tema                    |
| Autores e coautores        | Identificação dos principais autores em termos de contribuição para o tema |
| Temas e Palavras-chave     | Identificação das palavras-chave que mais representam o tema               |

Fonte: Autoria Própria (2022)

Ao ser encerrada a Análise Bibliométrica, a Análise de Conteúdo foi realizada para identificar as principais tendências de utilização do *Big Data* e seus impactos em indústrias tecnológicas. No presente trabalho, esta Análise Bibliométrica foi feita com o auxílio do programa Nvivo e do RStudio, com a função *biblioshiny* da biblioteca *bibliometrix* que permite a visualização de dados da pesquisa realizada. As métricas apresentadas foram aproveitadas para melhor interpretar o portfólio de artigos.

### 3.3.2 Análise de Conteúdo

Para Camara (2013), a análise de conteúdo permite produzir uma visão crítica de publicações científicas. Segundo Bardin (2016), a análise de conteúdo deve ser separada em três etapas:

- a) Pré-análise: leitura flutuante, para que o autor tenha um conhecimento do texto e seleção de documentos a serem utilizados;
- b) Exploração do material: etapa onde ocorre a categorização (sistema de codificação), em que são separadas unidades ou categorias pertinentes à pesquisa;
- c) Tratamento e interpretação dos resultados: condensação e ênfase na análise das informações coletadas por meio da leitura dos documentos.

A interpretação e análise de conteúdo dos artigos foram conduzidas considerando as unidades de parâmetros apresentadas no Quadro 6:

**Quadro 6 - Parâmetros para Análise Conteúdo**

| <b>Parâmetro</b> | <b>Definição</b>                                       |
|------------------|--|
| Metodologia      | Tipo de metodologia utilizada para realizar a pesquisa |

|                                  |   |
|----------------------------------|---|
| Objetivo                         | Mostrar os objetivos esperados dos artigos selecionados   |
| Resultados                       | Resultados e impactos gerados pelos artigos   |
| Ferramentas do <i>Big Data</i>   | Identificar as principais ferramentas do <i>Big Data</i> presentes nos artigos  |
| Explicação sobre <i>Big Data</i> | Apresentar um resumo de trechos dos artigos que auxiliam na compreensão sobre o que é o <i>Big Data</i> , suas características, impactos gerados pelo seu uso, barreiras entre outras informações que contribuem para o melhor entendimento do tema baseado em informações trazidas pelos autores |
| Temática de Pesquisa             | Definir os principais temas abordados nos artigos selecionados  |

Fonte: Autoria própria (2022)

Para o presente trabalho, após a etapa de pré-análise dos 49 artigos do portfólio (descrito na seção 3.2.2), 11 artigos foram excluídos por não apresentarem o conteúdo adequado para extrair as informações dos parâmetros abordados no Quadro 6.

A Análise de Conteúdo permitiu explorar como a visão dos autores e resultados apresentados pelos materiais selecionados puderam auxiliar no atingimento dos objetivos deste trabalho ao identificar tendências, ferramentas e impactos gerados pelo uso do *Big Data* em empresas orientadas por dados.

## 4 RESULTADOS E DISCUSSÕES

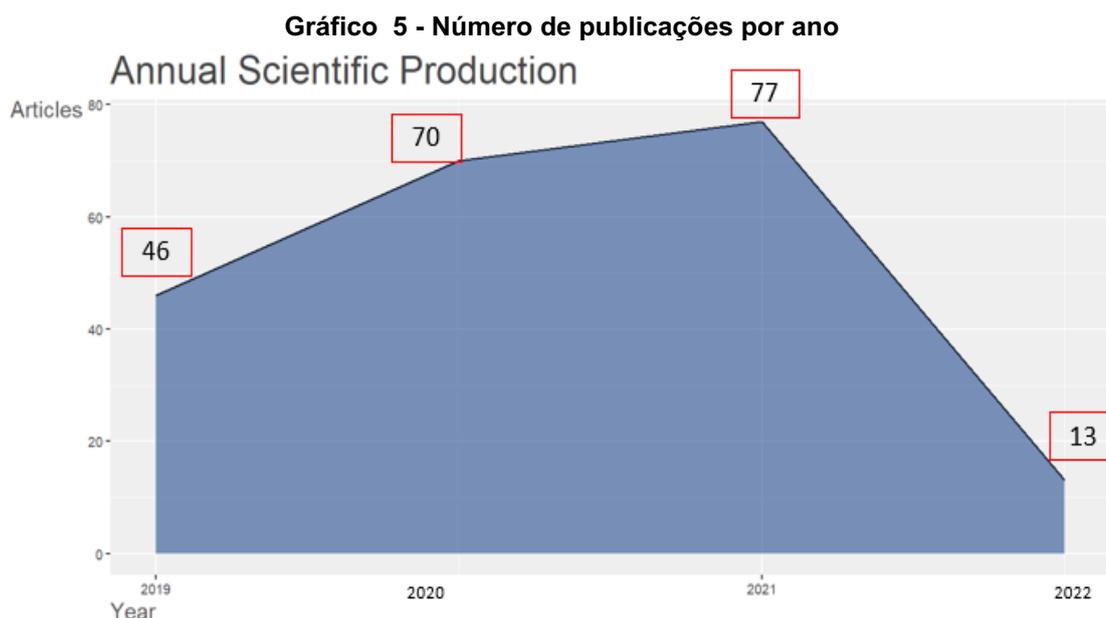
A seção de resultados foi desenvolvida levando em consideração as etapas e objetivos previamente propostos para a Análise Bibliométrica e Análise de Conteúdo.

### 4.1 ANÁLISE BIBLIOMÉTRICA

A Análise Bibliométrica deste trabalho toma como base os 206 artigos presentes na pesquisa realizada na base Scopus com o objetivo de adquirir uma melhor perspectiva do *Big Data* e como o tema influencia empresas inseridas em ambientes orientados a dados. As próximas seções apresentarão os principais resultados obtidos no *RStudio* e *Nvivo* com os parâmetros relatados no Quadro 5.

#### 4.1.1 Número de publicações por ano

O Gráfico 5 mostra a quantidade de artigos publicados de acordo com seu respectivo (iniciando em 2019):

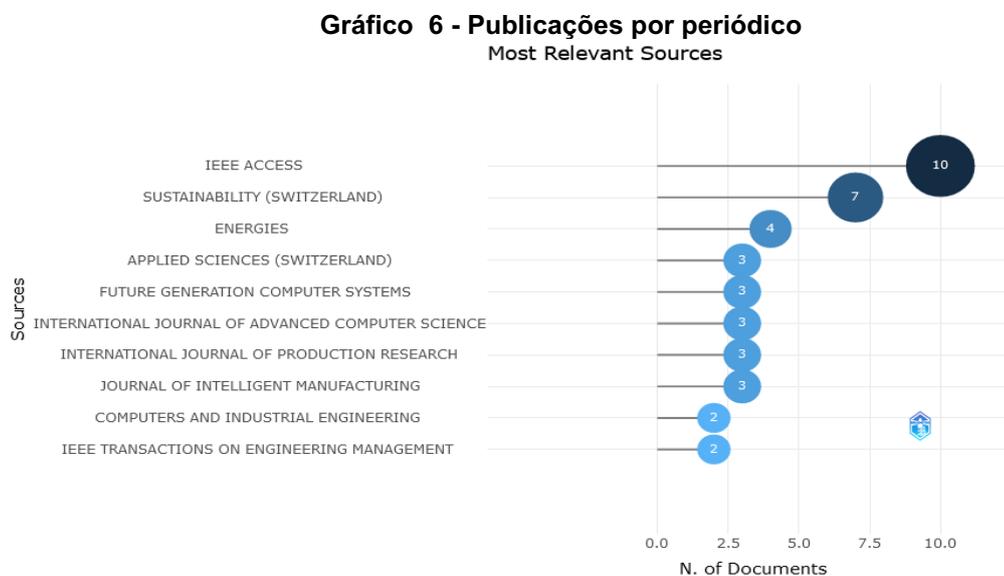


Fonte: Dados de pesquisa (2022)

Pode ser observado um crescimento no número de publicações por ano, desconsiderando o de 2022 que não havia se encerrado no momento da pesquisa.

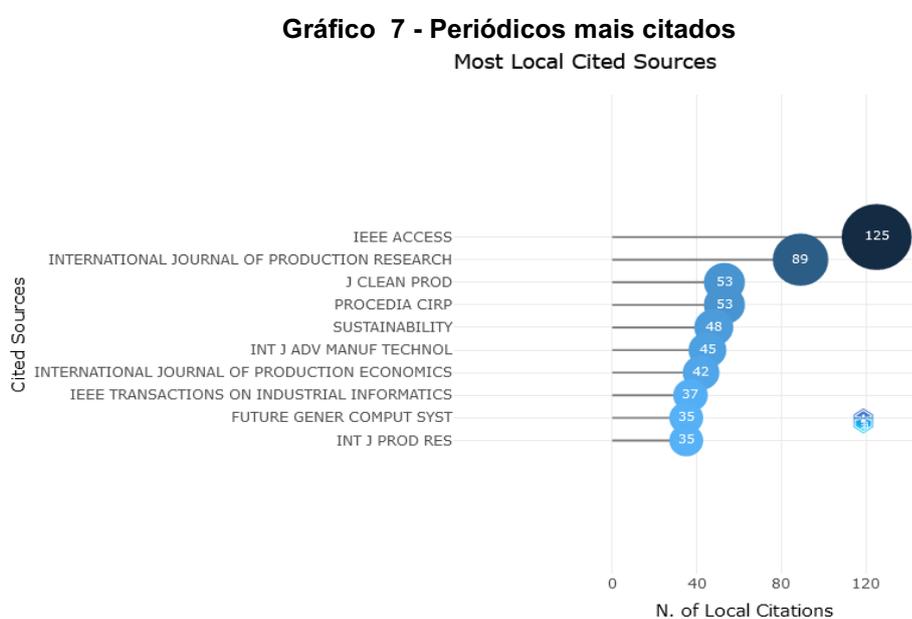
### 4.1.2 Periódicos

O Gráfico 6 mostra os principais periódicos de acordo com o número de publicações desde 2019:



Fonte: Dados de pesquisa (2022)

Tem-se como destaque a revista IEEE Access (*Institute of Electrical and Electronics Engineers*) com 10 publicações. O Gráfico 7 apresenta os periódicos mais citados no mesmo período:

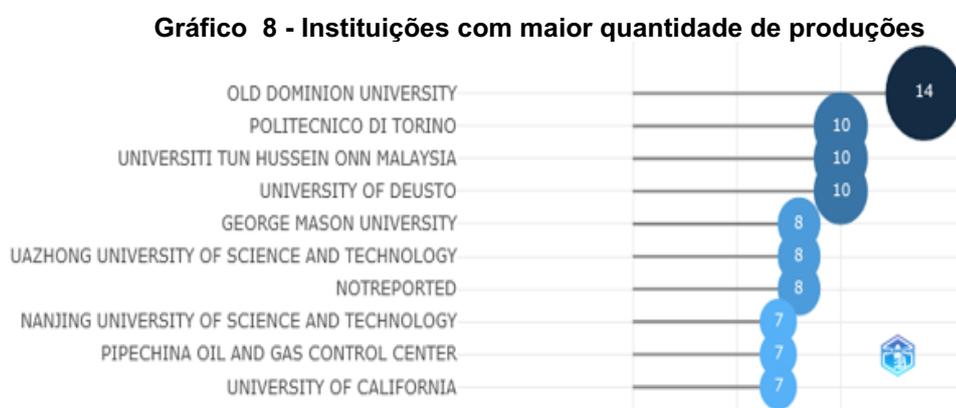


Fonte: Dados de pesquisa (2022)

Novamente, o destaque é da revista IIEE Access com 125 citações, 28,8% a mais que o segundo colocado (*International Journal of Production Research*) e 56,7% a mais comparado ao terceiro (*J Clean Prod.*). O IIEE Access pode ser destacado também por possuir um elevado *Citescore* (3,367) comparado aos outros periódicos da lista, sendo mais atrativo para publicações qualificadas.

#### 4.1.3 Instituições e dados por localização

Ao identificar as principais instituições que produziram material na temática, foram obtidos os resultados do Gráfico 8:



Fonte: Dados de pesquisa (2022)

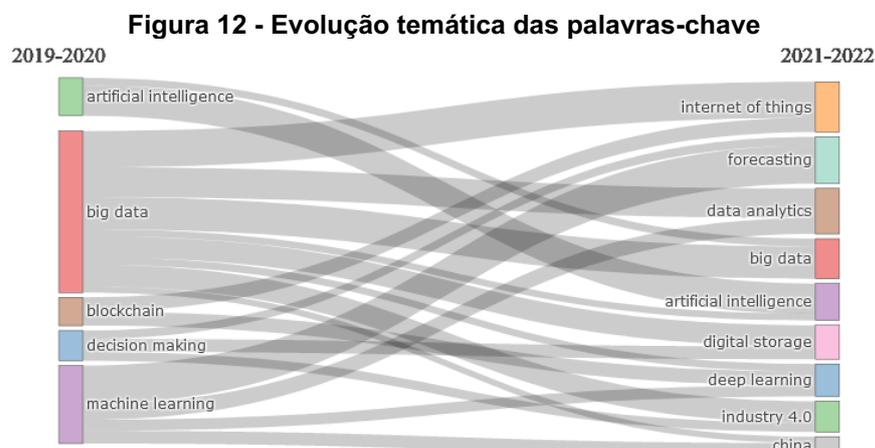
Nesse cenário, tem-se como destaque a *Old Dominion University*, uma instituição referência em modelagem e simulação, localizada nos Estados Unidos. Também tem destaque *Politecnico di Torino* (Italia), *Universiti Tun Hussien Onn Malaysia* (Malásia) e *University of Deusto* (Espanha).

Considerado a produção por países, foi gerado o resultado apresentado na Gráfico 9 com os dados presentes no Apêndice B:



Pode ser observado que os maiores *clusters* são formados pelos temas: *Big Data*, *IoT*, *Data Analytics*, *Industry 4.0* e *decision making*.

A biblioteca *bibliometrix* do RStudio permite visualizar a evolução temática com o passar dos anos, que pode ser observado na Figura 12:



Fonte: Dados de pesquisa (2022)

A Figura 12 mostra que os temas majoritariamente presentes no período 2019-2020 evoluíram para temas relacionado à *IoT*, *forecasting*, *data analytics*, *Big Data*, *artificial intelligence*, *digital storage*, *Aprendizagem Profunda*, *Industry 4.0* e *China* no período de 2021-2022. A Tabela 2 apresenta as principais palavras-chave que apareceram na pesquisa:

**Tabela 2 - Palavras-chave mais frequentes**

| Palavra-chave                  | Quantidade | Palavra-chave      | Quantidade |
|--------------------------------|------------|--------------------|------------|
| <b>Big Data</b>                | 60         | Data Mining        | 13         |
| <b>Decision Making</b>         | 24         | AM                 | 13         |
| <b>Industry 4 0</b>            | 23         | Human              | 11         |
| <b>Artificial Intelligence</b> | 20         | Manufacture        | 11         |
| <b>Data Analytics</b>          | 20         | Digital Storage    | 10         |
| <b>IoT</b>                     | 20         | Advanced Analytics | 9          |
| <b>Data Handling</b>           | 15         | Forecasting        | 9          |
| <b>Information Management</b>  | 15         | Machine Tools      | 9          |

|                              |    |                     |   |
|------------------------------|----|---------------------|---|
| <b>Aprendizagem profunda</b> | 14 | Industrial Research | 8 |
| <b>Article</b>               | 13 | Learning Systems    | 8 |

Fonte: Dados de pesquisa (2022)

Baseado nos resultados da Tabela 2, o Quadro 7 apresenta passagens de artigos da pesquisa pertinentes à algumas das principais palavras-chave apresentadas:

**Quadro 7- Palavras-chave e trechos de artigos**

| Palavras-Chave                 | Autor                         | Citação   |
|--------------------------------|-------------------------------|---|
| <i>Big Data</i>                | Rehman <i>et. al.</i> (2019)  | " <i>Big Data</i> se tornou uma área de pesquisa importante devido ao surgimento de enormes quantias de dados gerados de diferentes fontes como redes sociais, <i>Internet of Things</i> e aplicações multimídia. <i>Big Data</i> tem tido um papel importante nas áreas de tomada de decisão e previsão como sistemas de recomendação, análise de negócios, publicidade na internet, sistemas de saúde, transporte, detecção de fraudes, turismo e marketing." |
| <i>Industry 4.0</i>            | Sahal <i>et. al.</i> (2020)   | "O <i>IloT</i> (também conhecido como Indústria 4.0), que foi inicialmente concebido pelo governo alemão, é atualmente considerada a quarta revolução industrial. [...] Avanços em áreas como análise de dados, computação em nuvem e CPS tem alimentado a formação de atividades da Indústria 4.0 a fim de entregar flexibilidade, precisão e eficiência nunca vista em processos de manufatura."  |
| <i>Artificial Intelligence</i> | Mohamed <i>et. al.</i> (2019) | "Inteligência Artificial (IA) é a replicação digital de três habilidades humanas cognitivas: aprendizado, raciocínio e autocorreção. A maior parte dos sistemas de IA são derivados do aprendizado de máquina, aprendizado profundo, mineração de dados ou algoritmos baseados em regras."  |

Fonte: Dados de Pesquisa (2022)

As palavras "*Big Data*", "*Industry 4.0*" e "*Artificial Intelligence*" estão diretamente conectadas com o conceito e as ferramentas de *Big Data*, o que provavelmente incentivou os autores a escolherem estas como as principais palavras-chave dos artigos presentes na pesquisa realizada.

#### 4.1.5 Autores e coautores

Quanto aos autores com maior número de publicações, se destacam três pesquisadores chineses, listados na Tabela 3:

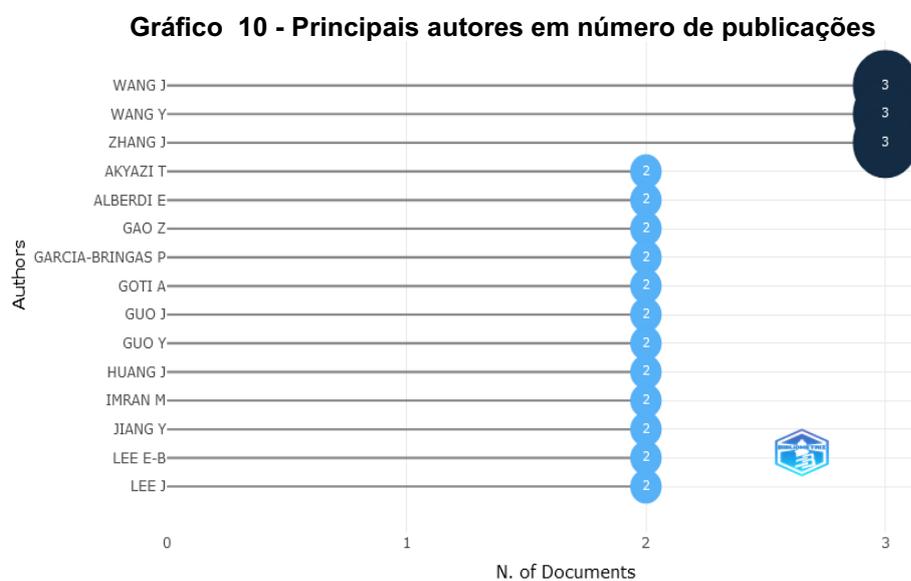
**Tabela 3 - Autores com maior número de publicações**

| <b>Autores</b>         | <b>Afiliação</b>                          | <b>h-index</b> | <b>Artigos</b> |
|------------------------|---|----------------|----------------|
| <b>Wang, JinJiang</b>  | China University of Petroleum-Beijing     | 27             | 3              |
| <b>Wang, Young</b>     | China Railway Seven Engineering Group Co. | 1              | 3              |
| <b>Zhang, Jiantong</b> | Tongji University                         | 7              | 3              |

Fonte: Dados de pesquisa (2022)

O *h-index* é um indicador que permite identificar o impacto do material produzido por pesquisadores relativo ao seu número total de publicações. Quanto maior o *h-index*, maior o impacto das pesquisas produzidas por determinado autor.

O Gráfico 10 apresenta a lista completa dos principais autores presentes na pesquisa realizada:



Fonte: Dados de pesquisa (2022)

A análise de coautoria fornece informações importantes sobre as redes colaborativas que moldam o capital intelectual de um tópico de estudo (Martínez-López *et al.* 2018). Inserindo os dados de pesquisa no *software* Nvivo, é possível identificar a relação entre autores e coautores do material bibliográfico. O *software* revelou que existem 721 ligações (relações de coautorias ou citações) entre autores que contribuíram para 2 ou mais artigos. Estas ligações podem formar grupos de ligações ou redes. A maior rede desta pesquisa contém três grupos com 24 ligações apresentadas na Figura 13:

**Figura 13 - Relação entre autores e coautores**



Fonte: Dados de pesquisa (2022)

Estes autores e coautores possuem em comum pesquisas nas áreas de Engenharia e Ciência da Computação, de acordo com seu histórico de publicações na base de dados *Scopus*.

#### 4.1.6 Discussão da análise bibliométrica

No momento em que a presente pesquisa foi realizada, não foram encontrados materiais de pesquisa bibliométrica que abordaram o tema *Big Data*, suas ferramentas, tecnologias e impactos em organizações dentro do período de 2019 até 2022. Porém existem materiais relevantes com temática similar apresentando resultados de recortes de anos diferentes que podem ser utilizados como parâmetro de comparação para as informações relatadas pela análise bibliométrica deste trabalho.

Khanra, Dhir e Mäntymäki (2020) realizaram um estudo bibliométrico com o tema de análise de *Big Data* em empresas onde afirmam que apenas em 2009 começaram a ser apresentadas publicações relevantes estudando o *Big Data* e seus impactos em organizações. O recorte de 2009 até 2019 traz um total de 3152 artigos na base de dados *Scopus*, sendo 1727 destes considerados para o estudo. Similarmente, Xu e Yu (2019) fizeram um estudo bibliométrico nas bases de dados *Science Citation Index* e *Social Science Citation Index* no período de 2009 até 2018 onde também pôde ser constatado um aumento no número de publicações com a temática *Big Data* ao longo do tempo como pode ser observado na Tabela 4:

| <b>Tabela 4 - Publicações de Big Data ao longo dos anos</b> |                    |
|---|--------------------|
| <b>Ano</b>  | <b>Publicações</b> |
| 2009  | 5                  |
| 2010  | 5                  |

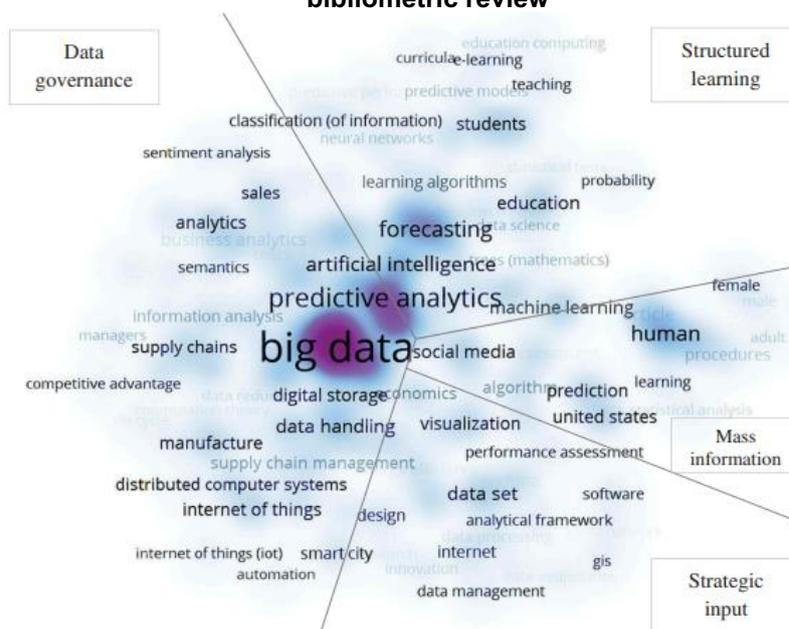
|       |       |
|-------|-------|
| 2011  | 15    |
| 2012  | 54    |
| 2013  | 278   |
| 2014  | 756   |
| 2015  | 1420  |
| 2016  | 2186  |
| 2017  | 2976  |
| 2018  | 3294  |
| Total | 10989 |

Fonte: Adaptado de Xu e Yu (2019)

Ao analisar o número de publicações nas pesquisas mencionadas e na secção 4.1.1 do presente trabalho, pode ser observado que a quantidade de publicações com esta temática aumentou ao longo dos anos.

Os temas presentes nas pesquisas da base de dados *Scopus* de Khanra, Dhir e Mäntymäki (2020), Figura 14, foram similares aos temas identificados no mapa apresentado na secção 4.1.4:

**Figura 14 - Mapa Temático da pesquisa Big data analytics and enterprises: a bibliometric review**



Fonte: Khanra, Dhir E Mäntymäki (2020)

Comparando os mapas dos temas de pesquisa, o maior destaque é no aumento dos clusters de *Industry 4.0* e *IoT* no mapa do presente trabalho comparado ao de Khanra, Dhir e Mäntymäki (2020).

Assandre e Martins (2019) realizaram uma análise bibliométrica a fim de estudar o desempenho do uso de *Big Data* em empresas. Eles analisaram um total de 398 materiais das bases de dados *Scopus* e *Web os Science*. O estudo afirma que há um crescimento no interesse por publicações com o tema *Big Data*, e usa como um argumento o aumento no número de publicações ao longo dos anos e a quantidade crescente de autores que publicaram material com este tema. Os principais autores deste estudo se encontram na tabela:

**Tabela 5 - Principais autores na pesquisa de Assandre e Martins**

| Autor      | Publicações |
|------------|-------------|
| Zhang J    | 6           |
| Chen L.    | 5           |
| Zhu M.     | 5           |
| Chen X.    | 4           |
| Lambert J. | 4           |
| Li J.      | 4           |
| Liu X.     | 4           |
| Wang Y.    | 4           |
| Zhang H.   | 4           |
| Kwon O.    | 4           |

Fonte: Assandre e Martins (2019)

É possível identificar Jiantong Zhang como o autor com maior número de publicações, e a relevância de Young Wang assim como foi destacado na secção 4.1.5. É possível destacar que Li J. Wang Y, presentes na rede de coautoria da secção 4.1.5 estão também presentes na lista de autores com maior número de publicações no estudo de Assandre e Martins (2019).

Sobre os dados apresentados na secção 4.1.2, Xu e Yu (2019) e Assandre e Martins (2019) também identificaram IIEE como o periódico com o maior número de documentos publicados com a temática estudada. Com relação aos países de origem destas pesquisas, Estados Unidos é país com o maior número de publicações nas

pesquisas de Khanra, Dhir e Mäntymäki (2020) e de Xu e Yu (2019), enquanto na secção 4.1.3 China é o país com maior número de artigos publicados e Estados Unidos em segundo.

## 4.2 ANÁLISE DE CONTEÚDO

A Análise de Conteúdo foi realizada considerando os parâmetros detalhados no Quadro 6: Objetivo, explanação sobre *Big Data*, ferramentas de *Big Data*, temática de pesquisa, resultados e metodologia.

### 4.2.1 Metodologia, objetivos e resultados

A fim de apresentar a metodologia, objetivos e resultados dos artigos contidos no portfólio, os materiais foram segmentados em 3 grupos distintos com os seguintes temas:

1. Utilização e impactos do uso de *Big Data*;
2. Descrição, uso ou criação de ferramentas de *Big Data*;
3. Tecnologias, processos e métodos de uso de *Big Data*.

Vale ressaltar que os três temas podem estar presentes nos artigos, porém sua segmentação foi selecionada considerado o tema principal abordado nas pesquisas em questão.

#### 4.2.1.1 Utilização e impactos do uso de *Big Data*

O Quadro 8 reúne a metodologia e os objetivos referentes aos artigos selecionados para a Análise de Conteúdo com a temática voltada para a utilização e impactos do uso de *Big Data*:

**Quadro 8 - Metodologia e Objetivos de Artigos Segmentados por Utilização e seus Impactos**

| Artigo  | Metodologia   | Objetivo  |
|---|---|---|
| <b>How Do Industry 4.0 Technologies Boost Collaborations in Buyer-Supplier Relationships?</b><br>(Patrucco <i>et. al.</i> , 2022) | Qualitativa<br>Aplicada<br>Descritiva<br>Pesquisa de levantamento | Averiguar como <i>Big Data</i> e computação em nuvem influenciam na relação entre compradores e fornecedores e o seu impacto na cadeia de suprimentos |

| Artigo   | Metodologia  | Objetivo   |
|--|--|--|
| <b>The Role of AI, AM, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities (Rathore et. al., 2021)</b> | Qualitativo<br>Aplicada<br>Descritivo<br>Bibliográfica                 | Realizar uma revisão bibliográfica analisando o papel da IA e AM na criação de DTs na indústria  |
| <b>The current sustainability scenario of Industry 4.0 enabling technologies in Indian manufacturing industries (Pasi et. al., 2021)</b>           | Qualitativo<br>Aplicada<br>Exploratória<br>Estudo de caso              | Determinar os fatores motivadores para a implementação da IIoT em empresas indianas, identificar as tecnologias capacitadoras da Industria 4.0 e explorar o impacto destas tecnologias nos pilares de sustentabilidade |
| <b>Technology-driven 5G enabled e-healthcare system during COVID-19 pandemic (Kamruzzaman et. al., 2021)</b>                                       | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Survey         | Realizar um levantamento de meios pelos quais tecnologias emergentes como 5G e <i>Big Data</i> podem auxiliar no combate a COVID-19  |
| <b>Mining open government data for business intelligence using data visualization: A two-industry case study (Gottfried, 2021)</b>                 | Quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso             | Explorar como dados públicos governamentais podem ser usados para gerar insights para a identificação de oportunidades de mercado e formulação de estratégias  |
| <b>How and when do Big Data investments pay off? The role of marketing affordances and service innovation (De Luca et. al., 2021)</b>              | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Survey         | Desenvolver uma estrutura para identificar os impactos no desempenho de serviços de investimentos empresariais em <i>Big Data</i>  |
| <b>Exploring the Factors Influencing Big Data Technology Acceptance (Rahman et. al., 2021)</b>   | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Fornecer aos profissionais de TI <i>insights</i> sobre como <i>Big Data</i> é capaz de aumentar o desempenho de decisões orientadas por dados das organizações   |
| <b>Big Data Impacting Dynamic Food Safety Risk Management in the Food Chain (Donaghy et. al., 2021)</b>  | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Fornecer uma visão do uso prospectivo de <i>Big Data</i> para gerenciamento dinâmico de riscos a partir de uma perspectiva de segurança alimentar  |
| <b>The state of the art and taxonomy of Big Data analytics: view from new Big Data framework (Mohamed et. al., 2021)</b>                           | Qualitativo<br>Aplicada<br>Exploratória<br>Estudo bibliográfico        | Compreender a tendência da pesquisa relacionada a <i>Big Data</i> e os quadros atuais das tecnologias de <i>Big Data</i> ; identificar tendências no uso ou pesquisa de ferramentas de <i>Big Data</i>                 |
| <b>Shared-use mobility competition: a trip-level analysis of taxi, bikeshare, and transit mode choice in Washington, DC (Welch et. al., 2021)</b>  | Quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso             | Aplicar ferramentas de <i>Big Data</i> para descobrir os preditores que motivam meio bilhão de viagens de transporte público, táxi e bicicletas compartilhadas em estações de trem de Washington.                      |

| Artigo   | Metodologia   | Objetivo  |
|--|---|---|
| <b>Integration of Lean practices and Industry 4.0 technologies:</b><br>(Shahin et. al., 2021)  | Qualitativo<br>Aplicada<br>Exploratória<br>Experimental             | Elaborar uma revisão sobre os vínculos entre as ferramentas Lean e as tecnologias da IIoT e como a implementação simultânea desses dois paradigmas afeta o desempenho operacional das fábricas. |
| <b>Big Data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case</b><br>(Sahal et. al., 2021) | Qualitativo<br>Aplicada<br>Exploratória<br>Bibliográfico            | Realizar uma revisão da literatura dos pontos fortes e fracos das tecnologias de processamento de fluxo de <i>Big Data</i> para a Indústria 4.0   |
| <b>Big Data analytics enhanced healthcare systems: a review</b><br>(Shafqat et. al., 2021)   | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Survey      | Apresentar uma pesquisa de diferentes sistemas integrados de saúde que utilizam <i>Big Data</i> e descrever algoritmos, técnicas e ferramentas que podem ser utilizados em sistemas de saúde    |
| <b>A Design-to-Device Pipeline for Data-Driven Materials Discovery</b><br>(Cole, 2021)   | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Experimento | Identificar como o <i>Big Data</i> pode reduzir o tempo de ciclo de descobrimento de novos materiais  |
| <b>The impact of Big Data analytics on company performance in supply chain management</b><br>(Oncioiu et. al., 2019)                                 | Quantitativo<br>Aplicada<br>Exploratória<br>Survey                  | Identificar como a análise de <i>Big Data</i> pode ajudar as empresas romenas a avaliar sua experiência, estratégias e capacidades na implementação da análise de <i>Big Data</i> ,             |
| <b>Barriers to Big Data analytics in manufacturing supply chains: A case study from Bangladesh</b><br>(Moktadir et. al., 2019)                       | Qualitativo<br>Aplicada<br>Exploratória<br>Bibliográfico            | Esta pesquisa explora as barreiras no uso de <i>Big Data</i> utilizando um processo de hierarquia analítica baseado em Delphi (AHP)   |

Fonte: Dados de pesquisa (2022)

De forma geral, os artigos exploram como o uso de *Big Data* pode impactar positivamente a sociedade e organizações por meio da geração de valor. Os resultados dos artigos estão apresentados no Quadro 9:

**Quadro 9 - Resultados de Artigos Segmentados por Utilização do Big Data e seus Impactos**

| Artigo   | Resultados  |
|--|---|
| <b>How Do Industry 4.0 Technologies Boost Collaborations in Buyer-Supplier Relationships?</b><br>(Patrucco et. al., 2022)          | A pesquisa identificou que existe uma grande amplitude de possibilidades de uso de DT e apresentou alguns pontos que os conectam  |
| <b>The Role of AI, AM, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges</b><br>(Rathore et. al., 2021) | A revisão permitiu destacar os desafios e potenciais de pesquisa em diversas áreas. Também foram identificados os critérios e ferramentas de DT que auxiliam no seu desenvolvimento bem executado. Por fim, foi projetado um modelo de referência para um sistema de DT |

| Artigo  | Resultados   |
|---|--|
| <p><b>The current sustainability scenario of Industry 4.0 enabling technologies in Indian manufacturing industries</b><br/>(Pasi <i>et. al.</i>, 2021)</p>          | <p>Após a coleta de dados por meio de questionários, verificou-se que o impacto das tecnologias capacitadoras I4.0 nos pilares da sustentabilidade é significativo. Foi proposto um roteiro para implementar tecnologias da Indústria 4.0 e melhorar a sustentabilidade nas indústrias manufatureiras indianas</p>   |
| <p><b>Technology-driven 5G enabled e-healthcare system during COVID-19 pandemic</b><br/>(Kamruzzaman <i>et. al.</i>, 2021)</p>                                      | <p><i>Big Data</i> e inteligência artificial ajudam a identificar pessoas infectadas e manter registros para uso posterior. Como o COVID-19 é um coronavírus de transmissão rápida, <i>Big Data</i> e inteligência artificial podem ajudar a identificar e controlar as áreas de transmissão do vírus de forma rápida</p>  |
| <p><b>Mining open government data for business intelligence using data visualization: A two-industry case study</b><br/>(Gottfried, 2021)</p>                       | <p>Os resultados fornecem evidências de que dados governamentais públicos podem ser uma valiosa fonte de informação para gerar inteligência de negócios e demonstrar como as ferramentas de modelagem e visualização de tópicos podem auxiliar as organizações na extração e análise de informações para a identificação de oportunidades</p>                                  |
| <p><b>How and when do <i>Big Data</i> investments pay off? The role of marketing affordances and service innovation</b><br/>(De Lucas <i>et. al.</i>, 2021)</p>     | <p>Após a inserção de dados de surveys com mais de 400 empresários de empresas de referência dos estados unidos, o modelo matemático proposto inferiu que os investimentos em <i>Big Data</i> compensam mais em indústrias altamente digitalizadas</p>   |
| <p><b>Exploring the Factors Influencing <i>Big Data</i> Technology Acceptance</b><br/>(Rahman <i>et. al.</i>, 2021)</p>   | <p>A análise do modelo estrutural revela que o os testes de hipóteses são significativos para as características: escalabilidade, processamento de dados, flexibilidade, confiabilidade, performance, qualidade de saída de dados e treinamentos</p>   |
| <p><b><i>Big Data</i> Impacting Dynamic Food Safety Risk Management in the Food Chain</b><br/>(Donaghy <i>et. al.</i>, 2021)</p>                                    | <p>Os autores identificaram que a utilização de <i>Big Data</i> e suas ferramentas podem prever riscos na cadeia de suprimentos de alimentos</p>   |
| <p><b>The state of the art and taxonomy of <i>Big Data</i> analytics: view from new <i>Big Data</i> framework</b><br/>(Mohamed <i>et. al.</i>, 2021)</p>            | <p>O artigo apresenta uma revisão da literatura que analisa o uso de ferramentas de <i>Big Data</i> e técnicas de análise de <i>Big Data</i> em áreas como saúde e assistência médica, redes sociais, governo, setor público, gestão de recursos naturais, setor econômico e empresarial para que os leitores tenham uma melhor noção de como e quais ferramentas utilizar</p> |
| <p><b>Shared-use mobility competition: a trip-level analysis of taxi, bikeshare, and transit mode choice in Washington, DC</b><br/>(Welch <i>et. al.</i>, 2021)</p> | <p>Os resultados do estudo indicam que o custo de viagem e os fatores ambientais afetam significativamente a probabilidade de um indivíduo viajar de táxi ou bicicletas compartilhadas em vez de trem</p>  |

| Artigo   | Resultados  |
|--|---|
| <b>Big Data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case</b><br>(Sahal et. al., 2021) | O artigo pôde identificar um conjunto de requisitos para casos de uso de manutenção preditiva, principalmente no setor de transporte ferroviário e o setor de energia de turbinas eólicas. Também. fornece um mapeamento abrangente dos requisitos de manutenção preditiva para tecnologias de streaming de <i>Big Data</i> de código aberto  |
| <b>Big Data analytics enhanced healthcare systems: a review</b><br>(Shafqat et. al., 2021)   | Foi formado um ponto de convergência de todas essas plataformas, técnicas e ferramentas em forma do sistema SmartHealth que pode resultar na padronização do uso de dados da área da saúde em um futuro próximo   |
| <b>A Design-to-Device Pipeline for Data-Driven Materials Discovery</b><br>(Cole, 2021)   | O artigo apresenta um guia de como a ciência de dados pode permitir a descoberta de materiais por meio de um pipeline de quatro etapas “design-to-device” que envolve (1) extração de dados, (2) enriquecimento de dados, (3) previsão de materiais e (4) validação experimental  |
| <b>Barriers to Big Data analytics in manufacturing supply chains: A case study from Bangladesh</b><br>(Moktadir et. al., 2019)                       | As conclusões desta pesquisa são: (i) as barreiras relacionadas à coleta de dados e tecnologia são as mais relevantes, (ii) os cinco componentes mais importantes dessas barreiras são falta de infraestrutura, complexidade da integração de dados, privacidade dos dados, falta de disponibilidade de ferramentas de análise e alto custo de investimento   |
| <b>The role of Big Data analytics in industrial IoT</b><br>(Rehman et. al., 2019)  | São apresentados estudos de caso das várias empresas que se beneficiaram da análise de <i>Big Data</i> . Também são enumeradas as oportunidades geradas pelo uso de <i>Big Data</i> na IIoT   |
| <b>The impact of Big Data analytics on company performance in supply chain management</b><br>(Oncioiu et. al., 2019)                                 | Os dados coletados foram analisados com o pacote Statistical Package for the Social Sciences (SPSS) utilizando tabelas de frequência, tabelas de contingência e análise de componentes principais. As principais contribuições desta pesquisa destacam o fato de as empresas se preocuparem em identificar novos métodos, ferramentas e abordagens estatísticas, como computação em nuvem e tecnologias de segurança, que precisam ser rigorosamente exploradas |
| <b>Integration of Lean practices and Industry 4.0 technologies: smart manufacturing for next-generation enterprises</b><br>(Shahin et. al., 2021)    | O artigo explora as soluções existentes e potenciais das práticas Lean habilitadas por tecnologias associadas ao <i>Big Data</i> , como redes sem fio, computação em nuvem e realidade virtual. Também foi desenvolvido um sistema de suporte à decisão Kanban baseado em computação em nuvem   |

Fonte: Dados de pesquisa (2022)

Pôde ser observado que 16 dos artigos presentes no portfólio exploraram meios de como utilizar *Big Data* e, todas as pesquisas desta seção identificaram que sua utilização pode impactar positivamente a sociedade e organizações. Rehman *et. al.* (2019) mostraram o papel do *Big Data* em empresas orientadas por dados, e como

seu uso aprimora a eficiência na tomada de decisão, utilização de recursos, segurança de dados, entre outros pontos positivos. Oncioiu *et. al.* (2019) e Patrucco (2022) mostram como empresas podem utilizar *Big Data* para otimizar processos relacionadas à cadeia de suprimentos, aumentando a integração de informações entre comprador-fornecedor. Rahman *et. al.* (2020) relatam que o uso de ferramentas de *Big Data* pode mitigar desafios relacionados a escalabilidade, processamento, flexibilidade, performance e qualidade de saída de dados. De Luca *et. al.* (2021) produziram um estudo que apresenta uma diretriz para que empresas possam analisar quando investimentos em *Big Data* podem gerar retorno ou não, realçando a importância da digitalização de empresas. Welch *et. al.* (2020) desenvolveram um modelo de AM que pode ser aplicado em diferentes empresas de transporte que utilizam *Big Data* para identificar a motivação dos passageiros em utilizarem um meio de transporte em questão (seja preço, questões ambientais, distância, entre outros).

#### 4.2.1.2 Descrição, uso ou criação de ferramentas de Big Data

O Quadro 10 reúne a metodologia e os objetivos referentes aos artigos que possuem como tema central a descrição, uso ou criação de ferramentas de *Big Data*:

**Quadro 10 - Metodologia e Objetivos de Artigos Segmentados por Descrição, Uso ou Criação de Ferramentas de Big Data**

| Artigo  | Metodologia  | Objetivo  |
|---|--|---|
| <b>State-of-the-art survey on digital twin implementations (Liu <i>et. al.</i>, 2022)</b>                           | Qualitativa<br>Aplicada<br>Descritiva<br>Pesquisa bibliográfica        | Esclarecer quais são os diferentes tipos de implementações de DT, destacar pesquisas que utilizaram RA para visualização de dados em DT e analisar os principais desafios na integração entre DT e RA |
| <b>Human health Big Data evaluation based on FPGA processor and Big Data decision algorithm (Sun e Yang, 2021)</b>  | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Propor um modelo de utilização de <i>Big Data</i> para analisar dados médicos   |
| <b>The architecture development of Industry 4.0 compliant smart machine tool system (Jeon <i>et. al.</i>, 2021)</b> | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Propor uma arquitetura funcional para versões Indústria 4.0 de máquinas-ferramenta (sistema de máquina-ferramenta inteligente)  |
| <b>REDTag: a Predictive Maintenance Framework for Parcel Delivery Services (Proto <i>et. al.</i>, 2021)</b>         | Quantitativo<br>Aplicada<br>Exploratória<br>Experimento                | Apresentar uma solução de entrega de fretes facilmente rastreável por meio de <i>Big Data</i> e IoT   |

| Artigo  | Metodologia  | Objetivo   |
|---|--|--|
| <b>BlockIoTelligence: A Blockchain-enabled Intelligent IoT Architecture with Artificial Intelligence (Singh et. al., 2021)</b>        | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Propor uma arquitetura Inteligente de IoT habilitada para Blockchain com Inteligência Artificial que fornece uma maneira eficiente de convergir blockchain e IA para IoT com técnicas e aplicativos atuais               |
| <b>A method of NC machine tools intelligent monitoring system in smart factories (Liu et. al., 2021)</b>                              | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Propor uma estrutura de dados bidirecionais e fluxos de controle entre máquinas-ferramenta, apresentando obstáculos importantes e as soluções correspondentes são posteriormente investigadas e analisadas profundamente |
| <b>DataWarrior: an evaluation of the open-source drug discovery tool (López-López et. al., 2019)</b>                                  | Qualitativo<br>Aplicada<br>Exploratória<br>Estudo de caso              | Realizar uma avaliação da ferramenta de descoberta de medicamentos de código aberto DataWarrior  |
| <b>Performance prediction of trace metals and cod in wastewater treatment using artificial neural network (Matheri et. al., 2021)</b> | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Utilizar um modelo de rede neural artificial para prever a qualidade de oligoelementos e demanda química de oxigênio em águas residuais  |
| <b>Government Construction Project Budget Prediction Using AM (Kusonkhum et. al., 2022)</b>   | Quantitativa<br>Aplicada<br>Exploratória<br>Experimental               | O estudo tem o objetivo de prever se projetos de construção passarão do limite orçamentário estabelecido por meio de um algoritmo de AM  |

Fonte: Dados de pesquisa (2022)

As ferramentas tratadas nestes artigos buscaram acelerar ou automatizar processos que com o uso do *Big Data* possuem o potencial de consumir menos tempo e reduzir o desperdício de recursos. Os resultados dos artigos apresentados no Quadro 10 estão presentes no Quadro 11:

**Quadro 11 - Resultados de Artigos Segmentados por Descrição, Uso ou Criação de Ferramentas de Big Data**

| Artigo   | Resultados   |
|--|--|
| <b>State-of-the-art survey on digital twin implementations (Liu et. al., 2022)</b> | A pesquisa identificou que existe uma grande amplitude de possibilidades de uso de DT e apresentou pontos que os conectam  |
| <b>Human health Big Data evaluation based on FPGA (Sun e Yang, 2021)</b>           | O trabalho resultou um modelo de três etapas: 1) processo de eliminação de ruído e dados irrelevantes 2) seleção de dados baseado no modelo de árvore 3) modelo de AM para analisar os dados |
| <b>The architecture development of Industry 4.0 SMT (Jeon et. al., 2021)</b>       | O procedimento de implementação e um estudo de caso ilustrativo são apresentados para a aplicação da arquitetura funcional   |

| Artigo  | Resultados   |
|---|--|
| <b>REDTag: a Predictive Maintenance Framework for Parcel Delivery Services (Proto et. al., 2021)</b>                                  | Os resultados experimentais confirmam a capacidade preditiva dos modelos extraídos do programa proposto. Os autores obtiveram sucesso em apresentar uma tecnologia de manutenção preditiva em logística de transporte inteligente para prever danos à pacotes a partir de dados do mundo real  |
| <b>BlockIoTelligence: A Blockchain-enabled Intelligent IoT Architecture with Artificial Intelligence (Singh et. al., 2021)</b>        | Os resultados mostram que a arquitetura proposta tem desempenho superior sobre as arquiteturas de IoT existentes para Blockchain e mitigam os desafios atuais  |
| Artigo  | Resultados   |
| <b>A method of NC machine tools intelligent monitoring system in smart factories (Liu et. al., 2021)</b>                              | Com base em um “data bar” uniforme, foi desenvolvido um sistema inteligente de monitoramento e processamento de dados de máquinas-ferramenta em fábricas inteligentes, que valida o conceito apresentado no artigo, e funciona bem no ambiente de fabricação   |
| <b>DataWarrior: an evaluation of the open-source drug discovery tool (López-López et. al., 2019)</b>                                  | O DataWarrior se destaca como uma tecnologia que combina predição de propriedades físico-químicas de interesse farmacêutico, cálculos quimioinformáticos, análise multivariada de dados e visualização interativa com gráficos dinâmicos. As ferramentas de quimioinformática bem estabelecidas implementadas no DataWarrior, bem como os algoritmos inovadores, tornam a tecnologia útil e atraente, conforme revelado pelo número crescente de aplicações documentadas |
| <b>Performance prediction of trace metals and cod in wastewater treatment using artificial neural network (Matheri et. al., 2021)</b> | O modelo proposto foi testado e aprovado em um estudo de caso como um método eficiente para analisar a qualidade de água residual  |
| <b>Government Construction Project Budget Prediction Using AM (Kusonkhum et. al., 2022)</b>   | De acordo com os resultados do teste, o modelo KNN atingiu uma precisão de 0,86, demonstrando que ele pode ser usado para prever os projetos de construção que estarão acima do orçamento de acordo com os dados fornecidos  |

Fonte: Dados de pesquisa (2022)

Independente se o artigo possui o objetivo de descrever ou desenvolver uma ferramenta que utiliza *Big Data*, IA e AM estão presentes em todos os materiais desta seção. A pesquisa de Jeon *et. al.* (2021) desenvolveu uma arquitetura funcional de *smart machines* em ambientes da IIoT, O *software* apresentado por Proto *et. al.* (2021) foi capaz de utilizar AM para prever danos à pacotes de entregas. López-López *et. al.* (2019) apresentaram o uso de um software para a indústria farmacêutica capaz de identificar com maior rapidez propriedades físico-químicas de medicamentos por meio de AM. Kusonkhum *et. al.* (2022) desenvolveram um algoritmo de AM capaz de analisar um conjunto de dados de projetos de construção civil e predizer se ele ficará acima do limite orçamentário.

#### 4.2.1.3 Tecnologias, processos e métodos de uso do Big Data

Sobre tecnologias, processos e métodos de uso do *Big Data*, o Quadro 12 expõe a metodologia e objetivos dos artigos presentes neste segmento:

**Quadro 12 - Metodologia e Objetivos de Artigos Segmentados por Tecnologias, Processos e Métodos de Uso do Big Data**

| Artigo  | Metodologia  | Objetivo   |
|---|--|--|
| <b>Designing early warning systems for detecting systemic risk: A case study and discussion (Wever et. al., 2022)</b>   | Quantitativa<br>Aplicada<br>Exploratória<br>Experimental               | Identificar as características dos sistemas de alerta precoce, descrever uma abordagem para projetar estes sistemas e elabora sobre o valor desta abordagem  |
| <b>Building a model to exploit association rules and analyze purchasing behavior based on rough set theory (Tran e Huh, 2022)</b>                                   | Quantitativa<br>Aplicada<br>Exploratória<br>Estudo de caso             | Neste artigo, os autores pesquisam o sistema de recomendação, a teoria dos conjuntos aproximados, e a teoria dos conjuntos <i>fuzzy</i> , construindo assim um modelo parcial de <i>Rough Set</i>              |
| <b>Operational time-series data modeling via LSTM network integrating principal component analysis based on human experience (Yang et. al., 2021)</b>               | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Propor a Análise de Componentes Principais Integrantes de Memória de Longo Prazo com Base na Experiência Humana, que utiliza dados de séries temporais operacionais para prognósticos de saúde de equipamentos |
| <b>Industry 4.0 technologies in tourism education: Nurturing students to think with technology (Bilotta et. al., 2021)</b>  | Qualitativo<br>Aplicada<br>Exploratória<br>Pesquisa de levantamento    | Identificar habilidades conceituais, metodológicas, tecnológicas e práticas relacionadas à <i>Big Data</i> a serem desenvolvidas em um currículo acadêmico para estudantes de Ciências do Turismo              |
| <b>Exploring IT/IS proactive and knowledge transfer on enterprise digital business transformation (EDBT): a technology-knowledge perspective (Li et. al., 2021)</b> | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br><i>Survey</i>  | Desenvolver um modelo de transferência de conhecimento de tecnologia da informação durante o processo de transformação de empresas na era digital  |
| <b>Artificial Intelligence and Technology in Teaching Negotiation (Dinnar et. al., 2021)</b>  | Qualitativo<br>Aplicada<br>Exploratória<br>Estudo de caso              | Descrever sistemas que auxiliam negociadores humanos a avaliar e aprender com simulações de IA   |

| <b>Artigo</b>  | <b>Metodologia</b>   | <b>Objetivo</b>  |
|--|--|--|
| <b>A <i>Big Data</i>-driven framework for sustainable and smart additive manufacturing</b> (Majeed et. al., 2021)  | Qualitativo<br>Aplicada<br>Exploratória<br>Estudo de caso              | O artigo visa desenvolver uma estrutura combinando análise de <i>Big Data</i> , manufatura aditiva e tecnologias sustentáveis de manufatura inteligente para empresas de manufatura aditiva    |
| <b>A <i>Big Data</i>-centric architecture metamodel for Industry 4.0</b> (López et. al., 2021)   | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Propor uma arquitetura de referência centrada em dados, distribuída e dinamicamente escalável, que serve como modelo para empresas em busca da digitalização vertical e horizontal             |
| <b>Towards data-driven approaches in manufacturing: an architecture to collect sequences of operations</b> (Farooqui et. al., 2021)                                      | Quantitativo<br>Aplicada<br>Exploratória<br>Experimento                | Apresentar uma arquitetura de dados baseada em eventos, que pode ser aplicada a sistemas pré-Indústria 4.0, bem como a novos sistemas de última geração, para coletar dados de chão de fábrica |
| <b>Lean 4.0: A new holistic approach for the integration of lean manufacturing tools and digital technologies</b> (Valamede e Akkari, 2021)                              | Qualitativo<br>Aplicada<br>Exploratória<br>Estudo de caso              | Este artigo propôs uma integração de ferramentas do Lean Manufacturing e tecnologias da indústria 4.0  |
| <b>Identification and classification of materials using machine vision and AM in the context of industry 4.0</b> (Penumuru et. al., 2021)                                | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Experimental   | Apresentar uma metodologia geral para identificação automatizada de materiais utilizando tecnologias AM para contribuir com as habilidades cognitivas de máquinas-ferramentas                  |
| <b>Creating a road map for industry 4.0 by using an integrated fuzzy multicriteria decision-making methodology</b> (Kaya et. al., 2021)                                  | Qualitativo-quantitativo<br>Aplicada<br>Exploratória<br>Estudo de caso | Desenvolver um método de tomada de decisão para definir a melhor metodologia de transição de uma empresa para a indústria 4.0  |
| <b>Application of industrial <i>Big Data</i> for smart manufacturing in product service system based on system engineering using fuzzy DEMATEL</b> (Zhang et. al., 2021) | Qualitativo<br>Aplicada<br>Exploratória<br>Estudo de caso              | Propor um modelo de sistema para <i>Big Data</i> industrial a partir da análise de engenharia de sistemas  |

Fonte: Dados de pesquisa (2022)

Os artigos agrupados nesta secção possuem o intuito de explorar as tecnologias que sustentam o *Big Data* e metodologias, boas práticas e soluções para sua implementação. Os resultados dos artigos do Quadro 12 são apresentados no Quadro 13:

**Quadro 13 - Resultados de Artigos Segmentados por Tecnologias, Processos e Métodos de Uso do Big Data**

| Artigo  | Resultados  |
|---|---|
| <b>Designing early warning systems for detecting systemic risk: A case study and discussion (Wever et. al., 2022)</b>   | O artigo consegue descrever com sucesso o sistema de alerta precoce e destaca a importância de realizar seu design no início dos processos  |
| <b>Building a model to exploit association rules and analyze purchasing behavior based on rough set theory (Tran e Huh, 2022)</b>                                   | O autor desenvolveu um <i>software</i> capaz de detectar padrões de consumo de clientes baseado em um histórico de compras e características do perfil do cliente   |
| <b>Operational time-series data modeling via LSTM network integrating principal component analysis based on human experience (Yang et. al., 2021)</b>               | O modelo proposto é validado em um caso prático e utiliza um método de análise de aprendizado profundo, LSTM, para processar totalmente os dados da série de monitoramento; ele extrai efetivamente os recursos envolvidos com a experiência humana e leva em consideração as atualizações dinâmicas. |
| <b>Industry 4.0 technologies in tourism education: Nurturing students to think with technology (Bilotta et. al., 2021)</b>  | O trabalho conseguiu identificar os principais pontos que devem ser abordados em um currículo tecnológico da área de Ciência do Turismo   |
| <b>Exploring IT/IS proactive and knowledge transfer on enterprise digital business transformation (EDBT): a technology-knowledge perspective (Li et. al., 2021)</b> | Foi apresentado um modelo que possui o objetivo de orientar os gerentes corporativos sobre como gerenciar seus recursos de conhecimento e maximizar as tecnologias emergentes para alcançar a transformação dos negócios digitais corporativos  |
| <b>Artificial Intelligence and Technology in Teaching Negotiation (Dinnar et. al., 2021)</b>  | Foram desenvolvidas ferramentas que permitirão melhorias não apenas no ensino da negociação, mas também no ensino de humanos em como programar e colaborar com sistemas de negociação baseados em tecnologia, incluindo avatares e agentes de negociação controlados por computador                   |

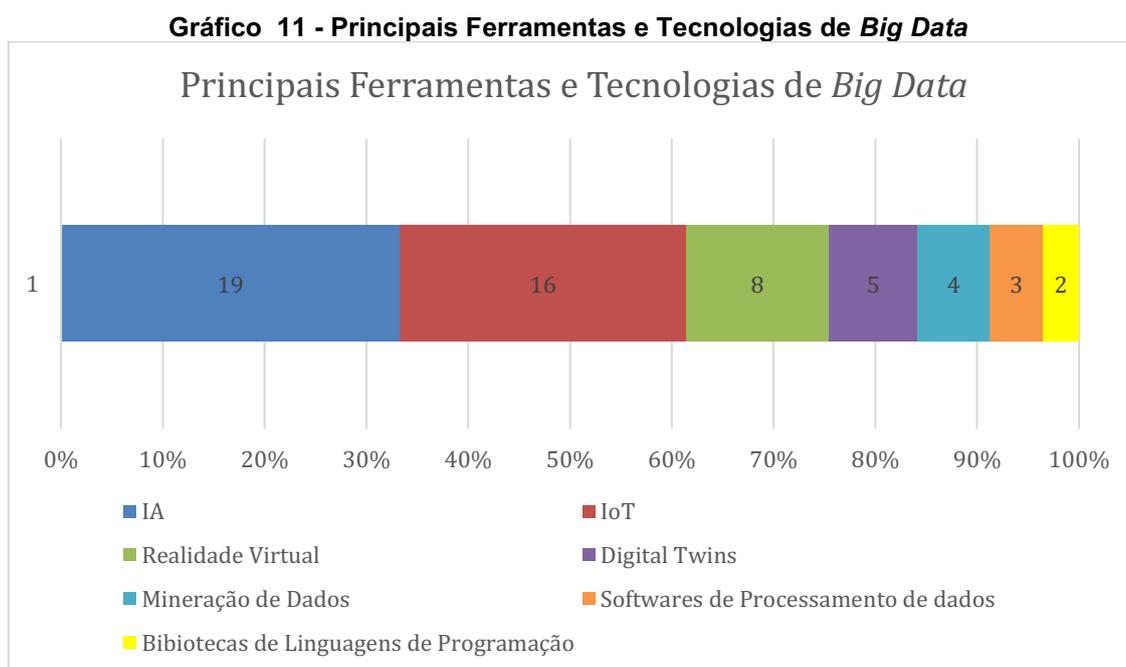
| Artigo   | Resultados   |
|--|--|
| <b>A <i>Big Data</i>-driven framework for sustainable and smart additive manufacturing (Majeed et. al., 2021)</b>  | Foi proposto uma estrutura de fabricação aditiva inteligente e sustentável orientada por <i>Big Data</i> que pode auxiliar líderes da indústria a tomar melhores decisões para o estágio de início de ciclo de vida do produto   |
| <b>A <i>Big Data</i>-centric architecture metamodel for Industry 4.0 (López et. al., 2021)</b>   | Foi criado um metamodelo que coleta a descrição de todos os elementos envolvidos em uma plataforma digital (dados, recursos, aplicações e métricas de monitoramento) bem como as informações necessárias para configurar, implantar e executar aplicações nela. O modelo foi testado com sucesso em uma simulação de distribuição de energia em uma cidade inteligente |
| <b>Towards data-driven approaches in manufacturing: an architecture to collect sequences of operations (Farooqui et. al., 2021)</b>                                      | A arquitetura foi desenvolvida com sucesso no <i>software</i> “ <i>Sequence Planner</i> ” e implementado como um projeto piloto em uma empresa do setor automobilístico  |
| <b>Lean 4.0: A new holistic approach for the integration of lean manufacturing tools and digital technologies (Valamede e Akkari, 2021)</b>                              | Com base em uma metodologia de três etapas, que incluiu mapeamento tecnológico e industrial, foram identificados 25 pontos de sinergia onde há relação entre Lean Manufacturing e Indústria 4.0. Foi comprovado que a utilização destas ferramentas permite uma integração maior dos processos e redução de desperdício  |
| <b>Identification and classification of materials using machine vision and AM in the context of industry 4.0 (Penumuru et. al., 2021)</b>                                | A metodologia apresentada foi validada através da realização de quatro experimentos para verificar a precisão de classificação do classificador utilizando no experimento materiais de cobre, alumínio, MDF e Aço  |
| <b>Creating a road map for industry 4.0 by using an integrated fuzzy multicriteria decision-making methodology (Kaya et. al., 2021)</b>                                  | Um modelo fuzzy foi criado e um roteiro de transição para a indústria 4.0 foi determinado em relação às prioridades das estratégias com base no contexto da empresa selecionada.   |
| <b>Application of industrial <i>Big Data</i> for smart manufacturing in product service system based on system engineering using fuzzy DEMATEL (Zhang et. al., 2021)</b> | Foi apresentado uma arquitetura como referência para implementação de <i>Big Data</i> Industrial. Como exemplo, ele foi implementado em um caso na área de engenharia naval  |

Fonte: Dados de pesquisa (2022)

Kava *et. al.* (2020) criaram um guia no qual é possível identificar a melhor estratégia e ferramentas a serem utilizados na implementação de *Big Data* a fim de prover tecnologias da Indústria 4.0 em uma empresa, de acordo com suas características e objetivos. Tran e Huh (2022) construíram um modelo capaz de prever comportamentos de consumidores baseado em um histórico de compras e perfil de clientes analisados por meio de *Big Data*. Dinnar *et. al.* (2021) utilizam modelos de inteligência artificial por meio de *Big Data* capazes de aprimorar a capacidade de negociação de vendedores humanos por meio de treinamentos e simulações.

#### 4.2.2 Ferramentas e tecnologias

Ao analisar o conteúdo dos materiais selecionados, foi desenvolvido o Gráfico 11, comparando o número de vezes em que as ferramentas e tecnologias relatadas foram exploradas nos artigos e sua respectiva porcentagem comparado com a total de ferramentas utilizadas nas pesquisas:



Fonte: Dados de pesquisa (2022)

Pode ser observado que IA, IoT e Realidade Virtual são as principais ferramentas presentes nos estudos, em que ao menos uma dessas tecnologias foi abordada em 90% dos artigos.

IA, presente em 33% dos artigos do portfólio, engloba conceitos como AM, aprendizagem profunda, linguagem natural de processamento e identificação de imagens. Segundo Kusunthum *et. al.* (2022), AM é o principal tópico dentro do estudo de análise de *Big Data*.

IoT foi a segunda tecnologia com maior presença nos artigos (28%), englobando, principalmente, a utilização de sensores e os conceitos de IIoT. Para Bilotta *et. al.* (2021), IoT refere-se a um conjunto de tecnologias (como sensores distribuídos no ambiente e conectados entre eles) permitindo que qualquer tipo de dispositivo ou sistema seja conectado à *Internet* com a finalidade de coletar, monitorar, controlar e transferir informações.

#### 4.2.3 Explicação sobre *Big Data*

Durante a produção do conteúdo dos artigos, os autores contribuíram de alguma forma com explicações que auxiliam na compreensão sobre o que é o *Big Data*, suas características, impactos gerados pelo seu uso, barreiras entre outras informações que auxiliam em um melhor entendimento sobre o tema. O Quadro 10 traz um resumo destes conceitos apresentados nos artigos.

**Quadro 14 - Explicação sobre *Big Data***

| <b>Artigo</b>   | <b>Autor</b>   | <b>Explicação sobre <i>Big Data</i></b>   |
|---|--|---|
| <b>State-of-the-art survey on digital twin implementations</b>                                | Liu, Y.K., Ong, S.K., Nee, A.Y.C.                    | O aumento da quantidade de dados fez com que surgisse a necessidade de o <i>Big Data</i> ser armazenado em tipos de servidores diferentes dos que eram tradicionalmente utilizados.   |
| <b>How Do Industry 4.0 Technologies Boost Collaborations in Buyer-Supplier Relationships?</b> | Patrucco, A., Moretto, A., Trabucchi, D., Golini, R. | A análise do <i>Big Data</i> e a computação em nuvem podem apoiar decisões habilitadas por máquinas com mínima ou nenhuma intervenção humana, melhorando assim o tempo e a profundidade dessas decisões, especialmente no caso de processos de decisão compartilhados |
| <b>Government Construction Project Budget Prediction Using AM</b>                             | Leungbootnak, N., Aksorn, P., Chaitongrat, T.        | Muitos estudos foram realizados sobre como usar <i>Big Data</i> e AM, porém seu uso depende da disponibilidade de dados e dos objetivos organizacionais   |

| Artigo  | Autor  | Explicação sobre <i>Big Data</i>  |
|---|--|---|
| <b>Designing early warning systems for detecting systemic risk: A case study and discussion</b>                                   | Wever, M., Shah, M., O'Leary, N.   | <i>Big Data</i> tem se tornado uma ferramenta importante na detecção preditiva de erros em sistemas   |
| <b>Building a model to exploit association rules and analyze purchasing behavior based on rough set theory</b>                    | Tran, D.T., Huh, J.-H.   | Existem muitos estudos sobre a aplicação de <i>Big Data</i> na previsão do comportamento de consumo do cliente. Como o comportamento do consumidor é variável ao longo tempo, a análise e coleta de dados de compras pode ajudar os vendedores a tomar melhores decisões de marketing e vendas. |
| <b>The Role of AI, AM, and <i>Big Data</i> in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities</b> | Rathore, M.M., Shah, S.A., Shukla, D., Bentafat, E., Bakiras, S.                 | A interação de <i>Big Data</i> com realidade aumentada resulta em monitoramentos de processos com maior confiabilidade, previsões mais precisas, diagnóstico de falhas na montagem e melhoria de processos de fabricação.   |
| <b>The current sustainability scenario of Industry 4.0 enabling technologies in Indian manufacturing industries</b>               | Pasi, B.N., Mahajan, S.K., Rane, S.B.  | <i>Big Data</i> oferece aos fabricantes uma enorme oportunidade de prever, inovar e implementar soluções. O desafio atual é integrar os dados de várias fontes e extrair insights valiosos  |
| <b>Technology-driven 5G enabled e-healthcare system during COVID-19 pandemic</b>  | Alshammari, N., Sarker, M.N.I., Kamruzzaman, M.M., Raihan, M.L., AlQahtani, S.A. | Tecnologia da informação pode dar suporte a sistemas de informação por meio da IoT, <i>Big Data</i> , Inteligência Artificial e aplicativos de computação em nuvem para garantir melhores atendimentos na área da saúde   |
| <b>Performance prediction of trace metals and cod in wastewater treatment using artificial neural network</b>                     | Matheri, A.N., Ntuli, F., Ngila, J.C., Seodigeng, T., Zvinowanda, C.             | Sistemas de gerenciamento de <i>Big Data</i> exigem abordagens multidisciplinares adequadas de diferentes engenharias   |
| <b>Operational time-series data modeling via LSTM network integrating principal component analysis based on human experience</b>  | Yang, K., Liu, Y.-L., Yao, Y.-N., Fan, S.-D., Mosleh, A.                         | Os sistemas de monitoramento modernos coletam dados de vários pontos de medição e de longo prazo que tornam a previsão da integridade de certos equipamentos um problema de <i>Big Data</i>   |
| <b>Mining open government data for business intelligence using data visualization: A two-industry case study</b>                  | Gottfried, A., Hartmann, C., Yates, D.   | <i>Big Data</i> (ou seja, volume, variedade e velocidade de dados não estruturados e indefinidos) impactou ainda mais a inteligência de negócios ao incluir análises visuais rápidas e ciência de dados   |
| <b>Industry 4.0 technologies in tourism education: Nurturing students to think with technology</b>                                | Bilotta, E., Bertacchini, F., Gabriele, L., Pantano, P.S., Romita, T.            | A definição comum de <i>Big Data</i> é baseada em suas características essenciais, descritas pelos 3Vs: Volume, Variedade e Velocidade. A análise de <i>Big Data</i> permite compreender o comportamento dos consumidores, as suas escolhas e melhorar as ofertas de marketing.                 |

| <b>Artigo</b>  | <b>Autor</b>  | <b>Explicação sobre <i>Big Data</i></b>   |
|--|---|---|
| <b>Human health <i>Big Data</i> evaluation based on FPGA processor and <i>Big Data</i> decision algorithm</b>                                    | Sun, K., Yang, D.   | Em termos de tratamento médico, o <i>Big Data</i> inclui principalmente vários registros médicos, dados corporais do paciente e uma rede de informações   |
| <b>How and when do <i>Big Data</i> investments pay off? The role of marketing affordances and service innovation</b>                             | De Luca, L.M., Herhausen, D., Troilo, G., Rossi, A.                 | A análise de <i>Big Data</i> é frequentemente descrita em termos de sua abordagem e contribuição para a tomada de decisão (ou seja, descritiva, diagnóstica, preditiva e prescritiva e com base em seu domínio de uso específico (ou seja, texto, fala, <i>web</i> , rede ou análise móvel)   |
| <b>Exploring the Factors Influencing <i>Big Data</i> Technology Acceptance</b>   | Rahman, N., Daim, T., Basoglu, N.                                   | As tecnologias de <i>Big Data</i> são escaláveis em termos de armazenamento, processamento e construção de um modelo robusto de AM.   |
| <b>Exploring IT/IS proactive and knowledge transfer on enterprise digital business transformation (EDBT): a technology-knowledge perspective</b> | Li, J., Saide, S., Ismail, M.N., Indrajit, R.E.                     | O <i>Big Data</i> pode maximizar as tecnologias emergentes a fim de alcançar a transformação de negócios digitais corporativos.   |
| <b><i>Big Data</i> Impacting Dynamic Food Safety Risk Management in the Food Chain</b>   | Donaghy, J.A., Danyluk, M.D., Ross, T., Farber, J., for ICMSF       | A capacidade do <i>Big Data</i> tem implicações em recursos tecnológicos, estratégicos e econômicos de empresas.  |
| <b>Artificial Intelligence and Technology in Teaching Negotiation</b>  | Dinnar, S.M., Dede, C., Johnson, E., Straub, C., Korjus, K.         | Inteligência artificial, AM, computação afetiva e técnicas de <i>Big Data</i> estão melhorando as maneiras pelas quais os humanos negociam e aprendem a negociar.   |
| <b>A <i>Big Data</i>-driven framework for sustainable and smart additive manufacturing</b>   | Majeed, A., Zhang, Y., Ren, S., Waqar, S., Yin, E.                  | O <i>Big Data</i> é classificado de acordo com diferentes estágios do ciclo de vida, que são beginning of life (BOL), middle of life (MOL) e end of life (EOL)  |
| <b>A <i>Big Data</i>-centric architecture metamodel for Industry 4.0</b>   | López Martínez, P., Dintén, R., Drake, J.M., Zorrilla, M.           | <i>Big Data</i> e computação em nuvem são atualmente considerados como os principais facilitadores que podem levar à construção do futuro ecossistema industrial, interligando os mundos cibernético e físico.  |
| <b>Towards data-driven approaches in manufacturing: an architecture to collect sequences of operations</b>                                       | Farooqui, A., Bengtsson, K., Falkman, P., Fabian, M.                | Para entender e gerenciar a complexidade da mudança da Indústria 4.0 com relação a tecnologias ultrapassadas, técnicas existentes, como AM, análise de <i>Big Data</i> e mineração de dados, podem ser empregadas.  |
| <b>The state of the art and taxonomy of <i>Big Data</i> analytics: view from new <i>Big Data</i> framework</b>                                   | Mohamed, A., Najafabadi, M.K., Wah, Y.B., Zaman, E.A.K., Maskat, R. | <i>Big Data</i> tornou-se uma área de pesquisa significativa devido ao nascimento de enormes dados gerados a partir de várias fontes, como mídias sociais, IoT e aplicativos multimídia. <i>Big Data</i> tem desempenhado um papel crítico em muitas tomadas de decisão e domínios de previsão, como sistemas de recomendação, análise de negócios, saúde, publicidade de <i>web</i> , transporte, detecção de fraudes e marketing no geral |

| Artigo   | Autor   | Explicação sobre <i>Big Data</i>   |
|--|---|--|
| <b>The architecture development of Industry 4.0 compliant smart machine tool system (SMTS)</b>                                     | Jeon, B; Yoon, J.-S.; Um, J; Suh, S.-H.;                          | O uso de <i>Big Data</i> , inteligência artificial e <i>Digital Twins</i> pode maximizar os indicadores dos sistemas de manufatura   |
| <b>REDTag: a Predictive Maintenance Framework for Parcel Delivery Services</b>   | Proto, S., Di Corso, E., apiletti, D., Malnati, G., Mazzucchi, D. | A forte integração das tecnologias da IoT e a solução de análise de <i>Big Data</i> tornou-se necessária para gerenciar efetivamente os processos industriais e prever antecipadamente falhas de produtos ou interrupções de serviços  |
| <b>Lean 4.0: A new holistic approach for the integration of lean manufacturing tools and digital technologies</b>                  | Valamede, L.S., Akkari, A.C.S.                                    | Análise de <i>Big Data</i> refere-se a uma grande quantidade de dados variados, que são processados em insights.   |
| <b>Integration of Lean practices and Industry 4.0 technologies: smart manufacturing for next-generation enterprises</b>            | Shahin, M., Chen, F.F., Bouzary, H., Krishnaiyer, K.              | A transformação do <i>Big Data</i> e Indústria 4.0 foi acelerada recentemente graças às soluções de <i>hardware</i> e <i>software</i> cada vez mais acessíveis, sensores e equipamentos mais baratos e eficazes, equipamentos de rede e plataformas mais poderosos, e outras tecnologias como redes sem fio e computação em nuvem. |
| <b>Identification and classification of materials using machine vision and AM in the context of industry 4.0</b>                   | Penumuru, D.P., Muthuswamy, S., Karumbu, P.                       | A indústria da manufatura passou por enormes mudanças da indústria 1.0 para a indústria 4.0 com o avanço de tecnologias como o <i>Big Data</i> .   |
| <b>Creating a road map for industry 4.0 by using an integrated fuzzy multicriteria decision-making methodology</b>                 | Kaya, İ., Erdoğan, M., Karaşan, A., Özkan, B.                     | A Indústria 4.0 é um novo conceito em sistemas de produção que se revela a partir da integração de tecnologias atuais, como análise de <i>Big Data</i> , IoT, simulação, segurança cibernética, computação em nuvem, realidade aumentada e troca de dados entre estes agentes  |
| <b>BlockIoTelligence: A Blockchain-enabled Intelligent IoT Architecture with Artificial Intelligence</b>                           | Singh, S.K., Rathore, S., Park, J.H.                              | Para análises de <i>Big Data</i> , a Inteligência Artificial desempenha um papel significativo como uma forte ferramenta analítica e fornece uma análise escalável e precisa de dados em tempo real.   |
| <b><i>Big Data</i> and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case</b> | Sahal, R., Breslin, J.G., Ali, M.I.                               | As tecnologias de <i>Big Data</i> e <i>Internet</i> da Coisas estão desempenhando um papel fundamental na criação de soluções orientadas a dados, como manutenção preditiva.   |
| <b><i>Big Data</i> analytics enhanced healthcare systems: a review</b>   | Shafqat, S., Kishwer, S., Rasool, R.U., Amjad, T., Ahmad, H.F.    | Há um interesse crescente em implementar tecnologias relacionadas à <i>Big Data</i> no setor de saúde para gerenciar coleções massivas de conjuntos heterogêneos de dados de saúde, como registros eletrônicos e dados de sensores, que estão aumentando em volume e variedade   |

| Artigo   | Autor   | Explicação sobre <i>Big Data</i>  |
|--|---|---|
| <b>Application of industrial <i>Big Data</i> for smart manufacturing in product service system based on system engineering using fuzzy DEMATEL</b> | Zhang, X., Ming, X., Yin, D.                                      | <i>Big Data</i> aparece em diferentes campos de aplicação, como <i>Big Data</i> financeiro, <i>Big Data</i> de telecomunicações, <i>Big Data</i> de mídia, <i>Big Data</i> médico, <i>Big Data</i> de turismo, <i>Big Data</i> de mercado, <i>Big Data</i> de governo, <i>Big Data</i> de agricultura, entre outros |
| <b>A method of NC machine tools intelligent monitoring system in smart factories</b>   | Liu, W., Kong, C., Niu, Q., Jiang, J., Zhou, X.                   | Nos últimos anos, o AM e a análise de <i>Big Data</i> alcançaram grandes avanços e se tornaram uma excelente ferramenta para extração de informações, estatísticas e descoberta de conhecimento   |
| <b>A Design-to-Device Pipeline for Data-Driven Materials Discovery</b>   | Cole, J.M.  | Os avanços recentes em <i>Big Data</i> também afetaram a descoberta de novos materiais, o que é chamado de “ <i>data-driven materials discovery</i> ”   |
| <b>The role of <i>Big Data</i> analytics in industrial IoT</b>   | Rehman, M.H., Yaqoob, I., Salah, K., Jayaraman, P.P., Perera, C.  | A produção de <i>Big Data</i> com IoT Industrial é evidente devido à implantação massiva de sensores e dispositivos da IoT. No entanto, o processamento de <i>Big Data</i> é um desafio devido aos recursos computacionais, de rede e de armazenamento limitados  |
| <b>The impact of <i>Big Data</i> analytics on company performance in supply chain management</b>   | Oncioiu, I., Bunget, O.C., Türkes, M.C., Rakos, I.-S., Hint, M.S. | A análise de <i>Big Data</i> pode agregar valor e fornecer uma nova perspectiva, melhorando a análise preditiva e as práticas de modelagem  |
| <b>Barriers to <i>Big Data</i> analytics in manufacturing supply chains: A case study from Bangladesh</b>  | Moktadir, M.A., Ali, S.M., Paul, S.K., Shukla, N.                 | A análise de <i>Big Data</i> está se tornando cada vez mais popular entre as empresas de manufatura. No entanto, existem muitas barreiras para a adoção de análise de <i>Big Data</i> nas cadeias de suprimentos de fabricação  |

Fonte: Dados de pesquisa (2022)

Muitas contribuições presentes no Quadro 14 estão diretamente conectados com conceitos apresentados na seção de Referencial Teórico. Gottfried, Hartman e Yates (2021) definem o *Big Data* de acordo com as características de volume, velocidade e variedade, associando este conceito com ferramentas de ciência de dados e *Business Intelligence*. Rehman *et. al.* (2019) apontam a importância dos sensores para a IoT e lista barreiras de processamentos de recursos computacionais de processamento de dados. Penumuru *et. al.* (2021) exploram a evolução da Indústria 4.0 com o uso de tecnologias e ferramentas do *Big Data*.

De forma geral, os autores apresentam o *Big Data* como uma tecnologia que afeta positivamente as organizações e o caracterizam como um agente importante na produção de *insights*. De Lucas *et. al.* (2021), Shahin *et. al.* (2021), Sahal *et. al.* (2021) e López *et. al.* (2021) mencionam como o fluxo e volume de produção de dados são

importantes na digitalização de empresas, principalmente, quando se trata da transformação da Indústria 4.0. Dentre os desafios, Rahman *et. al.* (2021) menciona barreiras como a dificuldade em armazenar um grande volume de dados e também o acesso a tecnologias de ponta requeridas para extrair o máximo potencial do *Big Data*

#### 4.2.4 Temática

A Figura 15 mostra uma nuvem de palavras-chave, com as principais temáticas do portfólio de artigos estudado. Quanto maior a palavra, maior a frequência da presença nos artigos:

Figura 15 - Nuvem de Palavras



Fonte: Dados de pesquisa (2022)

Os temas mais comumente tratados são “*Big Data*”, “*decision making*” (tomada de decisão), *data analytics* (análise de dados), *Internet of Things* (IoT), *Industry 4.0*, *artificial intelligence* (inteligência artificial) e *machine learning* (AM). A importância da inteligência artificial na análise de *Big Data*, como mencionado na secção 4.2.2, evidência a importância deste tema neste tipo de pesquisa. “*decision*

*making*” também é um tema significativo, já que uma grande parte da coleta e análise de *Big Data* é realizada para embasar a tomada de decisão.

#### 4.3 FERRAMENTAS, TECNOLOGIAS E IMPACTOS DO BIG DATA

Tomando como base todo o conteúdo elaborado após a realização da Análise Bibliométrica e Análise de Conteúdo, é possível evidenciar algumas das ferramentas e tecnologias do *Big Data* com maior significância nos artigos, e como podem gerar valor para organizações. O Quadro 15 apresenta um resumo das tecnologias mais frequentes.

**Quadro 15 - Ferramentas, tecnologias e impactos do Big Data**

| Ferramentas/Tecnologias | Definição  | Impactos  |
|-------------------------|--|---|
| IoT                     | Conexão entres objetos ou seres vivos que permite transporte de dados em tempo real (RIBEIRO, 2019). | Redução de desperdício de recursos e custos (Rehman <i>et. al.</i> , 2019)  |
|                         |  | Automação de processos (Singh <i>et. al.</i> , 2021)  |
|                         |  | Maior velocidade de transporte de informações (Proto <i>et. al.</i> , 2021)   |
| IA                      | Simulação do raciocínio humano por meio de máquinas (DOGAN; BIRANT, 2020).                           | Automação de processos e redução de custos (Rathore <i>et. al.</i> , 2021)  |
| Realidade Virtual       | Ambiente virtual onde a realidade física é simulada para um usuário (Shahin <i>et. al.</i> , 2021)   | Melhor percepção de um ambiente ou situação (Rathore <i>et. al.</i> , 2021)   |
|                         |  | Melhor experiência de usuário (Shahin <i>et. al.</i> , 2021)  |
| DT                      | Réplica virtual de um objeto físico (Rathore <i>et. al.</i> , 2021)                                  | Permite identificar falhas antes de um produto ser lançado para produção ou fazer testes sem ter perdas consideráveis de recursos (Liu <i>et. al.</i> , 2022) |
| Mineração de Dados      | Se trata do processo de exploração e coleta de dados (DE CASTRO; FERRARI, 2016).                     | Decisões organizacionais assertivas baseadas em dados (Gottfried, 2021)   |
|                         |  | Maior volume, velocidade e variedade de dados Rahman <i>et. al.</i> (2020)  |

Fonte: Dados de pesquisa (2022)

É importante ressaltar que os resultados mencionados neste trabalho não ocorrerão em todas as instancias onde essas tecnologias forem implementadas, porém, de forma geral, os materiais estudados ressaltam como o seu uso pode agregar valor e o uso do Big Data possui o potencial de ser um diferencial competitivo

de empresas dos mais diversos setores. O objetivo de identificar estas tecnologias é prover uma visão sobre o que esperar em uma organização voltada para dados (mais especificamente *Big Data*) e ajudar no planejamento e no conhecimento a respeito da temática.

#### 4.4 LACUNAS E SUGESTÕES PARA TRABALHOS FUTUROS

A principal lacuna identificada no material pesquisado e sugestão para trabalhos futuros envolve a constatação de que existem poucos trabalhos produzidos que investigam meios que possibilitem a mitigação das barreiras de uso do *Big Data*.

Os maiores desafios tecnológicos são os de armazenamento de dados e o alto custo de tecnologias que habilitam o uso de *Big Data* para os mais diversos fins. Outro trabalho futuro sugerido é um modelo capaz de mensurar os benefícios resultantes do uso cada uma das ferramentas em organizações de diferentes setores.

Um ponto de interesse para trabalhos futuros é o estudo de materiais publicados com o tema privacidade e o direito à informação. Privacidade de dados está se tornando um assunto interessante, já que um dos desafios apresentados nos artigos estudados é conseguir separar um dado privado do público.

No presente trabalho, apenas a base de dados *Scopus* foi utilizada como fonte de artigos. Para futuros trabalhos, pode ser utilizada uma base de dados diferente, ou até mesmo realizar um estudo contendo artigos de diversas bases.

## 5 CONCLUSÃO

O presente trabalho de revisão sistemática da literatura teve como objetivo geral identificar ferramentas, tecnologias e impactos do *Big Data* em organizações orientadas por dados, buscando gerar contribuições nos aspectos acadêmicos, sociais e econômicos.

Através da conclusão do primeiro objetivo específico, a Análise Bibliométrica pôde-se evidenciar o aumento considerável do número de materiais de pesquisa com o tema *Big Data* que foi produzido nos últimos anos. Foi identificado que Estados Unidos e China são países que se destacam em pesquisas com esta temática, não só em quantidade de publicações, mas também em termos de impacto gerado por estas produções científicas. Como autores relevantes, foram identificados JianJian Wang, Young Wang e Jiantong Zhang. Entre as instituições de destaque, foram apontadas *Old Dominion University*, *Politenico di Torino*, *Universiti Tun Hussein Onn Malaysia* e *University of Deusto*.

Sobre o segundo objetivo específico deste trabalho, por meio da Análise de Conteúdo foram apresentados os objetivos e resultados dos artigos selecionados, além da metodologia utilizada pelos autores para desenvolver a pesquisa. Este conjunto de material mostrou os principais conceitos, ferramentas, desafios e impactos positivos que a utilização de *Big Data* consegue produzir em organizações que utilizam este tipo de tecnologia. Também foram expostas ideias, conceitos e explicações sobre *Big Data* e ferramentas que os autores levantaram em seus trabalhos. Cumprindo o terceiro objetivo específico deste documento, foi observado que *IoT*, IA, realidade virtual, *digital twins* e mineração de dados estavam entre as tecnologias mais mencionados nos materiais pesquisados. O uso de *Big Data* aliado a estas ferramentas gera redução de desperdícios, diminuição de custos, automação de processos, melhor experiência de usuários, melhora nos 3Vs, entre outros impactos positivos.

De forma geral, este trabalho obteve êxito em expor conceitos de *Big Data* e mostrar o potencial de como organizações podem utilizar suas tecnologias a fim de obter os impactos positivos mostrados pelos trabalhos científicos explorados.

## REFERÊNCIAS

ALSHAMMARI, N. et al. **Technology-driven 5G enabled e-healthcare system during COVID-19 pandemic**. IET Communications, 5 jun. 2021.

AMARAL, F. **Big Data**: uma visão gerencial para executivos, consultores e gerentes de projetos. [S.l.]: Fernando Amaral, 2016. ASIN: B01MDSEP3V.

AL-BARHAMTOSHY, H. A Data Analytic Framework for Unstructured Text. **Life Science Journal**, v. 48, p. 339–350, 10 nov. 2014.

ASSANDRE, J. A.; MARTINS, R. A. **BIG DATA E MEDIÇÃO DE DESEMPENHO: UMA ANÁLISE BIBLIOMÉTRICA DESCRITIVA**. Anais do Encontro Nacional de Engenharia de Produção, 6 nov. 2019.

Banco Nacional de Desenvolvimento Econômico e Social – BNDES. (2018). **Internet das coisas**: um plano de ação para o Brasil (Cartilha de Cidades). Rio de Janeiro. Disponível em: <<https://www.bndes.gov.br/wps/wcm/connect/site/db27849e-dd37-4fbd-9046-6fda14b53ad0/produto-13-cartilha-das-cidades-publicada.pdf?MOD=AJPERES&CVID=m7tz8bf>>. Acesso em: 22 jan. 2022.

Bardin, L. **Análise de conteúdo**. Edição revista e ampliada. São Paulo: Edições 70 Brasil, [1977] 2016.

BATISTA, F. et al. **MongoDB**: Banco de dados orientado a documento. 2012. Disponível em: Acesso em: 5 abr. 2016.

BEAN, R. **How Companies Say They're Using Big Data** – Harvard Business Review, 2017. Disponível em: <<https://hbr.org/2017/04/how-companies-say-theyre-using-big-data>>. Acesso em: 04 nov. 2021.

BELCASTRO, Loris *et al.* Programming *Big Data* analysis: principles and solutions. **Survey Paper**, Rende, ano 2022, v. 9, n. 1, 6 jan. 2022. 4, p. 1-50.

BELCASTRO, Loris *et al.* *Big Data* Analysis on Clouds. **Handbook of Big Data Technologies**, Calabria, ano 2017, p. 101-142, 25 fev. 2017.

BERRY, M. **Major Computing Vendors**. IT Manager Daily. Disponível em <<http://www.itmanagerdaily.com/cloud-computing-vendors/>>. Acesso em: 24 de nov. 2021.

BILOTTA, E. et al. Industry 4.0 technologies in tourism education: Nurturing students to think with technology. *Journal of Hospitality, Leisure, Sport & Tourism Education*, p. 100275, out. 2020.

BLACKMAN, L. C. F.; DEWAR, M. J. S. 30. **Promoters for the dropwise condensation of steam**. Part IV. Discussion of dropwise condensation and testing of compounds. *Journal of the Chemical Society (Resumed)*, p. 171, 1957.

BOYES, Hugh; HALLAQ, Bil; CUNNINGHAM, Joe; WATSON, Tim. "The industrial IoT (IIoT): An analysis framework". **Computers in Industry**, v. 101, 2018, p. 1–12. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0166361517307285>>. Acesso em: 22 jan. 2022.

CHAKRAVORTI, Bhaskar *et al.* **Which Countries Are Leading the Data Economy?**. [S. l.], 24 jan. 2019. Disponível em: <https://hbr.org/2019/01/which-countries-are-leading-the-data-economy>. Acesso em: 28 fev. 2022.

CHANG, K.-H. et al. **Big data analytics energy-saving strategies for air compressors in the semiconductor industry** – an empirical study. *International Journal of Production Research*, v. 60, n. 6, p. 1782–1794, 12 jan. 2021.

CICHY, Radoslaw; KAISER, Daniel. **Deep Neural Networks as Scientific Models**. *Trends in Cognitive Sciences*, [S. l.], v. 23, n. 4, p. 305-317, 22 nov. 2019. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1364661319300348>. Acesso em: 22 jan. 2022.

COLE, J. M. A **Design-to-Device Pipeline for Data-Driven Materials Discovery**. *Accounts of Chemical Research*, v. 53, n. 3, p. 599–610, 25 fev. 2020.

CORALLO, A. et al. **Cybersecurity Challenges for Manufacturing Systems 4.0: Assessment of the Business Impact Level**. *IEEE Transactions on Engineering Management*, p. 1–21, 2021.

COSER, T. Contabilidade de gestão em sintonia com o Business Intelligence (BI): estudo de caso. **Brazilian Journals of Business**, Curitiba, v. 2, n. 3, p. 3093-3112.

COSTA, J.; VANZ, S.. **Indicadores da produção científica e co-autoria: análise do departamento de ciências da informação da UFRGS**. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, Florianópolis, v. 17, n. 33, p. 97-115, jan/abr. 2012. Disponível em: <<https://brapci.inf.br/index.php/res/v/271>> Acesso em: 08 fev. 2022.

CUNHA, T. M. de A. **Escalabilidade de Sistemas com Banco de Dados NoSQL: um Estudo de Caso Comparativo com MongoDB e MySQL**. 2011. 85 f. Centro Universitário da Bahia – Estácio, Salvador.

DE CASTRO, Leandro Nunes; FERRARI, Daniel Gomes. **Introdução à mineração de dados: Conceitos básicos, algoritmos e aplicações**. 1. ed. [S. l.]: Saraiva, 2016. 376 p. v. 1.

DE LUCA, L. M. et al. **How and when do big data investments pay off?** The role of marketing affordances and service innovation. *Journal of the Academy of Marketing Science*, 5 ago. 2020.

DINNAR, S. “MOOLY” et al. **Artificial Intelligence and Technology in Teaching Negotiation**. *Negotiation Journal*, v. 37, n. 1, p. 65–82, jan. 2021.

DOGAN, Alican; BIRANT, Derya. **AM and Data Mining in Manufacturing**. *Expert Systems with Applications*, [S. l.], p. 1-45, 24 set. 2020.

DONAGHY, J. A. et al. Big Data Impacting Dynamic Food Safety Risk Management in the Food Chain. **Frontiers in Microbiology**, v. 12, 21 maio 2021.

DUANA, Yanqing; EDWARDSB, John; DWIVEDIC, Yogesh. **Artificial intelligence for decision making in the era of Big Data** – evolution, challenges and research agenda. *International Journal of Information Management*, Bedfordshire, v. 48, p. 63-71, 1 out. 2019. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0268401219300581>. Acesso em: 22 jan. 2022.

FAROOQUI, A. et al. **Towards data-driven approaches in manufacturing: an architecture to collect sequences of operations**. *International Journal of Production Research*, v. 58, n. 16, p. 4947–4963, 9 mar. 2020.

GARTNER GROUP. **Key Issues for Analytics, Business Intelligence and Performance Management**, 2019. Disponível em: <https://www.gartner.com/en/research/methodologies/magic-quadrants-research>. Acesso em: 27 nov. 2021.

GENG, H. **The IoT and Data Analytics Handbook**. 1. Ed. Nova Jersey: John Wiley & Sons Inc, 2017. 800 p.

GÉRON, Aurélien. **Mãos à Obra: AM com Scikit-Learn & TensorFlow**. [S. l.]: Alta Books, 2019. 576 p.

GEYSER, W. **100+ Social Media Statistics for 2022 [+Internet in Real Time Live Infographic]**. Disponível em: <https://influencermarketinghub.com/social-media-statistics/#toc-1>. Acesso em: 19 fev. 2022.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. Editora Atlas SA, 2008.

Google Trends. Disponível em: <https://trends.google.com.br/trends/explore?date=2006-01-01%202022-02-19&q=Big%20data>. Acesso em: 19 fev. 2022.

GOTTFRIED, A.; HARTMANN, C.; YATES, D. **Mining Open Government Data for Business Intelligence Using Data Visualization: A Two-Industry Case**

Study. Journal of Theoretical and Applied Electronic Commerce Research, v. 16, n. 4, p. 1042–1065, 18 mar. 2021.

GUNDERSEN, Erik *et al.* **Do AM Platforms Provide Out-of-the-box Reproducibility?**. Future Generation Computer Systems, Trondheim, ano 2022, v. 126, p. 34-47, 8 jun. 2021.

GÜNTHER, Hartmut. **Pesquisa qualitativa versus pesquisa quantitativa: esta é a questão.** Psicologia: teoria e pesquisa, v. 22, n. 2, p. 201-210, 2006.

JEON, B. *et al.* **The architecture development of Industry 4.0 compliant smart machine tool system (SMTS).** Journal of Intelligent Manufacturing, v. 31, n. 8, p. 1837–1859, 29 jan. 2020.

KAYA, İ. *et al.* **Creating a road map for industry 4.0 by using an integrated fuzzy multicriteria decision-making methodology.** Soft Computing, 20 maio 2020.

KHANRA, S.; DHIR, A.; MÄNTYMÄKI, M. **Big data analytics and enterprises: a bibliometric synthesis of the literature.** Enterprise Information Systems, p. 1–32, 17 mar. 2020.

KIPPER, L. M. *et al.* **Scopus scientific mapping production in industry 4.0 (2011–2018): a bibliometric analysis.** International Journal of Production Research, v. 58, n. 6, p. 1605–1627, 9 out. 2019.

Como funciona o Business Intelligence? - **Know Solutions.** Disponível em: <<https://www.knowsolution.com.br/como-funciona-o-business-intelligence/>>. Acesso em: 24 jun. 2022.

KURU, K. **Management of geo-distributed intelligence: Deep Insight as a Service (DINSaaS) on Forged Cloud Platforms (FCP).** Journal of Parallel and Distributed Computing, v. 149, p. 103–118, mar. 2021.

KUSONKHUM, W. *et al.* **Government Construction Project Budget Prediction Using Machine Learning.** Journal of Advances in Information Technology, v. 13, n. 1, 2022.

LAVAGNOLI, Silvia. **AM ou Aprendizagem profunda?**. [S. l.], 12 set. 2019. Disponível em: <https://opencadd.com.br/machine-learning-ou-deep-learning/>. Acesso em: 22 jan. 2022.

LEDRO, C.; NOSELLA, A.; VINELLI, A. **Artificial intelligence in customer relationship management: literature review and future research directions.** Journal of Business & Industrial Marketing, v. 37, n. 13, p. 48–63, 30 mar. 2022.

LESK, M. **Big Data, Big Brother, Big Money**. IEEE Security & Privacy, v. 11, n. 4, p. 85–89, jul. 2013.

LI, J. et al. **Exploring IT/IS proactive and knowledge transfer on enterprise digital business transformation (EDBT): a technology-knowledge perspective**. Journal of Enterprise Information Management, v. ahead-of-print, n. ahead-of-print, 15 jun. 2021.

LIU, Y. K.; ONG, S. K.; NEE, A. Y. C. State-of-the-art survey on digital twin implementations. Advances in Manufacturing, 4 jan. 2022.

LIU, W. et al. **A method of NC machine tools intelligent monitoring system in smart factories**. Robotics and Computer-Integrated Manufacturing, v. 61, p. 101842, fev. 2020.

LOH, Stanley. **Volume, Velocidade, Variedade, Veracidade e Valor: Como os 5 Vs do Big Data estão impactando as Organizações e a Sociedade**. Porto Alegre, 2019.

LÓPEZ MARTÍNEZ, P. et al. **A big data-centric architecture metamodel for Industry 4.0**. Future Generation Computer Systems, v. 125, p. 263–284, dez. 2021.

LÓPEZ-LÓPEZ, E.; NAVEJA, J. J.; MEDINA-FRANCO, J. L. **DataWarrior: an evaluation of the open-source drug discovery tool**. Expert Opinion on Drug Discovery, v. 14, n. 4, p. 335–341, 26 fev. 2019.

LV, Zhihan *et al.* **Big Data Analytics for 6G-Enabled Massive IoT**. IEEE IOT JOURNAL, [S. l.], ano 2021, v. 8, n. 7, p. 5350-5359, 1 abr. 2021. Disponível em: <https://sci-hub.se/10.1109/jiot.2021.3056128>. Acesso em: 21 fev. 2022.

50 Sensor Applications for a Smarter World. **Libelium**, 2020. Disponível em: <https://www.libelium.com/libeliumworld/top-50-iot-sensor-applications-ranking/>. Acesso em: 22 jan. 2022.

MACHADO, F. N. R. **Big Data: o Futuro dos Dados e Aplicações**. 1. ed. São Paulo: Saraiva, 2018.

MARTÍNEZ-LÓPEZ, F. J. et al. **Fifty years of the European Journal of Marketing: a bibliometric analysis**. European Journal of Marketing, v. 52, n. 1/2, p. 439–468, 12 fev. 2018.

MAJEED, A. et al. **A big data-driven framework for sustainable and smart additive manufacturing**. Robotics and Computer-Integrated Manufacturing, v. 67, p. 102026, fev. 2021.

MARQUESONE, R. **Big Data: Técnicas e tecnologias para extração de valor dos dados**. [S.l.]: Casa do Código, 2016.

MATHERI, A. N. et al. **Performance prediction of trace metals and cod in wastewater treatment using artificial neural network**. Computers & Chemical Engineering, v. 149, p. 107308, jun. 2021.

MAZHAR RATHORE, M. et al. **The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities**. IEEE Access, p. 1–1, 2021.

MOHAMED, Azlinah *et al.* The state of the art and taxonomy of *Big Data* analytics: view from new *Big Data* framework. **Artificial Intelligence Review**, [S. l.], p. 989–1037, 1 fev. 2019. doi:10.1007/s10462-019-09685-9. Disponível em: <<https://link.springer.com/article/10.1007/s10462-019-09685-9>> Acesso em: 1 mar. 2022.

MOKTADIR, MD. A. et al. **Barriers to big data analytics in manufacturing supply chains: A case study from Bangladesh**. Computers & Industrial Engineering, v. 128, p. 1063–1075, fev. 2019.

NAÇÕES UNIDAS. **Roteiro para a Localização dos Objetivos de Desenvolvimento Sustentável: Implementação e Acompanhamento no nível subnacional**. 2016. Disponível em: <<http://www.br.undp.org/content/brazil/pt/home/library/ods/roteiro-para-a-localizacao-dos-objetivos-de-desenvolvimento-sust.html>>. Acesso em: 24 jun. 2022.

NETO, J. A. R. **Big Data para Executivos e Profissionais de Mercado**. Jose Antonio Ribeiro. 2 ed. 2021.

ONCIOIU et al. **The Impact of Big Data Analytics on Company Performance in Supply Chain Management**. Sustainability, v. 11, n. 18, p. 4864, 5 set. 2019.

PAGANI, Regina Negri; KOVALESKI, João Luiz; RESENDE, Luis Mauricio. **Methodi Ordinatio: a proposed methodology to select and rank relevant scientific papers encompassing the impact factor, number of citation, and year of publication**. Scientometrics, v. 105, n. 3, p. 2109-2135, 2015.

Pagani, R.N., Pedroso, B., Picinin, C.T., Bilynkiewicz, C.S., Kovaleski, J.L. **FIndex**. Sistema para Captura, Tratamento e Organização de Dados. UEPG. Registro de Programa de Computador Processo Nº: BR512021002567-9, INPI: 2021a.

Pagani, R.N., Pedroso, B., Picinin, C.T., Bilynkiewicz, C.S., Kovaleski, J.L. **RankIn**: Ferramenta para Operacionalização Quantitativa, Ranqueamento e Organização de Dados. UEPG. Registro de Programa de Computador Processo Nº: BR512021002568-7, INPI, 2021b.

**Gartner Top 10 Strategic Technology Trends for 2018**. Panneta, K. 2017. Disponível em <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2018/>. Acesso em: 22 jan. 2022

PASI, B. N.; MAHAJAN, S. K.; RANE, S. B. **The current sustainability scenario of Industry 4.0 enabling technologies in Indian manufacturing industries.** *International Journal of Productivity and Performance Management*, v. 70, n. 5, p. 1017–1048, 1 dez. 2020.

PATRUCCO, A. et al. **How Do Industry 4.0 Technologies Boost Collaborations in Buyer-Supplier Relationships?** *Research-Technology Management*, v. 65, n. 1, p. 48–58, 2 jan. 2022.

PENUMURU, D. P.; MUTHUSWAMY, S.; KARUMBU, P. **Identification and classification of materials using machine vision and machine learning in the context of industry 4.0.** *Journal of Intelligent Manufacturing*, v. 31, n. 5, p. 1229–1241, 16 nov. 2019.

PRANCKUTÉ, R. **Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World.** *Publications*, v. 9, n. 1, p. 12, 12 mar. 2021.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico.** *Novo Hamburgo: Feevale*, 2009.

PROTO, S. et al. **REDTag: A Predictive Maintenance Framework for Parcel Delivery Services.** *IEEE Access*, v. 8, p. 14953–14964, 2020.

RAHMAN, N.; DAIM, T.; BASOGLU, N. **Exploring the Factors Influencing Big Data Technology Acceptance.** *IEEE Transactions on Engineering Management*, p. 1–16, 2021.

RANJAN, J.; FOROPON, C. **Big Data Analytics in Building the Competitive Intelligence of Organizations.** v. 56, 1 fev. 2021.

RIBEIRO, C. J. S. **Big Data: Os Novos Desafios para o Profissional da Informação.** *Revista Informação & Tecnologia (ITEC)*, João Pessoa/Marília, v. 1, p. 96-105, jan./jun. 2014.

RIBEIRO, F. T. **IoT: da teoria à prática.** Monografia (Graduação em Engenharia de controle e automação) - Escola de Minas, UFOP. Ouro Preto, 2019.

Disponível em:

[https://monografias.ufop.br/bitstream/35400000/1794/1/MONOGRRAFIA\\_InternetCoisasTeoria.pdf](https://monografias.ufop.br/bitstream/35400000/1794/1/MONOGRRAFIA_InternetCoisasTeoria.pdf). Acesso em: 22 jan. 2022.

RIBEIRO, Marcello Batista; BRASILEIRO, Tania Suely Azevedo. **CONCEITOS E PRÁTICAS DE TI VERDE: UM BREVE ESTADO DA ARTE.** *EDUCAmazônia - Educação Sociedade e Meio Ambiente, Humaitá*, [S. l.], v. 15, n. 1, p. 31-35, 24 jan. 2022. Disponível em: <file:///C:/Users/Acer/Downloads/T2,+Vol+XV,+ano+15,+n%C3%BAmero+1,+jan-jun-2022,+p%C3%A1g+31-51.pdf>. Acesso em: 21 fev. 2022.

ROMERO, Carlos Andrés Tavera *et al.* Business Intelligence: Business Evolution after Industry 4.0. **Sustainability**, [S. I.], p. 1-12, 2 jul. 2021. DOI <https://doi.org/10.3390/su131810026>. Disponível em: <https://www.mdpi.com/2071-1050/13/18/10026>. Acesso em: 22 fev. 2022.

SAHAL, R.; BRESLIN, J. G.; ALI, M. I. **Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case**. Journal of Manufacturing Systems, v. 54, p. 138–151, jan. 2020.

SANTOS, Pedro Miguel Pereira. **IoT: O desafio da privacidade**. 2016. Tese de Doutorado. Instituto Politécnico de Setúbal. Escola Superior de Ciências Empresariais. Disponível em: <https://comum.rcaap.pt/bitstream/10400.26/17545/1/Disserta%C3%A7%C3%A3o%20Pedro%20Santos%20140313004%20MSIO.pdf>. Acesso em: 22 jan. 2022.

SHAFQAT, S. *et al.* **Big data analytics enhanced healthcare systems: a review**. The Journal of Supercomputing, v. 76, n. 3, p. 1754–1799, 3 fev. 2018.

SHAHIN, M. *et al.* **Integration of Lean practices and Industry 4.0 technologies: smart manufacturing for next-generation enterprises**. The International Journal of Advanced Manufacturing Technology, v. 107, n. 5-6, p. 2927–2936, mar. 2020.

SILVA, R. G. da. **Análise Comparativa entre os Modelos CNN e LSTM para Predição de Fluxo do Tráfego Urbano na Cidade de Recife**. 2020. 61 f. TCC (Graduação) - Curso de Ciência da Computação, Escola de Ciências Exatas e da Computação, Pontifícia Universidade Católica de Goiás, Goiânia, 2020.

SINGH, S. K.; RATHORE, S.; PARK, J. H. **BlockIoTelligence: A Blockchain-enabled Intelligent IoT Architecture with Artificial Intelligence**. Future Generation Computer Systems, set. 2019.

SODRÉ, L. **Big Data Estratégico: Um Framework para Gestão Sistêmica do Ecosistema Big Data**. 125 f. Dissertação (Mestrado) - Programa de Pós-graduação em Engenharia de Produção, COPPE, Universidade Federal do Rio de Janeiro, abr. 2016.

STATISTA. **Data Created Worldwide 2010-2025 | Statista**. Disponível em: <https://www.statista.com/statistics/871513/worldwide-data-created/>.

STEINGARTNER, W.; GALINEC, D.; KOZINA, A. **Threat Defense: Cyber Deception Approach and Education for Resilience in Hybrid Threats Model**. Symmetry, v. 13, n. 4, p. 597, 3 abr. 2021.

SULTAN, S. **Business Intelligence: Harness BI tools for the decision-making and business forecasting**. [S.I.]: Expert of Course, 2017. ASIN B076YS5P86.

SUN, K.; YANG, D. **Human health big data evaluation based on FPGA processor and big data decision algorithm**. *Microprocessors and Microsystems*, v. 81, p. 103793, mar. 2021.

SYED, Abdul Raheem; GILLELA, Kumar; VENUGOPAL, C. The Future Revolution on *Big Data*. *IJARCCCE*, 2013. Disponível em: . Acesso em: 30 de nov de 2021.

O QUE é a internet industrial das coisas (iot)? **Synnex Westcon**, 2020. Disponível em: <https://blogbrasil.westcon.com/o-que-a-internet-industrial-das-coisas-iiot>. Acesso em: 22 jan. 2022.

TAURION, C. *Big Data*. Rio de Janeiro: Brasport, 2013a. Não paginado.  
UNIVERSITY ALLIANCE. **What is Big Data?** Villanova, [201-]. Não paginado. Disponível em: <http://www.villanovau.com/resources/bi/what-is-big-data/#.VTmUaSFViko>>. Acesso em: 24 nov. 2021.

TRAN, D. T.; HUH, J.-H. **Building a model to exploit association rules and analyze purchasing behavior based on rough set theory**. *The Journal of Supercomputing*, v. 78, n. 8, p. 11051–11091, 7 fev. 2022.

UCKELMANN, D.; HARISSON, M.; MICHAHELLES, F. **Architecting the IoT**. 1. Ed. Nova York: Springer, 2011. 388p.

USMANI, R. S. A. et al. **Geographic Information System and Big Spatial Data: A Review and Challenges**. *International Journal of Enterprise Information Systems (IJEIS)*, v. 16, n. 4, p. 101–145, 1 out. 2020.

VALAMEDE, L. S.; AKKARI, A. C. S. **Lean 4.0: A New Holistic Approach for the Integration of Lean Manufacturing Tools and Digital Technologies**. *International Journal of Mathematical, Engineering and Management Sciences*, v. 5, n. 5, p. 851–868, 1 out. 2020.

VIEIRA, C. S.; MEIRELLES, F. S. Computação em Nuvem: Análise bibliométrica da produção científica sobre os fatores que influenciam as empresas no seu uso. **Revista Eletrônica Gestão e Serviços**, v. 6, n. 2, p. 1215-1230, jul./dez., 2015.

VISTRO, D. M. **IoT based Big Data Analytics for Cloud Storage Using Edge Computing**. *Journal of Advanced Research in Dynamical and Control Systems*, v. 12, n. SP7, p. 1594–1598, 25 jul. 2020.

TAURION, C. **O estágio atual do Big Data no Brasil**. *Revista Power channel*, São Paulo, n.20, p.5-7, jun. 2013b. Disponível em< <http://www.powerchannel.com.br/2013/06/03/cezar-taurion-o-estagio-atual-do-big-data-no-brasil/>>. Acesso em: 04 nov. 2021.

TRANFIELD, David; DENYER, David; SMART, Palminder. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. **British journal of management**, v. 14, n. 3, p. 207-222, 2003.

TREINTA, Fernanda Tavares et al. Metodologia de pesquisa bibliográfica com a utilização de método multicritério de apoio à decisão. **Production**, v. 24, n. 3, p. 508-520, 2014.

Empresas utilizam *Big Data* para obtenção de informações sobre negócios e ter impacto positivo no mercado. **Terra**, 2021. Disponível em: <<https://www.terra.com.br/noticias/empresas-utilizam-big-data-para-obtencao-de-informacoes-sobre-negocios-e-ter-impacto-positivo-no-mercado,d07f7431108cfdb4789f76fe4fb3bb82e21j1qy9.html>> .Acesso em: 10 jan. 2022.

VIANNA, W. B.; DUTRA, M. L. *Big Data* e gestão da informação: Modelagem do Contexto Decisional Apoiado pela Sistemografia. **Revista Informação e Informação**, Londrina, v. 21, n. 1, p. 185 – 212, jan./abr. 2016

VRANOPOULOS, Georgios; CLARKE, Nathan; ATKINSON , Shirley. Addressing *Big Data* variety using an automated approach for data characterization. **Journal of Big Data**, Plymouth, ano 2022, v. 9, n. 1, 10 jan. 2022. 8, p. 1-28.

WANG, C.; LAI, W. **A fuzzy model of wearable network real-time health monitoring system on pharmaceutical industry**. Personal and Ubiquitous Computing, 5 jul. 2019.

WELCH, T. F.; GEHRKE, S. R.; WIDITA, A. **Shared-use mobility competition: a trip-level analysis of taxi, bikeshare, and transit mode choice in Washington, DC**. Transportmetrica A: Transport Science, v. 16, n. 1, p. 43–55, 21 set. 2018.

WEVER, M.; SHAH, M.; O'LEARY, N. **Designing early warning systems for detecting systemic risk: A case study and discussion**. Futures, v. 136, p. 102882, fev. 2022.

XU, Z.; YU, D. **A Bibliometrics analysis on big data research (2009–2018)**. Journal of Data, Information and Management, v. 1, n. 1-2, p. 3–15, maio 2019.

YANG, H.; TATE, M. A Descriptive Literature Review and Classification of Cloud Computing Research. **Communications of the Association for Information Systems**, v. 31, a. 2, p. 35-60, jul., 2012

YANG, Minggao; SULAIMAN, Riza; YIN, Yahua; MALLAMACI, Valentina; ALRABIAH, Hussam. The effect of business intelligence, organizational learning and innovation on the financial performance of innovative companies located in Science Park. **Information Processing & Management**, [S. l.], v. 59, n. 2, p. 32-43, 1 mar. 2022. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S030645732100323X>. Acesso em: 21 fev. 2022.

YANG, C. et al. **Big Earth data analytics: a survey**. Big Earth Data, v. 3, n. 2, p. 83–107, 3 abr. 2019.

YANG, K. et al. **Operational time-series data modeling via LSTM network integrating principal component analysis based on human experience**. Journal of Manufacturing Systems, v. 61, p. 746–756, out. 2021.

ZHANG, X.; MING, X.; YIN, D. **Application of industrial big data for smart manufacturing in product service system based on system engineering using fuzzy DEMATEL**. Journal of Cleaner Production, v. 265, p. 121863, ago. 2020.

ZIJLSTRA, H; MCCULLOUGH, R>. **Scimago: Help**. [S. /], 6 jun. 2016. Disponível em: <https://www.scimagojr.com/help.php>. Acesso em: 8 fev. 2022.

**APÊNDICE A – PORTFÓLIO DE ARTIGOS SELECIONADOS**

| Artigo  | Ano  | InOrdinatio |
|---|------|-------------|
| State-of-the-art survey on digital twin implementations   | 2022 | 100,00      |
| How Do Industry 4.0 Technologies Boost Collaborations in Buyer-Supplier Relationships?  | 2022 | 100,00      |
| Government Construction Project Budget Prediction Using AM  | 2022 | 100,00      |
| Designing early warning systems for detecting systemic risk: A case study and discussion  | 2022 | 100,00      |
| Building a model to exploit association rules and analyze purchasing behavior based on rough set theory                                   | 2022 | 101,00      |
| Threat defense: Cyber deception approach and education for resilience in hybrid threats model   | 2021 | 96,00       |
| The Role of AI, AM, and <i>Big Data</i> in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities                | 2021 | 114,00      |
| The current sustainability scenario of Industry 4.0 enabling technologies in Indian manufacturing industries                              | 2021 | 98,00       |
| Technology-driven 5G enabled e-healthcare system during COVID-19 pandemic   | 2021 | 95,00       |
| Performance prediction of trace metals and cod in wastewater treatment using artificial neural network                                    | 2021 | 98,00       |
| Operational time-series data modeling via LSTM network integrating principal component analysis based on human experience                 | 2021 | 97,01       |
| Mining open government data for business intelligence using data visualization: A two-industry case study                                 | 2021 | 93,00       |
| Management of geo-distributed intelligence: Deep Insight as a Service (DINSaaS) on Forged Cloud Platforms (FCP)                           | 2021 | 94,00       |
| Industry 4.0 technologies in tourism education: Nurturing students to think with technology   | 2021 | 102,00      |
| Human health <i>Big Data</i> evaluation based on FPGA processor and <i>Big Data</i> decision algorithm                                    | 2021 | 93,00       |
| How and when do <i>Big Data</i> investments pay off? The role of marketing affordances and service innovation                             | 2021 | 120,01      |
| Exploring the Factors Influencing <i>Big Data</i> Technology Acceptance   | 2021 | 93,01       |
| Exploring IT/IS proactive and knowledge transfer on enterprise digital business transformation (EDBT): a technology-knowledge perspective | 2021 | 98,01       |
| Cybersecurity Challenges for Manufacturing Systems 4.0: Assessment of the Business Impact Level   | 2021 | 95,01       |
| <i>Big Data</i> Impacting Dynamic Food Safety Risk Management in the Food Chain   | 2021 | 93,01       |
| <i>Big Data</i> analytics energy-saving strategies for air compressors in the semiconductor industry—an empirical study                   | 2021 | 97,01       |
| Artificial Intelligence and Technology in Teaching Negotiation  | 2021 | 94,00       |
| A fuzzy model of wearable network real-time health monitoring system on pharmaceutical industry   | 2021 | 94,00       |
| A Aprendizagem profunda Perspective on Dropwise Condensation  | 2021 | 94,02       |
| A <i>Big Data</i> -driven framework for sustainable and smart additive manufacturing  | 2021 | 154,01      |
| A <i>Big Data</i> -centric architecture metamodel for Industry 4.0  | 2021 | 95,01       |
| Towards data-driven approaches in manufacturing: an architecture to collect sequences of operations                                       | 2020 | 98,01       |

|   |      |        |
|---|------|--------|
| The state of the art and taxonomy of <i>Big Data</i> analytics: view from new <i>Big Data</i> framework                                     | 2020 | 138,01 |
| The architecture development of Industry 4.0 compliant smart machine tool system (SMTS)   | 2020 | 117,01 |
| Shared-use mobility competition: a trip-level analysis of taxi, bikeshare, and transit mode choice in Washington, DC                        | 2020 | 102,00 |
| Scopus scientific mapping production in industry 4.0 (2011–2018): a bibliometric analysis   | 2020 | 174,01 |
| REDTag: a Predictive Maintenance Framework for Parcel Delivery Services   | 2020 | 96,00  |
| Privacy and security in the <i>Big Data</i> paradigm  | 2020 | 113,00 |
| Lean 4.0: A new holistic approach for the integration of lean manufacturing tools and digital technologies                                  | 2020 | 105,00 |
| Integration of Lean practices and Industry 4.0 technologies: smart manufacturing for next-generation enterprises                            | 2020 | 122,00 |
| Identification and classification of materials using machine vision and AM in the context of industry 4.0                                   | 2020 | 124,01 |
| Geographic Information System and Big Spatial Data: A Review and Challenges   | 2020 | 93,00  |
| Creating a road map for industry 4.0 by using an integrated fuzzy multicriteria decision-making methodology                                 | 2020 | 93,00  |
| BlockIoTIntelligence: A Blockchain-enabled Intelligent IoT Architecture with Artificial Intelligence  | 2020 | 298,01 |
| <i>Big Data</i> and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case                 | 2020 | 228,01 |
| <i>Big Data</i> analytics enhanced healthcare systems: a review   | 2020 | 138,00 |
| Application of industrial <i>Big Data</i> for smart manufacturing in product service system based on system engineering using fuzzy DEMATEL | 2020 | 110,01 |
| A method of NC machine tools intelligent monitoring system in smart factories   | 2020 | 121,01 |
| A Design-to-Device Pipeline for Data-Driven Materials Discovery   | 2020 | 111,02 |
| The role of <i>Big Data</i> analytics in industrial IoT   | 2019 | 239,01 |
| The impact of <i>Big Data</i> analytics on company performance in supply chain management   | 2019 | 100,00 |
| DataWarrior: an evaluation of the open-source drug discovery tool   | 2019 | 100,01 |
| Big Earth data analytics: a survey  | 2019 | 108,01 |
| Barriers to <i>Big Data</i> analytics in manufacturing supply chains: A case study from Bangladesh  | 2019 | 168,01 |

**APÊNDICE B – NÚMERO DE PUBLICAÇÕES POR PAÍS**

| <b>País</b>         | <b>Artigos</b> | <b>País</b>  | <b>Artigos</b> |
|---------------------|----------------|--------------|----------------|
| <b>CHINA</b>        | 127            | CANADA       | 18             |
| <b>USA</b>          | 112            | GERMANY      | 14             |
| <b>SPAIN</b>        | 41             | AUSTRALIA    | 13             |
| <b>INDIA</b>        | 40             | TURKEY       | 12             |
| <b>ITALY</b>        | 38             | IRAN         | 10             |
| <b>MALAYSIA</b>     | 36             | THAILAND     | 10             |
| <b>SOUTH KOREA</b>  | 29             | SOUTH AFRICA | 9              |
| <b>SAUDI ARABIA</b> | 24             | SWEDEN       | 9              |
| <b>UK</b>           | 24             | INDONESIA    | 8              |
| <b>BRAZIL</b>       | 18             | ROMANIA      | 8              |