

Big Data na Indústria 4.0: Tecnologias, Desafios e Oportunidades

Fabiane Florencio de Souza, Prof. Dra. Regina Negri Pagani, Giovanni Luigi Schiavon

Resumo:

O *Big Data* (BD) é um dos pilares da Indústria 4.0, considerada atualmente como a quarta revolução industrial. Diante da grande quantidade de dados envolvendo as conexões entre as duas áreas de estudos, aliado a constante mudança do cenário tecnológico, surgem nesse meio diversas oportunidades, desafios e tecnologias a serem exploradas continuamente. O objetivo dessa pesquisa é de realizar uma revisão de literatura no formato de bibliometria com a finalidade de identificar as interações entre Indústria 4.0 e BD, quais são as análises e ferramentas utilizadas, além de pontuar sobre as oportunidades e desafios encontrados por diversos autores. O estudo possibilitou uma visão das tecnologias utilizadas, linguagens de programação e análise de dados, assim como as oportunidades e desafios para estudos futuros, como novos modelos de compartilhamentos de dados, maior envolvimento *Humans-in-the-loop* entre outras fatores apresentados no decorrer do artigo.

Palavras chave: Indústria 4.0, *Big Data*, Tecnologia, bibliometria.

Big Data in Industry 4.0: Technologies, Challenges and Opportunities

Abstract: Big Data (BD) is one of the pillars of Industry 4.0, currently considered the fourth industrial revolution. Given a large amount of data involving the connections between the two areas of study, coupled with the constantly changing technological landscape, there are several opportunities, challenges, and technologies to be explored continuously. The objective of this research is to perform a literature review in a bibliometric format in order to identify the interactions between Industry 4.0 and BD, what the analyzes and tools used are, and to point out the opportunities and challenges found by several authors. The study provided insight into the technologies used, programming languages, and data analysis, as well as opportunities and challenges for future studies, such as new data sharing models, increased *Humans-in-the-loop* engagement, and other factors presented throughout the course of the article.

Key-words: Industry 4.0, Big Data, Technology, Bibliometrics.

1. Introdução

A Indústria 4.0, também conhecida como *Smart Manufacturing*, *Fourth Industrial Revolution*, *Smart Industry* e *Integrated industry* (SILVA et al., 2019; HERMANN et al., 2016), teve sua origem nos avanços tecnológicos (DRAGICEVIC et al., 2019). Alguns autores apontam que um projeto de manufatura avançada, apoiada pelo governo alemão no ano de 2011, também influenciou o surgimento desse conceito que, desde então, tornou-se muito utilizado (KAGERMANN et al., 2011; XU et al., 2018) e que, de acordo com Lu (2017), refere-se à quarta revolução industrial projetada para a produção descentralizada, visando obter personalização e eficiência de recursos (BRETTEL et al. 2014).

Segundo Baena et al. (2017), são observados alguns pilares, que deverão receber investimentos e serem implementados em toda a fábrica para que a Indústria 4.0 passe a ser uma realidade nos sistemas de fabricação. Ainda que existam alguns debates significativos sobre a quantidade e relevância desses pilares, é amplamente aceito que os mais importantes são: os *Cyber-Physical System*, *Cyber Physical Production Systems*, BD e *Cloud*.

BD tornou-se um termo genérico para definir qualquer técnica que processa, analisa, armazena, captura, monitora a segurança e a privacidade de uma grande quantidade de dados, e seu tamanho é um valor em constante mudança, em 2005 era classificado como *Terabytes*, a partir de 2010 cresceu para *Exabyte* e mais recentemente, em 2017, passou a ser classificado como *Zettabyte* (XU & DUAN, 2018).

O termo BD não é novo, desde a década de 1980 alguns problemas de grande volume de dados já existiam, como os dados do genoma humano e da física de partículas. Entretanto, nos últimos anos o tópico tem se tornado cada vez mais popular no setor 4.0 e em muitos outros setores, visto que as técnicas existentes estão se tornando mais maduras para lidar com BD e obter valor a partir disso (XU & DUAN, 2018).

O trabalho de Ren et al. (2019), busca por meio de uma revisão de literatura analisar como o BD pode apoiar a fabricação inteligente e sustentável do produtos em todo o ciclo de vida, entretanto, seu foco não está especificamente na indústria 4.0.

Pieroni, Scarpato e Brilli (2018), abordam BD e Indústria 4.0 com foco em apenas um setor, o de veículos autônomos e conectados, a pesquisa demonstra uma abordagem não convencional para gerenciar a grande quantidade de dados gerada nesse cenário.

Esses e outros autores já abordaram a temática Indústria 4.0 e BD. No entanto, nenhum deles proporcionou uma pesquisa com foco em identificar as ferramentas para tratar o BD dentro da Indústria 4.0, além de apresentar oportunidades e desafios nesse contexto.

Diante da lacuna apresentada, este estudo tem como objetivo realizar uma revisão de artigos com a finalidade de identificar as interações entre Indústria 4.0 e BD, quais são as análises e ferramentas utilizadas, além de pontuar sobre as oportunidades e desafios encontrados por alguns autores. O estudo possibilitará, uma visão das tecnologias empregadas, assim como as oportunidades para estudos futuros.

O artigo apresenta em sua primeira Seção a Introdução sobre os temas abordados; a Seção seguinte descreve os procedimentos metodológicos adotados para a construção deste trabalho; a Seção 3 apresenta os Resultados e Discussões; na quarta e última Seção as Considerações Finais sobre o trabalho são apresentadas.

2. Metodologia

Uma revisão bibliográfica foi utilizada para a construção desse artigo. Kovaleski, Coutinho e Horodyski (2018), bem como os autores de referência na área a ser estudada, descrevem ser esse o processo essencial para a compreensão do tema a ser explorado.

Baseando-se nos princípios da Methodi Ordinatio (PAGANI, et al., 2015; 2017), a primeira etapa da pesquisa consistiu em estabelecer a intenção da pesquisa, que nesse artigo, é de fazer uma revisão sistemática da literatura sobre a interação entre Indústria 4.0 e BD, e a partir disso, identificar desafios e oportunidades dentro desse campo.

Depois disso foi definida a combinação das palavras-chave e qual base de artigos seria utilizada. Nessa pesquisa a busca sobre o tema foi realizada na Scopus utilizando a seguinte combinação: ('big data' AND 'industry 4.0' OR 'smart manufacturing' OR 'fourth industrial revolution' OR 'smart industry'), procurando por artigos e artigos de revisão.

Por fim, a exportação das informações dos artigos na base escolhida resultou no total de 82 artigos. A discussão dos artigos estudados é exposta na Seção de Resultados e Discussões, juntamente com as ferramentas, desafios e oportunidades encontrados.

3. Resultados e Discussões

3.1 Análise Bibliométrica das Publicações

A Figura 1 apresenta um *dashboard* com o panorama das publicações analisadas neste artigo. Como pode ser observado, mas de 90% das publicações eram do tipo artigo completo, enquanto que o restante eram artigos de revisão, ambos abordavam os dois campos de estudo deste artigo: BD e Indústria 4.0.

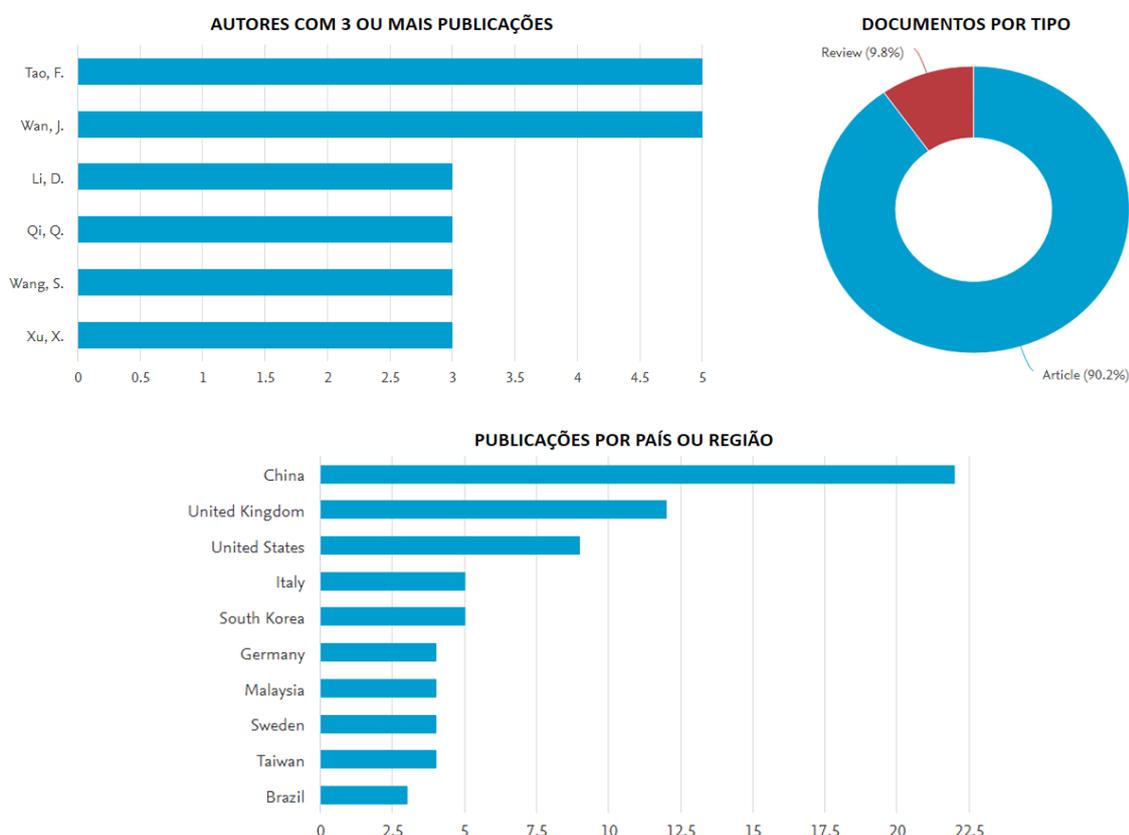


Figura 1 – Dashboard das publicações analisadas

Uma nuvem de palavras foi gerada a partir das palavras-chave encontradas nos artigos analisados, conforme demonstra a Figura 4, em que quanto maior a palavra, mais vezes ela apareceu no texto. Como esperado, as palavras BD e 4.0 aparecem em destaque, mas palavras como *Machine Learning*, *Digital*, *Revolution*, *Cloud*, *Intelligent*, *Data* e *Management* também aparecem em muitos artigos.

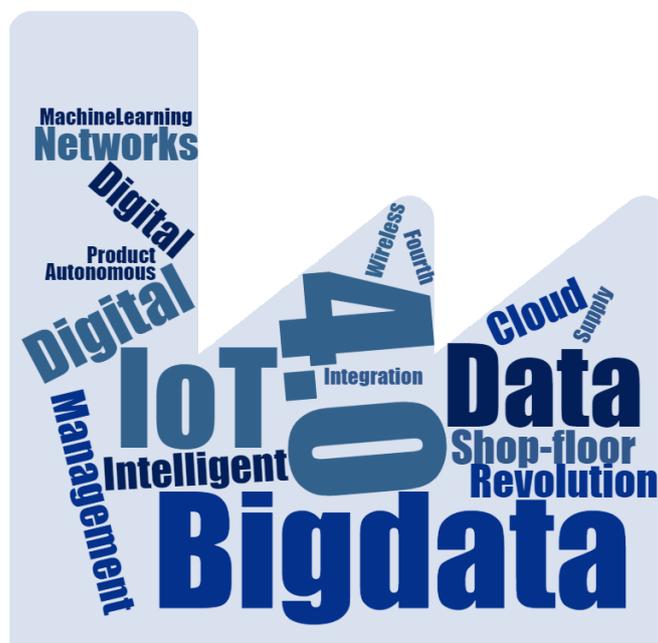


Figura 4 – Nuvem de Palavras-chave que mais aparecem nos artigos

Essa análise promove a visualização do panorama das publicações nessa área e pode servir para dar direcionamento aos pesquisadores.

3.2 Indústria 4.0 e Big Data

O surgimento de tecnologias disruptivas, como BD, inteligência artificial e computação em nuvem (CHRISTENSEN et al., 2015; RAMZI et al., 2019), aliado ao avanço da digitalização são considerados os principais facilitadores de uma mudança de paradigma observada na produção industrial (TERRAZAS et al., 2019). Essas tecnologias estão adentrando em boa parte das indústrias e, juntamente com a manufatura de ponta, reúne os mundos físicos e virtuais formando o *Cyber-Physical System* (RAMZI et al., 2019).

Esse cenário ficou conhecido como o Indústria 4.0 e propõe a integração das soluções de Tecnologia da Informação e Comunicação (TIC) para o monitoramento, adaptação, simulação e otimização de fábricas (TERRAZAS et al., 2019). No contexto da Indústria 4.0, várias fontes podem ser geradoras de dados, como por exemplo: controladores de máquinas, sensores, sistemas de manufatura, pessoas, entre muitas outras.

Esse grande volume de dados, que chega em alta velocidade e apresentam diferentes formatos, são chamados de *Big Data* (MAURO et al., 2019), mas também pode ser citado como *Huge data*, *Large scale data*, *Urban data management* e *Large volume of data* (CHAUHAN et al., 2016) e tem como características os '5Vs': volume, velocidade, variedade, variabilidade, veracidade, volatilidade e valor (MAVRAGANI et al., 2018)

Dentro da Indústria 4.0 o processamento de BD com a finalidade de identificar padrões, modelos e ideias é considerado a chave para a inovação sustentável (LEE et al., 2014). Sendo assim, analisar a grande quantidade de dados gerados é uma exigência tanto em termos de

potência de computação quanto em qualidade de sinal de *Internet*, além da necessidade de soluções altamente escaláveis de armazenamento (TERRAZAS et al., 2019).

Nesse contexto, o processo de *Big Data Analytics* (BDA) pode ser apontado como um examinador de grandes conjuntos de dados variados para descobrir padrões ocultos, correlações ainda não conhecidas, tendências de mercado, preferências dos clientes entre outras informações que podem ser úteis na tomada de decisões (ABELL et al., 2017). A análise dos dados é usada para descobrir futuras ameaças e novos problemas que possam ocorrer futuramente na produção, por exemplo (VAIDYA et al., 2018).

Os dados complexos também aparecem no estudo de ciclo de vida do produto, que influenciam na tomada de decisão das empresas de manufatura e portanto, necessitam de uma análise eficaz (AUSCHITZKY et al., 2014). A BDA juntamente com a tecnologia de mineração de dados podem ser usadas para fazer uma análise profunda do ciclo de vida do produto com dados históricos, por exemplo, detectar e capturar todos os dados em tempo hábil é um dos desafios nessa área (ZHANG et al., 2017).

Entretanto, o BDA é apenas uma análise que descobre o valor do BD, o envolvimento do conhecimento humano na tomada de decisão ainda é muito necessário para que as anseios dos clientes possam ser conhecidos. Essa colaboração é conhecida como *Humans-in-the-loop* (DRAGICEVIC et al., 2019). A Indústria 4.0 busca garantir que as pessoas envolvidas sejam capacitadas e com foco no ser humano ao invés de dividir e desumanizar (SCHWAB, 2017).

Um desafio identificado nessa interação é a disponibilidade de pessoas qualificadas em todos os níveis da organização, capazes de acompanhar a crescente complexidade do futuro dos sistemas de produção (AHMAD et al., 2019).

A implementação da virtualização e visualização da manufatura caracteriza vários desafios. Os objetos de fabricação devem ser visualizados em tempo real para garantir a qualidade e a segurança da produção. Aliado a isso, a produção de diferentes serviços deve ser virtualizada, para que se torne possível acompanhar todos da forma mais unificada possível. Abordagens virtualizadas e modelos de compartilhamento raramente são relatados e investigados (RAMZI et al., 2019), o que sinaliza um campo de estudo.

Algumas ferramentas são apontadas para armazenamento, manipulação e análise da grande quantidade de dados que compõem o BD. O Apache Storm (IQBAL & SOOMRO, 2015), Kafka, CouchDB, JSON, Rserve e ProActive, são utilizados por permitirem a interoperabilidade e cooperam eficientemente com o maquinário na automação da produção (TERRAZAS et al., 2018). Tecnologias como Hadoop (BORTHAKUR, 2007) e MapReduce fornecem as habilidades de computação paralela (CHU et al., 2016) para tratar a grande escala dados (DEAN & GHEMAWAT, 2004).

A *Elastic computing*, também entra nessa esfera, e pode ser definida como uma computação que tem capacidade de se adaptar às mudanças na carga de trabalho, ampliando e diminuindo os recursos de computação de maneira automática, de modo que os recursos de computação disponíveis correspondam às demandas atuais (COUTINHO, 2014). Dentro da computação, as Linguagens de programação como Java e linguagens de análise de dados como o R, são aliados no tratamento do BD (TERRAZAS et al., 2019).

O Quadro 1 apresenta um compilado das oportunidades, desafios e tecnologias encontradas durante a pesquisa.

DESAFIOS E OPORTUNIDADES	
Visualização e virtualização em tempo real	(RAMZI et al., 2019)
Abordagens virtualizadas e modelos de compartilhamento	(RAMZI et al., 2019)
Tratamento e análise do BD	(DEAN & GHEMAWAT, 2004)
Pessoas qualificadas em todos os níveis da organização	(AHMAD et. all. 2019)
Incentivos para Humans-in-the-loop	(DRAGICEVIC et al., 2019; SCHWAB, 2017)
Análise do BD do Ciclo de Vida do Produto	(AUSCHITZKY et al., 2014)
Soluções altamente escaláveis de armazenamento	(TERRAZAS et al., 2019)
Potência de computação e qualidade de sinal de Internet	(TERRAZAS et al., 2019)
Previsão de ameaças e futuros problemas na produção	(VAIDYA et al., 2018).
TECNOLOGIAS	
Apache Storm	(IQBAL & SOOMRO, 2015)
Kafka, CouchDB, JSON, Rserve e ProActive	(TERRAZAS et al., 2018)
Hadoop	(BORTHAKUR, 2007)
MapReduce	(CHU et al., 2016; DEAN & GHEMAWAT, 2004).
Linguagens: Java e R	(TERRAZAS et al., 2019)
<i>Elastic computing</i>	(COUTINHO, 2014).
<i>Big Data Analytics</i>	(ABELL et al., 2017; ZHANG et al., 2017).

Quadro 1 – Indústria 4.0 e BD: Oportunidades, Desafios e Tecnologias

O interesse pela Indústria 4.0 é crescente e é evidenciado pelas inúmeras pesquisas e iniciativas estratégicas propostas pelos principais países industriais, que tem o objetivo de desenvolver sistemas industriais sustentáveis e inteligentes. A China anunciou sua iniciativa de pesquisa "*Made in China 2025*", enquanto os Estados Unidos da América têm sua iniciativa "*Industrial Internet*" (DRAGICEVIC et al., 2019), demonstrando a importância de estudos nessa área.

4. Considerações Finais

A grande quantidade de dados gerados na Indústria 4.0 desafia as indústrias a pensarem em formas de coletá-los, analisa-los, entendê-los e extrair deles todas as informações possíveis e de forma rápida, com a finalidade de obter um processo mais inteligente e alinhado as demandas do mercado atual.

Esse BD, quando bem gerenciado, pode servir para a tomada de decisões mais assertivas, acompanhamento da produção em tempo real (RAMZI et al., 2019), provisionamento de insumos e até mesmo na previsão de dados futuros que podem ajudar no planejamento da indústria. Devido a essa importância o BD tornou-se um dos pilares mais importantes da Indústria 4.0.

A revisão bibliométrica possibilitou a visualização de alguns pontos a serem considerados, como o campo de estudo do ciclo de vida do produto, que gera uma grande quantidade de dados que precisam ser analisados (ZHANG et al., 2017), a necessidade do desenvolvimento de modelos de compartilhamento e visualização de dados (RAMZI et al., 2019) e o incentivo a participação humana dentro da Indústria 4.0 (DRAGICEVIC et al., 2019; SCHWAB, 2017), entre outros aspectos.

Entretanto, quando se fala em TIC as mudanças ocorrem de forma contínua, novas tecnologias surgem, a quantidade de dados aumenta e são criadas novas ferramentas para lidar com tudo isso, o que demonstra que o assunto deve estar em constante debate e

pesquisa.

As oportunidades e desafios listadas nesse artigo poderão servir de base para estudos posteriores na área. Outras bases de artigos podem ser incluídas nas buscas, além de outras palavras-chave menos utilizadas, como por exemplo, *huge data* ou até mesmo com foco em outro pilar da Indústria 4.0, como o *Cyber-Physical System*. Esse artigo limita-se a um mapeamento bibliométrico, e portanto, não realiza uma análise profunda do conteúdo. Para trabalhos futuros recomenda-se uma análise ainda mais detalhada dos temas abordados.

Referências

- ABELL, J. A. et al. Big Data-Driven Manufacturing Process Monitoring for Quality Philosophy. **Journal Of Manufacturing Science And Engineering**, [s.l.], v. 139, n. 10, p.1-12, 24 ago. 2017.
- AHMAD A. R. et al. Factors Influence The Students Readiness on Industrial Revolution 4.0. **International Journal of Recent Technology and Engineering**, Bhopal, v.8, n.2, p. 461-468. 2019.
- AUSCHITZKY, E.; HAMMER, M.; RAJAGOPAL, A. **How Big Data Can Improve Manufacturing**. Disponível em: <<https://www.mckinsey.com/business-functions/operations/our-insights/how-big-data-can-improve-manufacturing>>. Acesso em 27 set. 2019.
- BAENA, F. et al. Learning Factory: The Path to Industry 4.0. **Procedia Manufacturing**, [s.l.], v.9, n.1, p.73-80, 2017.
- BORTHAKUR, D. **The hadoop distributed file system**: architecture and design. Disponível em: <https://svn.apache.org/repos/asf/hadoop/common/tags/release-0.16.0/docs/hdfs_design.pdf>. Acesso em 27 set. 2019.
- BRETTEL, M. et al. “How Virtualization, Decentralization and Network Building Change the Manufacturing Landscape: An Industry 4.0 Perspective. **International Journal of Mechanical, Industrial Science and Engineering**, [s.l.], v.8, n.1, p.37–44. 2014.
- CHAUHAN, S.; AGARWAL, N.; KAR, A. K. Addressing big data challenges in smart cities: a systematic literature review. **Info**, [s.l.], v.18, n.4, p.73-90, 2016.
- CHRISTENSEN, C. M., RAYNOR, M. E., MCDONALD, R. “What is disruptive innovation”, **Harvard Business Review**, [s.l.], v.93, n.12, p.44-53, 2015.
- CHU, J.; WU, J.; SONG, M. An SBM-DEA model with parallel computing design for environmental efficiency evaluation in the big data context: a transportation system application. **Annals Of Operations Research**, [s.l.], v.270, n.1-2, p.105-124, 2016.
- COUTINHO, E. F. et al. Elasticity in cloud computing: a survey. **Annals Of Telecommunications**, [s.l.], v.70, n.7-8, p.289-309, 2014.

DEAN, J.; GHEMAWAT, S. MapReduce: simplified data processing on large clusters. **Communications Of The Acm**, New York, v.51, n.1, p.107-113, 2008.

DRAGICEVIC, N. et al. A conceptual model of knowledge dynamics in the industry 4.0 smart grid scenario. **Knowledge Management Research & Practice**, Londres, v.17, n.3, p.1-15, 2019.

HERMANN, M.; PENTEK, T.; OTTO, B. Design Principles for Industrie 4.0 Scenarios. *In*: INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 49, 2016, Havaí. **Anais [...]**. Havaí: IEEE Computer Society Washington, 2016. p. 3928-3937. Disponível em: <https://dl.acm.org/citation.cfm?id=2918269&picked=prox>. Acesso em: 26 set. 2019.

IQBAL, M. H.; SOOMRO, T. R. Big Data Analysis: Apache Storm Perspective. **International Journal Of Computer Trends And Technology**, [s.l.], v.19, n.1, p.9-14, 2015.

KAGERMANN, H., LUKAS, W. D.; WAHLSTER, W. **Industrie 4.0: Mit dem internet der dinge auf dem weg zur 4. industriellen revolution**. Disponível em: <https://www.ingenieur.de/technik/fachbereiche/produktion/industrie-40-mit-internet-dinge-weg-4-industriellen-revolution/>. Acesso em: 30 set. 2019.

KOVALESKI, A.; COUTINHO, H. M.; HORODYSKI, G. S. Discussão atual sobre o tema souvenir gastronômico no turismo: revisão de literatura. **Marketing & Tourism Review**, Belo Horizonte, v.3, n.3, p.1-24, 2018.

LEE, J.; KAO, H.; YANG, S. Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. **Procedia Cirp**, [s.l.], v.16, n.1, p.3-8, 2014.

LU, Y. Cyber Physical System (CPS)-Based Industry 4.0: A Survey. **Journal Of Industrial Integration And Management**, [s.l.], v.2, n.3, p.1750014-1750025, 2017.

MAURO, A.; GRECO, M.; GRIMALDI, M. Understanding Big Data Through a Systematic Literature Review: The ITMI Model. **International Journal Of Information Technology & Decision Making**, [s.l.], v.18, n.04, p.1433-1461, 2019.

MAVRAGANI, A.; OCHOA, G.; TSAGARAKIS, K. P. Assessing the Methods, Tools, and Statistical Approaches in Google Trends Research: Systematic Review. **Journal Of Medical Internet Research**, Toronto, v.20, n.11, p.270-282, 2018.

PAGANI, R. N.; KOVALESKI, J. L.; DE RESENDE, L. M. M.. Avanços na composição da Methodi Ordinatio para revisão sistemática de literatura. **Ciência da Informação**, Brasília, v.46, n.2, p.161-187, 2017.

PAGANI, R. N.; KOVALESKI, J. L.; DE RESENDE, L. M. M.. Methodi Ordinatio: a proposed methodology to select and rank relevant scientific papers encompassing the impact factor, number of citation, and year of publication. **Scientometrics**, [s.l.], v.105, n.3, p.2109-2135, 2015.

PIERONI, A.; SCARPATO, N.; BRILLI, M. Industry 4.0 Revolution in Autonomous and Connected Vehicle A non-conventional approach to manage Big Data. **Journal of Theoretical and Applied Information Technology**, Islamabad, v.96, n.1, p.10-18, 2018.

RAMZI, N.; AHMAD, H., ZAKARIA, N. A Conceptual Model on People Approach and Smart Manufacturing. **International Journal of Supply Chain Management**, Londres, v.8, n.4, p. 1102- 1107, 2019.

REN, S. et al. A comprehensive review of big data analytics throughout product lifecycle to support sustainable smart manufacturing: A framework, challenges and future research directions. **Journal Of Cleaner Production**, [s.l.], v.210, n.1, p.1343-1365, 2019.

SCHWAB, K. **The fourth industrial revolution**. New York: Ed. Crown Business, 2017.

SILVA, V. L.; KOVALESKI, J. L.; PAGANI, R. N. Technology Transfer and Human Capital in the Industrial 4.0 Scenario: A Theoretical Study. **Future Studies Research Journal: Trends and Strategies**, São Paulo, v.11, n.1, p.102-122, 2019.

TERRAZAS, G.; FERRY, N.; RATCHEV, S. A cloud-based framework for shop floor big data management and elastic computing analytics. **Computers In Industry**, [s.l.], v.109, n.1, p.204-214, 2019.

TERRAZAS, G.; SILVA, L.; RATCHEV, S. Towards a Cloud-Based Analytics Framework for Assembly Systems. **Precision Assembly In The Digital Age**, [s.l.], v.530, n.1, p.134-141, 2018.

VAIDYA, S.; AMBAD, P.; BHOSLE, S. Industry 4.0 – A Glimpse. **Procedia Manufacturing**, [s.l.], v.20, n.1, p.233-238, 2018.

XU, L. d.; Duan, L.. Big data for cyber physical systems in industry 4.0: a survey. **Enterprise Information Systems**, Londres, v.13, n.2, p.148-169, 2018.

XU, L. d.; Xu, E. L.; Li, L. Industry 4.0: state of the art and future trends. **International Journal Of Production Research**, Londres, v.56, n.8, p.2941-2962, 2018.

ZHANG, Y. et al. A framework for Big Data driven product lifecycle management. **Journal Of Cleaner Production**, [s.l.], v.159, n.1, p.229-240, 2017.