



ConBRepro

XI CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO



01 a 03
de dezembro 2021

Big Data Analytics no contexto Healthcare 4.0. Quais são os principais desafios desta interação?

Ma. Fabiane Florencio de Souza

Gestão de Transferência de Tecnologia – Universidade Tecnológica Federal do Paraná

Ma. Alana Corsi

Gestão de Transferência de Tecnologia – Universidade Tecnológica Federal do Paraná

Prof. Dra. Regina Negri Pagani

Gestão de Transferência de Tecnologia – Universidade Tecnológica Federal do Paraná

Prof. Dr. João Luiz Kovaleski

Gestão de Transferência de Tecnologia – Universidade Tecnológica Federal do Paraná

Resumo: Este artigo teve como objetivo reunir os principais desafios relacionados a Big Data Analytics na Healthcare 4.0. Para atingir este objetivo, foi realizada uma revisão sistemática de literatura utilizando a metodologia Methodi Ordinatio, construindo um portfólio de artigos para a análise quantitativa e qualitativa, com o auxílio do software VOSviewer. Os resultados indicaram que apesar do conceito Healthcare 4.0 ter sido abordado por vários artigos, esse termo ainda não é amplamente adotado pelos autores. Foi possível perceber que os desafios se concentram em questões de segurança, privacidade de dados e pacientes, custo com estrutura adequada de Tecnologia de Informação, entre outros. Por fim, uma agenda de pesquisas futuras foi proposta visando nortear futuros estudos nesse contexto.

Palavras-chave: Big Data; Big Data Analytics, Healthcare 4.0, Health Data, Desafios.

Big Data Analytics in the Healthcare 4.0 context. Which are the main challenges in this interaction?

Abstract: This article aimed to bring together the main challenges related to Big Data Analytics in Healthcare 4.0. To achieve this goal, a systematic literature review was carried out using the Methodi Ordinatio methodology, building a portfolio of articles for quantitative and qualitative analysis, with the help of the VOSviewer software. The results indicated that although the Healthcare 4.0 concept has been addressed in several articles, this term is still not widely adopted by the authors. It was possible to see that the challenges are focused on issues of security, data and patient privacy, cost with an adequate Information Technology structure, among others. Finally, an agenda for future research was proposed to guide future studies in this context.

Keywords: Big Data; Big Data Analytics, Healthcare 4.0, Health Data, Challenges.

1. Introdução

A Indústria 4.0 é representada pela crescente automação do ambiente de manufatura, buscando a eficácia dos processos, serviços e produtos por meio das Tecnologias de Informação e Comunicação (TIC) (LASI *et al.*, 2014; LIAO *et al.*, 2017).

Com isso, o “4.0” passou a ser adotado por variadas áreas visando a ampliação do uso de tecnologias aplicadas aos seus processos, como por exemplo em Total Quality Management 4.0 (SOUZA *et al.*, 2021), Education 4.0 (RAMÍREZ-MONTOYA *et al.*, 2021) e Healthcare 4.0 (CHUTE; FRENCH, 2019).

A área da saúde, representada pela Healthcare 4.0, já muito habituada ao uso das tecnologias para monitoramento e aferições dos pacientes em momentos anteriores, passou a ver um aumento de tecnologias capazes de captar dados relacionados a diversos aspectos de saúde do paciente no H4.0.

Essa captação resulta em uma grande quantidade de dados nomeada como Big Data, que se refere aos dados gerados continuamente a partir de várias fontes, apresentando características como: grande volume, alta velocidade de geração e variedade de formatos (ALI; ABDULLAH, 2019; WASIM *et al.*, 2019).

Para que o Big Data de saúde se transforme em informações que possam ser visualizadas de forma clara e o mais próximo possível do tempo real, é necessário que uma análise de dados seja realizada, sendo necessário então o uso da Big Data Analytics (BDA), que torna possível a compreensão de informações em uma quantidade massiva de dados, por meio de um conjunto de tecnologias, técnicas e ferramentas, capazes de recuperar os dados de forma correta, identificando padrões e servindo de base para tomada de decisões (HARIRI; FREDERICKS; BOWERS, 2019; CORSI *et al.*, 2021; GOMES *et al.*, 2021).

Quando a BDA é aplicada ao cenário da saúde, a análise desses dados busca compreender o uso de modelos estatísticos, cognitivos, preditivos, contextuais e quantitativos para tomada de decisão eficiente, quanto ao planejamento e gerenciamento de recursos, dentro da Healthcare 4.0 (BAHRI *et al.*, 2019).

Dessa forma, buscando reunir os desafios identificados na literatura, quanto a Big Data Analytics e o emergente cenário de Healthcare 4.0, esse artigo tem como objetivo realizar uma análise quantitativa dos artigos disponíveis nas bases de dados Science Direct, Scopus e Web of Science, além de uma análise qualitativa visando destacar quais são os principais desafios nesse contexto.

Para que o objetivo da pesquisa fosse atingido, foi realizado uma Revisão Sistemática de Literatura (SLR) por meio da metodologia Methodi Ordinatio, proposta por Pagani, Kovaleski e Resende (2015; 2017), a fim de construir um portfólio de artigos com relevância científica, o qual será fonte das coletas e análises dos dados.

2. Metodologia

Para explorar o emergente conceito de Healthcare 4.0 aliado a Big Data Analytics, foi realizado uma Revisão Sistemática de Literatura (SLR), baseada na metodologia multicritério Methodi Ordinatio, proposta por Pagani *et al.* (2015; 2017).

Essa metodologia permite a construção de um portfólio de artigos ranqueados por relevância científica, por meio da aplicação da Equação InOrdinatio, que classifica o portfólio de artigos com base em três critérios: Impact Factor (IF), Number of Citation (Ci), and Year of Publication. Ao todo são nove protocolos a serem seguidos para construir o portfólio de artigos, conforme descrito a seguir:

Passo 1 – Definir a intenção de pesquisa: a intenção foi explorar o Big Data Analytics aplicado ao emergente conceito de Health 4.0. Dessa forma, buscou-se encontrar documentos que abordassem esse relacionamento.

Passo 2 – Pesquisa preliminar: as pesquisas foram realizadas em diferentes bases de dados, buscando identificar as bases de dados com maior número de resultados, utilizando as seguintes palavras-chaves “big data analy*”, “healthcare 4.0”, “health 4.0”, “healthcare data” e “health data”.

Passo 3 – Definindo as combinações de palavra chave e bases de dados: tendo como base os resultados obtidos com as buscas preliminares, foram selecionadas as bases dados Scopus, Web of Science and Science Direct por apresentarem um número satisfatório de artigos. Foram definidas duas combinações de palavras chave para atender o objetivo desta pesquisa, conforme Tabela 1.

Passo 4 - Pesquisa final nas bases de dados: utilizando as combinações de palavras-chave definidas no Passo 3, foram realizadas as pesquisas finais, utilizando os seguintes parâmetros: Sem corte temporal; Pesquisa por título, resumo e palavra-chave; Artigo e revisão; e Uso de operadores booleanos. Os resultados podem ser visualizados na Tabela 1.

Tabela 1 – Resultado da pesquisa final nas bases de dados

Passos 1 - 4			
Combinações de palavras-chave	Scopus	WoS	Science
"big data analy*" AND ("healthcare data" OR "health data")	104	89	24
"big data analy*" AND ("healthcare 4.0" OR "health 4.0")	4	3	0
Total de artigos		224	
Passo 5			
Artigos duplicados deletados		98	
Artigos sem relação com tema		60	
Total de artigos restantes		66	
Passos 6 - 7			
Portfólio final		32	

Fonte: Os autores (2021)

Passo 5 – Filtragem: os procedimentos de filtragem visam eliminar artigos duplicados; artigos de conferências, livros e capítulos de livros; e artigos que abordem temas fora do escopo desta pesquisa, sendo eliminados pela leitura do título, resumo e palavras-chave. Os resultados obtidos nos procedimentos de filtragem foram, conforme Tabela 1.

Passo 6 - Identificando as variáveis: Fator de Impacto (IF), Número de Citação (Ci) e Ano de Publicação (PublishYear) foram identificados para cada um dos 66 artigos. Para a coleta do IF, utilizou-se a métrica Journal Citation Reports (JCR), disponível no portal da CAPES. A métrica Scimago (SJR) foi utilizada para artigos sem JCR. Para a coleta do número de citações, foi utilizada a plataforma Google Scholar.

Passo 7 – Classificação com o InOrdinatio: após a coleta das variáveis, foi aplicada a Equação InOrdinatio (1), resultando em um portfólio ordenado de artigos científicos, de acordo com a relevância científica.

$$\text{InOrdinatio} = (\text{IF}/1000) + \alpha * [10 - (\text{ResearchYear} - \text{PublishYear})] + (\text{Ci}) \quad (1)$$

A equação é composta pelos seguintes elementos: IF (fator de impacto); α (valor de alfa, variando de 1 a 10, definido pelo pesquisador de acordo com a importância da novidade do tema; neste estudo, o valor de α foi definido em 10, devido ao tema ser objeto do estudo em artigos mais recentes); ResearchYear (ano em que a pesquisa foi desenvolvida, neste caso, 2021); PublishYear (o ano em que o artigo foi publicado); e Ci (número de vezes que o artigo foi citado).

Com o portfólio ordenado conforme critérios da Equação 1, foi adotado mais um critério de exclusão, devido à grande quantidade de artigos, visto que esse artigo buscou analisar o portfólio completo de forma quantitativa e qualitativa. Portanto, definiu-se que todos os artigos que apresentassem InOrdinatio maior ou igual a 100 seriam selecionados para compor o portfólio final. Assim, o portfólio final, ordenado por relevância científica, foi composto por 32 artigos.

Passo 8 – Encontrando os artigos: todos os 32 artigos foram encontrados em sua versão completa, e recuperados usando o Mendeley®, possibilitando o início do último passo da Methodi Ordinatio (Pagani et al., 2015; 2017).

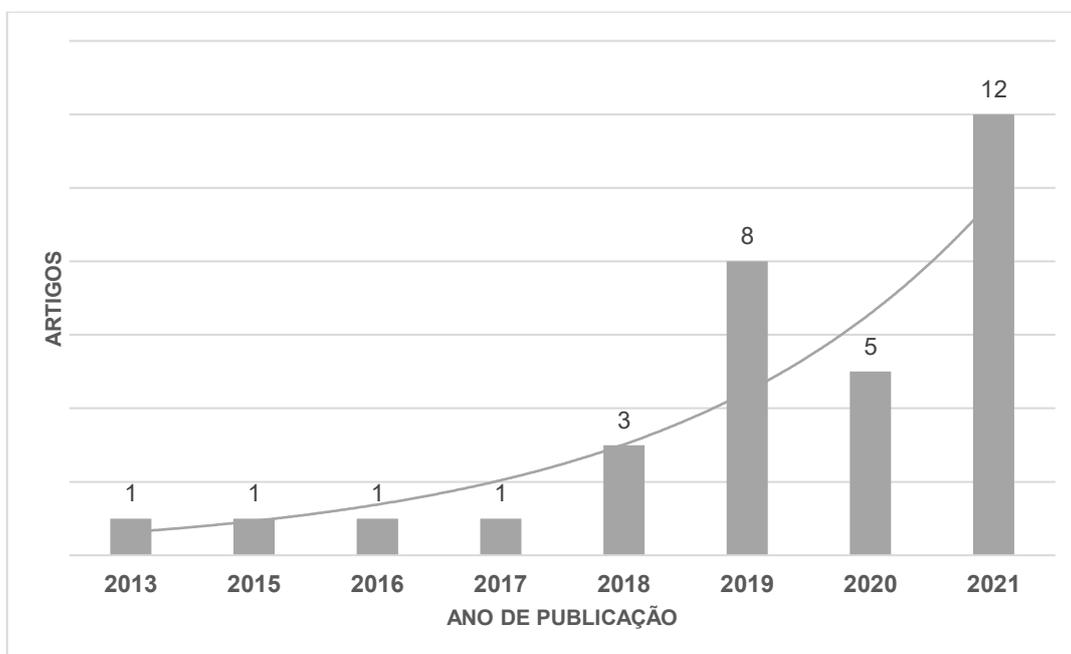
Passo 9 – Leitura sistemática: finalmente, teve início a leitura e análise sistemática do conteúdo, que se dividiu na análise bibliométrica por meio do software VOSviewer (3.1) e na análise de conteúdo realizada através da leitura dos artigos (3.2), conforme apresentado na próxima Seção.

3. Resultados e Discussões

3.1 Resultados da análise bibliométrica

Visando melhor compreender as perspectivas relacionadas a Big Data Analytics (BDA) e a Healthcare 4.0 (H4.0), essa Seção aborda os resultados bibliométricos. No Figura 1 é possível visualizar que o portfólio final conta com publicações de 2013 a 2021, sendo que 2019 a 2021 representam 78,1% do portfólio, evidenciando a atualidade dos artigos utilizados na construção deste artigo.

Figura 1 – Total de artigos do portfólio final publicados por ano



Fonte: Os autores (2021)

Foi realizada também a análise das principais palavras-chave presente no portfólio, sendo utilizada a funcionalidade de mapa de redes do software VOSviewer, conforme Figura 2.

Quadro 1 – Principais artigos que abordam desafios no contexto Big Data Analytics e Healthcare 4.0

Autor	Revista	Palavras-chave	Portfólio de artigos
Mehta e Pandit (2018)	Journal Of Medical Informatics	“big data” “big data analytics” “healthcare” “medicine” “biomedicine”	58
Razzak, Imran e Xu (2019)	Neural Computing And Applications	“disease prevention” “data analytics” “healthcare” “knowledge discovery” “prevention methodologies”	227*
Bahri <i>et al.</i> , (2019)	IEEE Access	“big data” “big data analytics” “healthcare”	50*
Shafqat <i>et al.</i> , (2020)	The Journal of Supercomputing	“healthcare analytics” “big data” “cloud computing” “knowledge management” “learning healthcare system”	66*
Aceto, Persico e Pescapé (2020)	Journal Of Industrial Information Integration	“industry 4.0” “internet of things” “big data” “cloud computing” “healthcare 4.0”	171*

* número total de referencias por não apresentar o número total do portfólio final de artigos

Fonte: Os autores (2021)

Como pode ser observado no Quadro 1, os artigos são um recorte do ano 2018 a 2020, publicados em diferentes revistas científicas, apresentando o portfólio de artigos entre 50 e 227 artigos no total, além de compartilharem de palavras-chave como: “big data”, “big data analytics” ou “data analysis” e termos relacionados a área da saúde como: “healthcare”, “medicine”, “biomedicine”, “disease prevention”, entre outros. Apenas o artigo de Aceto, Persico e Pescapé (2019), utiliza o a palavra “healthcare 4.0” para abordar esse conceito em sua pesquisa.

Mehta e Pandit (2018), publicaram uma revisão sistemática de literatura em que são apresentadas as aplicações e desafios de adoção da Big Data Analytics na área de saúde. Os autores destacaram que um dos principais desafios para a adoção da Big Data Analytics é a indisponibilidade de evidências de seus benefícios práticos na área de saúde, sendo assim, a necessidade de estudos que foquem nos resultados atingidos com o uso da BDA nos dados da saúde, dentro de um cenário real é destacada.

No ano seguinte, Razzak, Imran e Xu (2019), buscaram fornecer uma visão abrangente sobre o avanço dos métodos de análise de dados para a prevenção de doenças. A pesquisa ressaltou que os dados médicos apresentam como característica a dificuldade de serem analisados, portanto, entender como os provedores de saúde podem usar ferramentas de análise de dados para analisar e criar valor a partir de dados complexos ainda é um desafio.

Como foi visto na Seção 3.1, privacidade e segurança são palavras recorrentes no cenário de Big Data Analytics e Healthcare 4.0, e conforme abordado por Bahri *et al.* (2019), a garantia de que os dados dos pacientes estarão seguros e sem os menores riscos de dano a privacidade é um desafio importante para ser explorado em pesquisas futuras. Neste artigo é possível visualizar também os impactos do Big Data para os médicos, pacientes, farmacêuticas e seguradoras de saúde, além da análise dos diferentes desafios que deve

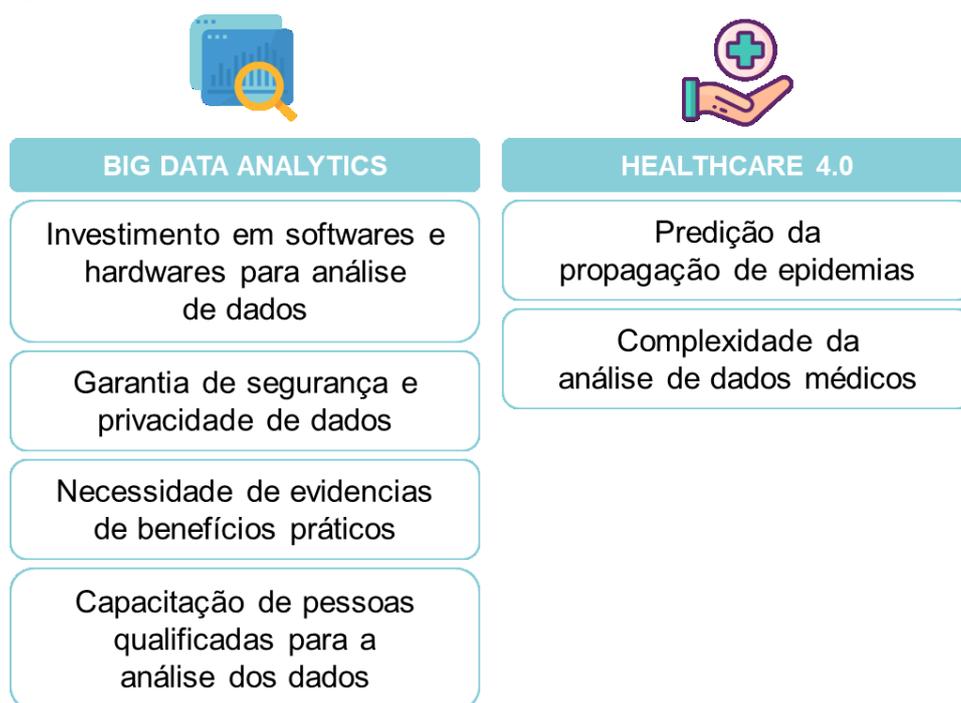
ser levado em consideração para obter benefícios através do Big Data Analytics. Para os autores, o desafio atual é o de desenvolver um sistema de decisão médica capaz de prever a propagação de epidemias em diferentes localizações geográficas, por meio da análise de dados de redes sociais como Facebook e Twitter.

Por meio de uma visão abrangente de diferentes sistemas de saúde integrados de Big Data Analytics, Shafqat *et al.* (2020), descreve algoritmos, técnicas e ferramentas de análise de dados de saúde que podem ser implantados em configurações Wireless, Cloud e Internet of Things. Esta pesquisa destaca que os sistemas de saúde modernos necessitam de uma revisão dos paradigmas tradicionais de software e hardware de saúde, que não estão suficientemente equipados para lidar com o volume e a diversidade dos dados de saúde que devem analisados com recursos de Big Data Analysis.

Por fim, Aceto, Persico e Pescapé (2020), fornecem uma descrição sobre as principais tecnologias e paradigmas em relação ao Healthcare 4.0 e Big Data, bem como cenários de utilização, desafios e benefícios. Ao entrelaçar os temas os autores concluíram que a extrema heterogeneidade dos dados da saúde dificultam a ampla adoção das técnicas de Big Data Analytics, devido a necessidade de investimentos em hardware e pessoas qualificadas para analisar.

Com a finalidade de resumir e reunir os principais desafios elencados pelos autores, foi construída Figura 3, em que é possível visualizar que existem desafios provenientes da BDA e outros mais relacionados a H4.0.

Figura 3 – Principais desafios no contexto Big Data Analytics e Healthcare 4.0



Fonte: Os autores (2021)

Os desafios atrelados a BDA são comuns também em outras áreas em que a análise de Big Data é uma necessidade, como no setor industrial (BONNARD *et al.*, 2021), por exemplo, em que o investimento em software e hardware e necessidade de capacitar as pessoas para executar tarefas cada vez mais analíticas também é um desafio.

No âmbito da saúde, os desafios provenientes da H4.0, destacam a predição de propagação de pandemias, que pode ser vinculado a recente pandemia de COVID-19, juntamente com a complexidade que a análise dos dados se tornou devido ao a Big Data.

Como visto, ainda existem diversos desafios que precisam ser explorados e melhor compreendidos para que a BDA possa ser amplamente aplicada aos dados da saúde, e principalmente, para que dessa análise resultem novos *insights* que apoiem a tomada de decisão baseada em dados reais. Portanto, os desafios reunidos na Figura 2, representam também uma agenda de pesquisas futuras, baseada em artigos científicos atuais, que podem nortear a busca de outros autores.

4. Considerações finais

Nesta revisão sistemática buscou-se reunir os desafios identificados na literatura, quando relacionados os termos Big Data Analytics (BDA) e Healthcare 4.0 (H4.0). Para isso, a metodologia Methodi Ordinatio foi utilizada, visando compor um portfólio de artigos recentes e com relevância científica conforme descrito na Seção 2, de Metodologia.

Com base no portfólio final de artigos, foi possível realizar a análise quantitativa que demonstrou a atualidade dos temas, bem como a pulverização de publicações ao longo dos anos. Dando sequência, a análise principais palavras-chave foi realizada e demonstrou o relacionamento entre a Big Data Analytics no contexto da saúde, atrelada a soluções para *internet, healthcare industry, medical data e healthcare system*.

Na sequência, teve início a análise qualitativa, que evidenciou cinco artigos que abordavam a temática e destacavam, entre outras questões, os desafios em relação a BDA e H4.0. Por meio disso, foi possível perceber que o termo Healthcare 4.0 não é amplamente adotado pelos autores, mesmo quando o conceito de tecnologias e saúde são explorados em suas publicações, mostrando que existe uma fragmentação nesse sentido.

Por fim, os desafios foram reunidos na Figura 3, e se concentraram em necessidade provenientes tanto da área da saúde, quanto da BDA. Em suma, os desafios se concentraram em questões como: segurança, privacidade, capacitação, predição de doenças, hardware, software e na necessidade de evidencia reais dos benéficos da BDA aplicada ao cenário H4.0.

Desta forma, além de reunir alguns dos principais desafios, este artigo apresenta também uma agenda de pesquisa, que pode ser explorada em pesquisas futuras visando ampliar a adoção BDA na H4.0, com foco na precisão dos resultados médicos baseados em dados reais.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradecemos também à UTFPR e ao programa de pós-graduação PPGEF pelo apoio e empenho no incentivo à pesquisa acadêmica.

Referências

ACETO, G.; PERSICO, V.; PESCAPÉ, A. Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0. **Journal Of Industrial Information Integration**, [s.l.], v.18, n.1, p.100129, 2020.

ALI, A. H.; ABDULLAH, M. Z. A Survey on Vertical and Horizontal Scaling Platforms for Big Data Analytics. **International Journal Of Integrated Engineering**, [s.l.], v.11, n.6, p.138-150, 2019.

BAHRI, S. et al. BIG DATA for Healthcare: A Survey. **IEEE Access**, [s.l.], v.7, n.1, p.7397-7408, 2019.

- BONNARD, R. et al. Big data/analytics platform for Industry 4.0 implementation in advanced manufacturing context. **The International Journal of Advanced Manufacturing Technology**, v.1, 2021.
- CHUTE, C.; FRENCH, T. Introducing Care 4.0: An Integrated Care Paradigm Built on Industry 4.0 Capabilities. **International Journal Of Environmental Research And Public Health**, v.16, n.12, p.2247, 2019.
- CORSI A. et al. Big data analytics as a tool for fighting pandemics: a systematic review of literature. **Journal of Ambient Intelligence and Humanized Computing**, v.12, p. 9163–9180, 2021.
- GOMES, M. A. S. et al. Big data analytics applied to health services: a literature review. **Exacta**, v.19, n.4, p.1-20, 2021.
- HARIRI, R. H.; FREDERICKS, E. M.; BOWERS, K. M. Uncertainty in big data analytics: survey, opportunities, and challenges. **Journal of Big Data**, v.6, n.44, p. 1-16, 2019.
- LASI, H.; FETTKE, P.; KEMPER, H.; FELD, T.; HOFFMANN, M. Industry 4.0. **Business & Information Systems Engineering**, v.6, n.4, p.239–242, 2014.
- LIAO, Y.; DESCHAMPS, F.; LOURES, E. F. R.; RAMOS, L. F. P. Past, present and future of industry 4.0: a systematic literature review and research agenda proposal. **International Journal of Production Research**, v.55, n.12, p.3609–3629, 2017.
- MEHTA, N.; PANDIT, A. Concurrence of big data analytics and healthcare: A systematic review. **International Journal Of Medical Informatics**, v.114, n.1, p.57-65, 2018.
- PAGANI, R. N.; KOVALESKI, J. L.; RESENDE, L. 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**, v.105, n.1, p.2109–2135, 2015.
- PAGANI, R. N.; KOVALESKI, J. L.; RESENDE, L. M. Tics na composição da methodi ordinatio: construção de portfólio bibliográfico sobre modelos de Transferência de Tecnologia. **Ciência da Informação**, v.46, n.2, p.161-187, 2017.
- RAMÍREZ-MONTOYA, M. et al. Characterization of the Teaching Profile within the Framework of Education 4.0. **Future Internet**, v.13, n.4, p.91, 2021.
- RAZZAK, M.; IMRAN, M.; XU, G. Big data analytics for preventive medicine. **Neural Computing And Applications**, v.32, n.9, p.4417-4451, 2019.
- SHAFQAT, S. et al. Big data analytics enhanced healthcare systems: a review. **The Journal of Supercomputing**, v.76, n.3, p.1754-1799, 2020.
- SOUZA, F.F. et al. Total quality management 4.0: adapting quality management to Industry 4.0. **The TQM Journal**, v. ahead-of-print, n. ahead-of-print, p.1-21, 2021.
- WASIM, A. et al. Big Data Analytics - Current Status, Challenges and Connection of unbounded data Processing platforms. **International Journal Of Innovative Technology And Exploring Engineering**, [s.l.], v. 8, n. 92, p. 698-700, 2019.