

## Revisão de Literatura: *Big Data Analytics* na Engenharia de Produção

João Gabriel de Sousa Vale (Universidade de Brasília) joaogabriel1988@hotmail.com

Júlia Souza Brito (Universidade de Brasília) juliabritobsb@gmail.com

Ari Melo Mariano (Universidade de Brasília) [arimariano@unb.br](mailto:arimariano@unb.br)

**Resumo:** O objetivo desta pesquisa é entender como as pesquisas de Big Data estão se desenvolvendo dentro da Engenharia de Produção, apresentando uma visão sistemática das principais contribuições da literatura de alto impacto no que tange o uso da análise de dados na Engenharia de Produção. Através da base Web of Science coleta de dados mostrou que há crescimento significativo no número de citações sobre o tema nos últimos anos. Este fato está de acordo com a crescente necessidade das empresas basearem cada vez mais suas decisões nas análises de dados, tendo em vista o seu alto impacto na geração de valor. Os resultados apontaram as principais abordagens através de análises tendo em vistas os núcleos mais fortes e principais fronts de pesquisa.

**Palavras chave:** Análise de Dados, Teoria do Enfoque Meta Analítico Consolidado, Engenharia de Produção.

### Big Data Analytics Literature Review in Production Engineering

**Abstract:** The aim of this research is to understand how Big Data research is developing within Production Engineering, presenting a systematic overview of the main contributions of high impact literature regarding the use of data analysis in Production Engineering. Through the Web of Science database, data collection has shown that there has been significant growth in the number of citations on the subject in recent years. This fact is in line with the growing need for companies to increasingly base their decisions on data analysis, with its high impact on value creation. The results pointed out the main approaches through analysis in view of the strongest nuclei and main research fronts.

**Key-words:** Data Analysis, Theory Consolidated of the Meta Analytical Approach, Production Engineering.

#### 1. Introdução

Uma pesquisa divulgada pelo IBM revelou que, com a difusão da internet, são gerados mais de 2 quintilhões de terabytes de dados por dia. De acordo com X. Wu (2013), nossa eficiência em gerar dados nunca foi tão grande. Essa enorme quantidade de dados gerados tornou popular o termo *Big Data*. Segundo Frisk e Bannister (2017), a evolução da tecnologia tem gerado novas maneiras de coletar e analisar esses dados. Mas nem todo tipo de banco de dados pode ser considerado como *Big Data*, o qual, de acordo com McAfee e Brynjolfsson (2012), é caracterizado pelos 3 Vs (Velocidade, Volume e Variedade). Ainda segundo McAfee e Brynjolfsson (2012), a humanidade está passando por uma revolução ocasionada pelo *Big Data* e entrando em uma era na qual uma grande quantidade de informação, que é do interesse de todo tipo de empresa, está presente no universo digital. Utilizar avançadas técnicas de análise no *Big Data* é denominado de *Big Data Analytics* de acordo com Russom (2011). Segundo Kohavi (2002), a utilização dessa análise de dados dentro de empresas tem

crescido significativamente nos últimos anos e tem possibilitado o aprimoramento das tomadas de decisões.

De acordo com Kohavi (2002), o sucesso da utilização da técnica de análise de banco de dados com foco nos negócios possibilitou novas aplicações, como, por exemplo, dentro da engenharia de produção nas áreas de cadeia de suprimentos, logística dos materiais e controle de inventário. Contudo, de acordo com Ardito (2018), as pesquisas na área de *Big Data* para negócios estão se desenvolvendo sem uma base teórica clara. Deste modo, compreender as principais contribuições até o presente momento e como vai continuar, se faz necessário. Assim, esta pesquisa busca responder: quais as principais abordagens a respeito do big data? Quais os principais fronts de pesquisa?

Uma vez apresentada a justificativa desta pesquisa e os principais problemas a serem respondidos, este estudo tem como objetivo: entender como as pesquisas de *Big Data* estão se desenvolvendo dentro da Engenharia de Produção, apresentando uma visão sistemática das principais contribuições da literatura de alto impacto no que tange o uso da análise de dados na Engenharia de Produção.

Assim, este estudo apresenta no capítulo 2, a metodologia, 3 revisão da literatura e resultados e finalmente as considerações finais. O uso da metodologia após a introdução é um artifice necessário em trabalhos de revisão da literatura, uma vez que o referencial teórico se trata do resultado em si.

## 2. Metodologia

A metodologia utilizada, de caráter exploratório e abordagem quantitativa, foi proposta por Mariano e Rocha (2017) e é intitulada Teoria do Enfoque Meta Analítico Consolidado – TEMAC. Esta teoria está dividida em 3 etapas: (i) Preparação da pesquisa, (ii) Apresentação e inter-relação dos dados e (iii) Detalhamento, modelo integrador e validação por evidências.

Na primeira etapa, foi realizada a pesquisa na base de dados Web of Science no dia 11/10/2019, utilizando a string “Big Data Analytics”, com raio de busca de 2009 - 2019, envolvendo as áreas de conhecimento “Engineering Industrial” e “Engineering Manufacturing”. Foram encontrados 175 resultados, os quais compõem a amostra desta pesquisa.

Para a segunda etapa, foram utilizadas as informações adquiridas na base de dados e com elas pode-se aplicar algumas leis de bibliometria referentes a inter-relações entre os dados, como a Lei de Brandford – que cadencia a relevância de um periódico em determinada área de conhecimento, Lei da Obsolescência da Literatura – que estima a decaída de registros em determinada área de conhecimento, Lei do elitismo e Lei do 80/20 – que apresenta a elite de um determinado tema e a Lei de Lokta – que informa o grau de relevância dos autores, conforme exposto por Mariano e Rocha (2017).

Para a apresentação e inter-relação dos dados, utilizou-se das informações coletadas na base de dados para a definição da evolução do tema ano a ano, dos autores e trabalhos mais citados, das revistas mais relevantes, e da frequência de palavras chaves.

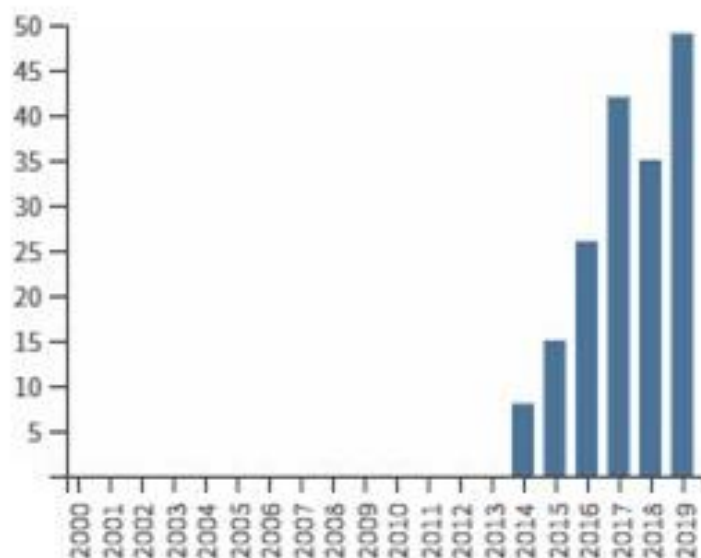
Na etapa de detalhamento, modelo integrador e validação por evidências os resultados apresentados onde se responderam as perguntas da pesquisa: Quais as linhas de abordagem sobre o tema? E quais os fronts de pesquisa? Para responder o primeiro questionamento foi utilizada a análise de *co-citation*, que inter-relaciona as citações dos documentos evidenciando documentos que são citados juntos e indicando linhas de pesquisa. Para o segundo questionamento, será utilizada a análise de *bibliographic coupling* que identifica futuros fronts de pesquisa. Para a etapa 3, o software *VOSviewer* foi utilizado com o intuito de realizar o detalhamento e validação por evidência, a partir de análises como o coupling e co-citation.

### 3. Revisão da Literatura e Resultados

#### 3.1. *Big Data Analytics* na literatura

Após o exame dos registros, constatou-se que o dado referente à pesquisa mais antigo presente no Web of Science é um artigo de 2014 intitulado "*Big Data Analytics for Supply Chain Management*", publicado por Leveling. No artigo, o autor lida com soluções baseadas em *Big Data* tendo foco na cadeia de suprimentos. Desde então, o total de citações sobre o tema alcançou um valor de 2900 citações, considerando-se o período de 2009 a outubro de 2019, com uma média de 16,57 citações por item. Os últimos 3 anos são responsáveis por 2594 citações, o que corresponde a 89% do total. A partir disso, é possível inferir o quão relevante o tema tem se tornado nos últimos anos e isso o torna um tema muito propício para o desenvolvimento de estudos.

Com o intuito de entender a evolução do tema no espaço de tempo pesquisado, foi criada a figura 1.



Fonte: própria. Extraído do web of science.

**Figura 1- Evolução do Tema**

Ao analisar a Figura 1, é possível perceber o desenvolvimento desse tema no meio acadêmico, que tende a continuar crescendo. Essa tendência de crescimento é evidenciada pelo ano de 2019, que ainda não chegou ao fim e já se tornou o ano com mais trabalhos publicados.

Dentro da base Web of Science, o trabalho mais citado foi “Internet of Things in Industries: A Survey”, escrito por Xu, Li Da, que resume sistematicamente a Internet das Coisas dentro da indústria. O trabalho foi publicado em 2014 e desde a sua publicação ele foi citado 1093 vezes. O segundo trabalho mais citado, com 181 citações, foi “Big Data Analytics in logistics and supply chain management: Certain investigations for research and applications”, publicado por Wang. Mesmo sendo o segundo mais citado, ele ainda possui 6 vezes menos citações que o primeiro lugar, o que reforça a importância do artigo “Internet of Things in Industries: A Survey”. Para descrever as principais contribuições dos artigos mais citados foi elaborado a Tabela 1 com os principais resultados:

<b>Autores</b>	<b>Título</b>	<b>Resumo</b>	<b>Citações</b>
Xu, Li Da; He, Wu; Li, Shancang	Internet of Things in Industries: A Survey	Este artigo analisa a pesquisa atual da IoT, as principais tecnologias facilitadoras, as principais aplicações da IoT nas indústrias e identifica tendências e desafios da pesquisa. A principal contribuição deste artigo de revisão é que ele resume sistematicamente a atual IoT de ponta em indústrias.	1093
Wang, Gang; Gunasekaran, Angappa; Ngai, Eric W. T.; Papadopoulos, Thanos	Big Data analytics in logistics and supply chain management: Certain investigations for research and applications	Este artigo revisa e classifica a literatura de Big Data business analytics no gerenciamento de logística e cadeia de suprimentos .	181
Ren, Steven Jifan; Wamba, Samuel Fosso; Akter, Shahriar; Dubey, Rameshwar; Childe, Stephen J.	Modelling quality dynamics, business value and firm performance in a Big Data analytics environment	O estudo identifica que a qualidade do sistema e a qualidade da informação são essenciais para aprimorar o valor comercial e o desempenho da empresa em um ambiente de Big Data.	122

Akter, Shahriar; Wamba, Samuel Fosso; Gunasekaran, Angappa; Dubey, Rameshwar; Childe, Stephen J.	How to improve firm performance using big data analytics capability and business strategy alignment?	Este estudo propõe um modelo de análise de Big Data baseado na teoria baseada em recursos (RBT) e na visão de entrelaçamento do sociomaterialismo.	87
Zhong, Ray Y.; Newman, Stephen T.; Huang, George Q.; Lan, Shulin	Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives	Este artigo fornece uma discussão da análise dos movimentos atuais no Big Data para gerenciamento da cadeia de suprimentos em serviços e manufatura em todo o mundo, incluindo América do Norte, Europa e região Ásia-Pacífico.	86

Fonte: própria. Extraído da base Web of Science

Tabela 1 – Artigos mais citados

Dentre os autores que mais publicaram estão Gunasekaran e Angappa com 9 publicações, sendo que seu artigo mais citado foi “*Big Data analytics in logistics and supply chain management: Certain investigations for research and applications*” com 181 citações, Wamba, Samuel Fosso com 5 publicações, tendo o artigo “*Modelling quality dynamics, business value and firm performance in a big data analytics environment*” com 122 citações, como seu trabalho mais citado e Wang, Lihui com 5 publicações, tendo como seu trabalho mais citado “*Big Data Analytics based fault prediction for shop floor scheduling*” com 22 citações.

Entre as revistas que mais publicaram sobre Big Data Analytics dentro das áreas de conhecimento escolhidas estão a *International Journal of Production Research*, com 23 registros, *Computers Industrial Engineering* com 15 registros *IEEE Transactions On Industrial Informatics*, com 10 publicações. A revista *IEEE Transactions On Industrial Informatics* se encontra em primeiro lugar em questão de fator de impacto na área pesquisada, o que mostra que a principal revista do tema já possui presença significativa sobre este tópico. A *International Journal Of Production Research* está no rank de número 25 do fator de impacto e a *Computers Industrial Engineering* não está entre as principais revistas quanto ao fator de impacto.

Com a finalidade de apresentar os registros obtidos de forma visual, foi criada uma nuvem de palavras através do software *TagCrowd* foram utilizadas as palavras chave mais frequentes das 175 publicações encontradas na base Web of Science.



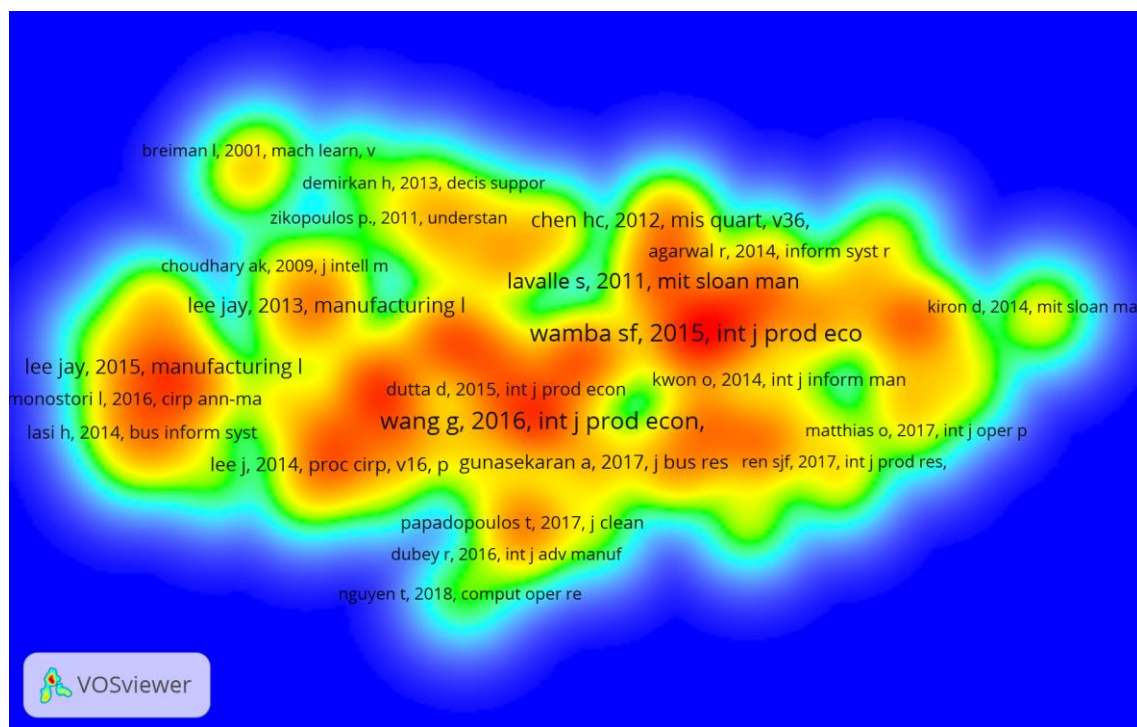
Fonte: própria. Extraído do software online TagCrowd.

Figura 2- Mapa de frequência de palavras chaves

Ao analisar a figura, que contém as 30 principais palavras-chave, é possível perceber que “*data*” e “*analytics*” possuem a maior recorrência, e, dessa forma, corroboram com o foco da pesquisa. Por meio da figura, é possível perceber que um dos principais focos do *Big Data Analytics* na Engenharia de Produção é a cadeia de suprimentos, fato comprovado pela recorrência das palavras “*suply*” e “*chain*”. Além disso, a palavra “*management*” também é constante e ressalta a possibilidade de aplicação do *Big Data Analytics* no referente ao gerenciamento. A palavra “*systems*” referente aos sistemas utilizados no *Big Data* também possui grande assiduidade .

### 3.2 Mapas de calor

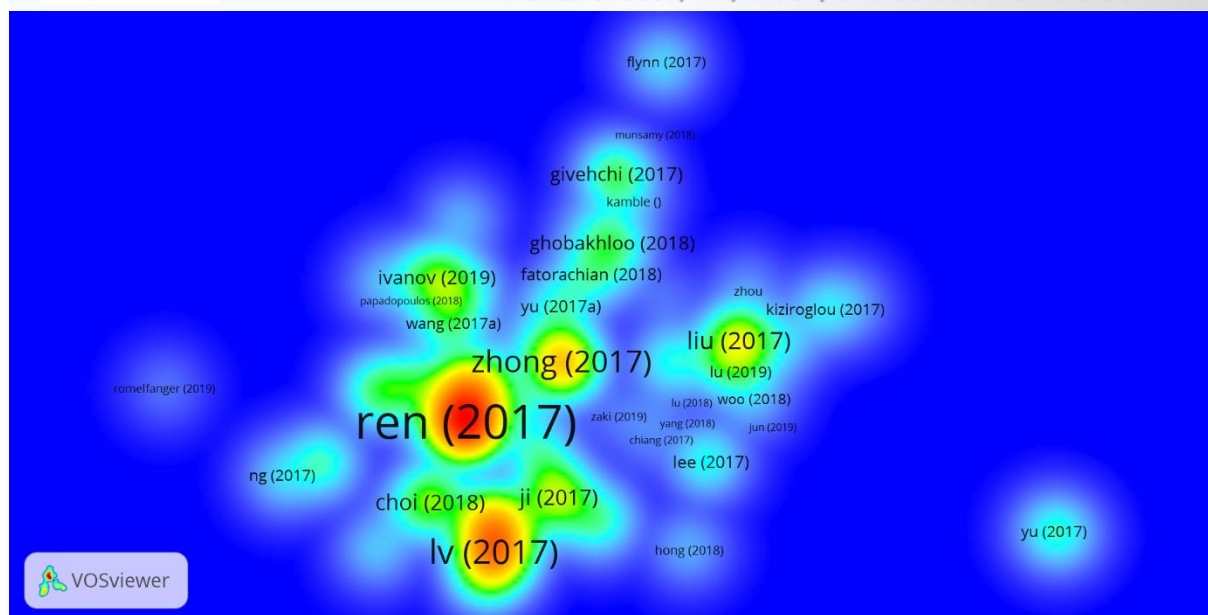
Com o intuito de ampliar o conhecimento acerca do tema da pesquisa, foram extraídos dois mapas de calor referentes a *co-citation*, *bibliographic coupling* , utilizando o software VOSviewer. A análise de *co-citation* permite identificar quais são as referências que as publicações encontradas na pesquisa usam em comum. A partir disso, é possível identificar as principais abordagens que o tema utiliza. Já o *bibliographic coupling* revela os fronts de pesquisa, ou seja, como os estudos estão se configurando e quais são as tendências do tema.



Fonte: própria. Extraído do software software VosViewer.

Figura 3- Mapa de calor de co-citações.

A figura 3 é um mapa de calor que possibilita a análise das co-citações. Segundo Grácio(2016) essa análise permite reconhecer através da frequência de ocorrência simultânea a ligação ou a afinidade de dois documentos citados. Tendo como referência a figura é possível notar 3 grandes grupos de co-citações, o primeiro com o trabalho dos autores Lee Jay (2015), Monostori (2016) e Lasi (2014), nos quais abordam Sistemas ciber-físicos na fabricação dentro da Indústria 4.0. O segundo grupo inclui os trabalhos Dutta (2015) e Wang (2016), os documentos tratam da análise de *Big Data* para o gerenciamento com intuito de aprimorar a tomada de decisões. Por último, o grupo composto por Wamba (2015) e Lavalle (2011), os quais discorrem acerca do impacto que o *Big Data* pode causar. Essas trabalhos são citados juntos pois, de acordo com Smiraglia (2011), eles possuem uma linha de abordagem similar.



Fonte: própria. Extraído do software software VosViewer.

Figura 4- Mapa de calor do acoplamento bibliográfico.

De acordo com Grácio (2016), o acoplamento bibliográfico acontece quando dois trabalhos citam pelo menos uma publicação em comum. Por meio da análise do coupling, é possível perceber como os estudos estão se configurando no momento. A figura 4 retrata uma análise coupling de 2017 até 2019, utilizando um mapa de calor. Com a figura, fica nítido a existência de 3 grandes frentes de estudos. A primeira com o trabalho de Ji-fan Ren e Steven (2017), que trata acerca da dinâmica da qualidade dentro Big Data relacionada ao crescimento do valor e desempenho da empresa. O segundo front é concentrado em Lv (2017), que possui o foco em pesquisas na área de Big Data com o intuito de prever tendências atuais e futuras. E no terceiro núcleo de pesquisa está Zhong(2017) que aborda a análise de Big Data para fábricas inteligentes.

Pode-se perceber assim, que por meio da análise de co-citation foi possível chegar aos resultados das principais abordagens e por meio da análise de coupling pode-se encontrar os principais fronts.

#### 4. Conclusão

O problema da pesquisa era responder: quais as principais abordagens a respeito do big data? Quais os principais fronts de pesquisa?

A análise de co-titation revelou três principais abordagens, uma aplicada aos sistemas ciber-físicos na fabricação dentro da Indústria 4.0, uma segunda relacionada ao Big Data para o gerenciamento com intuito de aprimorar a tomada de decisões e finalmente uma terceira, que aborda o impacto que o Big Data pode causar. Já os principais fronts de pesquisa foram descoberto via o coupling. Um primeiro front revela o uso do Big Data relacionada ao crescimento do valor e desempenho da empresa, mostrando seu papel estratégico. Um segundo front está centrado em prever tendências atuais e futuras, mostrando o interesse da pesquisa em modelos preditivos e prescritivos e o uso do Big Data para fábricas inteligentes,



que aparece como terceiro front. Deste modo pode-se observar que em uma perspectiva macro ou específica o Big data é um tema em crescimento e ligado ao contexto da Indústria 4.0. Para futuras linhas de pesquisa aconselha-se o estudo via entrevistas para compreender a visão dos profissionais.

## Referências

### DATA, IBM What Is Big. Bring Big Data to the Enterprise. 2012.

FRISK, Jane Elisabeth; BANNISTER, Frank. Improving the use of analytics and big data by changing the decision-making culture: a design approach. **Management Decision**, v. 55, n. 10, p. 2074-2088, 2017.

MARIANO, A. M.; ROCHA, M. S..Revisão da Literatura: Apresentação de uma Abordagem.

MCAFEE, Andrew et al. Big data: the management revolution. **Harvard business review**, v. 90, n. 10, p. 60-68, 2012.

Integradora. In:Anais XXVI Congreso Internacional AEDEM | 2017 AEDEM International Conference - Economy, Business and Uncertainty: ideas for a European and Mediterranean industrial policy?ISBN: 978-84-697-5592-1. Reggio Calabria- Italia. 2017.

WU, Xindong et al. Data mining with big data. **IEEE transactions on knowledge and data engineering**, v. 26, n. 1, p. 97-107, 2013.

RUSSOM, Philip et al. Big data analytics. **TDWI best practices report, fourth quarter**, v. 19, n. 4, p. 1-34, 2011.

KOHAVI, Ron; ROTHLEDER, Neal J.; SIMOUDIS, Evangelos. Emerging trends in business analytics. **Communications of the ACM**, v. 45, n. 8, p. 45-48, 2002.

ARDITO, Lorenzo et al. A bibliometric analysis of research on big data analytics for business and management. **Management Decision**, 2018.

LEVELING, Jens; EDELBROCK, Matthias; OTTO, Boris. Big data analytics for supply chain management. In: **2014 IEEE International Conference on Industrial Engineering and Engineering Management**. IEEE, 2014. p. 918-922.

DA XU, Li; HE, Wu; LI, Shancang. Internet of things in industries: A survey. **IEEE Transactions on industrial informatics**, v. 10, n. 4, p. 2233-2243, 2014.

WANG, Gang et al. Big data analytics in logistics and supply chain management: Certain investigations for research and applications. **International Journal of Production Economics**, v. 176, p. 98-110, 2016.

JI-FAN REN, Steven et al. Modelling quality dynamics, business value and firm performance in a big data analytics environment. **International Journal of Production Research**, v. 55, n. 17, p. 5011-5026, 2017.

AKTER, Shahriar et al. How to improve firm performance using big data analytics capability and business strategy alignment?. **International Journal of Production Economics**, v. 182, p. 113-131, 2016.

ZHONG, Ray Y. et al. Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. **Computers & Industrial Engineering**, v. 101, p. 572-591, 2016.

Ji, Wei; WANG, Lihui. Big data analytics based fault prediction for shop floor scheduling. **Journal of Manufacturing Systems**, v. 43, p. 187-194, 2017.

GRÁCIO, Maria Cláudia Cabrini. A coplamente bibliográfico e análise de cocitação: revisão teórico-conceitual. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 21, n. 47, p. 82-99, 2016.

LV, Zhihan et al. Next-generation big data analytics: State of the art, challenges, and future research topics. **IEEE Transactions on Industrial Informatics**, v. 13, n. 4, p. 1891-1899, 2017.

ZHONG, Ray Y. et al. Big data analytics for physical internet-based intelligent manufacturing shop floors. **International journal of production research**, v. 55, n. 9, p. 2610-2621, 2017.

Lasi, H., Fettke, P., Kemper, H.G., Feld, T. and Hoffmann, M. (2014) Industry 4.0. **Business & Information Systems Engineering**, 6, 239-242.

MONOSTORI, László et al. Cyber-physical systems in manufacturing. **Cirp Annals**, v. 65, n. 2, p. 621-641, 2016.

LEE, Jay; BAGHERI, Behrad; KAO, Hung-An. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. **Manufacturing letters**, v. 3, p. 18-23, 2015.

DUTTA, Debprotim; BOSE, Indranil. Managing a big data project: the case of ramco cements limited. **International Journal of Production Economics**, v. 165, p. 293-306, 2015.

WAMBA, Samuel Fosso et al. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. **International Journal of Production Economics**, v. 165, p. 234-246, 2015.

LAVALLE, Steve et al. Big data, analytics and the path from insights to value. **MIT sloan management review**, v. 52, n. 2, p. 21-32, 2011.