



ConBRepro

XI CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO



01 a 03
de dezembro 2021

O impacto da Indústria 4.0 e do Big Data em programas de melhoria Six Sigma

Gabriela Dall'Agnol

Programa de Pós Graduação em Engenharia de Produção - UFSCar

José Carlos de Toledo

Programa de Pós Graduação em Engenharia de Produção - UFSCar

Roberto Fernandes Tavares Neto

Programa de Pós Graduação em Engenharia de Produção - UFSCar

Resumo: Não é novidade que as empresas buscam diariamente melhorar seus processos e produtos para garantir sua permanência em um mundo globalizado e competitivo, lançando mão de programas de melhoria mundialmente famosos e consolidados. Com o avanço tecnológico inserido nas organizações a partir do advento da Indústria 4.0, tais programas, como o Seis Sigma (SS) ou o Lean Seis Sigma (LSS), enfrentam o desafio de extrair informação útil de grandes e complexos conjuntos de dados, chamados de Big Data (BD), que podem proporcionar insights valiosos, mas dependem de ferramentas analíticas muito mais desafiadoras. O presente estudo desenvolveu uma revisão de literatura, contendo 10 artigos recentes sobre a ligação de programas SS ao advento do BD. Através da leitura e discussão de trabalhos com diferentes metodologias e autores, foi possível evidenciar a importância e os desafios enfrentados neste cenário. Além disso, foi possível identificar em qual fase da melhoria é mais comum o uso de técnicas para lidar com BD, quais são as principais ferramentas de análise de dados utilizadas e quais habilidades são esperadas do profissional da área. Estes achados podem nortear pesquisas futuras na busca de tendências e lacunas a serem preenchidas, buscando o aperfeiçoamento da implementação de projetos SS, em um contexto de BD.

Palavras-chave: Big Data, Indústria 4.0, Seis Sigma, Lean Six Sigma, DMAIC.

The impact of Industry 4.0 and Big Data on Six Sigma improvement programs

Abstract: It is not new that companies seek daily to improve their processes and products to ensure their permanence in a globalized and competitive world, making use of world-famous and consolidated improvement programs. With the technological advance inserted in organizations since the advent of Industry 4.0, such programs as Six Sigma (SS) or Lean Six Sigma (LSS), face the challenge of extracting useful information from large and complex data sets, called Big Data (BD), which can provide valuable insights but rely on much more challenging analytics tools. The present study developed a literature review, containing 10 recent articles on the connection of SS programs to the advent of BD. Through the reading and discussion of works with different methodologies and authors, it was possible to highlight the importance and challenges faced in this scenario. In addition, it was possible to identify in which phase of improvement the use of techniques to deal with BD is

more common, which are the main data analysis tools used and which skills are expected from the professional in the area. These findings can guide future research in the search for trends and gaps to be filled, seeking to improve the implementation of SS projects, in a BD context.

Keywords: Big Data, Industry 4.0, Six Sigma, Lean Six Sigma, DMAIC.

1. Introdução

A inserção de novas tecnologias integradas a ambientes produtivos está transformando a indústria tradicional, proporcionando melhores resultados competitivos (SANTOS et al., 2018). Em meio a esse cenário de avanços tecnológicos, encontra-se a chamada Indústria 4.0 (I4.0), que é considerada por muitos como o quarto estágio da revolução industrial e baseia-se na implementação de fábricas, produtos e serviços inteligentes, que trocam informações de forma autônoma (STOCK; SELIGER, 2016).

A estrutura da I4.0 é viabilizada pela adoção da Internet das Coisas, Internet of Things (IoT), que conecta todos esses dispositivos a uma rede de internet e faz a troca de informações acontecer em tempo real, o que leva à construção de enormes bancos de dados com alta heterogeneidade, contendo números inteiros, caracteres, áudios, imagens e vídeos (CAI et al., 2016), conhecidos como *Big Data*, especialmente por conter essa característica heterogênea e de grande volume.

Contudo, impulsionados pela Indústria 4.0, os processos de manufatura estão gerando uma avalanche de dados, cada vez mais complexos e heterogêneos (BELHADI et al., 2019; KOPEL; CHANG, 2020). Esses autores definem o termo *Big Data* (BD) como um conjunto de dados atípicos, sob os quais não são aplicáveis as ferramentas tradicionais de captura, armazenamento, gerenciamento e análise estatística, devido à sua grande complexidade, que é caracterizada pelos cinco V's: volume, velocidade, variedade, veracidade, variabilidade e volatilidade.

Uma das principais dificuldades em lidar com o BD nas empresas é o desconhecimento das técnicas adequadas para analisar essa grande quantidade de dados, o que pode levar perca de informações valiosas. Segundo Arcidiacono et al. (2016), mesmo com o aumento significativo de dados fornecidos pela adoção da Internet of Things (IoT) e tendo BDA (*Big Data Analytics*) como ferramental para a extração e identificação de informações relevantes, apenas 10% do potencial dos dados gerados são realmente utilizados para melhorar o nível da produtividade de gestão em sistemas de manufatura.

Com o crescente avanço da tecnologia em ambientes de manufatura e o surgimento do *Big Data*, os programas de melhoria da qualidade encontram um cenário desafiador para desempenhar seu papel na busca de novas soluções e oportunidades de crescimento nas empresas (AMINE et al., 2019). Segundo Dogan e Gurcan (2018) a maioria dos métodos de melhoria de qualidade requerem coleta e análise de dados para solucionar problemas. Alcançar e, sobretudo, manter altos padrões de qualidade requer constante monitoramento do desempenho dos sistemas produtivos, que geram grandes volumes de dados, necessitando de análises estatísticas baseadas em técnicas mais complexas do que as usualmente empregadas pela área de gestão de qualidade.

Ao longo dos anos, muitos métodos foram propostos para aperfeiçoar a performance das empresas com relação aos 5 objetivos de desempenho: qualidade, rapidez, confiabilidade, flexibilidade e custos (SLACK et al., 1999). Para os principais autores de Administração de Produção, a gestão de qualidade é um tópico extensivamente discutido e que possui diversos aspectos e diferentes linhas de pesquisa (CHASE; AQUILANO; JACOBS, 2006; SLACK et al., 1999). Uma das mais estudadas é a metodologia de melhoria Seis Sigma (SS), inicialmente proposta pelo engenheiro da Motorola, Bill Smith em 1986, para melhorar a competitividade no mercado internacional.

Tradicionalmente, o SS usa uma abordagem baseada em dados buscando a redução da variabilidade dos processos, utilizando controle estatístico de processos e análise de capacidade do processo para a redução de erros (SHAH; CHANDRASEKARAN; LINDERMAN, 2008; KOPEL; CHANG, 2020). O *Lean Six Sigma* (LSS) une a abordagem SS com a cultura de eliminar desperdícios da Manufatura Enxuta, do inglês *Lean Manufacturing*, identificando e eliminando atividades do processo que não adicionam valor ao produto final, também tem grande popularidade entre pesquisadores e profissionais da área de qualidade (DOGAN; GURCAN, 2018). Apesar de programas de SS serem amplamente reconhecidos pelo mundo como uma estratégia de sucesso, capaz de produzir efeitos positivos na melhoria de processos e produtos, reduzindo custos e aumentando os lucros, estima-se que cerca de 62% das iniciativas SS falham em seus objetivos após um período de implementação bem-sucedida (ANTONY; SONY; GUTIERREZ, 2020).

Internamente, o SS e o LSS carregam os ciclos DMAIC/DMADV, reconhecidos como peça fundamental para suas aplicações, atuando como um dos mais eficientes métodos estruturados para a solução de problemas já existentes, com o objetivo de encontrar futuras oportunidades de melhoria ou gerenciar projetos (DOGAN; GURCAN, 2018).

O DMAIC faz referência aos cinco passos de execução para a resolução estruturada de problemas com foco na melhoria contínua, que são: “Definir” (*Define*), “Medir” (*Measure*), “Analisar” (*Analyse*), “Melhorar” (*Improve*). Já o DMADV é utilizado em alguns casos para a criação de novos projetos, utilizando o mesmo ciclo DMA (Definir, Medir e Analisar) mas acrescentando os passos “Projetar” (*Design*) e “Verificar” (*Verify*). Independente da escolha da abordagem DMAIC ou DMADV, a coleta e análise de dados tem papel fundamental em todas as etapas do processo (GEORGE; ROWLANDS; KASTLE, 2004).

Para a realização do DMAIC/DMADV há a necessidade de coleta de dados buscando uma base sólida para tomada de decisão, tarefa que tradicionalmente consistia na construção de um banco de dados com cerca de 30 a 1000 observações, o que era considerado suficiente para descrever a performance de um processo. Neste cenário, poucas variáveis seriam facilmente analisadas através de ferramentas estatísticas tradicionais, como regressão linear, análise de variância e teste T de *Student* (ZWETSLOOT et al., 2018).

Ainda que métodos como DMAIC ou DMADV sejam empregados com sucesso há décadas em diversos setores da indústria e serviços, as empresas podem encontrar barreiras ao ter que aplicar essas ferramentas a conjuntos de dados grandes e complexos como o Big Data, gerado pela era da I4.0. É inevitável uma adaptação técnica e estrutural da organização para lidar de maneira mais eficiente com esses desafios.

Dessa forma, para garantir a implementação e a vitalidade de projetos SS nas empresas, se faz necessário entender as tendências emergentes relacionadas a mudanças tecnológicas ligadas ao ambiente produtivo. Adaptar-se a essa nova realidade do Big Data contribui de maneira significativa para que os programas de melhoria que usam dados como suporte à decisão, como em ciclos DMAIC dentro do SS, continuem exercendo seu papel no aumento da performance das organizações.

Este trabalho tem como objetivo explorar as relações entre os conceitos apresentados, assim como responder às seguintes perguntas de pesquisa:

P1: Como ferramentas de Big Data são aplicadas dentro de programas SS ou LSS?

P2: Quais técnicas de Big Data são utilizadas dentro de programas SS ou LSS?

P3: Qual é o perfil do profissional para lidar com SS ou LSS e Big Data?

Foi construído um referencial teórico sobre o assunto, discutindo os trabalhos mais relevantes que relacionam projetos SS ou LSS ao conceito de Big Data, para explorar esse

tema bastante inovador e importante para os profissionais da área de qualidade das indústrias em geral.

A seguir, a seção 2 descreve a metodologia utilizada para a busca dos artigos da revisão. A seção 3 apresenta o referencial teórico encontrado a partir dos artigos coletados. A seção 4 fornece uma discussão sobre o referencial teórico encontrado. Por fim, a seção 5 apresenta as considerações finais, próximos passos e as limitações da pesquisa.

2. Metodologia

Para a construção desta pesquisa e a fim de investigar as relações existentes entre BD e SS, uma revisão de literatura foi conduzida. Foram utilizadas as os termos “*six sigma*” + “*big data*” e a busca executada sobre a plataforma *Google Scholar*.

Com a intenção de buscar os trabalhos mais recentes desenvolvidos na área, foi definida a busca de documentos publicados entre o ano de 2015 à março de 2021. Além disso, foram considerados apenas os 20 trabalhos mais relevantes da busca, não sendo possível acessar 4 deles devido a restrições adversas.

A partir da leitura dos resumos de 16 trabalhos coletados, 6 não se encaixavam no escopo por não tratarem simultaneamente de Big Data e SS ou LSS. Portanto, 10 trabalhos foram considerados totalmente relacionados ao tema pretendido, os quais foram revisados e discutidos a fundo na construção do referencial teórico.

3. Referencial Teórico

Em sua revisão de literatura, Antony, Sony e Gutierrez (2020) avaliam que o uso de BD oferece aos profissionais novas oportunidades para resolver problemas que antes eram considerados insolúveis. Apesar do gerenciamento de grandes e complexos conjuntos de dados ser uma tarefa bem consolidada em áreas como marketing digital e pesquisa médica, ainda é pouco desenvolvida em sistemas produtivos, que poderiam usar BD para alavancar a melhoria de produtos, serviços e qualidade de processo.

Com o objetivo de identificar e avaliar as tendências emergentes da área Antony, Sony e Gutierrez (2020) realizaram uma pesquisa empírica com 307 experts globais da área de SS, concluindo que a integração entre BD e SS é a tendência mais importante considerada pelos países da Ásia, América do Sul e África, pelo setor de serviços, pelos profissionais *Black* e *Green Belts* e também aqueles com menos de cinco anos de experiência em projetos SS.

O uso da abordagem SS combinada a grandes conjuntos de dados pode fornecer vantagens competitivas às organizações. Uma nova estratégia para o SS aplicado a um ambiente de manufatura, utilizando BD, é proposto por (KOPEL; CHANG, 2020) através de um modelo de simulação. O novo método chamado MDAIC procura inverter as etapas Definir e Medir que ocorrem em um ciclo tradicional DMAIC, primeiramente coletando e monitorando todos os dados para a posterior identificação de áreas que necessitam de um maior esforço de melhoria.

Segundo os autores, o método proposto monitora os parâmetros do processo e de qualidade do produto através de um sistema de “semáforos”, mostrando de maneira mais visual em quais áreas os indicadores medidos estão longe de seus alvos, facilitando a identificação de áreas críticas para as equipes SS voltarem seus esforços. Os resultados se mostraram importantes para facilitar o monitoramento do sistema produtivo como um todo.

O tema de Curadoria Digital, abordado no artigo de Arcidiacono et al. (2016), se refere a um processo que desenvolve repositórios de longo prazo com arquivos digitais para fins de

pesquisa que atualmente as empresas estão utilizando como ferramenta para melhorar o armazenamento e gerenciamento dos dados gerados pelos seus processos operacionais e estratégicos. Os autores buscaram na literatura trabalhos existentes relacionando BD e Curadoria Digital com foco em abordagens LSS e concluíram que o LSS, por meio das fases do ciclo DMAIC, pode ser aplicado para otimizar funções como, categorização, classificação, agrupamento e digitalização de Big Cultural Data, auxiliando no ciclo de vida da Curadoria Digital e otimizando seus resultados.

Um problema comum em projetos de aprimoramento como SS é que as análises estatísticas de dados usualmente realizadas são limitadas e muitas vezes não aplicáveis em monitoramento de processos de produção modernos inseridas em um contexto de BD, que possuem centenas de variáveis importantes, como por exemplo, a fabricação de semicondutores, que gera mais de 50.000 gráficos de controle de processo para monitoramento de mais de 300 etapas de produção (STOJANOVIC; DINIC; STOJANOVIC, 2015). Uma abordagem de BDA aplicada em dados históricos de teste de máquinas de lavar, foi realizada pelos autores em uma fábrica da *Whirpool* localizada na Itália. Os autores encontraram resultados promissores para o melhoramento do controle de processos desse produto, evidenciando como o uso de BD pode fornecer suporte à filosofia SS. Afirmam, ainda, que o uso de BDA potencializa os resultados de projetos SS, sendo encarado como um caminho promissor para a nova geração de ferramentas de gestão de qualidade guiada pelo processamento e análise de grandes conjuntos de dados.

Por meio de sua revisão de literatura, Sodhi (2020) afirma que o SS pode precisar evoluir ao ter que lidar com processos cada vez mais automatizados criados pela I4.0, nos quais um número muito grande de sensores é usado em ambientes de manufatura para medir diversas características importantes do processo ou produto. O autor avalia que as técnicas estatísticas tradicionais utilizadas pelos profissionais especialistas em SS não serão capazes de lidar com esse tipo de banco de dados. Dado esse cenário, o autor reafirma a importância do uso de BDA para a criação de modelos de gestão e controle baseados técnicas de mineração de dados, como agrupamento, classificação, associação e predição, que podem aparecer em todas as fases do SS.

Em seu trabalho, Belhadi et al. (2019) realizaram uma *survey* em que foram aplicados questionários a 201 profissionais experientes de empresas localizadas no norte da África a fim de entender se as capacidades voltadas as técnicas de BDA podem influenciar positivamente projetos de LSS. Os resultados demonstram que o uso de BDA atua como agente facilitador na implementação de programas LSS em empresas de manufatura e aumenta seu desempenho em metas de melhoria no longo prazo.

Os autores Dogan e Gurcan (2018) e Gupta et al. (2019) enumeram uma série de ferramentas utilizadas pela área de BDA e como estas são dispostas dentro de cada fase dos ciclos DMAIC ou DMADV. Em seu trabalho os autores reforçam que o uso de BDA como apoio para os ciclos DMAIC e DMADV em programas SS ou LSS, contribui para a geração de resultados mais confiáveis, ajudando na identificação mais assertiva de padrões que seriam dificilmente encontrados sem uma análise de dados mais elaborada. Os autores propõem uma revisão de literatura para sumarizar as técnicas estatísticas e ferramentas da qualidade, mineração de dados, BD e mineração de processos com relação aos 7 elementos encontrados no DMAIC e no DMADV: definir; medir; analisar; melhorar; controlar; projetar e verificar. As técnicas previstas em cada fase estão dispostas na Tabela 1.

É possível notar que técnicas de BDA são usadas em quase todas as fases do ciclo DMAIC ou DMADV, com exceção da fase Medir, onde os autores não identificaram oportunidade de aplicações de Data Mining ou de BD.

Segundo Gupta et al. (2019), o desenvolvimento de habilidades de BDA pode ajudar as organizações a descobrirem importantes padrões ocultos nos dados, como tendências de mercado e preferências de clientes. Dados que são coletados desde o fornecedor até o consumidor final, chegam em grandes quantidades e diferentes formatos, como texto, números, vídeos e tweets, por exemplo, e necessitam de técnicas de BDA apropriadas. Os autores realizaram uma revisão de literatura para entender como a aplicação de BDA pode auxiliar projetos de LSS a tomar decisões mais confiáveis, propondo também uma estrutura com possíveis aplicações de técnicas de BDA de acordo com cada fase do ciclo DMAIC, inspirada no trabalho de Moeuf et al. (2018). A disposição das técnicas de BDA usadas de acordo com as fases DMAIC são apresentadas na Figura 1.

Tabela 1 – Métodos usados em ciclos de LSS

| | Estatísticas | Ferramentas de Qualidade | Mineração de Dados *** | Big Data | Mineração de Processo |
|------------------|--|---|---|---|---|
| Definir | Descritiva* | <i>Brain Storming</i> , NGT, Análise de Pareto, Diagrama de Matriz, QFD, FMEA, SIPOC, Matriz de priorização, Diagrama de Ishikawa | | Mineração de Texto, Mineração de Vídeo | Descoberta de Processo |
| Medir | Descritiva*, gráfico de contagem, teste Z, Intervalos de Confiança, Preditiva** | Análise de Pareto, Processo Sigma | | | Verificação de Conformidade |
| Analisar | Correlação, Teste T, Teste F, Teste de Hipóteses, ANOVA, Historgama, Preditiva** | SPC | Regras de associação, Agrupamento, Classificação | Aprendizado de Máquina, Árvores de Decisão, Mineração de Texto, Mineração de Vídeo, Inteligência Artificial | Descoberta de Processo, Verificação de Conformidade |
| Melhorar | Teste de Hipóteses, Análise Multivariada | TRIZ, DOE | Predição | Aprendizado de Máquina, Inteligência Artificial | Diagramas de Fluxo, Aprimoramento |
| Controlar | | FMEA, Diagrama de Controle, Padronização, SPC | | Aprendizado de Máquina, Inteligência Artificial | Diagramas de Fluxo, Verificação de Conformidade |
| Projetar | Descritiva*, Preditiva | QDF, DOE | Análise de Cesta de Compras, Regras de Associação | | |
| Verificar | Correlação, Causalidade | | | Aprendizado de Máquina, Inteligência Artificial | Gráficos, Visualização |

* Descritiva: inclui alguns métodos estatísticos como média, frequência e desvio padrão.

** Preditiva: inclui alguns métodos estatísticos como testes de hipótese, análise de variância e regressão.

*** Note que muitos métodos de mineração de dados são usados por *Big Data* com algoritmos avançados.

ANOVA: Análise de Variância

DOE: *Design of Experiments* (Planejamento de Experimentos)

FMEA: *Failure Mode and Effects Analysis* (Modo de Falha e Análise de Efeitos)

NGT: *Nominal Group Technique* (Técnica de Grupo Nominal)

QFD: *Quality Function Deployment* (Desdobramento da Função Qualidade)

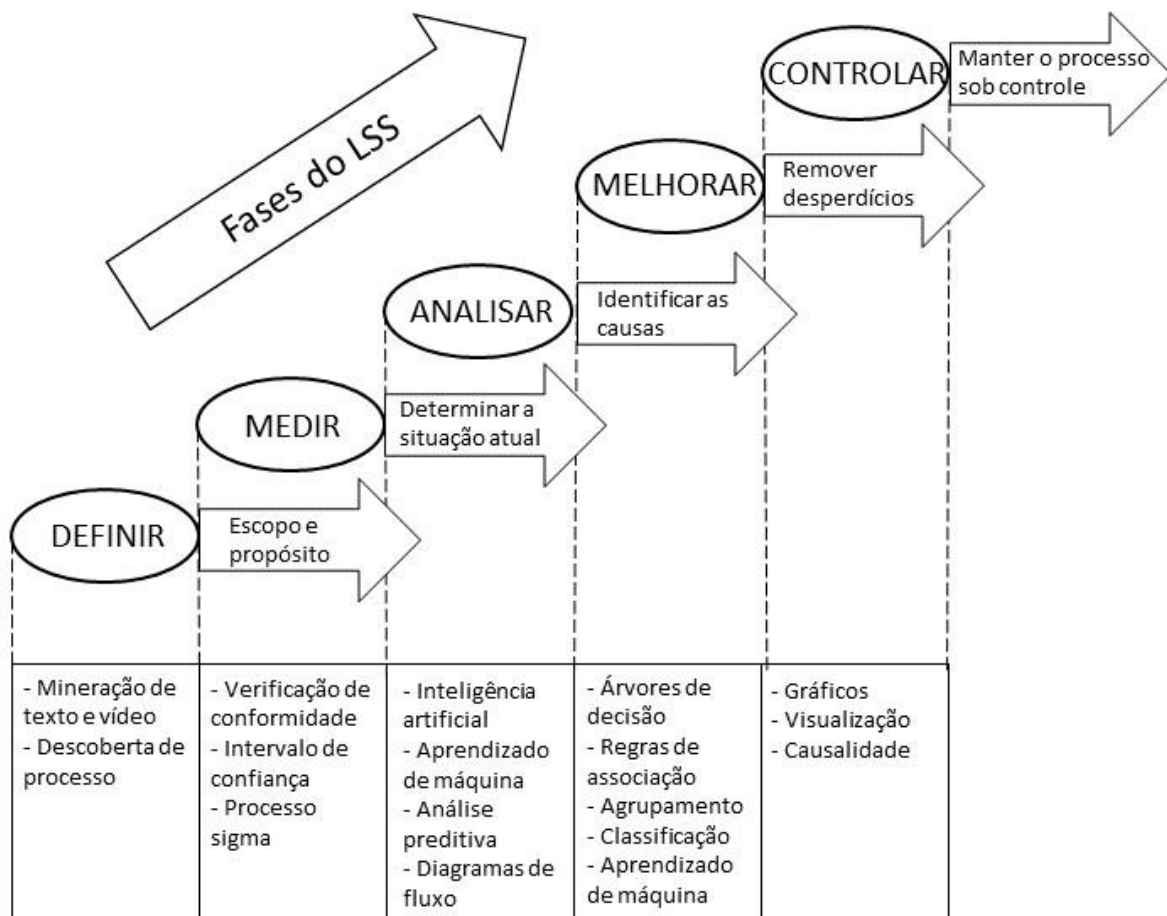
SIPOC: *Supplier, Inputs, Process, Outputs, Customers* (Fornecedor, Entrada, Processo, Saída, Consumidores)

SPC: *Statistical Process Control* (Controle Estatístico de Processo)

TRIZ: *Theory of Inventive Problem Solving* (Teoria da Solução Inventiva de Problemas)

Fonte: Adaptado de Dogan e Gurcan (2018)

Figura 1 – Aplicação de BDA em diferentes fases do LSS



Fonte: Adaptado de Gupta et al. (2019)

Segundo os autores, as fases Definir, Analisar e Melhorar são as que oferecem maiores oportunidades de aplicação de técnicas de BDA, como mineração de texto e vídeo, classificação, associação, entre outras, auxiliando o LSS nas tarefas de determinar o escopo do projeto, avaliar a atual situação do processo, identificar as possíveis causas de defeitos no sistema e remover desperdícios para obter resultados melhores.

Zwetsloot et al. (2018) apresentaram em seu artigo um estudo de caso qualitativo para responder quais modificações seriam necessárias em uma empresa para que a metodologia LSS seja compatível com projetos de BD. Segundo os autores, uma das principais mudanças que devem ser feitas é relacionada à composição dos profissionais envolvidos nos projetos de melhoria, que devem conter colaboradores com dois perfis distintos, mas trabalhando em conjunto: o especialista em SS (um Black Belt, por exemplo) que faz algum treinamento para ter conhecimento da área de dados, e o Cientista de Dados, que é altamente capacitado para lidar com análise de grandes conjuntos de dados, mas que possua treinamento relacionado a ferramentas de melhoria contínua como SS. Os autores chamam essa união de habilidades como “Data Science Black Belt”, afirmando que ela pode ser mais eficaz na liderança de um projeto LSS baseado em BD.

A Tabela 2 exibe todos os artigos analisados por este trabalho, os classificando de acordo com a metodologia utilizada, em qual programa de melhoria está focado e verificando se o trabalho aborda os temas DMAIC/DMADV e BD/BDA.

Tabela 2 – Trabalhos avaliados pela Revisão de Literatura

| | Ano | Autor | Metodologia | LSS ou SS | DMAIC/DMADV | BDA/BD |
|----|------|--------------------------|-----------------------|-----------|-------------|--------|
| 1 | 2015 | Stojanovic et al. | Estudo de Caso | LSS | Não | Sim |
| 2 | 2016 | Arcidiacono et al. | Revisão de Literatura | LSS | Sim | Sim |
| 3 | 2018 | Zwetsloot et al. | Estudo de Caso | LSS | Sim | Sim |
| 4 | 2018 | Dogan, O.; Gurcan, O. F. | Revisão de Literatura | LSS | Sim | Sim |
| 5 | 2019 | Gupta et al. | Revisão de Literatura | LSS | Sim | Sim |
| 6 | 2020 | Belhadi et al. | Survey | LSS | Sim | Sim |
| 7 | 2020 | Koppel, S.; Chang, S. | Modelagem e Simulação | SS | Sim | Sim |
| 8 | 2020 | Antony et al. | Survey | SS | Não | Sim |
| 9 | 2020 | Sodhi, H. S. | Revisão de Literatura | LSS | Não | Sim |
| 10 | 2020 | Antony et al. | Revisão de Literatura | LSS | Não | Sim |

Fonte: Elaborado pela autora

É possível notar que grande parte dos autores realizam *surveys* e revisões de literatura, demonstrando o caráter exploratório das pesquisas que relacionam BD e SS/LSS. Apesar de diversos autores evidenciarem em seus trabalhos a importância de saber lidar de maneira assertiva com BD e como essa atitude pode desempenhar um papel significativo no sucesso de programas de melhoria como o SS e LSS, o único caso de modelagem prática foi encontrado no trabalho de Kopel e Chang (2020). Essa constatação pode validar a ideia de Antony, Sony e Gutierrez (2020), que indica a pouca maturidade de gerenciamento de BD em ambientes de manufatura se comparados com outras áreas.

Uma das principais motivações apontados por Kopel e Chang (2020), Stojanovic, Dinic e Stojanovic (2015) e Sodhi (2020) para que habilidades de BDA sejam aprimoradas na indústria é a falta de compatibilidade que as técnicas estatísticas tradicionais para extrair conhecimento útil de conjuntos de dados muito grandes e complexos. Sem o uso de técnicas de análise de dados mais elaboradas, dificilmente seria possível extrair padrões assertivos de BD para orientar a tomada de decisão dos programas de melhoria (DOGAN; GURCAN, 2018; GUPTA; MODGIL; GUNASEKARAN, 2019).

Outro ponto a ser considerado a partir da leitura dos artigos é a grande variedade de técnicas de BDA que pode ser aplicada em todas as fases dos ciclos DMAIC ou DMADV (DOGAN; GURCAN, 2018; GUPTA; MODGIL; GUNASEKARAN, 2019; ARCIDIACONO et al., 2016). Porém, para corresponder à dificuldade de implementação de tais técnicas e a sua integração com os programas SS e LSS, a capacitação dos colaboradores e da empresa é fundamental (ZWETSLOOT et al., 2018).

Como o esperado para um tema que ainda está em ascensão na literatura e com poucos trabalhos retratando sua aplicabilidade na vida real, a definição de quais habilidades técnicas seriam exigidas dos profissionais da área ainda é pouco abordada. Apenas o trabalho de Zwetsloot et al. (2018) tratou do assunto, no qual os autores afirmam que a habilidade principal esperada de um profissional que lida com BD em programas SS ou LSS seria a chamada "*Data Science Black Belt*", pessoas altamente capacitadas em cada uma das duas áreas de atuação.

A partir da análise dos trabalhos presentes na revisão de literatura, os apontamentos levantados serviram como base para atender às perguntas de pesquisa.

Com relação à pergunta P1, os trabalhos mostraram que o BD é encontrado no interior dos ciclos DMAIC ou DMADV, por se tratarem de abordagens que lidam com ferramentas para análise de dados.

As técnicas citadas pelos autores que são implementadas nos programas de melhoria SS ou LSS para lidar com BD são: Regras de Associação, Agrupamento, Classificação, Mineração de Texto, Mineração de Vídeo, Aprendizado de Máquina, Árvores de Decisão, Inteligência Artificial, Análise Preditiva, respondendo à questão de pesquisa P2.

Para responder à pergunta P3, a maioria dos autores menciona ser necessário um treinamento mais adequado às análises de dados avançadas, pois as técnicas tradicionais possivelmente aprendidas pelos profissionais de SS ou LSS, não seriam adequadas para lidar com BD. Portanto, o perfil ideal seria de um expert em programas de melhoria que ao mesmo tempo seja expert em análise de dados. Possivelmente, a solução seria ter uma equipe multidisciplinar, que contenha os dois perfis colaborando entre si.

4. Considerações Finais e Próximos Passos

Com o advento da I4.0 e o surgimento do BD, se faz necessário compreender quais as técnicas apropriadas para lidar com esses dados tão complexos e como elas se relacionam com as metodologias do SS/LSS, dada a importância que os programas de melhoria de processos e produtos desempenham no sucesso competitivo de qualquer ambiente de produção.

Pode-se dizer que a análise de dados popularmente feita pelos especialistas de SS dentro de programas de melhoria é baseada em regressão, análise de variância, correlação, técnicas estatísticas empregadas em conjuntos de dados tradicionalmente encontrados na prática ou academia, com poucas variáveis numéricas.

Para enfrentar os desafios da I4.0, será necessário extrair informações e tomar decisões baseadas em dados muito mais complexos, com múltiplas variáveis de características distintas. O BD traz consigo muitos desafios técnicos para as organizações, mas em contrapartida fornece uma oportunidade de identificação de padrões que não seriam enxergados em pequenos conjuntos de dados de monitoramento.

O surgimento da necessidade de lidar com Big Data pode acabar impulsionando algumas empresas a desenvolverem maiores e melhores capacidades de BDA. Esta pesquisa buscou encontrar respostas a questionamentos que seriam comuns a quaisquer gestores que queiram se adaptar a este cenário.

A revisão de literatura proposta concedeu um panorama geral sobre a união de BD com SS e conseguiu elencar conceitos e tendências importantes para compreender a relevância de extrair informação útil de grandes e complexos conjuntos de dados para um programa de melhoria. Como próximos passos da pesquisa, a busca de trabalhos poderia ser realizada em múltiplas bases de dados, como Web of Science e Scopus. Além disso, uma metodologia mais robusta de revisão sistemática poderia ser adotada, incluindo mais artigos e avaliando outros aspectos, como dados bibliométricos.

Apesar de encontrar diversos trabalhos relevantes sobre o tema, o caráter exploratório das pesquisas avaliadas demonstra que a área ainda está em constante expansão e aprimoramento. Pesquisas futuras poderiam explorar casos práticos de aplicações de técnicas de Big Data no mundo real ou simulado, para que as dificuldades da prática sejam também analisadas.

Referências

ANTONY, J.; HOERL, R. SNEE; R. An Overview of Lean Six Sigma. **Lean Six Sigma in Higher Education**. s.l.:Emerald Publishing Limited, p. 1-11, 2020.

ANTONY, J.; SONY, M. GUTIERREZ; L. An Empirical Study Into the Limitations and Emerging Trends of Six Sigma: Findings From a Global Survey. **IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT**, p. 1-14, 2020.

ARCIDIACONO, G.; DE LUCA, E. W.; FALLUCCHI, F.; ALESSANDRA, P. THE USE OF LEAN SIX SIGMA METHODOLOGY IN DIGITAL CURATION. **Germany, 1st Workshop on Digital Humanities and Digital Curation**, 2016.

BELHADI, A.; KAMBLE, S. S.; ZKIK, K.; CHERRAFI, A.; TOURIKI, F. E. The integrated effect of Big Data Analytics, Lean Six Sigma and Green Manufacturing on the environmental performance of manufacturing companies: The case of North Africa. **Journal of Cleaner Production**, v. 252, 2019.

CAI, H.; XU, B.; JIANG, L.; VASILAKOS, A. V. IoT-Based Big Data Storage Systems in Cloud Computing: Perspectives and Challenges. **IEEE Internet of Things Journal**, 4(1), p. 75-87, 2016.

CHASE, R. B.; AQUILANO, N. J.; JACOBS, F. R. **Administração da Produção Para Vantagem Competitiva**. Porto Alegre: Bookman, 2006.

DOGAN, O.; GURCAN, O. F. Data Perspective of Lean Six Sigma in Industry 4.0 Era: A Guide To Improve Quality. **Proceedings of the International Conference on Industrial Engineering and Operations Management**, July, 2018.

GEORGE, M., ROWLANDS, D. & KASTLE, B., 2004. **What is Lean Six Sigma**. New York: McGraw-Hill.

GUPTA, S.; MODGIL, S.; GUNASEKARAN, A. Big data in lean six sigma: a review and further research directions. **International Journal of Production Research**, p. 947-969, 2019.

KOPEL, S.; CHANG, S. MDAIC – a Six Sigma implementation strategy in big data environments. **International Journal of Lean Six Sigma**, v. 12 n. 2, p. 432-449, 2020.

LEE, J., BAGHERI, B.; KAO, H.-A. Recent advances and trends of cyber-physical systems and big data analytics in industrial informatics. Porto Alegre, Brazil, **IEEE**, p. 1-6, , 2014.

MOEUF, A.; PELLERIN, F.; LAMOURI, S.; TAMAYO-GIRALDO, S.; BARBARAY, R. The Industrial Management of SMEs in the era of Industry 4.0. **International Journal of Production Research**, p. 1118-1136, 2018.

SANTOS, B. P.; ALBERTO, A.; LIMA, T.; CHARRUA-SANTOS, F. INDÚSTRIA 4.0: DESAFIOS E OPORTUNIDADES. **Revista Produção e Desenvolvimento**, 4(1), p. 111-124, 2018.

SLACK, N.; CHAMBERS, S.; HARLAND, C.; HARRISON, A.; JOHNSTON, R. **Administração da Produção**. São Paulo: Atlas, 1999.

SODHI, H. S. When Industry 4.0 meets Lean Six Sigma: A review. **Industrial Engineering Journal**, v. 13, p. 1-12, 2020.

STOCK, T.; SELIGER, G. Opportunities of Sustainable Manufacturing in Industry 4.0. **13th Global Conference on Sustainable Manufacturing - Decoupling Growth from Resource Use**, p. 536-541, 2016.

STOJANOVIC, N.; DINIC, M.; STOJANOVIC, L. Big Data Process Analytics for Continuous Process Improvement in Manufacturing. **IEEE International Conference on Big Data**, p. 1398-1407, 2015.

ZWETSLOOT, I. M.; KUIPER, A.; AKKERHUIS, T. S.; KONING, H. Lean Six Sigma meets data science: Integrating two approaches based on three case-studies. **Quality Engineering**, 1-13, 2018.