

Abstract

A fabricação industrial está atualmente no meio de uma revolução baseada em dados, a qual promete transformar fábricas tradicionais em fábricas “smart”, altamente otimizadas. Essas fábricas estão focadas em criar processos de fabricação inteligentes baseados em dados em tempo real para que as tomadas de decisão sejam mais rápidas e precisas causando um impacto positivo na organização como um todo. Para implementar estas tecnologias emergentes como a Internet das Coisas (IoT) e Cyber Physical Systems (CPS) serão incorporados fisicamente em processos para medir e monitorar dados em tempo real os dados de toda a fábrica, os quais darão origem a níveis de dados de produção sem precedentes. Assim sendo, instalações de produção devem ser capazes de gerenciar as exigências de aumento da produção de dados, bem como possuir as técnicas analíticas necessárias para extrair significado desses grandes conjuntos de dados. Mais especificamente, as organizações devem ser capazes de trabalhar com as tecnologias de big data para atender à demanda de maneira mais efetiva. Contudo, como big data é um fenômeno relativamente novo e o potencial das aplicações industriais são diversificadas e amplas, existe uma falta óbvia de pesquisa secundária realizada na área. Sem uma investigação secundária, é difícil identificar as lacunas das áreas, bem como alinhar com outras áreas para desenvolver uma pesquisa consistente. Dentro deste estudo, utilizamos a metodologia de pesquisa de mapeamento formal que fornece uma ampla visão das tecnologias de big data no setor de manufatura.

Introdução

Instalações modernas de produção são ambientes ricos em dados que suportam a transmissão, compartilhamento e análise de informações através de redes difusas para produzir Inteligência. Os benefícios potenciais da inteligência na fabricação incluem melhorias na eficiência operacional, inovação de processos e no impacto ambiental, citando apenas alguns. No entanto, semelhante a outras indústrias e domínios, as informações atuais sistemas que suportam informações comerciais e de fabricação estão sendo encarregados a responsabilidade de armazenar conjuntos de dados cada vez maiores (Big Data), bem como o processamento em tempo real deste "Big Data" usando análises avançadas. O crescimento exponencial previsto na produção de dados será resultado de um aumento do número de equipamentos que registram medidas de ambientes físicos e processos, bem como um aumento na frequência com que estes dispositivos registram e gravam essas medições. As tecnologias que transmitem esses dados brutos incluirão automação e redes de sensores, em acréscimo de paradigmas novos e emergentes, tais como a Internet das Coisas (IoT) e Cyber Physical Systems (CPS). Os Dados de baixo nível capturados por essas tecnologias podem ser processados por aplicações de análise de dados e modelar aplicações que permitam aos fabricantes compreender melhor as suas atividades e processos para obter idéias que possam melhorar as operações existentes.

O foco em tecnologias de big data em ambientes de manufatura é uma nova área de pesquisa interdisciplinar que incorpora automação, engenharia, tecnologia da informação e análise de dados, para citar alguns. Neste momento, é importante compreender o estado atual da pesquisa referente às tecnologias de big data no setor de manufatura, e identificar áreas onde pesquisas futuras e esforços devem ser focados para apoiar, futuramente, a infra-estrutura e tecnologias para fabricação. Portanto, este estudo tem por objetivo classificar pesquisas, derivar temas de pesquisa proeminentes e identificar lacunas atuais.

Este estudo emprega o método de pesquisa secundário bem conhecido e formal de mapeamento sistemático para captar as linhas de pesquisa amplas e diversificadas atualmente relacionadas com tecnologias big data na manufatura. A contribuição deste estudo é um relatório abrangente sobre o estado atual das tecnologias big data na manufatura, incluindo (a) o tipo de pesquisas em andamento, (b) as áreas de fabricação onde o big data é focado, e (c) os resultados dessas pesquisas. A metodologia de pesquisa empregada neste estudo é guiado pelo processo de mapeamento sistemático descrito por Peterson et al.

O restante deste artigo é descrito a seguir. Na seção 2, a metodologia de pesquisa e o processo utilizado no estudo (mapeamento sistemático) são descritos. Na seção 3, os resultados do estudo são apresentados. Na seção 4, os resultados do estudo são discutidos em detalhes. Na seção 5, são consideradas as ameaças relativas à validade dos resultados. Finalmente, na seção 6, as conclusões da pesquisa são apresentadas e futuras áreas de pesquisa são identificadas.

Metodologia de Pesquisa

Este estudo utilizou um mapeamento sistemático para captar o estado atual das pesquisas big data na manufatura. Em comparação com outros métodos de pesquisas, tais como revisões de literatura tradicional, um estudo de mapeamento facilita uma investigação de grande amplitude, mas sacrifica a profundidade. Dentro do contexto de um método de mapeamento sistemático foi considerado apropriado, pois proporcionou uma abordagem bem estruturada para a análise. Essa estrutura também serviu para uma base para reduzir o viés e harmonizar os esforços de revisão da literatura por toda a equipe. Além disso, a perspectiva de largura-primeira que pode ser derivada pelo mapeamento sistemático foi especialmente útil para informar sobre uma nova e abrangente área de pesquisa (grandes manufaturas) que atualmente não tem teorias proeminentes e consistentes. Na verdade, é a falta de bons temas de pesquisa que faz uma primeira revisão mais profunda da literatura da área um desafio.

O processo de mapeamento

As principais etapas do processo são mostradas na parte superior, com os resultados de cada etapa mostrado na parte inferior. À medida que os esforços de pesquisa progredem no processo, de cada etapa forma a entrada para a próxima etapa.

No início do estudo, as questões iniciais de pesquisa foram acordadas para fornecer um escopo geral para o estudo. Com base neste escopo, os termos e frases primárias foram identificados e usados para encontrar trabalhos de pesquisa em vários bancos de dados digitais. Após os resultados dessas pesquisas foram gravados, cada papel foi selecionado manualmente usando um conjunto de inclusão e exclusão de critérios que tentam identificar papéis alinhados com o tema e escopo do estudo. Os artigos que foram considerados relevantes para o estudo foram analisados para determinar palavras-chave e frases proeminentes que poderiam ser usadas para pesquisa realizada na área. Finalmente, os trabalhos classificados foram agregados, visualizados e mapeados de forma que nos permita responder às questões de pesquisa levantadas neste estudo.

Questões de pesquisa

O objetivo desta pesquisa é classificar a pesquisa atual e identificar os temas pertinentes e as tendências na literatura, que se relacionam diretamente com tecnologias de big data na indústria de manufatura. A principal questão de pesquisa é intencionalmente ampla, uma vez que atualmente não existem estudos substanciais ou levantamentos que abordam a questão, e é intenção deste estudo estabelecer uma mapa que irá transmitir os diversos temas associados com a área. Portanto, o questão de pesquisa do estudo é:

"Como as tecnologias big data estão sendo usadas na indústria de manufatura?"

Para responder à principal questão de pesquisa, cinco questões de pesquisa auxiliares de vários aspectos de big data na indústria de manufatura foram identificados. Decomposição e caracterização da questão principal de pesquisa em questões menores e mais específicas permite que o tópico seja considerado a partir de múltiplas perspectivas, ao mesmo tempo que fornece os resultados necessários para a principal questão de pesquisa. As perguntas adicionais da pesquisa são descritas abaixo.

RQ1: Qual é a publicação relacionada à big data na indústria de manufatura?

Fundamento: A intenção desta questão é ilustrar o interesse na área de pesquisa ao longo do tempo, bem como identificar as fontes primárias de literatura no campo. Este estudo pressupõe que a taxa de publicação é indicativa de interesse de pesquisa na área, as fontes mais proeminentes de pesquisa no campo são as revistas e conferências que têm a maior frequência de publicação da literatura relevante.

RQ2: Que tipo de pesquisa está sendo realizada na área de big data na indústria de manufatura?

Fundamento: A intenção desta questão é destacar o tipo de pesquisa formal que está sendo empreendidas na área, desde perspectivas filosóficas até avaliações do mundo real. Ao responder a esta pergunta, o estudo visa compreender o nível de maturidade da área, com o pressuposto de que os esforços de pesquisa que não apresentam uma validação rigorosa avaliação pode ser indicativa de um campo que ainda está amadurecendo e focado no desenvolvimento metodologias para apoiar futuros esforços de pesquisa.

RQ3: Que tipo de contribuições estão sendo feitas para a área de big data na indústria de manufatura?

Fundamento: A intenção desta questão é compreender o tipo de contribuições e resultados dos esforços de pesquisa no campo. Estas saídas podem variar muito e variar de arquiteturas de sistemas de informação, ferramentas analíticas e métodos para otimização de processos. Ao responder a esta questão, o estudo visa avaliar ainda mais o nível de maturidade, com o pressuposto de que os primeiros esforços de pesquisa podem se concentrar em diretrizes e metodologias, e áreas de pesquisa mais maduras podem se concentrar na implementação, avaliação e validação esses métodos. Além disso, a identificação de tendências e padrões nos resultados também irá fornecer uma compreensão para resolver desafios específicos na área.

RQ4: Que tipo de análise está sendo usado na área de big data na indústria de manufatura fabricação?

Fundamento: A intenção desta questão é identificar a proeminência do big data na investigação relativa às grandes na indústria de manufatura, bem como o tipo específico de análise que está sendo usado. Nos últimos anos, o termo analítica tornou-se sinônimo de big data. Ao responder a esta pergunta, o estudo visa entender a relação entre análise de dados e big data no contexto da manufatura. Além disso, a classificação dos diferentes tipos de análises utilizados na pesquisa podem fornecer uma compreensão dos tipos de problemas que estão sendo endereçados.

RQ5: Que áreas de manufatura estão aplicando tecnologias de big data?

Fundamento: A intenção desta questão é destacar as diferentes áreas de produção onde os pesquisadores estão utilizando big data. Ao responder a esta pergunta, o estudo tem como objetivo destacar temas de pesquisa específicos, bem como identificar as áreas de operações que estão se esforçando para enfrentar os desafios de dados em larga escala e processamento.

Pesquisa principal

Devido ao foco deste estudo, os termos de pesquisa 'big data' e 'manufatura' foram considerados para ser os termos de pesquisa principais mais óbvios. A Tabela 1 mostra os termos de

pesquisa do candidato identificados através da exploração de resumos que foram devolvidos pela consulta de pesquisa usando "big data" e "manufatura".

Tabela 1 - Principais e candidatos termos de pesquisa para grandes dados na fabricação

Principais termos	Termos do candidato
Manufatura	Fabricação Inteligente, Manufatura Avançada, Indústria 4.0, Cyber Sistemas Físicos, Cadeia de Fornecimento, Fábricas, Fábrica, Produção e Processo.
Big Data	dados de grande escala, computação em nuvem, aprendizado de máquina, grande análise de dados, virtualização de dados e Gerenciamento de dados master.

Tabela 2 - String primária usada para o estudo

(Manufatur* OR Factor* OR Industry 4.0) AND (Big Data)
--

Depois de avaliar diferentes combinações dos termos de pesquisa, a pesquisa principal foi alterada (Quadro 2) para incluir «Fábrica» e «Indústria 4.0». Nossa avaliação de diferentes pesquisas mostrou que os resultados que pareciam ser mais relevantes para o estudo consistentemente continham o termo "big data", bem como um ou mais dos termos "fabricao", 'Fábrica' ou 'fábricas', na seção de título, resumo ou palavras-chave. O símbolo do (*) Foi utilizado para captar as mudanças na pluralidade e no contexto das populações relevantes. O racional da seleção de sequências primárias era manter a pesquisa ampla para capturar muitos temas e tendências de pesquisa quanto possível, ao mesmo tempo em que omitindo documentos que não estão alinhados com o escopo ou tema do estudo.

A pesquisa principal escolhida foi utilizada como critério de pesquisa em sete bancos de dados digitais. Estas bases de dados foram escolhidas em conjunto por todos pesquisadores envolvidos e considerados uma fonte relevante de tecnologia, engenharia e pesquisa. Como a facilidade de pesquisa de cada banco de dados é diferente, a string de pesquisa principal foi adaptada para de cada banco de dados. Além disso, os resultados da pesquisa foram limitados publicação de revistas e conferências com base no pressuposto de que essas publicações mais prováveis de apresentar o conteúdo do que outras fontes, tais como guias e capítulos de livros. A Tabela 3 apresenta as bases de dados digitais pesquisadas durante o estudo, juntamente número de publicações retornadas usando a string de pesquisa principal. A única base de dados que não tinha a facilidade de restringir os critérios de pesquisa tanto ao título como ao resumo, era o Google Scholar. Portanto, dois números foram incluídos para o Google Scholar - 9540 refere-se a o número de documentos para a string de procura principal, enquanto 14 refere-se à número de documentos retornados para a string de pesquisa principal limitado aos títulos. Tendo em conta os resultados das outras bases de dados digitais, o número parecem ser um resultado mais realista.

Tabela 3 Resultados da pesquisa de repositórios digitais

Banco de dados digital	Número de publicações
Biblioteca Digital ACM	283
Engenharia Village	128
Google Scholar	14 (9540)
IEEE Xplore	86
Science Direct	14
Scopus	96

Triagem da pesquisa

Os resultados de pesquisa primária foram filtrados usando um conjunto de critérios de inclusão e exclusão para identificar a pesquisa mais relevante para o estudo. Em primeiro lugar, a string principal foi usada em cada um dos repositórios digitais mostrados na Tabela 3, o que resultou em 661 publicações. Em segundo lugar, estes documentos foram processados utilizando quatro filtros com a intenção de omitir publicações que não eram altamente relevantes para o estudo, o que resultou em 65 publicações restantes. Finalmente, o método snowballing foi usado para extrair as referências de cada uma dessas publicações, sendo vasculhado cada referência selecionada para saber se a pesquisa deve ser adicionada ao estudo. Isso resultou em 75 publicações para serem utilizadas como base para os dados do estudo.

Critério de inclusão

Para ser considerado para inclusão no estudo, a pesquisa avaliada teve que se originar de uma fonte acadêmica, como uma revista ou conferência, e mostrar claramente que sua contribuição foi focada em big data na indústria de manufatura, que foi determinada principalmente pela presença dos termos de pesquisa primários. Nas publicações que atendiam a esses critérios foram processadas critérios de exclusão (filtros), com a intenção de destacar as pesquisas mais relevantes na área de big data na indústria de manufatura.

Critério de exclusão

Para as publicações que passaram pelos critérios de inclusão, quatro filtros foram aplicados para reduzir apenas para aquelas que foram consideradas diretamente alinhadas com o foco do estudo.

Estes filtros são descritos como se segue:

- Filtro 1: Remove as publicações que não contenham "fabricação", "fábrica", "fábricas" ou 'Industry 4.0' no título, resumo ou meta-dados do documento.
- Filtro 2: Remove as publicações que não contêm 'big data' no título, resumo ou meta-dados do documento.
- Filtro 3: Remove os documentos que só se referem a 'fabricação', 'fábrica', 'fábricas', 'Industry 4.0', ou 'big data' como um ponto de referência. Por exemplo, muitos documentos relativos ao big data citam as aplicações potenciais à fabricação, sem investigar exclusivamente a área.
- Filtro 4: Rever as seções de introdução e discussão de cada publicação, e remover aqueles que não se concentram ou contribuem para a área de big data em indústrias de manufatura.

Classificação da pesquisa

Todas as publicações do estudo foram classificadas em quatro dimensões. Essas dimensões foram escolhidas para fornecer diferentes perspectivas sobre o estado atual da pesquisa na área, ao mesmo tempo em que se constrói um conjunto de dados que pode ser utilizado para responder a cada uma das questões destacadas no estudo.

Tipo de pesquisa

Para classificar o tipo de abordagens de investigação será usado um padrão de classificação existente foi escolhido. Este padrão foi definido por Wieringa et al. para a classificação dos requisitos de pesquisa de engenharia. A Tabela 4 fornece um resumo de cada tipo de pesquisa.

Área de fabricação

Para classificar as diferentes áreas de fabricação, um esquema de classificação foi escolhido. Este esquema foi definido por Meziane et al. ao identificar diferentes áreas de fabricação onde a Inteligência Artificial (AI) estava sendo aplicada. A Tabela 5 um resumo de cada área.

Tipo de contribuição

Para classificar o tipo de contribuição feita por cada publicação o método de pesquisa qualitativa conhecido como palavra-chave foi escolhido. O resumo de cada publicação foi sintetizado para destacar as principais palavras-chave que descreveram a contribuição da pesquisa. Essas palavras foram agregadas e sinônimos foram amalgamados para simplificar os resultados e promover uma melhor visibilidade das tendências. Em termos de cruzamentos entre palavras-chave, semelhanças significativas entre a investigação relativa a sistemas e plataformas, que resultaram em ambos serem rotulados como plataformas. Da mesma forma, pesquisas rotuladas como diretrizes, roadmaps e os requisitos foram unificados como teoria. A Tabela 6 apresenta um resumo de cada tipo de Investigação. Além disso, cada publicação usada no estudo é classificada.

Apêndice.

Para classificar o tipo de análise foi escolhido um esquema de classificação existente. Este esquema foi definido por Delen et. al como uma taxonomia de alto nível para análise de negócios. A tabela 7 fornece um resumo de cada tipo de análise.

Tabela 4 Tipos de investigação

Classificação	Descrição
Validação	Pesquisa que investiga técnicas novas e únicas, mas que ainda não foram implementadas em mundo real
Avaliação	Pesquisa que inclui uma implementação significativa de uma determinada técnica, juntamente com uma avaliação.
Solução	Pesquisa que inclui uma ilustração ou exemplo de uma solução para um problema específico.
Pesquisa filosófica	que fornece uma maneira conceitual de olhar para um determinado problema ou campo.
Opinião	Pesquisa que expressa uma opinião pessoal sobre se uma determinada técnica é boa ou mau, sem se concentrar em trabalho relacionado ou métodos de pesquisa padrão.
Experiência	Pesquisa que é escrita a partir da experiência pessoal do pesquisador, e descreve como algo foi feito.

Resultados

RQ1 - Qual é a publicação que se relaciona-se com big data na indústria de manufatura?

A Figura 3 ilustra o crescimento homólogo das publicações relacionadas com os grandes dados na indústria transformadora. A primeira publicação identificada neste estudo é de 2009, sem mais atividade Até 2012. De 2012 a 2014 as publicações no campo aumentaram quase dez vezes, com um aumento de 180% nas publicações entre 2012 e 2013, e um aumento de 242,9% entre 2013 e 2014. Note-se que a diminuição das publicações para 2015 está incompleta, uma vez que os dados deste estudo incluem apenas pesquisas publicadas em janeiro e fevereiro de 2015. A Figura 4 apresenta a distribuição das publicações por revista e conferência. Excluindo o dados parciais para 2015, as publicações da conferência foram maiores do que as para cada ano que foi ilustrado. Com efeito, à medida que o interesse na área começou a 2012 a 2014, a proporção de publicações de conferências para revistas aumentou de 60% em 2012 para 75% em 2014. Uma vez que esta área de investigação é relativamente nova, existe uma limitação a quantidade de dados históricos

disponíveis para identificar longas e fortes correlações, entretanto, com base nos dados disponíveis existe uma forte correlação linear entre a publicação de periódicos.

A Figura 5 destaca a distribuição de publicações por ano e revista. Existem quatro revistas que são responsáveis pela publicação de 45,84% da pesquisa na área. O *International Journal of Production Economics* é a principal fonte de pesquisa na área com 16,67% das publicações, com o *Journal of Business Logistics* publicando 12,5%, enquanto *Manufacturing Letters* e *Studies in Computational Intelligence* publicaram 8,34 e 8,33%, respectivamente. Além do fato de que o *Journal of Production Economics* tem a maioria das publicações relativas a grandes dados na indústria transformadora, também pode ser atual, com todas as suas pesquisas publicadas em 2015.

A Figura 6 apresenta um resumo visual das publicações por conferências e por ano. Os três primeiros fontes identificadas neste estudo constituem 30,77% do total de publicações da conferência. O IEEE a conferência sobre Big Data é a principal fonte de pesquisa com 11,54% das a Winter Simulation conference é a terceira fonte mais proeminente com 7,69%. O segunda fonte mais proeminente de pesquisa é o Procedia CRIP com 11,54% de todas as conferências Publicações - mas deve-se notar que Procedia CRIP é uma fonte agregada de publicações de todas as conferências CRIP.

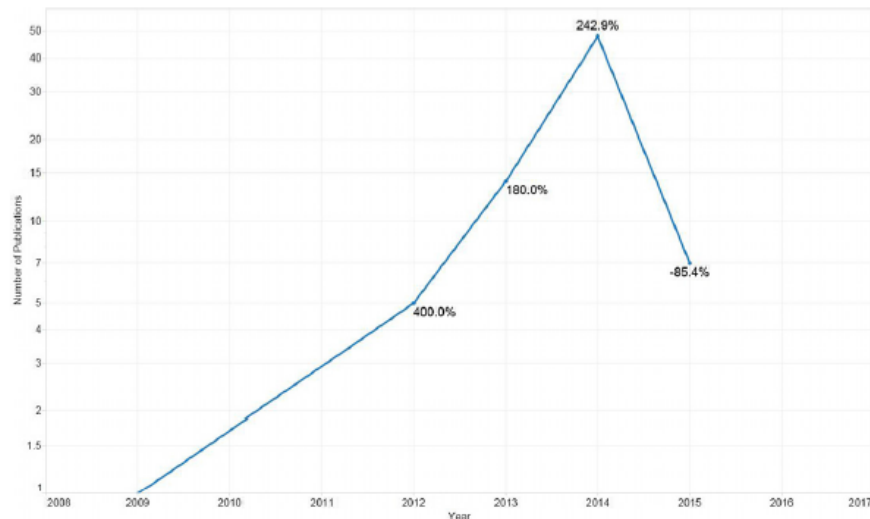


Fig. 3 Year-on-year publication growth for big data in manufacturing

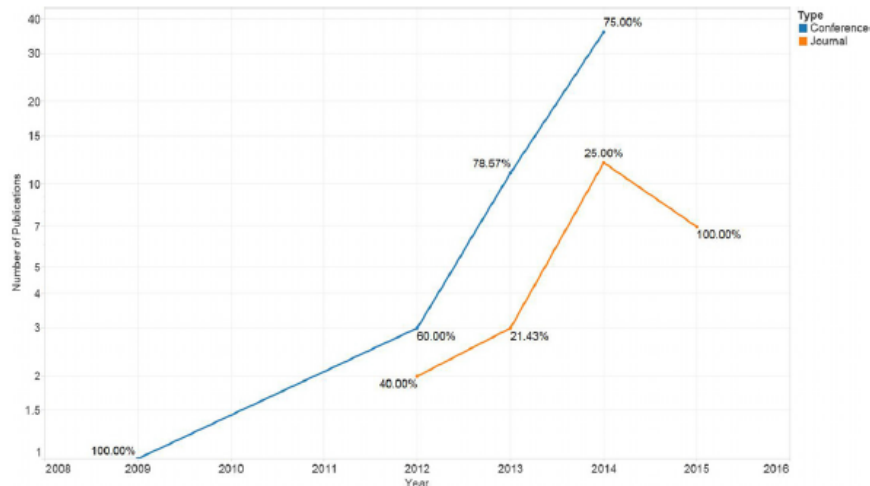


Fig. 4 Comparison of publications in conferences and journals

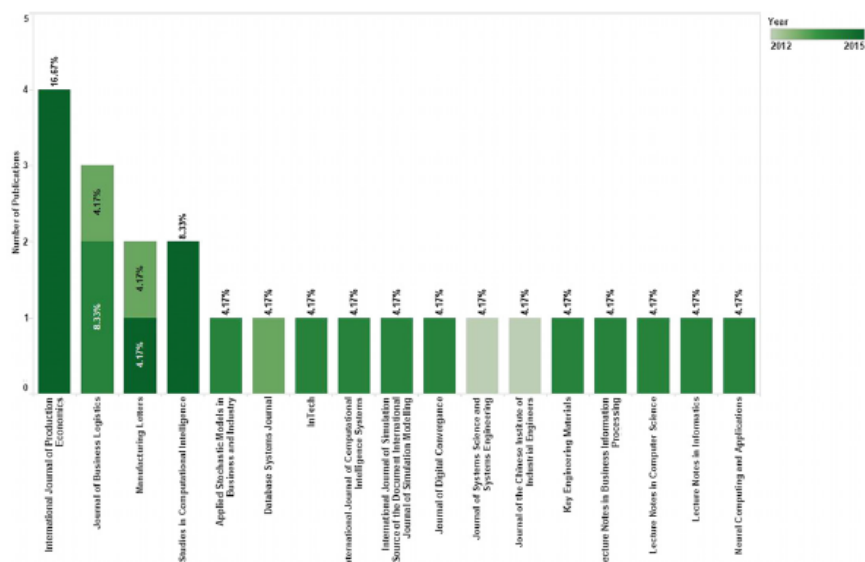


Fig. 5 Distribution of publications by journal

Tabela 5 Áreas na manufatura

Classificação	Descrição
Design	Investigação centrada na concepção de produtos para atividades de fábrica.
Processos e planeamento	Investigação centrada em todos os aspectos do processo e do planeamento, com ênfase na redução de resíduos e no aumento da produção.
Gestão da qualidade	Investigação centrada na gestão da qualidade em ambientes de produção.
Manutenção e diagnóstico	Investigação centrada nas máquinas das operações de fábrica, desde a manutenção preditiva até o diagnóstico em tempo real.
Programação	Pesquisa centrada no planejamento, gerenciamento e otimização de atividades e processos em ambientes de produção.
Controle	Investigação centrada no controle, gestão e otimização de operações e processos em ambientes de produção.

Ambiente, Saúde e Segurança	A investigação centrada nos fatores relacionados com o ambiente, energia, bem estar, saúde e segurança.
Fabricação virtual	Pesquisa com foco na realização de fábricas e processos virtuais.

Tabela 6 Tipos de contribuição

Classificação	Descrição
Arquitetura	Pesquisa que fornece uma visão teórica de como vários componentes em uma solução irá sentar Juntos e interagir.
Framework	Research que descreve o encapsulamento de múltiplas bibliotecas de software que problema, sendo também extensível.
Teoria	Pesquisa que desenvolve diretrizes de alto nível e roteiros para um problema específico.
Metodologia	Pesquisa que apresenta abordagens de baixo nível para resolver um problema específico.
Model	Research que produz modelos matemáticos para resolver problemas específicos.
Platform	Research que fornece um sistema com componentes de hardware e software, que aplicativos a serem executados.
Process	Research que apresenta processos de baixo nível para resolver um problema específico.
Ferramenta	Pesquisa que desenvolve utilitários de software bem definidos que abordam um subconjunto de um problema maior.

Tabela 7 Tipos de análise

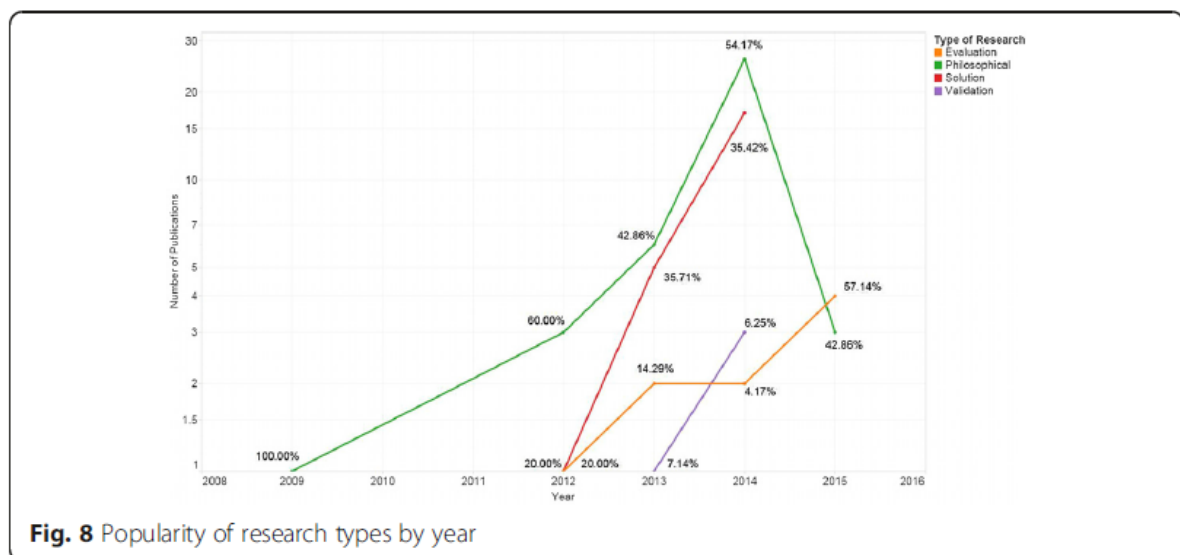
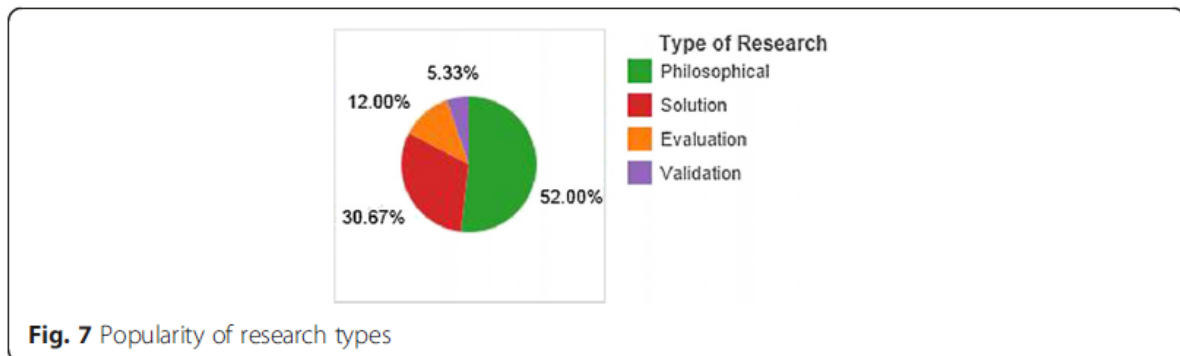
Classificação	Descrição
Descritiva	Pesquisa que se concentra em descrever a estrutura, as relações e o significado dos dados.
Preditiva	Pesquisa que se concentra na previsão de um resultado utilizando os dados disponíveis.
Prescritiva	Investigação que se concentra na prescrição de ações utilizando os dados disponíveis.

RQ2 - Que tipo de pesquisa está sendo realizada na área de big data na indústria de manufatura?

A Figura 7 apresenta uma análise do tipo de pesquisa que está sendo conduzida na área de dados na fabricação. Em 52,63%, a maioria das pesquisas realizadas é filosófica, que é considerado como uma pesquisa que não tem como objetivo demonstrar o(s) conceito(s) apresentado(s) no papel. O próximo tipo de pesquisa mais comum é a pesquisa baseada em soluções, que resolve um problema particular e apresenta um pequeno exemplo da solução na prática. Menos os tipos proeminentes de pesquisa incluem a avaliação e a validação, que compreendem cada um 11.84% e 5,26% das publicações, respectivamente. A investigação baseada na avaliação pode ser tipo avançado de solução, que explora todas as consequências, benefícios e desvantagens, método ou abordagem propostos.

A Figura 8 ilustra a popularidade do tipo de pesquisa por ano. Previsivelmente, devido para a popularidade da pesquisa de avaliação e solução destacada na Fig. 7, o primeiro e os segundos tipos de trabalhos de pesquisa publicados entre 2012 e 2014 são a avaliação e solução respectivamente. Os primeiros esforços de pesquisa em 2012 tiveram foco na pesquisa filosófica, compreendendo 60% dos trabalhos publicados. Como interesse na área cresceu entre 2013 e 2014, o percentual de

trabalhos voltados para desenvolver filosofias diminuiu globalmente, declinando a 42.86% em 2013, e então subindo novamente para 55,1% em 2014. A tendência nas publicações relativas à investigação estão intimamente alinhados com os da investigação filosófica, 35.71 e 34.69% das publicações globais em 2013 e 2014. Também aparece um atraso visível na pesquisa baseada em avaliações, que representava apenas 4,08% do em 2014, mas os dados parciais para 2015 mostram que o número de publicações no primeiro trimestre de 2015 é o dobro de 2014.



RQ3 - Que tipo de contribuições estão sendo feitas para a área de grandes dados na fabricação?

A Figura 9 ilustra os tipos de resultados e contribuições da pesquisa, bem como a decomposição dessas contribuições por conferência e publicações de periódicos. Os três tipos principais das contribuições constituem 61,33% de todas as publicações. Em 26,67%, o mais proeminente tipo de contribuição é a teoria. Na verdade, esses dados se alinham bem com os resultados anteriores da Fig. 7, que destacou a pesquisa filosófica como o tipo mais comum das pesquisas realizadas. As próximas contribuições mais significativas são e plataformas, com cada tipo de contribuição associada a 17,33% de todas as publicações. no entanto, uma estranheza que é evidente nos dados é a quantidade desproporcional de conferência para jornais associados a plataformas. Esta tendência não é uma presente em qualquer dos tipos de contribuição significativa, onde se espera previsivelmente a favor de conferências, mas com uma distribuição saudável entre ambos os tipos de publicação. A Figura 10 ilustra a popularidade das contribuições de pesquisa por ano. Semelhante ao tendência nas contribuições da pesquisa mostradas na Fig. 9, a tendência anual de contribuição para a pesquisa reforça a visão de que a teoria é a contribuição mais proeminente

na área, com sendo a contribuição mais proeminente de 2012 para 2014. Uma tendência notável é a Aumento da diversidade e da distribuição das contribuições da investigação, em dois tipos de contribuição em 2012, para um total de oito em 2014. Este é presumivelmente um subproduto taxas de publicação, mas a distribuição fornece uma visão equilibrada da pesquisas atuais no campo.

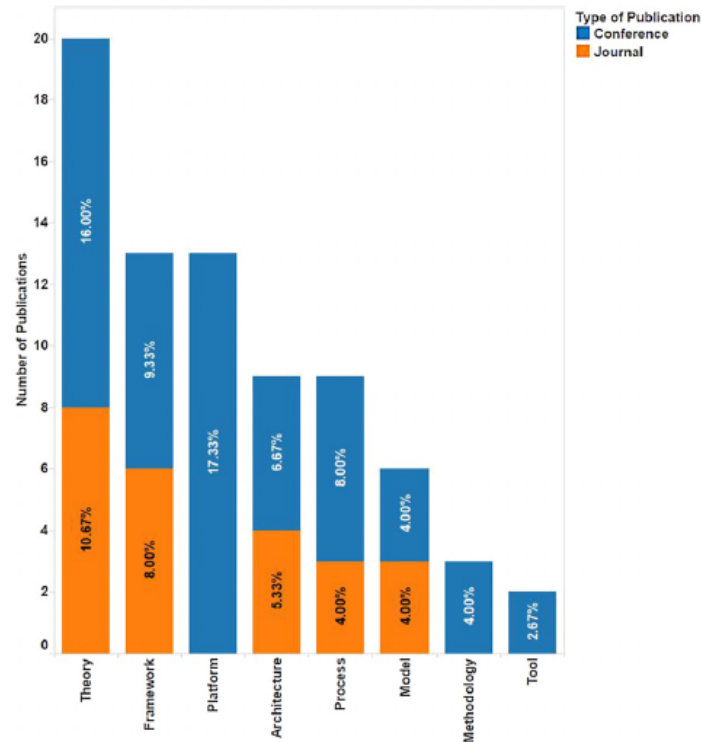


Fig. 9 Comparison of research contributions in journals and conferences

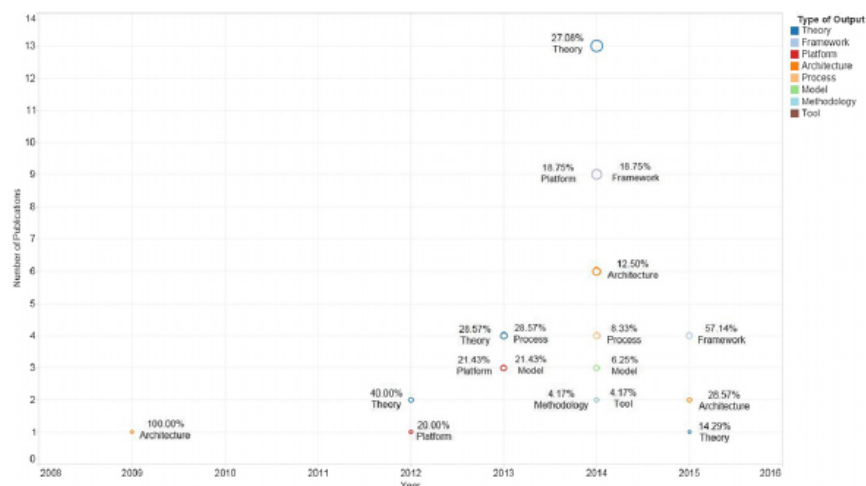
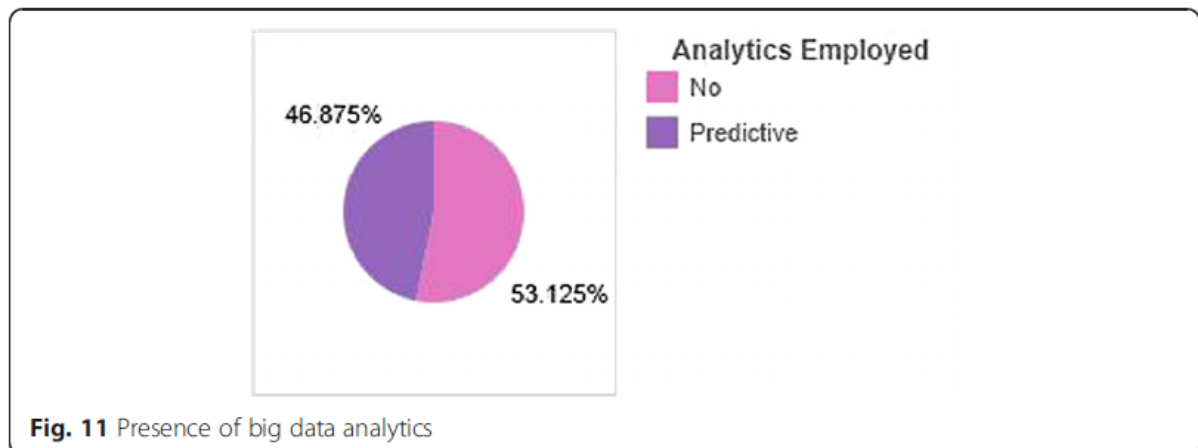


Fig. 10 Popularity of research contribution by year

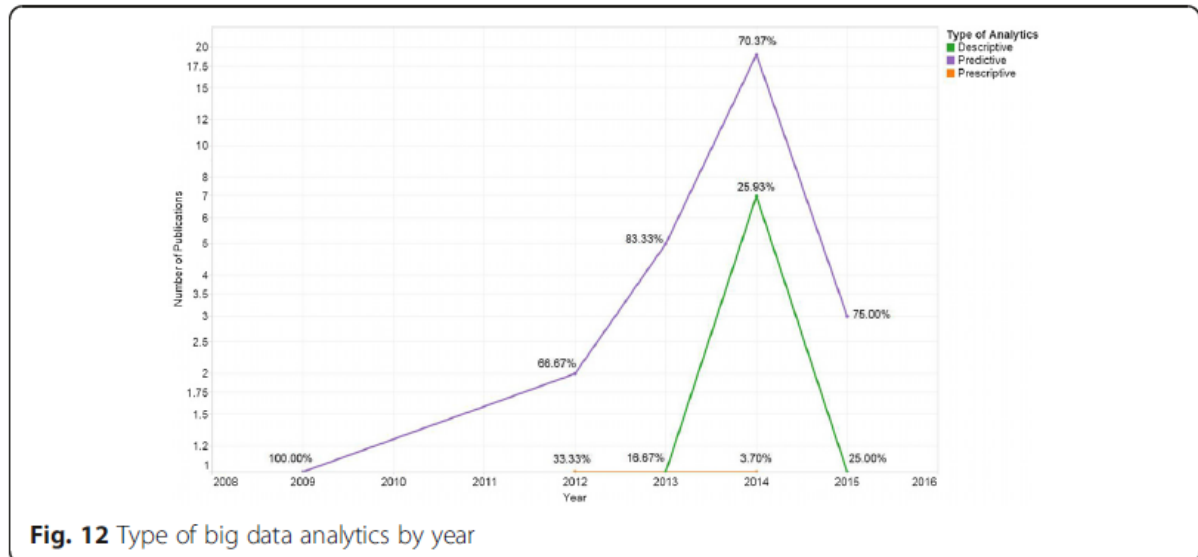
RQ4 - Que tipo de análise está sendo usado na área de big data na indústria de manufatura?

A Figura 11 mostra a porcentagem de pesquisa incluída neste estudo que possuía uma referência para análise e grandes dados na fabricação. Os resultados são relativamente uniformes,

com 47,69% dos os artigos empregando alguma forma de análise, e 52,31% focando-se apenas em grandes tecnologias de dados sem abordar a análise diretamente.



A Figura 12 ilustra a popularidade dos principais tipos de análise em uma base ano-a-ano. A tendência de 2009 a 2015 é notavelmente clara. A maioria das análises foca em análises preditivas, com uma minoria focada em análise descritiva. Na publicação de pico nível em 2014, alguma forma de análise preditiva foi evidente em 71,43% das publicações, comparado com análise descritiva a 25%. Para além destas tendências ascendentes, tendência estagnada é a falta de foco na análise prescritiva, que respondeu por 3,57% das publicações em 2014. Na verdade, apenas um único artigo foi publicado em cada ano entre 2012 E 2014, que se concentrou na análise prescritiva.



RQ5 - Quais são as áreas de fabricação em que tecnologias de big data estão sendo aplicadas?

A Figura 13 mostra áreas de manufatura em que os esforços de pesquisa estão usando tecnologias de big data, bem como identificar contribuições significativas dentro de cada área. Como é evidente resultados, processo e planejamento é a área mais proeminente de fabricação para pesquisa pertencentes às grandes tecnologias de dados. No processo e no planejamento, as contribuições mais arquiteturas, plataformas, estruturas e teorias, constituindo mais de 60% do publicações nesta área específica. O processo e planejamento a seguir é a empresa, que é rótulo de

alto nível que representa pesquisas que abrangem várias áreas da indústria que se concentra em uma única área, mas não pode ser claramente atribuído a qualquer classe única. Contribuições associadas à empresa compartilham uma distribuição similar para planejamento, onde as estruturas, teorias e processos são contribuições significativas pesquisa na área. Dado que a empresa é um agregado de tipos, manutenção e diagnóstico é possivelmente a área mais proeminente após o processo e planejamento. As contribuições relativas manutenção e diagnóstico são um pouco diferentes das áreas teóricas, modelos e arquiteturas a saída mais comum da pesquisa.

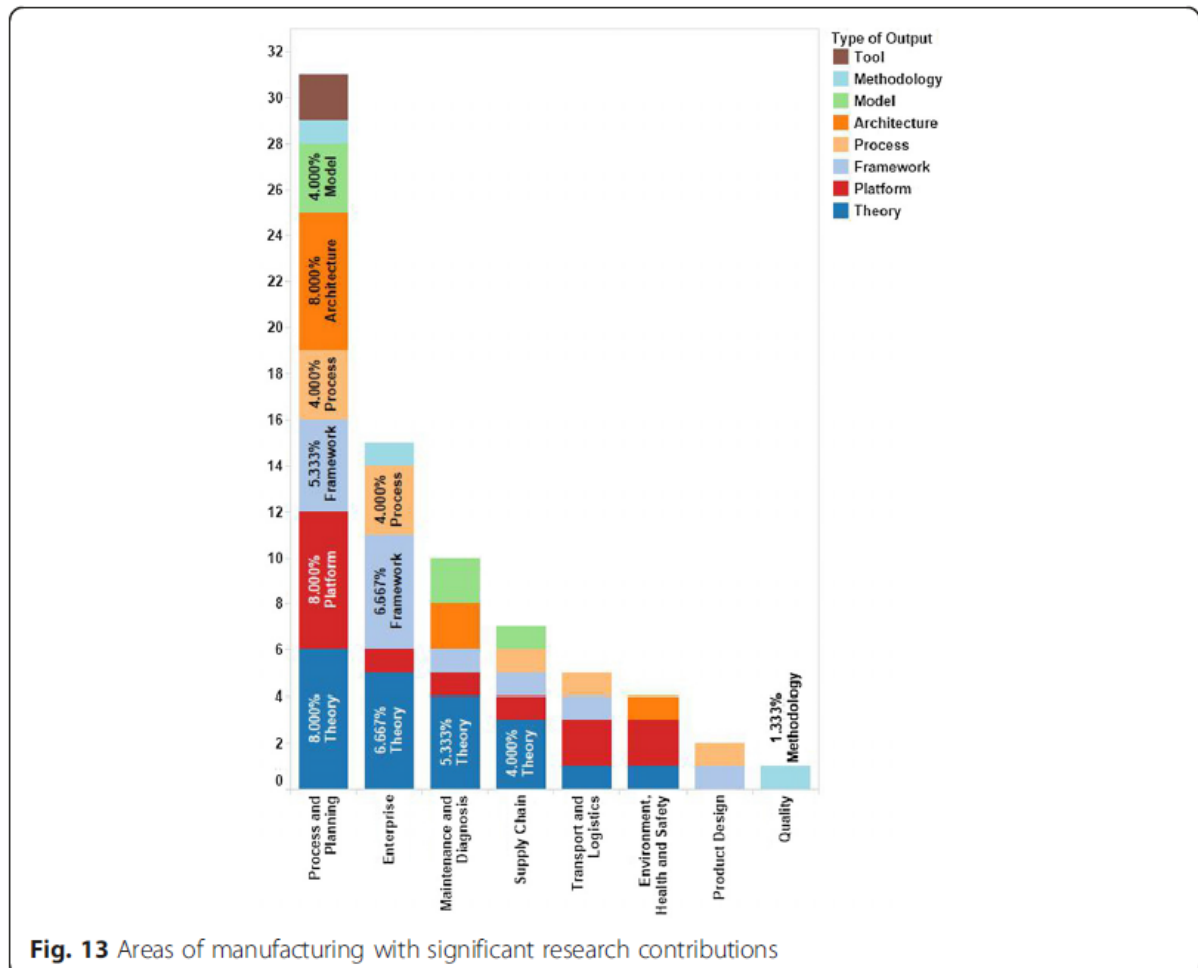


Fig. 13 Areas of manufacturing with significant research contributions

Discussão

RQ1 - Qual é a publicação é relacionada com big data na manufatura?

O raciocínio subjacente a esta questão de investigação era determinar o nível de interesse Área, bem como destacar fontes proeminentes de pesquisa primária. Os resultados mostram que os grandes dados na fabricação são uma área de pesquisa que está experimentando crescimento, com publicações sobre o tema aumentando em um múltiplo de dez entre 2014. Olhando para os resultados da publicação em maior profundidade, há uma correlação ano a ano em publicações de conferências e publicações de periódicos. Esta correlação pode ser um resultado dos primeiros esforços de investigação centrados no desenvolvimento de pequenos trabalhos de investigação para conferências, e em uma data mais atrasada, desenvolvendo aqueles documentos em documentos detalhados. Os resultados identificou um número de fontes de pesquisa mais proeminentes relacionadas a tecnologias de big data na manufatura, com o Journal of Production Economics e a

IEEE conferencias em Big Data, sendo as principais fontes de publicações de revistas e conferências, respectivamente. Atualmente, o interesse da pesquisa na área de big data é claramente ilustrado pelo crescimento exponencial ano-a-ano das publicações dos últimos anos.

RQ2 - Que tipo de pesquisa está sendo realizada na área de big data na indústria de manufatura?

O raciocínio subjacente a esta questão de investigação era compreender a maturidade da área de pesquisa. Baseou-se no pressuposto de que a pesquisa é filosófica, sem aplicação ou implementação, pode ser indicativo de uma área que é relativamente imatura e requer o desenvolvimento da teoria para apoiar futuras aplicações. Este tipo de processo de feed-forward é evidente nos resultados e tipo de abordagens de pesquisa empregadas, onde a pesquisa filosófica é a mais comum, com um atraso cronológico na pesquisa de solução e avaliação. Mais notavelmente, avaliação, que é considerado neste estudo como o tipo de pesquisa mais maduro relativa às implementações tecnológicas, está associada ao mesmo número de publicações no primeiro trimestre de 2015 como foi em 2013 e 2014 combinados. Enquanto tecnologias de big data no setor de manufatura é uma nova área de pesquisa em termos cronológicos, o crescimento exponencial mostrados nos resultados relacionados ao RQ1, juntamente com a cascata natural de pesquisas filosóficas, com pesquisa rigorosa, empírica e demonstrável, indica que a área está se desenvolvendo rapidamente. Contudo, com base nos resultados deste estudo, a área pode ainda ser classificada como sendo um pouco imatura devido à alta proporção da base filosófica, aliada à baixa quantidade de investigação rigorosa baseada na avaliação.

RQ3 - Que tipo de contribuições estão sendo feitas para big data na manufatura?

O raciocínio subjacente a esta questão de investigação foi o de identificar melhor a área, classificando o tipo de resultados oriundos da pesquisa, proeminentes temas de pesquisa atuais e tendências. Os dados anuais mostram um crescimento na distribuição e equilíbrio na área. Isto pode ser interpretado como sendo indicativo de uma vibrante comunidade de pesquisa que está amadurecendo e evoluindo. Parece ser uma forte relação entre os resultados em RQ2 e RQ3. Em particular, os mais proeminentes as classificações em ambos os conjuntos de resultados são em grande parte análogas, nomeadamente pesquisas filosóficas e contribuições teóricas. Como essa área de pesquisa é relativamente nova e imatura, há uma ênfase no desenvolvimento de teorias que podem ser futuros esforços de investigação para resolver problemas específicos na área. Na verdade, os próximos resultados de pesquisa após a teoria são estruturas e plataformas. Esses tipos podem ser vistos como um ponto intermediário entre teoria e aplicação, como eles são desenvolvidos numa base teórica (design ou arquitetura) e facilitar o desenvolvimento de aplicações e sistemas. Uma anomalia nos resultados mostrou que foi a falta de artigos em periódicos identificando plataformas como sua produção, quando conferências. Isto poderia simplesmente ser o resultado do termo "plataforma" ser mais proeminente em uma comunidade (conferências de indústria), versus outra comunidade (revistas acadêmicas). No entanto, investigar mais a anomalia não é este estudo, dado que não é crítico responder à pergunta da pesquisa. Como anteriormente os temas principais das contribuições da investigação se sobrepõem, até certo ponto, as conclusões de RQ2. Com base nestes resultados, parece que existe uma base teórica em desenvolvimento na área de pesquisa, com tecnologias sendo desenvolvidas para implementar essas teorias.

RQ4- Que tipo de análise está sendo usado em big data na manufatura?

O raciocínio subjacente a esta pergunta de investigação era identificar a extensão sendo utilizadas com tecnologias big data na manufatura, bem como a compreensão dos tipos dos problemas que estão sendo resolvidos. Os resultados mostram que cerca de metade das grandes pesquisa emprega alguma forma de análise. Isto é interessante no sentido de que confirma que as tecnologias de big data estão sendo usadas independentemente da análise, e os termos não devem

ser utilizados como sinônimo. O tipo mais comum de problemas tratados pelo big data é precisão de previsão, que é uma qualidade desejável na tomada de decisão. A proeminência da análise preditiva pode ser atribuída à presença de teorias e métodos relacionados à predição de outros campos (estatísticas) e a aplicabilidade da análise de previsão para problemas do mundo real. Em contraste, a falta de análise é evidente a partir dos resultados. No entanto, isso pode ser atribuído à dificuldade na construção de aplicações prescritivas. As aplicações prescritivas são inerentemente quando comparado com a análise descritiva e preditiva, dada a necessidade de alinhar tecnologia, modelagem, previsão, otimização e conhecimento do assunto. Portanto, big data na manufatura ainda está em sua infância, é pouco surpreendente que apenas alguns desses aplicativos analíticos prescritivos altamente complexos tenham emergido.

RQ5- Que áreas de fabricação são as grandes tecnologias de dados sendo aplicadas? O raciocínio subjacente a esta questão de investigação foi o de destacar temas de big data na manufatura, com ênfase particular na compreensão tipo de problemas de fabricação que estão sendo abordados por grandes dados. Os resultados processo e planejamento em, fabricação está recebendo atualmente a maioria dos interesses de investigação, com diversas aplicações empresariais inter-departamentais e manutenção e diagnóstico representando 74,6% das publicações deste estudo. No entanto, dada a investigação classificada como "empresa" é composta por diversas aplicações que abordam uma ampla gama de tópicos Não deve ser classificado como um importante tema de pesquisa para big data na manufatura. De acordo com os resultados do RQ2 e do RQ3, há uma forte ênfase em pesquisas teóricas, bem como o desenvolvimento de frameworks, plataformas e arquiteturas para realizar esses fundamentos teóricos.

Ameaças à validade

Como com qualquer metodologia de pesquisa secundária, o processo de mapeamento infalível, e há de fato uma série de ameaças à validade deste estudo. Para este fim, todos os esforços foram feitos para mitigar riscos potenciais ao longo do processo. As ameaças à validade que foram identificados estão descritos nesta seção.

Critérios de pesquisa

Os critérios de pesquisa utilizados para a obtenção de trabalhos para este estudo foram escolhidos pelo pesquisadores participantes. As escolhas relativas aos critérios de pesquisa foram impulsionadas pelo (a) escopo da pesquisa, (b) as questões de pesquisa que precisavam ser respondidas, bem como (c) a relevância dos artigos retornados de testar várias combinações de pesquisa. Contudo, enquanto maior cuidado foi tomado para escolher as cadeias de pesquisa mais adequadas para o estudo, existe o risco inerente de que isso possa restringir a descoberta de documentos que não os critérios de pesquisa. Embora com base na experiência da equipe de pesquisa, a sofisticação das bases de dados digitais modernas, juntamente com a disponibilidade de meta-dados de publicação, bem como a uniformidade dos termos incorporados em nossa pesquisa, sugere que o risco de omitir trabalhos relevantes foi pelo menos minimizado.

Bases de dados digitais

A equipe de pesquisa selecionou as bases de dados digitais para adquirir os trabalhos para o estudo. Estas bases de dados foram selecionadas utilizando uma combinação de conhecimentos

prévios de investigação tecnológica e tecnológica, bem como a identificação de bases de dados campos. Portanto, se um determinado repositório digital não foi pesquisado, existe o risco de documentos relevantes não seriam incluídos no estudo. No entanto, a quantidade de sobreposição experiência nos resultados de pesquisa em diferentes tipos de repositórios digitais nível de redundância. Mais precisamente, se um repositório digital específico não fosse estudo, existe uma possibilidade realista de que a investigação indexada por outra fonte que está sendo usada ou, de fato, descoberta seguindo a referências de cada um dos artigos do estudo (snowballing).

Google Scholar

Conforme especificado na metodologia de pesquisa, houve um problema com a construção de um pesquisa para o Google Scholar. Enquanto outros bancos de dados de pesquisas para interrogar títulos, resumos e palavras-chave usando a lógica booleana, O Google Scholar se limitou à pesquisa por título ou texto completo. Quando uma pesquisa de texto completo foi realizado um total de 9540 registros, o que obviamente são muitos dados para analisar para um estudo deste tamanho. Portanto, a opção de busca por título foi escolhida como ele retornou um gerenciável 14 publicações. Ao escolher esta abordagem de pesquisa para o Google Acadêmico, existe o risco de que publicações com resumos e palavras-chave que os critérios de pesquisa do estudo podem ter sido omitidos.

Critérios de inclusão e exclusão

Os critérios definidos para inclusão e exclusão neste estudo resultaram de discussões dentro da equipe de pesquisa, onde as regras e condições que foram consideradas como alinhados com o escopo do estudo foram identificados. Criação de regras para identificar a literatura para revisão significa que há o risco de que a pesquisa relevante possa ser omitida se utiliza terminologia diferente da dos critérios de inclusão / exclusão. Contudo, os termos de pesquisa principais do estudo, a saber, fabricação e grandes dados, são convencionais, termos bem definidos e aceites, o que deverá reduzir o número de publicações omitido devido a autores usando termos sinônimos. Além disso, como o estudo focada na identificação das principais pesquisas na área de grandes dados na indústria, não há tanto uma preocupação com a captura de pesquisa que é muito vagamente relacionados para o tópico.

Precisão de classificação

Existe o risco de as equipes de pesquisa rotularem e categorizarem a pesquisa no estudo pode ser diferente do de outro pesquisador. Reduzir o viés individual e ganhar confiança na precisão de nosso processo de classificação, cada pesquisador da equipe para classificar cada publicação. Os resultados deste processo de classificação foram então analisados, com as publicações que foram classificadas como sendo rotuladas imediatamente, e aquelas com diferentes classificações sujeitas a uma reunião de revisão para determinar as classificação.

Conclusão e trabalho futuro

No momento da redação, este é o único esforço de pesquisa focado no mapeamento sistemático das grandes tecnologias de dados na indústria de transformação. A pesquisa

apresentada neste artigo forneceu uma ampla revisão da pesquisa relacionada aos grandes dados na indústria promover uma melhor compreensão de uma nova e abrangente área. Em particular, questões de pesquisa relevantes para os atuais esforços de pesquisa de big data na indústria foram respondidas, ao mesmo tempo que forneceu uma excelente plataforma para investigação na área. Em particular, é lógico que trabalho deve se concentrar no desenvolvimento de revisões sistemáticas e de literatura alinhadas com as áreas de manufatura identificadas neste estudo, como a criação de uma revisão sistemática de grandes dados na indústria de manufatura que se concentra na diagnóstico. A combinação dessas revisões, aliada ao mapeamento sistemático apresentado pesquisa, pode servir para fornecer uma perspectiva completa da pesquisa primária relativa ao big data na indústria de manufatura.