



University of
Texas Libraries



e-revist@s



Centro Unversitário Santo Agostinho

revistafsa

www4.fsnet.com.br/revista

Rev. FSA, Teresina, v. 22, n. 1, art. 7, p. 133-157, jan. 2025

ISSN Impresso: 1806-6356 ISSN Eletrônico: 2317-2983

<http://dx.doi.org/10.12819/2025.21.1.7>

DOAJ DIRECTORY OF
OPEN ACCESS
JOURNALS

WZB
Wissenschaftszentrum Berlin
für Sozialforschung



Zeitschriftendatenbank



MIAR



Diadorim

O Impacto Operacional e Social dos Grandes Modelos de Linguagem e Automação na Indústria

The Operational and Social Impact of Large Language Models and Automation in the Manufacturing Industry

Gabriel José Souza e Silva

Graduando em Engenharia de Controle e Automação pela Universidade Federal do Rio de Janeiro

E-mail: gabrieljsss@poli.ufrj.br

Bruno Pereira Fraga

Graduando em Engenharia de Controle e Automação pela Universidade Federal do Rio de Janeiro

E-mail: brunofraga20@poli.ufrj.br

Moacir Porto Ferreira

Doutor em Administração pela Universidade do Grande Rio

Professor da Escola Politécnica da Universidade Federal do Rio de Janeiro - Poli/UFRJ

E-mail: moacir.ferreira@poli.ufrj.br

Endereço: Gabriel José Souza e Silva

Prca. Serzedelo Correia 15-A; Copacabana; Rio de Janeiro; RJ; 22040050; Brasil.

Endereço: Bruno Pereira Fraga

Praça Seca, Rua barão 450, bloco 2, Aplt 503; Rio de Janeiro; RJ; 21211300; Brasil.

Endereço: Moacir Porto Ferreira

Av. Athos da Silveira Ramos, 149, Bloco F- Sala F101. Cidade Universitária. Rio de Janeiro - RJ, 21941-485, Brasil.

Editor-Chefe: Dr. Tonny Kerley de Alencar Rodrigues

Artigo recebido em 06/12/2024. Última versão recebida em 19/12/2024. Aprovado em 20/12/2024.

Avaliado pelo sistema Triple Review: a) Desk Review pelo Editor-Chefe; e b) Double Blind Review (avaliação cega por dois avaliadores da área).

Revisão: Gramatical, Normativa e de Formatação



RESUMO

O objetivo deste trabalho é analisar o impacto operacional e social dos Grandes Modelos de Linguagem (LLMs) e da automação na indústria, identificando os desafios técnicos, legais e éticos envolvidos na integração dessas tecnologias em sistemas industriais. A pesquisa utiliza uma revisão integrativa da literatura, seguindo as diretrizes de Torracco (2005), com buscas realizadas nas bases Scopus e Perplexity AI. Foram selecionados e analisados estudos sobre a evolução histórica da automação e da inteligência artificial, enfatizando o surgimento dos LLMs e sua aplicação no contexto industrial. Os resultados destacam que, embora os LLMs representem um avanço significativo com potencial para revolucionar os processos produtivos, existem obstáculos como limitações técnicas, riscos de segurança e a necessidade de supervisão humana devido a limitações dos modelos, como alucinações. Conclui-se que a adoção dos LLMs na indústria requer um equilíbrio entre os benefícios e os riscos, enfatizando a importância da conformidade regulatória e da participação humana contínua para garantir a segurança e a eficiência dos sistemas industriais.

Palavras-chave: LLM. Automação Industrial. Inteligência Artificial. Engenharia de Produção. Impactos Sociais.

ABSTRACT

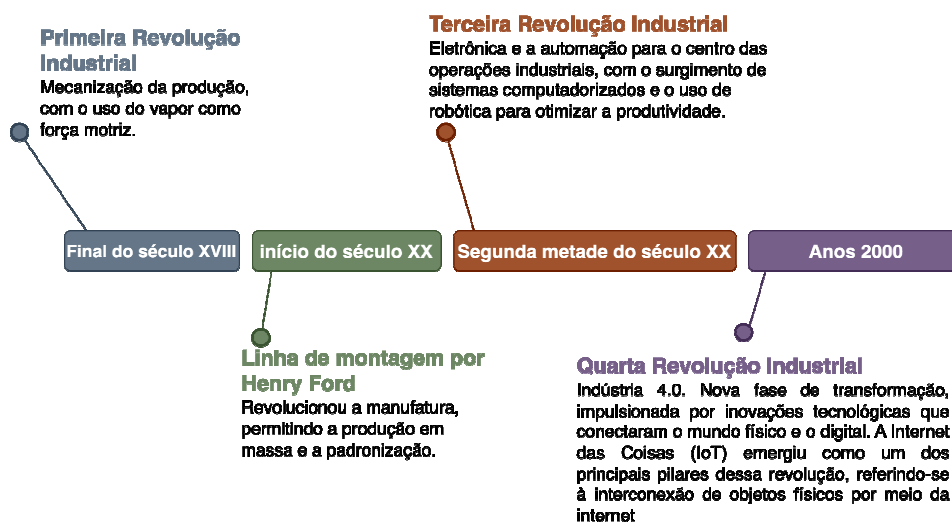
The objective of this work is to analyze the operational and social impact of Large Language Models (LLMs) and automation in the industry, identifying the technical, legal, and ethical challenges involved in integrating these technologies into industrial systems. The research employs an integrative literature review, following Torracco's (2005) guidelines, with searches conducted in the Scopus and Perplexity AI databases. Studies on the historical evolution of automation and artificial intelligence were selected and analyzed, emphasizing the emergence of LLMs and their application in the industrial context. The results highlight that, although LLMs represent a significant advancement with the potential to revolutionize production processes, there are obstacles such as technical limitations, security risks, and the need for human supervision due to model limitations like hallucinations. It is concluded that the adoption of LLMs in the industry requires a balance between benefits and risks, emphasizing the importance of regulatory compliance and continuous human participation to ensure the safety and efficiency of industrial systems.

Keywords: LLM. Industrial Automation. Artificial Intelligence. Production Engineering. Social Impacts

1 INTRODUÇÃO

Ao longo da história, o setor industrial tem sido moldado por uma série de marcos tecnológicos que transformaram profundamente a forma como os produtos são concebidos, fabricados e entregues, conforme ilustrado na linha do tempo da Figura 1.

Figura 1 – Linha do tempo da automação



Fonte: os autores.

A partir da década de 1990, o conceito de *big data* começou a ganhar destaque, com John Mashey sendo frequentemente creditado por popularizar o termo. Inicialmente, o *big data* referia-se a conjuntos de dados que ultrapassam a capacidade das ferramentas tradicionais de *software* de capturar, gerenciar e processar eficientemente. Durante os anos 2000, o *big data* tornou-se crucial para o avanço da digitalização nas indústrias, impulsionado pela crescente conectividade proporcionada pela Internet das Coisas (IoT). Isso permitiu que empresas coletassem vastos volumes de dados em tempo real, possibilitando análises preditivas, otimização de processos e decisões estratégicas mais informadas (ARAQUE GONZÁLEZ *et al.*, 2022).

Os Modelos de Linguagem de Grande Escala (LLMs), por sua vez, começaram a emergir a partir do final dos anos 1990, com a introdução de redes como *Long Short-Term Memory* (LSTM), que permitiram o processamento de grandes volumes de dados de forma mais eficiente. Em 2010, ferramentas como o CoreNLP da Stanford ajudaram a aprimorar tarefas de processamento de linguagem natural (NLP), como análise de sentimentos e

reconhecimento de entidades. No entanto, o grande avanço ocorreu em 2017 com o surgimento dos modelos baseados em transformadores, como o GPT (*Generative Pre-trained Transformer*) e o BERT (*Bidirectional Encoder Representations from Transformers*). Esses modelos revolucionaram o NLP ao utilizar aprendizado autossupervisionado para compreender o contexto em grandes conjuntos de dados não estruturados (VARSHNEY; SINGHAL, 2024).

Desde então, o desenvolvimento de LLMs tornou-se exponencial, especialmente após o lançamento do BERT pela Google em 2019, que redefiniu a forma como sistemas entendem e processam linguagem natural. Na Indústria 5.0, esses modelos têm sido fundamentais para promover a colaboração entre humanos e máquinas, utilizando interfaces mais intuitivas que facilitam a automação inteligente. Eles permitem, por exemplo, transformar dados não estruturados, como relatórios de manutenção e feedback de clientes, em *insights* valiosos para a melhoria contínua dos processos (AL-ABASSI *et al*, 2020). A adoção de LLMs na automação industrial representa, segundo o WORLD ECONOMIC FORUM (2024), o futuro da indústria de manufaturas, onde a customização e a eficiência são impulsionadas pela integração avançada de IA em processos produtivos.

2 REFERENCIAL TEÓRICO

O World Economic Forum (2024) destaca uma nova transformação substancial na automação industrial impulsionada pela aplicação de LLMs; Porém, para analisar plenamente as possibilidades introduzidas por esses grandes modelos de linguagem, é fundamental revisitar as origens da automação de processos.

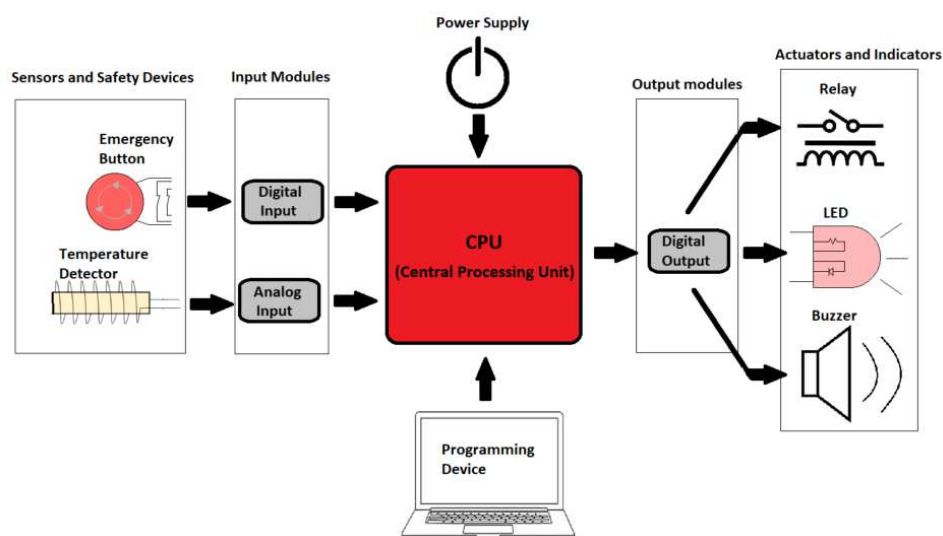
2.1 Automação

A perspectiva histórica da automação remonta ao século XX, surgindo com as revoluções industrial e agrícola, momentos que impulsionaram a mecanização e a automatização de processos produtivos. Inicialmente, a automação estava ligada à introdução de máquinas simples para otimizar o trabalho humano, mas evoluiu com o avanço dos sistemas de controle digital. Esses sistemas permitiram que os processos industriais fossem gerenciados automaticamente, alcançando resultados desejados sem a necessidade de intervenção manual constante.

O termo “automação” foi formalmente definido pela Radio-Electronics-Television Manufacturers Association (RETMA) em 1956, como a técnica de aumentar a produtividade humana no processamento de materiais, energia e informações, utilizando, em diferentes graus, elementos de autocontrole. Isso envolve o uso de tecnologias que permitem a execução automática de tarefas, sem intervenção manual constante, por meio de sistemas que programam e controlam a produção de forma autônoma, promovendo eficiência nos processos industriais (RETMA, 1956).

De um ponto de vista mais moderno, automação pode ser definida de forma geral como o controle assistido por computador de processos industriais, onde diversos sistemas de controle são empregados para operar máquinas, processos fabris e outras tarefas complexas sem a necessidade de intervenção humana direta. A anatomia de um sistema de automação industrial envolve diversos componentes essenciais que trabalham juntos para garantir o controle e a execução eficiente de processos. Hetmanczyk e Swider (2012) destacam que os sistemas modernos de automação industrial geralmente consistem em controladores (PLCs, reguladores, computadores industriais), atuadores (elétricos, pneumáticos, hidráulicos), equipamentos de controle e medição (sensores e detectores), redes industriais e sistemas SCADA/HMI, que trabalham de forma integrada para garantir o controle eficaz e a operação dos processos. A anatomia de um sistema de automação industrial é ilustrada na Figura 2.

Figura 2 – Anatomia de um sistema de automação



Fonte: os autores.

De acordo com a revisão apresentada por Singh e Namekar (2020), os sistemas de automação são compostos por dispositivos como Controladores Lógicos Programáveis (PLCs) e Controladores de Automação Programáveis (PACs), que são responsáveis por gerenciar e instruir as máquinas a realizar tarefas específicas. O PLC, por exemplo, é um computador digital especializado que opera sem exibição ou teclado, instalado em painéis de controle no chão de fábrica. Esses dispositivos são programados para controlar processos eletromecânicos, como o funcionamento de maquinário industrial. O sistema também inclui sensores, que monitoram o estado do processo, atuadores, que realizam ações com base nas instruções do controlador, e *softwares* de supervisão, que permitem o monitoramento e ajustes necessários por parte dos operadores humanos. Esse conjunto de elementos possibilita a automação de processos, reduzindo a necessidade de intervenção humana e aumentando a precisão e eficiência na produção industrial (SINGH; NAMEKAR, 2020).

Além disso, a automação promove a melhoria da qualidade e precisão dos produtos finais, garantindo que os padrões sejam mantidos com consistência ao longo do tempo (FLOYD, 2011). Um dos grandes benefícios da automação é a capacidade de realizar tarefas perigosas ou repetitivas, reduzindo a exposição dos trabalhadores a riscos e contribuindo para a segurança no ambiente industrial. Isso também abre espaço para que os trabalhadores humanos se concentrem em funções mais analíticas e criativas. Com a automação, as empresas conseguem não apenas aumentar sua competitividade, mas também responder de maneira mais rápida e eficiente às demandas de um mercado global em constante mudança.

Nesse contexto, a robótica tem um papel central na automação, permitindo a execução de tarefas com precisão e velocidade superiores à capacidade humana, além de possibilitar operações em ambientes de risco (IVANOV, KUYUMDZHIEV; WEBSTER, 2020).

De acordo com a International Federation of Robotics (2019), 422.271 novos robôs industriais foram instalados globalmente em 2018, elevando o número total de robôs em operação para 2.439.543. Esses sistemas automatizados podem ser classificados em três categorias: automação fixa, programável e flexível, cada uma atendendo a diferentes necessidades produtivas. A automação fixa, por exemplo, é caracterizada por altas taxas de produção e baixa flexibilidade, sendo usada para produtos com alta demanda (IVANOV, KUYUMDZHIEV & WEBSTER, 2020). Por outro lado, a automação programável permite ajustes nos processos de produção e é ideal para produção em lotes, enquanto a automação flexível possibilita alterações nos produtos sem interrupções, atendendo mercados com alta demanda por variedade (IVANOV; KUYUMDZHIEV; WEBSTER, 2020).

A decisão de automatizar muitas vezes se dá pela necessidade de redução de custos, aumento da produtividade ou eliminação de riscos à segurança humana (IVANOV, KUYUMDZHIEV; WEBSTER, 2020). A automação forçada ocorre quando o trabalho humano precisa ser substituído devido aos perigos da atividade (BAINBRIDGE, 1983), enquanto a automação economicamente justificada busca melhorar a eficiência e a qualidade, reduzindo desperdícios e custos operacionais (HUANG; RUST, 2018).

No campo da robótica industrial, segundo Ivanov, Kuyumdzhev & Webster, (2020), há uma distinção clara entre grande e pequena automação. A grande automação envolve robôs industriais que podem manipular grandes cargas e operar em alta velocidade, o que é comum em linhas de produção em massa. Já a pequena automação utiliza robôs colaborativos, que trabalham em proximidade com humanos, e são mais adequados para processos flexíveis.

Frey e Osborne (2013) estimaram que 47% dos empregos nos Estados Unidos estão em risco de automação nos próximos anos, especialmente nas áreas de transporte, logística e suporte administrativo. Esses dados refletem a crescente preocupação com a substituição do trabalho humano por máquinas, sobretudo em setores onde as tarefas são altamente repetitivas e previsíveis (FREY; OSBORNE, 2017). No entanto, outros estudos, como o de Autor (2015), apontam que a automação também pode criar novas oportunidades de emprego, especialmente em áreas que exigem habilidades mais avançadas, como a gestão e a programação de sistemas automatizados. Autor (2015) ressalta também que a automação não elimina necessariamente postos de trabalho, mas transforma a natureza do trabalho, deslocando a força de trabalho para tarefas mais complexas e menos repetitivas.

2.2 Inteligência Artificial

Segundo Jones *et al* (2018), os anos 1980 marcaram um ressurgimento das redes neurais, impulsionado pela introdução da retropropagação, permitindo o treinamento eficiente de redes complexas e renovando o interesse no campo.

Abraham (2002) destaca que a IA teve suas bases estabelecidas em 1943 por Warren McCulloch e Walter Pitts, com a criação de um modelo de neurônios artificiais que combinava lógica e neurologia para descrever a atividade neural. Abraham (2002) ainda salienta que esse modelo marcou um avanço na cibernética e ciência cognitiva.

Segundo Mira (2008), em 1949, Donald Hebb introduziu o conceito de aprendizado Hebbiano, associando a força das sinapses à repetição de estímulos, consolidando um pilar essencial para a teoria de redes neurais e para a compreensão do aprendizado adaptativo.

Hartman (1987) relata que a formalização da IA ocorreu na Conferência de Dartmouth em 1956, quando John McCarthy cunhou o termo “inteligência artificial”. Hartman (1987) ainda menciona que nesse evento McCarthy, junto a Marvin Minsky, Allen Newell e Herbert Simon, discutiu a viabilidade de simular aspectos da inteligência humana em máquinas.

Abraham (2002) observa que, apesar dos avanços nos anos 1960, Minsky e Papert, em 1969, destacaram as limitações dos perceptrons em resolver funções lógicas não lineares, como o XOR, o que levou a um período de estagnação nas pesquisas em redes neurais. Hartman (1987) observa que a década de 1970 trouxe uma mudança de foco para sistemas baseados em conhecimento, exemplificados pelos sistemas especialistas DENDRAL e MYCIN. Jones *et al* (2018) apontam que a década de 2010 foi significativa pela disponibilidade de grandes volumes de dados e pelo uso de GPUs, o que possibilitou o avanço das redes neurais profundas e consolidou a IA em diversos setores.

Dörrich *et al* (2023) apontam que, com a crescente complexidade dos modelos de IA, o consumo de energia e os custos computacionais aumentaram significativamente. Estes pesquisadores observam que técnicas como o treinamento em precisão mista oferecem um ganho de até 1,9× na eficiência de treinamento e uma aceleração de até 6× na inferência, especialmente ao utilizar TPUs Edge para modelos menores, mantendo a precisão e reduzindo o impacto ambiental. Esse progresso reflete uma tendência atual de integrar eficiência energética no desenvolvimento de IA, como observado na "Green AI".

Hoje, a combinação de grandes modelos de linguagem e GPUs avançadas permite processar vastas quantidades de dados rapidamente, aumentando a precisão em tarefas complexas. Dörrich *et al* (2023) também destacam que o uso de dispositivos como TPUs Edge e técnicas de quantização têm potencial para tornar a IA mais acessível e sustentável, mantendo alta performance mesmo com recursos limitados.

El-Had (2023) descreve os seis principais domínios da Inteligência Artificial que são organizados no mapa conceitual da Figura 3.

Figura 3 – Domínios da inteligência artificial



Fonte: os autores, baseados em El-Had (2023)

2.3 Grandes Modelos de Linguagem (LLMs)

A transformação impulsionada pelo aumento dos grandes conjuntos de dados e da capacidade computacional permitiu o desenvolvimento de modelos mais avançados. Um dos principais beneficiários desses avanços foi o campo do Processamento de Linguagem Natural (PLN), onde os Modelos de Linguagem de Grande Escala (LLMs) se tornaram amplamente utilizados. HOU *et al* (2023) destacam que esses modelos são treinados com dados de diversas fontes como livros, artigos e websites e são capazes de resolver tarefas de linguagem porque, durante o treinamento, aprendem a identificar padrões e relações entre palavras e frases.

Thangarasa *et al* (2023) destacam que modelos de grande escala, como o GPT-3, apresentam desempenho impressionante em processamento de linguagem natural devido ao aumento no tamanho dos modelos, dos conjuntos de dados e da capacidade computacional, embora seu treinamento exija recursos significativos. A computação paralela, utilizando frameworks como Hadoop e Storm, revolucionou o processamento de grandes volumes de texto, permitindo análises em tempo real. O Storm distribui operações de processamento em

diferentes máquinas, executando módulos de PLN em paralelo e aumentando a eficiência (AGERRI *et al*, 2015). Experimentos demonstraram que paralelizar módulos intensivos, como a rotulagem de papéis semânticos, pode aumentar a velocidade de processamento em até 63%, o que é essencial para aplicações que exigem alta performance e baixa latência.

Segundo Hou *et al* (2023), os Modelos de Linguagem de Grande Escala (LLMs) classificam-se em três arquiteturas principais. Os modelos *encoder-only*, como o BERT, utilizam apenas a parte *encoder* do transformador para compreender textos, sendo adequados para tarefas de compreensão, como classificação e análise semântica. Os modelos *encoder-decoder*, como o T5, empregam ambas as partes para transformar sequências de entrada em saídas, sendo úteis em tarefas como tradução e sumarização; Wang *et al* (2023) destacam a flexibilidade do CodeT5 em tarefas de programação. Já os modelos *decoder-only*, como o GPT-3 e GPT-4, concentram-se na geração sequencial de texto, prevendo o próximo *token* com base nos anteriores, ideais para criação de conteúdo. Fu *et al* (2023) exploram as vantagens e desafios de aplicar modelos *decoder-only* a tarefas complexas, apontando limitações na manutenção da atenção em sequências longas.

Na engenharia de software, inicialmente predominavam modelos *encoder-only*, mas desde 2021 houve um aumento no uso de modelos *decoder-only*, sendo amplamente utilizados a partir de 2023, principalmente para geração de código. Modelos como GPT e Codex são exemplos notáveis dessa tendência, refletindo o contínuo desenvolvimento de arquiteturas de LLMs na área. Espera-se que novos modelos, como LLaMA e Claude, também sejam integrados em tarefas específicas de engenharia de software (HOU *et al*, 2023).

A exposição à automação por sistemas baseados em inteligência artificial tornou-se um assunto extensivamente discutido com a popularização em massa do chat GPT, sendo o produto com maior crescimento na base de usuários na história, de acordo com publicação de Hu (Reuters) (2023).

No relatório de Matt Schendstok e Sydney Schreiner Wertz do U.S. Department of the Treasury (2024), é analisada a exposição à IA, focando especialmente na intensidade com que diferentes áreas de conhecimento são afetadas pela inteligência artificial. Os resultados mostram que indivíduos com pelo menos um diploma de bacharel estão mais expostos à IA em suas ocupações do que aqueles sem esse nível de educação. Entre os graduados, há variações significativas na exposição dependendo do campo de estudo.

Áreas relacionadas a STEM (Ciência, Tecnologia, Engenharia e Matemática) e cursos com maiores retornos salariais estão entre as mais expostas. Por exemplo, 66% dos graduados em Ciências Atuariais estão em ocupações altamente expostas à IA. Da mesma forma, 64%

dos formados em Farmácia e Ciências Farmacêuticas e 63% em Engenharia Aeroespacial enfrentam altos níveis de exposição. Outros cursos com elevada exposição incluem Arquitetura (61%), Engenharia Civil (60%) e Engenharia Mecânica (60%). Em contraste, áreas como Educação e Enfermagem apresentam menor exposição. Apenas 9% dos formados em Educação Especial e Administração e Supervisão Educacional estão em ocupações altamente expostas à IA. Enfermagem registra uma exposição ainda menor, com apenas 6% de seus graduados trabalhando em ocupações altamente expostas. A razão pela qual algumas áreas são mais expostas que outras está ligada à natureza das atividades desempenhadas nessas ocupações. (SCHENDSTOK; WERTZ, 2024)

Profissões em campos como engenharia, tecnologia da informação e ciências exatas envolvem tarefas analíticas, processamento de dados e resolução de problemas complexos, áreas nas quais a IA tem demonstrado grande potencial de atuação e substituição (WANG *et al*, 2024). Adicionalmente, países com economias avançadas, como os Estados Unidos e o Reino Unido, têm um número maior de trabalhadores em ocupações com alta exposição à IA, enquanto economias emergentes apresentam menor exposição devido à sua composição ocupacional, com maior presença em setores agrícolas e serviços menos impactados pela IA (PIZZINELLI *et al*, 2023). A categorização das habilidades cognitivas, como processamento visual e interação sensório-motora, permite compreender melhor as áreas onde a IA está direcionando maior esforço de pesquisa, afetando mais intensamente ocupações que demandam habilidades analíticas e resolução de problemas (TOLAN *et al*, 2021). Essas tarefas são mais facilmente automatizadas ou complementadas por sistemas de IA, aumentando a exposição dos profissionais a possíveis mudanças no mercado de trabalho.

2.4 Integrando LLMs em sistemas industriais

Devido à natureza recente dos produtos baseados em Modelos de Linguagem de Grande Porte, ainda não existem protocolos estabelecidos para integrá-los em sistemas industriais. Nesse contexto, o estudo de Urlana *et al* (2024) aborda a lacuna na integração de LLMs em aplicações industriais, adotando uma abordagem em duas etapas: inicialmente, conduzem uma pesquisa com profissionais da indústria para entender como os LLMs estão sendo utilizados na prática e quais desafios enfrentam; em seguida, analisam 68 artigos relevantes para responder a questões-chave sobre o uso de LLMs no contexto industrial.

O estudo identifica quatro desafios principais na adoção de LLMs pela indústria: os elevados requisitos computacionais dos modelos; preocupações com privacidade e segurança

dos dados; a falta de protocolos e padrões claros para a implementação e integração desses modelos em sistemas industriais existentes; e desafios regulatórios e de infraestrutura decorrentes de regulamentações como o GDPR na Europa e o CCPA na Califórnia, que impõem restrições ao uso de LLMs. Além disso, observa-se uma distribuição desigual nas aplicações dos LLMs, com maior foco em tarefas padrão de processamento de linguagem natural, geração de código e desenvolvimento de ferramentas, enquanto áreas como segurança e impacto social recebem menos atenção. Os autores também destacam a importância de considerar atributos de confiança, como interpretabilidade, robustez e mitigação de vieses (URLANA *et al*, 2024).

Em resposta a esses desafios, Lee e Su (2024) discutem um *framework* unificado para Modelos de Conhecimento Industrial de Grande Escala (ILKMs) na manufatura inteligente. Os ILKMs combinam o poder dos LLMs com conhecimento específico do domínio industrial, adaptando e especializando os modelos para aplicações práticas na indústria. Essa combinação ocorre por meio de um processo estruturado que inclui a construção de uma Grande Biblioteca de Conhecimento (LKL) com dados específicos da indústria, a preparação de instruções do domínio e o desenvolvimento de um LLM ajustado com esse conhecimento especializado.

Além disso, Lee e Su (2024) introduzem o “Princípio 6S” como diretriz para o desenvolvimento de ILKMs, englobando: Conhecimento Especializado (*Specialized Domain Knowledge*), Scrutabilidade, Segurança (*Safety*), Escalabilidade (*Scalability*), Sustentabilidade (*Sustainability*) e Sistematização & Padronização (*Systematization & Standardization*). Esses princípios são essenciais para garantir que os ILKMs possam ser integrados de forma eficaz nos sistemas industriais, atendendo às necessidades específicas da indústria e assegurando conformidade com normas e regulamentos. Com a aplicação do Princípio 6S, os ILKMs não apenas oferecem soluções precisas e relevantes para desafios complexos—como manutenção preditiva, otimização de processos e controle de qualidade—mas também promovem confiabilidade, segurança e eficiência operacional. (LEE; SU, 2024)

Xia *et al* (2023) propõem uma estrutura que integra LLMs, gêmeos digitais e sistemas de automação industrial para o planejamento e controle inteligentes de processos de produção. Eles adaptam o sistema de automação de uma instalação modular para criar interfaces executáveis de funcionalidades em diferentes granularidades, registradas em um gêmeo digital com informações descritivas do sistema de produção. Agentes baseados em LLMs interpretam essas descrições e controlam o sistema físico por meio das interfaces, planejando

e executando a produção de forma autônoma. O protótipo demonstrou a capacidade de lidar com tarefas não previstas, evidenciando o potencial dos LLMs em fábricas inteligentes.

Complementando essa abordagem, González-Santamarta *et al* (2023) exploram a integração de LLMs na arquitetura cognitiva MERLIN2 para robôs autônomos. A MERLIN2 é uma arquitetura híbrida que combina sistemas deliberativos, responsáveis por planejamento e tomada de decisões antecipadas, e sistemas reativos, que respondem rapidamente a mudanças no ambiente. Tradicionalmente, a MERLIN2 utiliza planejadores simbólicos baseados em PDDL (*Planning Domain Definition Language*), mas esse processo pode ser complexo e propenso a erros. Os autores substituem o PDDL por LLMs, permitindo a geração de planos a partir de comandos em linguagem natural, o que simplifica a definição de tarefas e amplia a adaptabilidade do robô a diferentes contextos. Isso destaca o potencial dos LLMs em tornar sistemas robóticos mais flexíveis e fáceis de usar em ambientes industriais.

Wang *et al* (2023) investigam o uso de LLMs como controladores de baixo nível para robôs, focando na capacidade de gerar comandos de movimentação de forma autorregressiva. Nos experimentos, o modelo foi testado em diferentes robôs e ambientes simulados, incluindo terrenos planos e acidentados, para avaliar sua capacidade de interpretar dados sensoriais e ajustar as posições das juntas em tempo real. Os *prompts* textuais usados detalhavam a tarefa de caminhar e incluíam informações como a descrição da tarefa, especificação das entradas (velocidades e posições) e saídas (posições-alvo das juntas), além da ordem das juntas do robô. O LLM foi integrado com um controlador Proporcional-Derivativo, que opera em alta frequência para rastrear as posições geradas. Os resultados mostraram que a abordagem permitiu ao robô caminhar de forma autônoma, demonstrando a viabilidade dos LLMs como controladores de feedback em sistemas complexos, sem a necessidade de treinamento especializado.

Apesar dos resultados promissores encontrados por Wang *et al* (2023), os autores apontam que o design do *prompt* é sensível a pequenas variações, o que pode impactar o desempenho do robô. A abordagem também não foi testada em *hardware* real devido à baixa velocidade de inferência do GPT-4, que obriga a pausar a simulação e limita a frequência de operação a 10 Hz, reduzindo a precisão em comparação com controladores mais rápidos. Além disso, o limite de *tokens* do LLM restringe o uso de um histórico de observações e ações mais longo, o que limita a capacidade de controle em cenários complexos.

A aplicação de LLMs na automação industrial, especialmente na geração de código para PLCs, tem ganhado destaque. Fakhri *et al* (2024) propuseram a metodologia LLM4PLC, que integra LLMs com ferramentas de verificação externa e *feedback* do usuário para gerar

código verificável e confiável de forma iterativa. Nessa abordagem, o LLM recebe especificações em linguagem natural e gera um plano de design baseado em modelos, produzindo código em *Structured Text* (ST), conforme a norma IEC 61131-3. O código passa por verificações sintáticas e formais, sendo refinado até atender às especificações. Os experimentos mostraram um aumento na taxa de sucesso de geração de código de 47% para 72%, evidenciando o potencial dos LLMs em auxiliar engenheiros em sistemas industriais complexos.

Complementando essa pesquisa, Koziolk *et al* (2023) exploraram o uso do ChatGPT no desenvolvimento de lógica de programação para PLCs e Sistemas de Controle Distribuído (DCSs) na linguagem IEC 61131-3. Em um estudo com 100 exemplos de *prompts*, descobriram que o ChatGPT foi capaz de gerar código funcional para tarefas como conversão de valores decimais em hexadecimais e controle de semáforos com transições entre modos. Apesar de algumas limitações devido à especificidade dos *prompts*, o código gerado foi considerado adequado. O ChatGPT também demonstrou competência na geração de código para algoritmos PID de controle e forneceu códigos esqueleto úteis para aplicações mais complexas.

Por essa capacidade de geração de código, debates sobre a substituição de engenheiros de controle pelo uso de LLMs emergem em publicações como a de Levanduski (2023), que questiona se o ChatGPT pode ser utilizado para programar sistemas de controle como PLCs. Em testes práticos, Levanduski (2023) solicitou a geração de lógicas de controle em diferentes linguagens, como *ladder logic* e texto estruturado. Embora o ChatGPT tenha demonstrado a capacidade de fornecer código básico, os resultados apresentaram limitações que evidenciam a necessidade de validação por engenheiros experientes. Por exemplo, em um circuito de travamento em *ladder logic*, o código gerado não estava totalmente correto, especialmente em relação à estrutura de endereçamento compatível com versões mais antigas de software, indicando que, apesar do potencial de apoio, o ChatGPT ainda não substitui o conhecimento especializado dos engenheiros.

Assim, conforme discutido por Levanduski (2023), embora o ChatGPT consiga gerar lógicas básicas e sirva como ferramenta de aprendizado e suporte, há limitações claras em tarefas complexas que demandam conhecimento prático e ajustes específicos. A conclusão é que, apesar de não eliminar a necessidade de supervisão humana, os LLMs podem acelerar a aquisição de conhecimento e a resolução de problemas mais simples, atuando como aliados dos engenheiros de controle na programação de sistemas industriais.

Uma interseção entre os artigos discutidos é a frequente pontuação à necessidade de supervisão ou validação humana da saída (seja ela um código, texto, comando) de um LLM. Segundo Dowdeswell *et al* (2020), sistemas industriais que utilizam IA, especialmente em ambientes de diagnóstico de falhas, precisam de técnicas robustas de supervisão para garantir que decisões tomadas por IA não comprometam a segurança operacional. Dowdeswell *et al* (2020) destacam que, embora técnicas baseadas em IA tenham avançado, a confiança nas suas saídas ainda depende de sistemas híbridos que integram supervisão humana contínua (Dowdeswell *et al*, 2020).

Dessa forma, faz-se necessário estudar a confiabilidade dessa tecnologia em cenários sensíveis como sistemas de automação industrial. Alucinações em LLMs referem-se a respostas imprecisas ou afirmações não fundamentadas que os modelos geram, muitas vezes sem indicar qualquer incerteza. Isso é particularmente preocupante em cenários críticos onde a precisão é essencial. Por exemplo, Friel e Sanyal (2023) observam que, em tarefas como responder a perguntas sobre a COVID-19 ou resolver problemas matemáticos complexos, os LLMs podem fornecer respostas erradas que parecem confiáveis. Para mitigar esse problema, Friel e Sanyal (2023) propuseram o *ChainPoll*, uma metodologia que utiliza cadeias de raciocínio geradas pelo próprio modelo para avaliar a veracidade das respostas.

No contexto financeiro, Roychowdhury *et al* (2023) destacam os desafios de aplicar LLMs devido ao risco de alucinações em dados numéricos críticos. Modelos como o GPT-3 podem gerar respostas incorretas sobre métricas financeiras, como atribuir erroneamente a taxa de crescimento do PIB de um país a outro, o que pode levar a decisões equivocadas. Para mitigar esse problema, eles desenvolveram um framework que transforma tabelas em “blocos de dados” textuais hierárquicos, permitindo que o LLM acesse informações estruturadas mais efetivamente. Além disso, implementaram um módulo que avalia a qualidade das respostas em seis métricas, incluindo detecção de alucinações numéricas e coerência contextual, reduzindo significativamente a incidência de erros e aumentando a confiança nas informações apresentadas.

Em sistemas de automação industrial, a confiabilidade das saídas dos LLMs é crucial, pois erros podem levar a falhas operacionais ou riscos à segurança. Ogundare *et al* (2023) estudam a resiliência e eficiência de sistemas de controle automatizados gerados por LLMs, utilizando a teoria da percolação para modelar e estimar a resiliência do sistema. Aplicando técnicas de otimização estocástica e análise de arrependimento, buscam soluções próximas ao ideal com desempenho comprovado. No estudo de caso, avaliam um sistema projetado pelo ChatGPT, composto por bombas acopladas, um tanque de mistura e válvulas de água. A

metodologia compara a solução gerada pelo LLM com os objetivos estabelecidos, destacando a necessidade de métodos robustos para garantir a confiabilidade de sistemas gerados automaticamente na indústria.

Yang *et al* (2021) exploram as limitações dos LLMs em comparação a bases de conhecimento estruturadas, como os grafos de conhecimento (KGs). Os LLMs, embora sejam capazes de aprender padrões linguísticos probabilísticos, enfrentam dificuldades significativas na geração de conteúdos que exijam precisão factual. Isso se deve, em parte, à incapacidade de acessar diretamente conhecimentos armazenados durante o treinamento, o que resulta em limitações na lembrança de fatos específicos e gera desafios na recuperação e aplicação precisa dessas informações em contextos diversos. Os autores destacam que os LLMs são particularmente propensos a alucinações (definidas pelos autores como gerar informações que não estão fundamentadas em fatos concretos ou que podem contradizer fontes confiáveis). Isso acontece porque esses modelos aprendem com base em correlações de dados ao invés de raciocínios explícitos e estruturados. Como resultado, a falta de uma estruturação clara de informações dificulta o uso de LLMs para tarefas que exigem precisão factual, especialmente quando comparados com os KGs, que organizam o conhecimento de forma explícita, estabelecendo relações claras entre entidades.

Para abordar essas limitações, a proposta discutida por Yang *et al* (2021) envolve a integração dos KGs durante diferentes estágios de treinamento dos LLMs. Esse processo pode ocorrer em três fases principais: antes do treinamento, durante o treinamento ou após o treinamento dos modelos. Cada uma dessas abordagens visa mitigar problemas específicos. Por exemplo, a utilização de dados provenientes de KGs antes do treinamento ajuda a aumentar a qualidade semântica e a capacidade de raciocínio dos LLMs sem necessariamente aumentar o tempo de treinamento. No entanto, também apresenta desafios, como a introdução de ruídos no processo, caso o conhecimento adicionado não seja diretamente relevante ao contexto original.

Outra vertente que limita as aplicações de LLMs é a incapacidade desses modelos de gerar conteúdo original ou inovar de maneira autônoma. Segundo Everitt *et al* (2018), embora a Inteligência Artificial Geral (AGI) seja uma aspiração no campo da IA, os modelos atuais são especializados em tarefas restritas e carecem da capacidade de raciocinar ou inventar autonomamente. Os LLMs operam com base em padrões pré-existentes de dados e, embora possam ser eficazes em tarefas específicas, não são capazes de “inventar” ou propor soluções novas além do que já foi registrado em seus dados de treinamento (EVERITT *et al*, 2018).

Nessa mesma linha, Girotra *et al* (2023) destacam que, apesar da alta produtividade dos LLMs na geração de ideias, essas ferramentas frequentemente apresentam respostas inconsistentes e sem verificação de precisão. A geração de ideias por LLMs se baseia em padrões pré-existentes nos dados de treinamento, resultando em conteúdo que é estatisticamente provável, mas que pode não ser factualmente correto. No entanto, essa inconsistência pode ser uma vantagem em contextos criativos, onde a variabilidade e a quantidade são mais valorizadas do que a consistência. Em estudos comparativos, ideias geradas por LLMs, como o ChatGPT, mostraram maior qualidade média quando comparadas a ideias de estudantes universitários, embora apresentassem menor novidade. Essa evidência reforça que, embora os LLMs possam gerar soluções eficazes em contextos específicos, sua capacidade de inovação autônoma continua limitada aos dados existentes e à falta de julgamento crítico inerente ao seu funcionamento.

Em complemento, Bazett (2024) discute que, embora modelos como o ChatGPT tenham aprimorado significativamente suas habilidades em problemas de matemática, eles ainda apresentam limitações em questões que exigem raciocínio matemático não convencional. Um exemplo específico abordado por Bazett (2024) ocorre em problemas que envolvem a escolha do menor inteiro cujo quadrado está dentro de um intervalo definido. Quando solicitado a encontrar o menor inteiro cujo quadrado está entre 15 e 30, modelos como ChatGPT e Gemini frequentemente respondem com “4”, uma resposta incorreta, pois ignoram a possibilidade de inteiros negativos. Nesse caso, o menor inteiro que satisfaz a condição é -5 que está dentro do intervalo especificado. Essa limitação ilustra que, apesar de seu treinamento em vastos bancos de dados matemáticos, como o GSM 8K, os LLMs ainda têm dificuldade em aplicar raciocínio fora do padrão aprendido, especialmente quando o problema exige a consideração de alternativas não convencionais (BAZETT, 2024).

3. METODOLOGIA

A metodologia aplicada nesta pesquisa seguiu as diretrizes para revisão integrativa da literatura, conforme proposto por Torraco (2005), que define esta abordagem como uma forma de pesquisa que sintetiza e critica a literatura existente, proporcionando novos entendimentos e perspectivas sobre um tema. A revisão integrativa é especialmente valiosa por permitir a combinação de diferentes métodos de pesquisa, tanto qualitativos quanto quantitativos, o que a torna uma ferramenta robusta para a prática baseada em evidências

(WHITTEMORE *et al*, 2005). Para assegurar o rigor metodológico, a revisão foi estruturada em etapas sequenciais de busca, leitura, anotação, análise crítica e redação, garantindo uma abordagem sistemática na construção do conhecimento (HOPIA; LATVALA; LIIMATAINEN, 2016).

3.1 Busca

A busca por literatura relevante foi realizada utilizando as plataformas Scopus e Perplexity AI. No Scopus, foram utilizadas combinações das palavras-chave “LLM”, “*Manufacturing*”, “*Industrial Engineering*”, “*Automation*”, “*Artificial Intelligence*” e “*Social Impacts*” para identificar estudos que abordassem os impactos de modelos de linguagem em contextos industriais. A inclusão de plataformas adicionais, como o Perplexity AI, possibilitou a ampliação do escopo da busca, utilizando prompts baseados no abstract e palavras-chave de artigos previamente selecionados. Essa abordagem diversificada permitiu uma coleta mais abrangente de dados, alinhando-se às recomendações de Whittemore e Knafl (2005) para maximizar a precisão e a exaustividade na fase de busca em revisões integrativas.

3.2 Leitura, anotação e distribuição em equipe

Os artigos coletados foram submetidos a uma triagem inicial através da leitura dos resumos para confirmar sua relevância. Em seguida, os estudos selecionados foram distribuídos entre os membros da equipe para leitura integral e anotação detalhada. Este processo colaborativo visou garantir uma análise diversificada e abrangente, permitindo uma triangulação de dados que melhora a validade dos achados (WHITTEMORE; KNAFL, 2005). A divisão de tarefas entre os membros da equipe buscou garantir que diferentes perspectivas fossem consideradas.

3.3 Análise crítica

A análise crítica dos artigos foi realizada com base nas diretrizes de Whittemore e Knafl (2005), que sugerem a avaliação da robustez metodológica dos achados principais e das contribuições teóricas de cada estudo. Essa etapa permitiu identificar tanto os benefícios, como melhorias em produtividade e eficiência, quanto os desafios, como questões éticas e barreiras técnicas associadas à aplicação de inteligência artificial em ambientes industriais. A

análise crítica foi essencial para identificar lacunas na literatura e propor novas direções de pesquisa.

4. RESULTADOS E DISCUSSÕES

A revisão bibliográfica apresentada oferece um panorama abrangente sobre a evolução da automação industrial e a integração dos Modelos de Linguagem de Grande Escala (LLMs) nesse contexto. No entanto, é necessário destacar pontos críticos que podem limitar a aplicabilidade prática dos LLMs em cenários de automação crítica (URLANA *et al*, 2024).

Em primeiro lugar, observa-se que a pesquisa em LLMs tem se direcionado majoritariamente para o auxílio humano em tarefas complexas, em vez de buscar a substituição total do trabalho humano. Embora esses modelos tenham demonstrado capacidade de gerar ideias e soluções, eles frequentemente carecem de precisão factual e inovação autônoma. Isso limita sua eficácia em tarefas que exigem julgamento crítico e criatividade, reforçando a ideia de que os LLMs atuam como ferramentas complementares que ampliam as capacidades humanas, mas não as substituem (GIROTRA *et al*, 2024; EVERITT *et al*, 2018).

Além disso, os LLMs enfrentam desafios técnicos significativos, como o limite de tokens e a capacidade de manter um contexto coerente em interações prolongadas. Em sistemas de controle em tempo real, a baixa velocidade de processamento e as restrições de memória podem comprometer a continuidade e a precisão, fatores cruciais para a operação segura e eficiente em ambientes industriais (WANG *et al*, 2024).

Questões legais e de segurança também emergem como obstáculos importantes na integração de LLMs em sistemas industriais. Regulamentações rigorosas impõem restrições ao uso de dados e algoritmos de inteligência artificial, levantando preocupações relacionadas à privacidade e à proteção de informações sensíveis. Em setores onde a segurança e a confidencialidade são vitais, depender exclusivamente de LLMs sem supervisão humana pode ser arriscado (URLANA *et al*, 2024).

A confiabilidade das respostas geradas pelos LLMs é outro ponto de atenção. A ocorrência de respostas imprecisas ou infundadas pode levar a falhas operacionais graves. Em contextos onde a margem para erro é mínima, como na automação industrial, é indispensável a supervisão e validação por profissionais qualificados para garantir a funcionalidade e a segurança dos sistemas (FRIEL, SANYAL, 2023; DOWDESWELL *et al*, 2020).

Adicionalmente, apesar de serem eficazes em tarefas baseadas em padrões existentes, os LLMs apresentam limitações quanto à capacidade de inovação e adaptação em situações que exigem soluções originais. A ausência de inteligência geral impede que esses modelos raciocinem ou criem além dos dados nos quais foram treinados, o que é uma desvantagem em ambientes industriais dinâmicos que demandam criatividade e flexibilidade (EVERITT *et al*, 2018; BAZETT, 2024).

5. CONCLUSÃO

A incorporação de Modelos de Linguagem de Grande Escala na automação industrial tem potencial para revolucionar os processos de produção e redefinir as dinâmicas laborais. Equipando máquinas com processamento avançado de linguagem, os LLMs podem otimizar operações, melhorar decisões e facilitar interações mais intuitivas entre humanos e sistemas automatizados, aumentando a eficiência operacional e promovendo colaboração sinérgica entre a expertise humana e a precisão das máquinas.

Contudo, a adoção desses modelos em ambientes industriais críticos apresenta desafios técnicos, como a necessidade de altos recursos computacionais e dificuldades na manutenção de contextos prolongados, que podem afetar a confiabilidade e a segurança das operações. Preocupações legais e éticas relacionadas à privacidade de dados e conformidade regulatória exigem uma implementação cuidadosa. A incapacidade atual dos LLMs de inovar autonomamente reforça a necessidade de supervisão humana contínua, garantindo que atuem como aliadas, não substitutas, dos profissionais especializados.

É fundamental equilibrar os benefícios dos LLMs com precauções para mitigar riscos. Investir na requalificação da força de trabalho e promover práticas éticas assegura que a integração dos LLMs contribua positivamente para a sociedade. Ao enfrentar os desafios técnicos e éticos proativamente, a indústria pode aproveitar o potencial dos LLMs, impulsionando eficiência e inovação sem comprometer a segurança e integridade dos sistemas.

Futuras pesquisas devem explorar soluções que superem as limitações técnicas dos LLMs, especialmente em eficiência computacional (buscando reduzir os custos de treinamento dos LLMs) e manutenção de contexto (para evitar alucinações). Também é importante investigar como aprimorar a integração entre LLMs e robótica colaborativa, promovendo interações mais seguras e eficientes entre humanos e máquinas. Estudos que ampliem palavras-chave em revisões da literatura, incluindo termos como "robótica",

"colaborativo" e "cobots", serão valiosos para abarcar a crescente importância dos robôs colaborativos na automação industrial. Pesquisas focadas em desenvolver frameworks que facilitem a colaboração entre LLMs, robôs e operadores humanos, bem como em estabelecer diretrizes éticas e regulatórias específicas, serão necessárias para maximizar os benefícios e minimizar os riscos dessa tecnologia emergente.

REFERÊNCIAS

ARAQUE-GONZÁLEZ, G *et al.* Sustainable manufacturing in the fourth industrial revolution: A big data application proposal in the textile industry. **Journal of Industrial Engineering and Management**, v. 15, n. 4, p. 614-636, 2022. Disponível em: <https://doi.org/10.3926/jiem.3922>.

VARSHNEY, A. K. *et al.* Development of Manufacturing Industry Revolutions from 1.0 to 5.0. **Journal of Informatics Education and Research**, v. 4, n. 1, p. 1230, 2024. ISSN 1526-4726.

ABDULRAHMAN, Al-Abassi *et al.* Industrial Big Data Analytics: Challenges and Opportunities. In: 2019. p. 37-61. doi: 10.1007/978-3-030-38557-6_3.

YOON, A; KIM, K. Y. World Economic Forum: Why Large Language Models are the future of manufacturing. World Economic Forum, 23 abr. 2024. Disponível em: <https://www.weforum.org/stories/2024/04/why-large-language-models-are-so-important-for-the-future-of-the-manufacturing-industry/>. Acesso em 06 set. de 2024.

TORRACO, R. Writing. Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review*, v. 4, p. 356-367, 2005. Disponível em: <https://doi.org/10.1177/1534484305278283>.

WHITTEMORE, R.; KNAFL, K. The integrative review: updated methodology. **Journal of Advanced Nursing**, v. 52, n. 5, p. 546-553, 2005. Disponível em: <https://doi.org/10.1111/j.1365-2648.2005.03621.x>.

HOPIA, H.; LATVALA, E.; LIIMATAINEN, L. Reviewing the methodology of an integrative review. **Scandinavian Journal of Caring Sciences**, v. 30, n. 4, p. 662-669, 2016. Disponível em: <https://doi.org/10.1111/scs.12327>.

RETMA gives definition of "automation". *Electrical Engineering*, v. 75, n. 8, p. 770-771, ago. 1956. doi: 10.1109/EE.1956.6442099

SINGH, S; NAMEKAR, S. A review on automation of industries. **International Journal of Engineering Applied Sciences and Technology**, v. 4, n. 12, p. 298-300, abr. 2020. ISSN 2455-2143.

FLOYD, L. Opportunities in Automation and Control [From the Editor's Desk]. **IEEE Industry Applications Magazine**, v. 17, n. 2, p. 2, mar./abr. 2011. doi: 10.1109/MIAS.2010.939805.

IVANOV, S; KUYUMDZHIEV, M; WEBSTER, C. Automation fears: Drivers and solutions. *Technology in Society*, v. 63, 2020, p. 101431. ISSN 0160-791X. Disponível em: <https://doi.org/10.1016/j.techsoc.2020.101431>.

INTERNATIONAL FEDERATION OF ROBOTICS. Executive Summary World Robotics 2019: Industrial Robots. 2019. Disponível em: <https://ifr.org/downloads/press2018/Executive%20Summary%20WR%202019%20Industrial%20Robots.pdf>.

IVANOV, S; KUYUMDZHIEV, M; WEBSTER, C. Automation fears: Drivers and solutions. *Technology in Society*, v. 63, 2020, p. 101431. ISSN 0160-791X. Disponível em: <https://doi.org/10.1016/j.techsoc.2020.101431>.

BAINBRIDGE, L. Ironies of automation. **Automatica**, v. 19, n. 6, p. 775-779, 1983. ISSN 0005-1098. Disponível em: [https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8).

HUANG, M.-H.; RUST, R. T. Artificial Intelligence in Service. **Journal of Service Research**, v. 21, n. 2, p. 155-172, 2018. Disponível em: <https://doi.org/10.1177/1094670517752459>.

FREY, C. B; OSBORNE, M. A. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, v. 114, p. 254-280, 2017. ISSN 0040-1625. Disponível em: <https://doi.org/10.1016/j.techfore.2016.08.019>.

AUTOR, David; SALOMONS, Anna. Is Automation Labor-Displacing? Productivity Growth, Employment, and the Labor Share. **NBER Working Paper Series**, n. 24871, jul. 2018. Disponível em: <http://www.nber.org/papers/w24871>.

JONES, L.; GOLAN, D.; HANNA, Sammy; RAMACHANDRAN, Manoj. Artificial intelligence, machine learning and the evolution of healthcare: A bright future or cause for concern? **Bone & Joint Research**, v. 7, p. 223-225, 1 mar. 2018. doi: 10.1302/2046-3758.73.BJR-2017-0147.R1.

ABRAHAM, T. H. (Physio)logical circuits: The intellectual origins of the McCulloch–Pitts neural networks. **Journal of the History of the Behavioral Sciences**, v. 38, n. 1, p. 3-25, 2002. Disponível em: <https://doi.org/10.1002/jhbs.1094>.

MIRA, J. Symbols versus connections: 50 years of artificial intelligence. *Neurocomputing*, v. 71, n. 4–6, p. 671-680, 2008. ISSN 0925-2312. Disponível em: <https://doi.org/10.1016/j.neucom.2007.06.009>.

HARTMAN, P. J. Practical applications of artificial intelligence in naval engineering. **Naval Engineers Journal**, v. 100, n. 6, p. 32-40, nov. 1988. Disponível em: <https://doi.org/10.1111/j.1559-3584.1988.tb00825.x>.

DÖRRICH, M; FAN, M; KIST, A. Impact of Mixed Precision Techniques on Training and Inference Efficiency of Deep Neural Networks. *IEEE Access*, v. 11, p. 57627-57634, 2023. doi: 10.1109/ACCESS.2023.3284388

EL-HAD, Mohamed. Artificial Intelligence Background, Definitions, Challenges and Benefits. *مجلة الجمعية المصرية لنظم المعلومات وتكنولوجيا الحاسبات*, v. 31, p. 124-132, 2023. doi: 10.21608/jstc.2023.297957.

HOU, Xinyi; ZHAO *et al.* Large Language Models for Software Engineering: A Systematic Literature Review. arXiv preprint arXiv:2308.10620v6 [cs.SE], 10 abr. 2024. Disponível em: <https://arxiv.org/abs/2308.10620v6>.

THANGARASA, V *et al.* SPDF: Sparse Pre-training and Dense Fine-tuning for Large Language Models. In: *Conference on Uncertainty in Artificial Intelligence*, 2023.

AGERRI, R *et al.* Big data for Natural Language Processing: A streaming approach. *Knowledge-Based Systems*, v. 79, p. 36-42, 2015. ISSN 0950-7051. Disponível em: <https://doi.org/10.1016/j.knosys.2014.11.007>.

BADRINARAYANAN, V; KENDALL, A; CIPOLLA, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 39, p. 2481-2495, 2015. Disponível em: <https://doi.org/10.1109/TPAMI.2016.2644615>.

WANG, Yue; LE, Hung; GOTMARE, Akhilesh; BUI, Nam Dinh; LI, Jia; HOI, Steven C. CodeT5+: Open Code Large Language Models for Code Understanding and Generation. In: **Conference on Empirical Methods in Natural Language Processing**, 2023.

FU, Z *et al.* Decoder-Only or Encoder-Decoder? Interpreting Language Model as a Regularized Encoder-Decoder. arXiv preprint arXiv:2304.04052, 2023. Disponível em: <https://arxiv.org/abs/2304.04052>.

HU, K. Reuters: ChatGPT sets record for fastest-growing user base - analyst note. Reuters, 2 fev. 2023. Disponível em: <https://www.reuters.com>.

SCHENDSTOK, M; WERTZ, S. S. Occupational exposure to artificial intelligence by geography and education. Office of Economic Policy Working Paper 2024-02, U.S. Department of the Treasury, abr. 2024. Disponível em: <https://home.treasury.gov/system/files/136/AI-Combined-PDF.pdf>.

WANG, Q; LIU, X; HUANG, K-W. Investigating employees' occupational risks and benefits resulting from artificial intelligence: An empirical analysis. *Information & Management*, v. 61, n. 8, p. 104036, 2024. ISSN 0378-7206. Disponível em: <https://doi.org/10.1016/j.im.2024.104036>.

PIZZINELLI, C *et al.* Longji. Labor Market Exposure to AI: Cross-country Differences and Distributional Implications. IMF Working Paper No. 2023/216, International Monetary Fund, 4 out. 2023. ISBN 9798400254802.

TOLAN, S *et al.* Measuring the occupational impact of AI: tasks, cognitive abilities and AI benchmarks. *Journal of Artificial Intelligence Research*, v. 71, p. 191-236, 2021. Disponível em: <https://doi.org/10.1613/jair.1.12345>.

URLANA, A *et al.* LLMs with Industrial Lens: Deciphering the Challenges and Prospects -- A Survey. arXiv preprint arXiv:2402.14558 [cs.CL], 22 fev. 2024. Disponível em: <https://doi.org/10.48550/arXiv.2402.14558>.

LEE, J; SU, H. A Unified Industrial Large Knowledge Model Framework in Smart Manufacturing. arXiv preprint arXiv:2312.14428v2 [cs.LG], 15 maio 2024. Disponível em: <https://doi.org/10.48550/arXiv.2312.14428>.

XIA, Y *et al.* Towards autonomous system: flexible modular production system enhanced with large language model agents. arXiv preprint arXiv:2304.14721v4 [cs.RO], 24 jul. 2023. Disponível em: <https://doi.org/10.48550/arXiv.2304.14721>.

GONZÁLEZ-SANTAMARTA, M. Á *et al.* Integration of Large Language Models within Cognitive Architectures for Autonomous Robots. arXiv preprint arXiv:2309.14945v2 [cs.RO], 23 mar. 2024. Disponível em: <https://doi.org/10.48550/arXiv.2309.14945>.

WANG, Y-J *et al.* Prompt a Robot to Walk with Large Language Models. arXiv preprint arXiv:2309.09969v3 [cs.RO], 15 out. 2024. Disponível em: <https://doi.org/10.48550/arXiv.2309.09969>.

FAKIH, M *et al.* LLM4PLC: Harnessing Large Language Models for Verifiable Programming of PLCs in Industrial Control Systems. arXiv preprint arXiv:2401.05443v1 [cs.SE], 8 jan. 2024.

KOZIOLEK, H; GRUENER, S; ASHIWAL, V. ChatGPT for PLC/DCS Control Logic Generation. arXiv preprint arXiv:2305.15809 [cs.SE], 25 maio 2023.

LEVANDUSKI, M. Can ChatGPT Be Used For Control System Programming? *Control*, 25 abr. 2023. Disponível em: <https://control.com/technical-articles/can-openai-chatgpt-generative-ai-be-used-for-control-system-programming>.

DOWDESWELL, B.; SINHA, R.; MACDONELL, S. G. Finding faults: A scoping study of fault diagnostics for industrial cyber-physical systems. **Journal of Systems and Software**, v. 168, p. 1-16, 2020. doi: 10.1016/j.jss.2020.110638.

FRIEL, R; SANYAL, A. C: A high efficacy method for LLM hallucination detection. arXiv preprint arXiv:2310.18344 [cs.CL], 22 out. 2023. Disponível em: <https://doi.org/10.48550/arXiv.2310.18344>.

ROYCHOWDHURY, S. *et al.* Hallucination-minimized Data-to-answer Framework for Financial Decision-makers. arXiv preprint arXiv:2311.07592 [cs.CL], 9 nov. 2023. Disponível em: <https://doi.org/10.48550/arXiv.2311.07592>.

OGUNDARE, O *et al.* Resiliency Analysis of LLM generated models for Industrial Automation. arXiv preprint arXiv:2308.12129 [cs.SE], 23 ago. 2023. Disponível em: <https://doi.org/10.48550/arXiv.2308.12129>.

YANG, L. *et al.* Give Us the Facts: Enhancing Large Language Models with Knowledge Graphs for Fact-aware Language Modeling. arXiv preprint arXiv:2306.11489v2 [cs.CL], 30 jan. 2024. Disponível em: <https://doi.org/10.48550/arXiv.2306.11489>.

EVERITT, T.; LEA, G.; HUTTER, M. AGI Safety Literature Review. arXiv preprint arXiv:1805.01109v2 [cs.AI], 21 maio 2018. Disponível em: <https://doi.org/10.48550/arXiv.1805.01109>.

MEINCKE, L. *et al.* Using Large Language Models for Idea Generation in Innovation. **The Wharton School Research Paper**, set. 2024. Disponível em: <https://ssrn.com/abstract=4526071>.

BAZETT, T. ChatGPT is destroying my math exams. YouTube, 2 jul. 2024. Disponível em: <https://www.youtube.com/watch?v=example-link>.

Como Referenciar este Artigo, conforme ABNT:

SILVA, G. J. S; FRAGA, B. P; FERREIRA, M. P. O Impacto Operacional e Social dos Grandes Modelos de Linguagem e Automação na Indústria. **Rev. FSA**, Teresina, v. 22, n. 1, art. 7, p. 133-157, jan. 2025.

Contribuição dos Autores	G. J. S. Silva	B. P. Fraga	M. P. Ferreira
1) concepção e planejamento.	X	X	X
2) análise e interpretação dos dados.	X	X	X
3) elaboração do rascunho ou na revisão crítica do conteúdo.	X		X
4) participação na aprovação da versão final do manuscrito.	X	X	X