

Arquiteturas de Sistemas Multiagentes: Desafios, Aplicações e Tendências na Transformação Digital

LUIZ GUSTAVO FERREIRA DA SILVA TRUFILHO
UNIVERSIDADE FEDERAL DE SÃO PAULO (UNIFESP)

FÁBIO LUÍS FALCHI DE MAGALHÃES
UNIVERSIDADE FEDERAL DE SÃO PAULO (UNIFESP)

JOÃO DE PAULA RIBEIRO NETO
UNIVERSIDADE MUNICIPAL DE SÃO CAETANO DO SUL (USCS)

ARQUITETURAS DE SISTEMAS MULTIAGENTES: DESAFIOS, APLICAÇÕES E TENDÊNCIAS NA TRANSFORMAÇÃO DIGITAL

1 INTRODUÇÃO

Com o avanço da transformação digital, cresce a demanda por soluções computacionais descentralizadas e adaptativas; nesse cenário, os sistemas multi-agentes (SMA) têm se destacado por sua capacidade de integrar diversos agentes inteligentes em arquiteturas cooperativas voltadas à resolução de problemas em tempo real. Tais sistemas vêm sendo aplicados em setores estratégicos como energia, transporte, saúde e segurança, impulsionando inovações em ambientes críticos (DAHLLING et al., 2021; OLIVEIRA et al., 2023). Agentes inteligentes são entidades computacionais capazes de perceber seu ambiente, tomar decisões autônomas e executar ações de forma a atingir objetivos definidos; tradicionalmente, são caracterizados por atributos como autonomia, percepção, capacidade de ação e aprendizado (RUSSELL; NORVIG, 2020; ZHOU, 2024).

Mais recentemente, estudos como Edge et al. (2024) e Summers et al. (2023) ampliaram essa definição ao considerar agentes baseados em aprendizado profundo e modelos de linguagem de grande porte (LLMs), destacando a capacidade desses sistemas em lidar com tarefas complexas, aprender com interações passadas e adaptar-se dinamicamente a novos contextos. À medida que a sociedade caminha para uma maior digitalização e automação de processos, torna-se essencial compreender os fundamentos arquiteturais desses sistemas e seus impactos práticos de modo a orientar decisões técnicas e políticas, além de contribuir para o avanço de aplicações mais seguras, éticas e interoperáveis.

Apesar do progresso, o campo ainda enfrenta importantes lacunas de pesquisa, a exemplo de padronização nas arquiteturas, interoperabilidade entre agentes heterogêneos, e integração com tecnologias emergentes, como computação quântica, redes neurais e aprendizado distribuído (KIRKE, 2020; NAJDEK et al., 2024). A capacidade de acomodar crescimento de agentes, tarefas e dados sem degradação de desempenho, esbarra em limitações, com implicações em custos operacionais e restrição a aplicações em cenários de crescimento dinâmico, demandando planejamento gerencial robusto para mitigação de gargalos (Gu e Jiang (2023). A sobrecarga de comunicação gerada por interações em tempo real entre múltiplos agentes induz lentidão, perda de pacotes e conflitos decisórios, reduzindo a eficácia sistêmica especialmente sob restrições de banda (LIMA; AGUIAR, 2024)

Diante da crescente importância dos SMA para construção de soluções computacionais escaláveis, resilientes e eficientes, especialmente em ambientes críticos, esses desafios ampliam a complexidade de desenvolvimento e manutenção, exigindo investimentos substanciais em previsão de falhas, adaptação a ambientes dinâmicos e integração de soluções parciais. Assim, este estudo apresenta a seguinte pergunta de pesquisa: quais são os principais modelos arquiteturais utilizados em sistemas multiagentes e quais são seus desafios e aplicações em setores estratégicos? Parte-se da hipótese de que tais arquiteturas, embora promissoras, ainda carecem de amadurecimento técnico e normativo para garantir sua adoção em larga escala; ao sistematizar as abordagens existentes, será possível identificar caminhos para aprimorar sua escalabilidade, coordenação e integração tecnológica.

O objetivo principal desta pesquisa é identificar modelos arquiteturais predominantes em sistemas multiagentes, analisar suas aplicações práticas em ambientes críticos e discutir os principais desafios associados à sua implementação. Os objetivos específicos incluem: I) mapear as arquiteturas utilizadas entre 2020 e 2024; II) avaliar suas vantagens e limitações; III) apontar direções futuras para o desenvolvimento de SMA.

2 FUNDAMENTAÇÃO TEÓRICA

Agentes inteligentes, no campo da Inteligência Artificial, podem ser considerados como entidades dotadas da capacidade de perceber o ambiente no qual estão inseridos e de manter algum tipo de interação com o meio. Possuem autonomia para realizar tarefas em ambientes diversos, baseados em experiências passadas e com a capacidade de tomar decisões alinhadas com objetivos predefinidos (RUSSEL; NORVIG, 2020). Os agentes podem ser categorizados em cinco tipos: Agentes de Reflexo Simples, Agentes de Reflexo Baseados em Modelo, Agentes Baseados em Objetivos, Agentes Baseados em Utilidade e Agentes de Aprendizado (RUSSEL; NORVIG, 2020).

Agentes baseados em Aprendizado por Reforço (agentes RL-based) e Agentes baseados em Modelos de Linguagem de Grande Porte (agentes LLM-based) se enquadram na categoria de Agentes de Aprendizado. Um aspecto fundamental dos Agentes de Aprendizado é sua capacidade de aprender e aprimorar seu comportamento com base na experiência, seus processos de tomada de decisão ao longo do tempo, observando seu ambiente e os resultados de suas ações (EDGE et al., 2024). Estudos recentes destacam proficiência do processamento natural da linguagem das LLM em áreas como raciocínio, resposta a perguntas gerais, programação e geração de texto (ZHAO et al, 2023; ZHOU 2024).

O uso de agentes baseados em LLM pode trazer vantagens, tais como (i) potente processamento de linguagem natural e conhecimento abrangente, aproveitando experiências cultivadas em treinamento; (ii) aprendizado *zero-shot* ou *few-shot*, devido ao conhecimento adquirido, necessitando de amostras menores em seu treinamento; e (iii) interação humana-computador, por compreender a linguagem natural. Ao mesmo tempo, tais estudos relatam limitações ao lidar com tarefas pragmáticas, como restrições a (i) comprimento de contexto, onde pode ocorrer a perda central do texto, (ii) atualização prolongada de conhecimento, onde necessitam de consideráveis recursos computacionais, (iii) a trabalhar como elementos externos (ZHAO et al, 2023; ZHOU 2024).

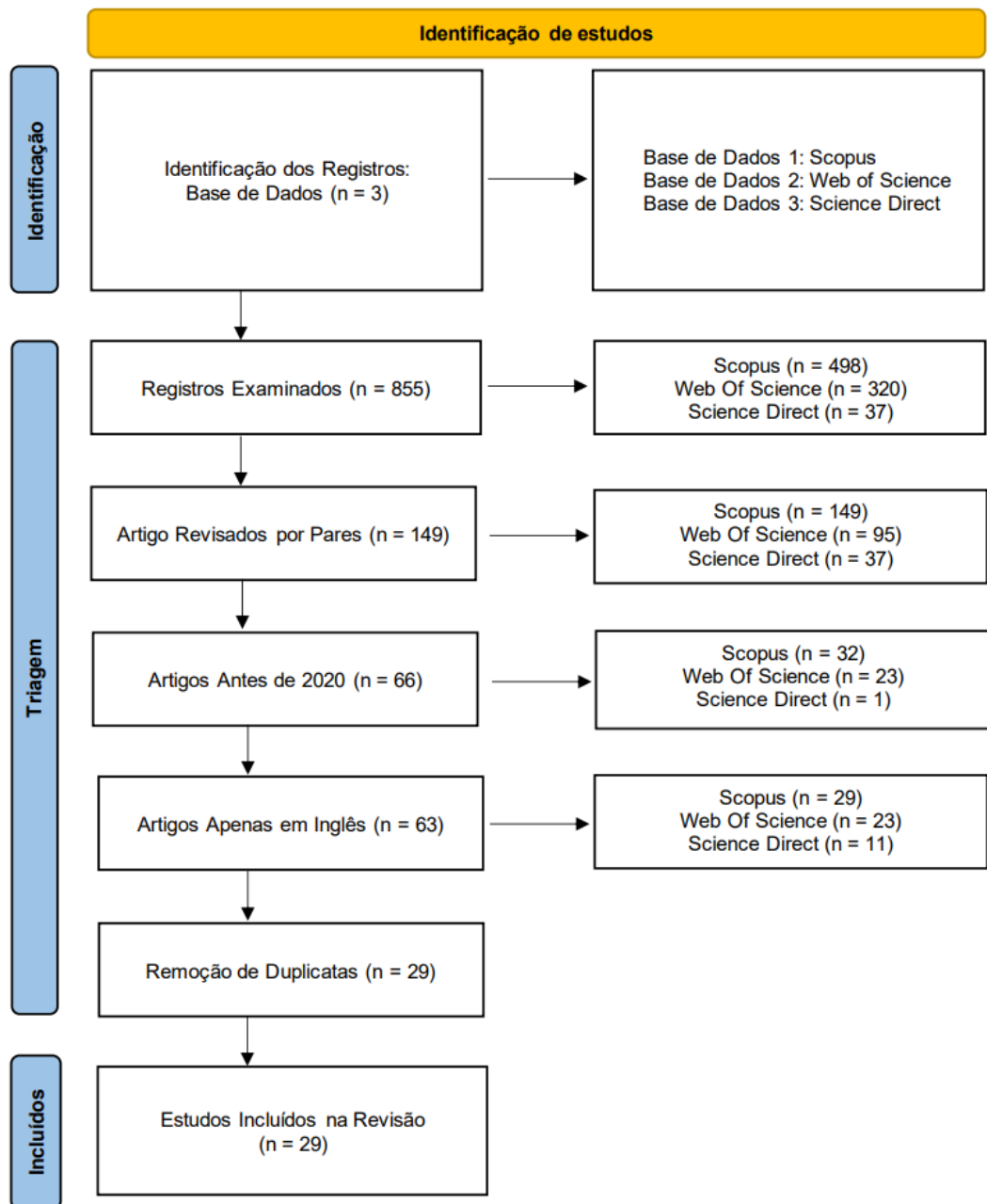
Agentes baseados em LLM, podem ser classificados em agentes únicos ou multiagentes; o sistema multiagente é composto por vários agentes inteligentes que interagem entre si, com expertises diferentes, realizando tarefas de forma coordenada em múltiplos domínios de conhecimento (DING et al., 2021). Esses sistemas, inspirados no comportamento cooperativo de agentes autônomos, têm se mostrado eficazes na resolução de problemas em diversos domínios, como energia, transporte, saúde, suportados principalmente por inteligência artificial e computação distribuída (LIU et al., 2024). Dentre suas aplicações encontram-se redes de energia inteligente, manutenção de infraestruturas críticas e sistemas de aprendizado distribuídos (OLIVEIRA et al., 2023), ilustrando como a colaboração entre agentes pode gerar soluções inovadoras para desafios complexos.

No entanto, a construção de arquiteturas robustas e escaláveis que atendam às demandas de ambientes dinâmicos permanece um desafio significativo (NAJDEK et al., 2024), como a complexidade de desenvolvimento e a sobrecarga de comunicação entre agentes (DAHLING et al., 2021). O problema central abordado neste estudo é a busca de arquiteturas mais adaptáveis e eficientes que permitam a coordenação de múltiplos agentes em cenários de alta complexidade, combinando autonomia, escalabilidade e interoperabilidade. Com base nessa perspectiva, a pesquisa busca identificar os principais modelos arquiteturais utilizados em SMA, analisar aplicações práticas em ambientes críticos e discutir desafios futuros, como escalabilidade e ética em sistemas autônomos.

3 METODOLOGIA

Esta pesquisa adota uma abordagem qualitativa e utiliza o método de Revisão Sistemática de Literatura (RSL) (MALLETT et al., 2012); a seleção das publicações foi realizada em bases de dados comunidade acadêmica Web of Science, Scopus, IEEE Xplore e Science Direct. Essas bases foram escolhidas devido à sua cobertura interdisciplinar e ao rigor dos periódicos indexados, permitindo uma análise abrangente dos modelos arquiteturais aplicados em sistemas multiagentes. Como sugerem Kitchenham e Charters (2007), um protocolo bem definido deve ser seguido para garantir a transparência e a reprodutibilidade do estudo, incluindo a definição de critérios rigorosos de inclusão e exclusão (Figura 1)

Figura 1 Modelo de Revisão Sistemática da Literatura aplicado a esta pesquisa.



Fonte: os Autores (2025)

3.1 Estratégia de Busca

A estratégia de busca foi estruturada para garantir a recuperação de estudos relevantes, conforme metodologias de revisão sistemática descritas por Petticrew e Roberts (2006); foram utilizadas palavras-chave e operadores booleanos para maximizar a abrangência e precisão dos resultados.

String de busca: (Multi) AND (Agent) AND (System) AND (Architecture)

Essa formulação das *strings* de busca foi projetada para abranger conceitos fundamentais relacionados à inovação, inteligência artificial e agentes inteligentes, com ênfase em sistemas multiagentes e suas arquiteturas. O objetivo foi capturar artigos relevantes que explorem modelos, abordagens, aplicações e desafios nesse campo, garantindo uma análise abrangente nas bases de dados selecionadas.

3.2 Seleção dos Artigos

O processo de seleção dos artigos seguiu uma abordagem rigorosa baseada em diretrizes metodológicas para revisões sistemáticas (Kitchenham & Charters, 2007). Inicialmente, a busca retornou 498 artigos na Scopus, 320 na Web of Science e 37 no Science Direct, totalizando 855 estudos. A aplicação dos critérios de inclusão e exclusão ocorreu em etapas progressivas para refinar os resultados.

A primeira filtragem considerou apenas artigos publicados em periódicos revisados por pares, reduzindo os resultados para 149 artigos na Scopus, 95 na Web of Science e 37 no Science Direct. Em seguida, foram excluídos artigos publicados antes de 2020, resultando em 32 artigos na Scopus, 23 na Web of Science e 11 no Science Direct. A terceira etapa eliminou estudos que não estavam disponíveis em inglês, consolidando 29 artigos na Scopus, 23 na Web of Science e 11 no Science Direct.

Por fim, para garantir a qualidade dos estudos incluídos, foram removidos artigos sem acesso ao texto completo e aqueles com menos de quatro páginas, culminando em 25 artigos da Scopus, 20 da Web of Science e 11 do Science Direct, totalizando 56 artigos. Todos os estudos foram importados para o *software* Mendeley Reference Manager, onde foi realizada a remoção de duplicatas, resultando em um conjunto final de 29 artigos únicos.

Esse processo rigoroso de seleção garantiu que a revisão sistemática considerasse apenas estudos de alta qualidade e relevância, contribuindo para uma análise robusta sobre arquiteturas de sistemas multiagentes.

3.3 Critérios de inclusão e exclusão

Para assegurar a qualidade e relevância da literatura analisada, foram adotados critérios rigorosos de inclusão e exclusão, conforme diretrizes estabelecidas por Petticrew e Roberts (2006), os critérios de inclusão e exclusão dos estudos são apresentados no Quadro 1.

Quadro 1 Critérios de Inclusão / Exclusão

Critério	Descrição	Condição
CI1	Apenas Artigos Científicos publicado em periódico	Inclusão
CI2	Publicações entre 2020 e 2024	Inclusão
CI3	Idioma do Artigo: Inglês	Inclusão
CE1	Artigos com até quatro páginas	Exclusão
CE2	Artigos não acessíveis	Exclusão

Fonte: os autores (2025)

Após a aplicação desses critérios, foi possível garantir que a literatura analisada fosse composta por estudos relevantes e atualizados, permitindo uma visão abrangente sobre os avanços e desafios enfrentados no desenvolvimento de sistemas multiagentes; essa triagem rigorosa contribui para a validade e confiabilidade dos achados desta revisão sistemática.

3.4 Perguntas de Pesquisa

A formulação de perguntas de pesquisa bem definidas é essencial para orientar a revisão sistemática, conforme recomendado por Kitchenham e Charters (2007); nesta pesquisa, foram elaboradas as questões do Quadro 1 para nortear a investigação.

Quadro 2 Perguntas de Pesquisa

Pergunta	Descrição da Pergunta
P1	Quais são os principais modelos arquiteturais utilizados em sistemas multiagentes e suas aplicações?
P2	De que forma os sistemas multi-agentes são empregados em ambientes críticos, como saúde, transporte e segurança?
P3	Quais desafios de escalabilidade e interoperabilidade são enfrentados pelas arquiteturas de sistemas multiagentes?
P4	Quais são as principais vantagens e limitações das arquiteturas baseadas em agentes em comparação com outras abordagens computacionais?
P5	Quais são os desafios futuros no desenvolvimento de arquiteturas para sistemas multiagentes?

Fonte: os autores (2025)

4 RESULTADOS

P1 - Quais são os principais modelos arquiteturais utilizados em sistemas multiagentes e suas respectivas aplicações?

A revisão sistemática identificou uma variedade de modelos arquiteturais empregados em sistemas multiagentes (SMA), com destaque para abordagens baseadas em redes neurais, contêineres, neuro arquiteturas e até computação quântica. Essas arquiteturas visam aprimorar a coordenação entre agentes, a capacidade de adaptação e o processamento distribuído. Um dos modelos mais recorrentes envolve a integração de redes *Long Short-Term Memory (LSTM)* a arquiteturas multiagentes, como proposto por Bouziane e Khadir (2024), aplicada à gestão de energia. Essa combinação permite maior precisão na previsão de demanda energética e na alocação de recursos, sendo especialmente útil em redes inteligentes.

Outro modelo relevante é a arquitetura baseada em containers Docker, apresentada por Lima e Aguiar (2024), que proporciona modularidade e escalabilidade em sistemas abertos. Essa arquitetura tem sido aplicada em contextos como cidades inteligentes e Internet das Coisas (IoT), permitindo que agentes sejam implantados, escalados e atualizados com maior flexibilidade. Além disso, essa abordagem facilita a interoperabilidade entre componentes heterogêneos, um dos desafios centrais dos SMA.

A integração entre sistemas multiagentes e computação quântica foi explorada por Kirke (2020), por meio de uma arquitetura híbrida voltada à geração de música interativa. Embora experimental, essa proposta demonstra o potencial das arquiteturas multiagentes em explorar recursos computacionais emergentes para aplicações criativas e de alto desempenho. Já Nagoev et al. (2023) propuseram arquiteturas neurocognitivas baseadas na combinação de inteligência simbólica e subsimbólica, com aplicações em diagnóstico médico, processamento de linguagem

natural e robótica autônoma.

Modelos descentralizados com algoritmos de consenso, como o proposto por Quan e Xi (2024), também ganharam destaque na literatura, sendo aplicados à distribuição de energia em redes sustentáveis. Tais arquiteturas reforçam a autonomia dos agentes ao distribuir a tomada de decisão e reduzir a dependência de sistemas centralizados. Em sistemas industriais, a modularidade tem sido fundamental para lidar com a variabilidade dos processos produtivos.

Os principais modelos arquiteturais identificados compartilham características como descentralização, escalabilidade, capacidade de aprendizado e modularidade. As aplicações mais frequentes concentram-se em domínios onde há necessidade de tomada de decisão em tempo real, gestão de recursos e adaptação a contextos dinâmicos, como energia, saúde, transporte e infraestrutura crítica.

P2 - De que forma os sistemas multiagentes são empregados em ambientes críticos, como saúde, transporte e segurança?

A aplicação de sistemas multiagentes (SMA) em ambientes críticos tem se mostrado promissora, principalmente pela capacidade desses sistemas de atuar de forma autônoma, distribuída e adaptativa. A revisão sistemática revelou que áreas como saúde, transporte e segurança têm se beneficiado de arquiteturas multiagentes voltadas à otimização de processos, resposta em tempo real e melhoria da tomada de decisão.

No setor de transporte, Jiménez-Bravo et al. (2024) desenvolveram uma arquitetura multiagente voltada para a manutenção de estradas durante o inverno, aplicada a um caso real na Espanha. O sistema demonstrou capacidade de monitorar variáveis como clima e condições de tráfego, otimizando a alocação de recursos como sal para o gelo e veículos de manutenção. Como resultado, houve melhorias significativas em termos de segurança viária e eficiência operacional, demonstrando o potencial dos SMA para ambientes com alta variabilidade e urgência na resposta.

Na área da saúde, os SMA têm sido empregados em diagnósticos médicos assistidos por inteligência artificial, sobretudo com o uso de arquiteturas neurocognitivas, como apontado por Nagoev et al. (2023). Essas arquiteturas permitem que os agentes aprendam com dados médicos, adaptem sua atuação de acordo com novos padrões e colaborem com profissionais de saúde em decisões clínicas complexas. Tais soluções são especialmente úteis em sistemas hospitalares descentralizados e em processos de triagem e análise de risco.

No campo da segurança e infraestrutura crítica, os SMA são aplicados em monitoramento e resposta a eventos em tempo real. Lima e Aguiar (2024) propuseram uma arquitetura baseada em Docker para SMA abertos, aplicada a sistemas de cidades inteligentes. A arquitetura permitiu maior interoperabilidade entre agentes responsáveis por diferentes aspectos da segurança urbana, como vigilância, controle de tráfego e resposta a emergências.

Além disso, Rajput e Sikka (2021) apresentaram uma arquitetura especializada na recuperação de falhas em sistemas autossuficientes, relevante para ambientes críticos como redes elétricas e indústrias de processos contínuos. Agentes especializados monitoram falhas e executam ações corretivas automaticamente, reduzindo o tempo de resposta e aumentando a resiliência do sistema.

Esses exemplos demonstram que, em ambientes críticos, os sistemas multiagentes não apenas substituem processos manuais, mas atuam como elementos autônomos capazes de antecipar, decidir e agir, elevando o grau de segurança, eficiência e adaptabilidade dos sistemas. Contudo, a adoção dessas soluções exige atenção especial à confiabilidade, à ética e à integração com sistemas já existentes.

P3 - Quais desafios de escalabilidade e interoperabilidade são enfrentados pelas arquiteturas de sistemas multiagentes?

Embora os sistemas multiagentes (SMA) apresentem inúmeras vantagens em termos de descentralização e adaptabilidade, sua implementação em larga escala esbarra em desafios técnicos significativos, especialmente no que se refere à escalabilidade e à interoperabilidade. Esses aspectos impactam diretamente a eficiência, a integração e a confiabilidade dos sistemas, sendo temas recorrentes nos estudos analisados.

A escalabilidade refere-se à capacidade do sistema de lidar com o aumento no número de agentes, tarefas ou dados sem perda de desempenho. De acordo com Gu e Jiang (2023), à medida que a quantidade de agentes cresce, surgem dificuldades na coordenação, no gerenciamento de mensagens e no controle de conflitos. A arquitetura baseada em liderança e observabilidade proposta pelos autores busca mitigar esse problema em ambientes com múltiplas escalas temporais, como redes de sensores e sistemas logísticos. Ainda assim, a sincronização entre agentes permanece um obstáculo, especialmente em cenários distribuídos.

Já a interoperabilidade diz respeito à capacidade de diferentes agentes – possivelmente desenvolvidos com linguagens, padrões ou objetivos distintos – se comunicarem e colaborarem eficazmente. A falta de padronização nos protocolos de comunicação e nos modelos de dados compromete a integração entre agentes heterogêneos. Lima e Aguiar (2024) enfrentaram esse desafio em sua proposta de arquitetura baseada em Docker para sistemas abertos. Apesar de a modularidade dos contêineres facilitar a integração inicial, a heterogeneidade dos sistemas e a ausência de um framework unificado para SMA ainda dificultam a escalabilidade funcional dessas soluções.

Outro ponto crítico é a sobrecarga de comunicação, citada por vários autores. À medida que mais agentes interagem em tempo real, o volume de mensagens pode se tornar excessivo, resultando em lentidão, perda de pacotes ou conflitos de decisão (Quan; Xi, 2024; Edge et al., 2024).. Essa sobrecarga reduz a eficácia do sistema e limita seu uso em aplicações que exigem respostas rápidas ou operam com restrições de largura de banda.

Além disso, a complexidade no desenvolvimento e na manutenção das arquiteturas multiagentes escaláveis é elevada. Isso ocorre porque, além de projetar os comportamentos individuais dos agentes, os desenvolvedores precisam prever como eles interagirão em diferentes cenários, considerando falhas, atrasos e mudanças no ambiente (Nagoev et al., 2023; Wooldridge, 2009).

Apesar de soluções como contêineres, algoritmos de consenso e arquiteturas baseadas em liderança contribuírem para reduzir esses obstáculos, os desafios de escalabilidade e interoperabilidade seguem como barreiras para a adoção ampla de SMA. Para superá-los, a literatura aponta a necessidade de novas abordagens arquiteturais, padronização de interfaces e protocolos, e mecanismos mais robustos de coordenação e auto-organização.

P4 - Quais são as principais vantagens e limitações das arquiteturas baseadas em agentes em comparação com outras abordagens computacionais?

As arquiteturas baseadas em agentes apresentam um conjunto de vantagens notáveis quando comparadas a abordagens computacionais tradicionais, especialmente aquelas centradas em arquiteturas monolíticas ou pipelines lineares. Tais vantagens estão diretamente relacionadas à autonomia, flexibilidade, robustez e capacidade de adaptação ao ambiente. Entretanto, essas arquiteturas também impõem limitações importantes, sobretudo no que tange à complexidade de desenvolvimento, coordenação e validação formal dos comportamentos emergentes (WOOLDRIDGE, 2009; RAJPUT; SIKKA, 2021).

Dentre os benefícios, destaca-se a autonomia dos componentes: agentes são projetados para tomar decisões localmente, sem a necessidade de um controlador central. Isso favorece a resiliência e permite respostas rápidas a mudanças no ambiente, característica essencial em sistemas dinâmicos e distribuídos. Segundo Russell e Norvig (2020), essa descentralização permite maior escalabilidade organizacional e redução de gargalos computacionais. Além disso,

sistemas multiagentes são especialmente eficazes em ambientes onde múltiplos objetivos ou papéis coexistem, sendo mais adequados do que sistemas centralizados rígidos. Kirke (2020) complementa essa visão ao demonstrar que, em contextos criativos e imprevisíveis, como geração de música interativa, arquiteturas distribuídas são superiores em adaptabilidade.

Outra vantagem das arquiteturas baseadas em agentes é sua capacidade de lidar com incertezas e adaptar-se ao contexto, especialmente quando combinadas com técnicas de aprendizado de máquina, como redes neurais. Bouziane e Khadir (2024), por exemplo, demonstraram a eficácia de uma arquitetura de agentes com redes LSTM para prever consumo energético e reconfigurar a alocação de recursos em tempo real. A modularidade também é uma característica positiva: como cada agente é projetado com uma função específica, a manutenção e a atualização do sistema podem ser realizadas de forma segmentada, como reforçado por Lima e Aguiar (2024) em sua proposta baseada em contêineres.

Contudo, essas arquiteturas não estão isentas de limitações. Uma das principais é a complexidade do design e da implementação, sobretudo quando se trata de sistemas com muitos agentes interdependentes. Os desenvolvedores enfrentam dificuldades em modelar as interações e antecipar os efeitos emergentes do comportamento coletivo, o que dificulta a validação e a verificação formal do sistema (NAJDEK et al., 2024; GU; JIANG, 2023).

Outro desafio é a coordenação entre agentes, que exige mecanismos robustos de comunicação, consenso e resolução de conflitos. Em muitas situações, a sobrecarga de mensagens e a latência na tomada de decisão comprometem a eficiência, especialmente em comparação com soluções baseadas em algoritmos centralizados otimizados (QUAN; XI, 2024).

A interoperabilidade, como discutido anteriormente, também se apresenta como uma limitação relevante, especialmente quando não há um padrão comum de desenvolvimento entre os agentes (LIMA; AGUIAR, 2024).

Além disso, quando comparadas a arquiteturas baseadas em pipelines tradicionais de processamento de dados, como os usados em machine learning centralizado, os SMA podem apresentar desempenho inferior em tarefas bem estruturadas e com dados homogêneos, justamente por conta da sobrecarga inerente à descentralização. Nesses contextos, modelos centralizados com inferência em lote ou aprendizado supervisionado direto são mais eficientes (EDGE et al., 2024).

As arquiteturas baseadas em agentes são particularmente vantajosas em ambientes complexos, dinâmicos e distribuídos, onde a adaptabilidade e a robustez são cruciais. No entanto, sua adoção exige um cuidado maior com o design, a validação e a integração, sendo menos recomendadas para problemas estáticos, simples ou que exigem respostas determinísticas altamente otimizadas (WOOLDRIDGE, 2009; RAJPUT; SIKKA, 2021).

P5 - Quais são os desafios futuros no desenvolvimento de arquiteturas para o desenvolvimento de sistemas multiagentes?

À medida que os sistemas multi-agentes (SMA) se consolidam como soluções promissoras para ambientes distribuídos, crescem também os desafios que precisam ser enfrentados para garantir sua adoção em larga escala, eficiência e confiabilidade. A revisão sistemática aponta uma série de desafios futuros que envolvem tanto aspectos técnicos quanto éticos, metodológicos e regulatórios, exigindo esforços conjuntos da academia, da indústria e dos formuladores de políticas públicas (RAJPUT; SIKKA, 2021; WOOLDRIDGE, 2009).

Um dos principais desafios está na integração com tecnologias emergentes, como computação quântica, modelos baseados em inteligência artificial generativa e computação neuromórfica. Tais tecnologias possuem arquiteturas e paradigmas distintos, o que exige que os SMA evoluam para serem compatíveis com esses novos ambientes. Estudos como o de Kirke (2020) e Bouziane e Khadir (2024) já demonstram movimentos nesse sentido, ao explorar

aplicações híbridas de SMA com redes neurais profundas e ambientes interativos, mas o nível de maturidade dessas soluções ainda é limitado.

Outro desafio futuro é o desenvolvimento de mecanismos de interoperabilidade padronizados entre agentes heterogêneos. Atualmente, a ausência de padrões consolidados dificulta a integração de agentes desenvolvidos por diferentes equipes, com linguagens e objetivos distintos. Essa limitação impede o avanço de ecossistemas abertos e colaborativos baseados em SMA. Lima e Aguiar (2024), ao proporem uma arquitetura baseada em Docker, reforçam a importância de modularidade e isolamento como formas de mitigar esse problema, mas reconhecem que ainda é necessária uma padronização em nível mais amplo.

A aplicabilidade e a governança das decisões autônomas também se impõem como grandes desafios. Em contextos sensíveis, como saúde e segurança, é essencial que as ações dos agentes possam ser compreendidas, auditadas e justificadas. Rajput e Sikka (2021) destacam a importância de desenvolver arquiteturas que combinem autonomia com transparência, especialmente diante do uso crescente de aprendizado de máquina em SMA, onde decisões muitas vezes ocorrem em ‘caixas-pretas’. Edge et al. (2024) reforçam que a rastreabilidade das decisões é uma demanda crítica para aumentar a confiança dos usuários e reguladores.

A segurança cibernética dos sistemas multi-agentes é outro ponto de atenção. À medida que agentes autônomos passam a operar em ambientes críticos, torna-se fundamental protegê-los contra ataques externos, falsificação de identidade, manipulação de mensagens e comportamentos maliciosos. Gu e Jiang (2023) argumentam que a ausência de mecanismos robustos de verificação e a descentralização excessiva podem abrir brechas de segurança, especialmente em arquiteturas que dependem fortemente da troca de mensagens entre agentes.

Além disso, há desafios relacionados ao desempenho em tempo real e ao uso eficiente de recursos computacionais, especialmente em ambientes com restrições severas, como dispositivos móveis, sensores IoT e veículos autônomos. Najdek et al. (2024) destacam que arquiteturas otimizadas para tempo real ainda enfrentam dificuldades de escalabilidade e sincronização, exigindo novas estratégias de balanceamento de carga e coordenação distribuída.

Há ainda o desafio da validação e certificação formal desses sistemas. Comportamentos emergentes e imprevisíveis são característicos dos SMA, o que torna difícil garantir que o sistema funcione corretamente em todos os cenários possíveis. Nagoev et al. (2023) defendem o desenvolvimento de ferramentas formais de simulação e verificação baseadas em lógica temporal, que permitam testar os sistemas sob diferentes condições e garantir sua conformidade com padrões técnicos e regulatórios.

Os desafios futuros no desenvolvimento de arquiteturas para SMA exigem uma abordagem multidisciplinar que combine engenharia de *software*, inteligência artificial, ética computacional e ciência de dados. Avançar nessas frentes permitirá o uso mais amplo e seguro desses sistemas no enfrentamento de problemas complexos da sociedade contemporânea (Wooldridge, 2009; Russell; Norvig, 2004).

Quadro 3 Principais Resultados da Revisão de Literatura

Proposição (P)	Resultados	Autores
P1 – Principais modelos arquiteturais	SMA com redes LSTM para previsão de energia; containers Docker para modularidade; arquiteturas neurocognitivas e quânticas em aplicações emergentes; modelo baseado em liderança e observabilidade para ambientes de múltiplas escalas temporais	Bouziane; Khadir (2024); Lima; Aguiar (2024); Kirke (2020); Nagoev et al. (2023); Gu; Jiang (2023); Quan; Xi (2024)

Proposição (P)	Resultados	Autores
P2 – Aplicações em ambientes críticos	Transporte (manutenção de rodovias); saúde (diagnóstico com agentes neurocognitivos); segurança urbana (monitoramento inteligente); recuperação de falhas autônomas em infraestrutura crítica.	Jiménez-Bravo et al. (2024); Nagoev et al. (2023); Lima; Aguiar (2024); Rajput; Sikka (2021)
P3 – Desafios de escalabilidade e interoperabilidade	Sobrecarga de comunicação; dificuldade na coordenação entre agentes; ausência de padronização de protocolos; manutenção complexa; integração limitada de agentes heterogêneos.	Gu; Jiang (2023); Lima; Aguiar (2024); Najdek et al. (2024); Rajput; Sikka (2021); Edge et al. (2024)
P4 – Vantagens e limitações em relação a outras abordagens	Vantagens: autonomia, adaptabilidade, modularidade; Limitações: validação formal difícil, sobrecarga, integração limitada, coordenação complexa.	Kirke (2020); Bouziane; Khadir (2024); Quan; Xi (2024); Najdek et al. (2024)
P5 – Desafios futuros	Integração com IA generativa e computação quântica; padronização; aplicabilidade; segurança cibernética; restrições em IoT e tempo real; validação e certificação formal.	Kirke (2020); Bouziane; Khadir (2024); Rajput; Sikka (2021); Edge et al. (2024); Gu; Jiang (2023); Najdek et al. (2024); Nagoev et al. (2023)

Fonte: os autores (2025)

5 DISCUSSÃO

Os achados desta revisão sistemática de literatura mostram uma notável diversidade de modelos arquiteturais empregados em sistemas multiagentes (SMA), refletindo o dinamismo e o estágio de maturação ainda em curso da área. Essa diversidade é confirmada em estudos como o de Bouziane e Khadir (2024), que propõem a integração entre SMA e redes LSTM para previsão de consumo energético, e o de Lima e Aguiar (2024), que destacam o uso de *containers* Docker para garantir modularidade e facilidade de implantação. Embora empreguem tecnologias distintas, ambos os trabalhos apontam para um mesmo objetivo: ampliar a flexibilidade e a adaptabilidade dos sistemas em contextos dinâmicos.

A coexistência de arquiteturas baseadas em paradigmas diferentes como as propostas neurocognitivas (Nagoev et al., 2023) e as abordagens com computação quântica (Kirke, 2020) reforça a ideia de que o campo ainda busca convergência quanto às melhores práticas arquiteturais. Essa constatação está alinhada com o que Gu e Jiang (2023) discutem sobre a necessidade de arquiteturas orientadas à liderança e observabilidade para garantir escalabilidade sem comprometer o desempenho em ambientes multiagentes com múltiplas escalas temporais. No entanto, mesmo essas soluções reconhecem limitações no controle de sobrecarga de comunicação e no sincronismo entre agentes.

Em termos de aplicação prática, Jiménez-Bravo et al. (2024) demonstram que SMA podem ser utilizados com sucesso na manutenção preditiva de estradas em condições climáticas adversas, enquanto Rajput e Sikka (2021) aplicam a tecnologia à recuperação autônoma de falhas em sistemas industriais. Esses estudos evidenciam o potencial dos SMA em ambientes críticos, mas também apontam a dificuldade de transpor soluções teóricas para o mundo real dificuldade essa também percebida em Lima e Aguiar (2024), ao indicarem a necessidade de ambientes de teste controlados antes da adoção em larga escala.

A discussão sobre escalabilidade e interoperabilidade foi recorrente nos resultados e

encontra respaldo em autores como Najdek et al. (2024), que destacam os desafios de manutenção e integração de SMA em arquiteturas complexas. A falta de padronização, já sinalizada por Edge et al. (2024), limita a construção de ecossistemas interconectados e impede que agentes desenvolvidos de forma independente atuem de forma cooperativa e segura. Além disso, a latência causada pela sobrecarga de mensagens em sistemas com muitos agentes, conforme destacado por Quan e Xi (2024), segue sendo um obstáculo prático à expansão desses sistemas.

Do ponto de vista comparativo, os resultados confirmam a análise de Wooldridge (2009) e Russell e Norvig (2004), que apontam SMA como superiores em adaptabilidade e robustez frente a arquiteturas monolíticas. Entretanto, também ficam evidentes as limitações em relação à explicabilidade, validação formal e complexidade de desenvolvimento. Essas limitações, embora conhecidas, ainda carecem de soluções padronizadas e replicáveis o que dificulta sua adoção em ambientes regulados, como saúde ou finanças.

Finalmente, os autores mais recentes, como Edge et al. (2024) e Bouziane e Khadir (2024), chamam atenção para a importância de se repensar os SMA diante de novas demandas sociais, como transparência algorítmica e responsabilidade na tomada de decisão autônoma. No entanto, mesmo entre as propostas mais recentes, observa-se pouca ênfase na integração de mecanismos de governança, o que evidencia uma lacuna entre o avanço técnico e as preocupações éticas.

Diante disso, este estudo reforça que, embora os SMA estejam em evolução constante e apresentem alto potencial de impacto, ainda há um distanciamento entre os avanços técnicos e a sua implementação prática em larga escala. A articulação crítica com os autores analisados permite compreender que esse distanciamento não se dá apenas por questões de performance, mas também pela ausência de diretrizes, validações empíricas e arcabouço normativo que sustentem o uso seguro e responsável dessas tecnologias.

6 CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo principal identificar os modelos arquiteturais mais relevantes utilizados em sistemas multiagentes (SMA), analisando suas aplicações práticas, desafios técnicos e perspectivas futuras (WOOLDRIDGE, 2009). A partir de uma revisão sistemática da literatura recente, foi possível mapear avanços significativos na área, com destaque para arquiteturas baseadas em redes LSTM para gerenciamento de energia (BOUZIANE; KHADIR, 2024), o uso de contêineres Docker para garantir escalabilidade (LIMA; AGUIAR, 2024), a aplicação de sistemas neurocognitivos para a resolução de problemas de fundamentação de símbolos (NAGOEV et al., 2023) e propostas híbridas que exploram a computação quântica (KIRKE, 2020).

Essas abordagens vêm sendo aplicadas de forma eficaz em setores críticos como energia, saúde, transporte e segurança, evidenciando o potencial dos SMA para resolver problemas complexos em ambientes dinâmicos e distribuídos (RAJPUT; SIKKA, 2021). Na manufatura, por exemplo, as arquiteturas multiagentes são fundamentais para a implementação de sistemas Ciber-Físicos (CPS) e para a Indústria 4.0, permitindo a rápida adaptação a mudanças na demanda (NAJDEK et al., 2024). Em sistemas de geração de energia sustentável, as MAS descentralizadas otimizam o despacho de comandos em tempo real, promovendo operações econômicas e de baixo carbono (QUAN; XI, 2024).

Além de identificar os principais modelos, a pesquisa revelou que a aplicação de SMA promove ganhos em adaptabilidade, autonomia e modularidade em relação a arquiteturas computacionais tradicionais (FRANKLIN; GRAESSER, 1996). Uma de suas maiores forças reside na capacidade de descentralização e controle distribuído, que não apenas minimiza os riscos de pontos únicos de falha, mas também garante maior robustez e flexibilidade

(DÄHLING; RAZIK; MONTI, 2021). No entanto, desafios como a escalabilidade, a interoperabilidade entre agentes com características heterogêneas e a complexidade de desenvolvimento ainda limitam a difusão mais ampla dessas soluções (GU; JIANG, 2023). A ausência de padrões de comunicação, a sobrecarga de mensagens e a dificuldade de validação formal são aspectos recorrentes que comprometem a eficiência em sistemas distribuídos (OLIVEIRA et al., 2022).

As implicações teóricas indicam que a evolução das arquiteturas multiagentes pode contribuir significativamente para o desenvolvimento de sistemas computacionais mais resilientes e flexíveis (RUSSEL; NORVIG, 2004). Do ponto de vista gerencial, o uso desses sistemas pode otimizar processos decisórios e operacionais, como demonstrado em estudos de caso para manutenção de estradas no inverno (JIMÉNEZ-BRAVO et al., 2024). Essa capacidade é fortalecida por agentes de aprendizagem que, ao assimilarem experiências passadas, otimizam o desempenho em diversas tarefas (NAGOEV et al., 2021). Contudo, é preciso considerar limitações importantes, como a necessidade de maior governança, padronização e regulamentação para garantir a aplicação eficaz em diferentes contextos (MALLET et al., 2012).

Esta revisão, realizada conforme as diretrizes para revisões sistemáticas (KITCHENHAM; CHARTERS, 2007), trouxe contribuições relevantes ao sistematizar o estado da arte sobre SMA, mas apresenta limitações que abrem espaço para investigações futuras. O recorte temporal (2020–2024) e o foco exclusivo em artigos de bases indexadas podem ter deixado de fora iniciativas práticas ou estudos técnicos não publicados em periódicos científicos (SHAFFRIL et al., 2020). Além disso, aspectos éticos e sociais relacionados à tomada de decisão autônoma foram mencionados, mas ainda carecem de exploração aprofundada, especialmente diante do surgimento de modelos com capacidades emergentes ().

No campo da otimização de processos, as SMA são fundamentais para a implementação da Indústria 4.0, permitindo a rápida adaptação a mudanças na demanda e a alocação eficiente de recursos de produção (NAJDEK et al., 2024). Em setores como o de energia, arquiteturas multiagentes descentralizadas viabilizam o despacho de comandos em tempo real para sistemas de energia sustentável, promovendo operações mais econômicas e com menor emissão de carbono (QUAN; XI, 2024). Da mesma forma, na logística, os SMA são aplicados para otimizar operações complexas, como a manutenção de estradas no inverno, gerando rotas eficientes e reduzindo custos (JIMÉNEZ-BRAVO et al., 2024). A capacidade de auto-organização desses sistemas permite uma recuperação ágil de falhas, garantindo que as operações sejam mantidas mesmo com a ocorrência de problemas em unidades individuais (RAJPUT; SIKKA, 2021).

Do ponto de vista gerencial, as arquiteturas multiagentes fortalecem a tomada de decisões estratégicas ao incorporar agentes de aprendizagem que otimizam o desempenho com base em experiências passadas (NAGOEV et al., 2021). Essa capacidade é evidente em sistemas de gerenciamento de saúde para equipamentos críticos, onde a interação entre agentes melhora drasticamente a detecção de falhas e a manutenção preditiva (BOUZIANE; KHADIR, 2024).

Além disso, a utilização de contêineres e tecnologias nativas da nuvem permite que essas arquiteturas sejam escaláveis, adaptando-se ao crescimento do negócio sem a necessidade de grandes reestruturações (LIMA; AGUIAR, 2024). A modularidade inerente também facilita a integração de novas tecnologias, como a computação quântica, para resolver problemas de alta complexidade (KIRKE, 2020).

Apesar dos benefícios evidentes, a implementação de SMA no ambiente corporativo enfrenta desafios que podem impactar o retorno sobre o investimento, como a interoperabilidade entre agentes heterogêneos e a complexidade de desenvolvimento (GU; JIANG, 2023). A ausência de padrões de comunicação e a dificuldade de validação formal são barreiras que podem comprometer a confiabilidade e a eficiência em aplicações de missão crítica (OLIVEIRA et al., 2022). A governança desses sistemas autônomos também exige novas

abordagens regulatórias e éticas, especialmente com o avanço de modelos de inteligência artificial cada vez mais capazes (BUBECK et al., 2023).

Diante disso, pesquisas futuras devem priorizar o desenvolvimento de soluções que melhorem a coordenação entre agentes por meio de representações formais da linguagem natural (MAKOEVA; NAGOEVA; GURTUEVA, 2022). Por fim, torna-se cada vez mais necessário aprofundar o debate sobre os aspectos éticos e os impactos sociais da autonomia dos agentes, garantindo que o avanço tecnológico caminhe lado a lado com a responsabilidade e a segurança, evitando vieses e comportamentos indesejados (ZHOU et al., 2023).

REFERÊNCIAS

BOUZIANE, SeifEddine; KHADIR, Mohamed Tarek. Towards an energy management system based on a multi-agent architecture and LSTM networks. **Journal of Experimental & Theoretical Artificial Intelligence**, v. 36, n. 4, p. 469-487, 2024.

BUCHINGER, Diego; CAVALCANTI, GA de S.; HOUNSELL, M. da S. Mecanismos de busca acadêmica: uma análise quantitativa. **Revista brasileira de computação aplicada**, v. 6, n. 1, p. 108-120, 2014.

DÄHLING, Stefan; RAZIK, Lukas; MONTI, Antonello. Enabling scalable and fault-tolerant multi-agent systems by utilizing cloud-native computing. **Autonomous Agents and Multi-Agent Systems**, v. 35, n. 1, p. 10, 2021.

EDGE, Darren et al. Do local ao global: Uma abordagem de gráfico rag para sumarização focada em consulta. arXiv preprint arXiv:2404.16130, 2024.

FRANKLIN, Stan; GRAESSER, Art. A taxonomy for autonomous agents. In: **In proc. of the Third International Workshop on Agent Theories, Architectures, and Language**. 1996.

GOMA, Edwardus Iwantri et al. Pelatihan mendeley reference dalam mengelola referensi karya tulis ilmiah. **Jurnal Pengabdian Masyarakat Ilmu Keguruan Dan Pendidikan (JPM-IKP)**, v. 5, n. 1, p. 1-8, 2022.

GU, Mengqi; JIANG, Guo-Ping. Observability of Discrete-Time Two-Time-Scale Multi-Agent Systems with Heterogeneous Features under Leader-Based Architecture. **Mathematics**, v. 11, n. 8, p. 1907, 2023.

JIMÉNEZ-BRAVO, Diego M. et al. Multi-agent system architecture for winter road maintenance: a real Spanish case study. **Knowledge and Information Systems**, v. 66, n. 9, p. 5409-5427, 2024.

KEELE, Staffs et al. **Guidelines for performing systematic literature reviews in software engineering**. Technical report, ver. 2.3 ebse technical report. ebse, 2007.

KIRKE, Alexis. Testing a hybrid hardware quantum multi-agent system architecture that utilizes the quantum speed advantage for interactive computer music. **Journal of New Music Research**, v. 49, n. 3, p. 209-230, 2020.

LIMA, G. L. D.; AGUIAR, M. S. D. Towards a docker-based architecture for open multi-agent systems. **IAES Int. J. Artif. Intell.(IJ-AI)**, v. 13, p. 45, 2024.

MAKOEVA, Dana; NAGOEVA, Olga; GURTUEVA, Irina. Formal representation of natural

language elements in multi-agent system based of self-organization of distributed neurocognitive architectures. **Procedia Computer Science**, v. 213, p. 631-635, 2022.

MALLETT, Richard et al. The benefits and challenges of using systematic reviews in international development research. **Journal of development effectiveness**, v. 4, n. 3, p. 445-455, 2012.

NAGOEV, Zalimhan et al. Learning algorithm for an intelligent decision making system based on multi-agent neurocognitive architectures. **Cognitive systems research**, v. 66, p. 82-88, 2021.

NAGOEV, Zalimkhan et al. The symbol grounding problem in the system of general artificial intelligence based on multi-agent neurocognitive architecture. **Cognitive systems research**, v. 79, p. 71-84, 2023.

NAJDEK, Mateusz et al. Three new design patterns for scalable agent-based computing and simulation. **Informatica**, v. 35, n. 2, p. 379-400, 2024.

OLIVEIRA, Gustavo HBS et al. A Multi-Agent Architecture for Distributed Data Mining Systems. In: **Brazilian e-Science Workshop (BreSci)**. SBC, 2022. p. 1-8.

PATI, Debajyoti; LORUSSO, Lesa N. How to write a systematic review of the literature. **HERD: Health Environments Research & Design Journal**, v. 11, n. 1, p. 15-30, 2018.

PETTICREW, Mark; ROBERTS, Helen. **Systematic reviews in the social sciences: A practical guide**. John Wiley & Sons, 2008.

QUAN, Yue; XI, Lei. Smart generation system: A decentralized multi-agent control architecture based on improved consensus algorithm for generation command dispatch of sustainable energy systems. **Applied Energy**, v. 365, p. 123209, 2024.

RAJPUT, Pushpendra Kumar; SIKKA, Geeta. Multi-agent architecture for fault recovery in self-healing systems. **Journal of Ambient Intelligence and Humanized Computing**, v. 12, n. 2, p. 2849-2866, 2021.

RUSSELL, Stuart Jonathan; NORVIG, Peter. **Artificial intelligence: A modern approach**. 4th ed. Prentice Hall, 2020.

SHAFFRIL, Hayrol Azril Mohamed et al. Systematic literature review on adaptation towards climate change impacts among indigenous people in the Asia Pacific regions. **Journal of cleaner production**, v. 258, p. 120595, 2020.

STONE, P. A Survey from a Machine Learning Perspective. **Technical Report CMU-CS-97-193, Carnegie Mellon University**, <http://www.cs.cmu.edu/~pstone/pstone-home.html>, 1997.

ZHOU, Hao et al. Large language model (llm) for telecommunications: A comprehensive survey on principles, key techniques, and opportunities. **IEEE Communications Surveys & Tutorials**, 2024.