Revista ft www.revistaft.com.br

Governança, Qualidade e Uso de Dados em Sistemas de Software com IA Generativa

André Luiz Silva Teotonio[1]
Josineide Teotonia[2]

RESUMO

Este artigo estuda o impacto da inteligência artificial generativa (GenAI) e dos grandes modelos de linguagem (LLMs) no Ciclo de Vida do Desenvolvimento de Software (SDLC). Embora a automação inteligente proporcione taxas de aumento de produtividade de até 55%, a confiança insuficiente, os riscos de segurança e os problemas de governança representam obstáculos significativos para a execução bem-sucedida de projetos. O estudo destaca a mudança necessária de paradigmas centrados em modelos para a IA Centrada em Dados, observando que a qualidade da saída de um sistema é resultado direto da qualidade dos dados de entrada. Para prevenir falhas como alucinações e envenenamento de dados, propomos o uso do framework AI TRiSM (Gestão de Confiança, Risco e Segurança), que cria o núcleo para soluções explicáveis e seguras. Seis princípios fundamentais de prontidão de dados são descritos: diversidade, pontualidade, precisão, segurança, identificabilidade e consumibilidade por máquinas. A governança responsável, conclui-se, precisa unificar desenvolvedores, líderes e comunidades impactadas, para fazer da conformidade ética uma vantagem competitiva estratégica para construir software sustentável e confiável.

---

[1] Graduação em Análise e Desenvolvimento de Sistemas (Faculdade de Tecnologia IBRATEC-2015); especialista em Inteligência Artificial com Ênfase em Ciências dos Dados (Faculdade de Tecnologia IBRATEC- 2019); especialista em Inteligência Artificial na Prática (Faculdade Metropolitana- 2026).

[2] Doutora em Ciências da Educação (Universidade Del Sol/ PY- 2022); Mestre em Ciências da Educação (Universidade da Madeira/ PT- 2018); especialista em Formação de Professores da Educação Básica (UNINASSAAU- 2012); especialista em Psicopedagogia Clínica e Institucional (UNIVERSO-2011); Graduação em Licenciatura Plena em Pedagogia (Universidade Vale do Acaraú-2009).

Palavras-chave: Inteligência Artificial Generativa; Governança de Dados; SDLC; AI TRiSM; IA Centrada em Dados.

INTRODUÇÃO

O desenvolvimento de software está mudando rapidamente por causa da Inteligência Artificial (IA). Com isso, o processo de desenvolvimento, do modelo em cascata ao ágil e DevOps, está levando a uma nova fronteira de automação inteligente e aprendizado de modelos de linguagem de grande escala (LLMs). Soluções como o GitHub Copilot já mostraram aumentos impressionantes na produtividade, pois permitem que os programadores produzam o código 55% mais rápido do que poderiam de outra forma. No entanto, a promessa da IA generativa (GenAI) pode acabar sendo questionada porque, devido à falta de confiança, ameaças à segurança e problemas de governança, os projetos podem começar a falhar.

Para serem considerados "prontos para IA", os dados devem seguir estas seis diretrizes críticas: *serem diversos* (para evitar viés), *oportunos* (atualizados em tempo real), *precisos, seguros, identificáveis e facilmente consumíveis por máquinas*. Na ausência de tais inovações técnicas, a integração da IA no ciclo de vida do desenvolvimento de software (SDLC) é limitada.

A falta dessa perspectiva de visão geral apresenta uma lacuna para as organizações lidarem com alucinações de IA, envenenamento de dados, ataques adversários e o declínio nas habilidades dos desenvolvedores. Nesse contexto, o framework AI TRiSM (Gestão de Confiança, Risco e Segurança) torna-se central para a conformidade e governança. O TRiSM é construído em torno dos pilares fundamentais de confiança, risco e segurança, mantendo as decisões de aprendizado de máquina explicáveis, integrais e seguras contra vulnerabilidades específicas da indústria de aprendizado de máquina. Este artigo pretende preencher o vazio geral da integração de ferramentas GenAI ao longo do SDLC.

Isso nos leva a propor uma abordagem de governança com alto equilíbrio entre a utilidade da automação e a responsabilidade e ética humanas. Ao examinar os princípios da qualidade dos dados e através de conceitos como o AI TRiSM (framework de gestão

de risco), visa um roteiro para projetar sistemas de software que sejam eficientes, sustentáveis, seguros e confiáveis.

1. FUNDAMENTAÇÃO TEÓRICA

O framework AI TRiSM (Artificial Intelligence TRiSM — que governa a aplicação da inteligência artificial da Gartner, para a gestão de confiança, risco e segurança da IA. Esta abordagem é usada para o desenvolvimento responsável e seguro da IA TRiSM, especificamente desenvolvida pela Gartner. De acordo com National Institute of Standards and Technology (NIST):

> Criar IA confiável exige o equilíbrio entre cada uma dessas características com base no contexto de uso do sistema de IA. Embora todas as características sejam atributos sociotécnicos do sistema, a responsabilidade e a transparência também se relacionam aos processos e atividades internos a um sistema de IA e ao seu ambiente externo. Negligenciar essas características pode aumentar a probabilidade e a magnitude de consequências negativas. (NIST, 2026).

Uma abordagem metódica que nos permite garantir segurança - os sistemas de IA funcionam de forma segura e confiável, responsável e ética do início ao fim - desde o início para manter práticas éticas até a operação, a governança torna-se uma vantagem competitiva no mercado, transformando o checkbox[3] de conformidade em uma prerrogativa concorrente.

A arquitetura de camadas na qual o framework funciona com *quatro* mecanismos básicos, de interconexões e de ligações que trabalham juntas para apoiar o núcleo do framework para a governança de IA:

> 1. *Governança de IA*: Alinha-se com a aplicação de políticas, supervisão de conformidade e estratégia.
> 2. *Inspeção em tempo de execução*: Monitoramento em tempo real, detecção de anomalias e validação de desempenho.

[3] Checkbox, ou caixa de seleção, é um elemento de interface gráfica que permite ao usuário selecionar uma, várias ou nenhuma opção de uma lista, ou alternar entre ligado/desligado. Cada opção é independente, tornando-o ideal para seleções não exclusivas (por exemplo, "quais estilos você gosta?"). Possui estados marcado, desmarcado ou indeterminado. (MICROSOFT LEARN, 2026).

> 3. *Governança da Informação*: Qualidade dos dados, rastreamento de linhagem e proteção de privacidade.
> 4. *Segurança da Infraestrutura*: Consiste no fortalecimento de plataformas, acesso e operabilidade de ferramentas e sistema de controle no lado do hardware. (CISCO, 2025).

Enquanto Gartner (2025), especifica as várias camadas, a aplicação prática do TRiSM se divide em *cinco* pilares concretos que cobrem os desafios da IA, desde o aprendizado contínuo e autonomia até a tomada de decisão autônoma e que abordam desafios únicos da IA:

1. *Confiança*: Foca na interpretabilidade e transparência das decisões da IA em relação às partes interessadas. As estratégias para implementação incluem o uso de ferramentas de interpretabilidade de modelos como LIME ou SHAP, bem como o desenvolvimento de trilhas de auditoria de decisões para aplicações de alto risco.
2. *Gestão de Risco*: Enfatiza a identificação estruturada, avaliação e mitigação de ameaças relacionadas à IA antes dos impactos operacionais. Isso consiste em verificações frequentes de vulnerabilidade do modelo e testes baseados em cenários para casos extremos.
3. *Integridade*: Combate o viés algorítmico, garantindo resultados justos para todos em um ambiente equitativo. A integridade é preservada com algoritmos de detecção de viés durante o treinamento e auditorias de equidade para vários grupos demográficos.
4. *Segurança*: Emprega diretrizes de cibersegurança para garantir que os protocolos de segurança específicos da IA que usamos sejam adaptados para atacar fraquezas, que os meios tradicionais não conseguem enfrentar. Isso pode incluir coisas como detecção de ataques adversariais e proteção contra envenenamento de modelos (em que o atacante injeta dados maliciosos para subverter a IA).
5. *Monitoramento*: Facilita o monitoramento regular do desempenho para que desvios de dados e degradação do modelo possam ser detectados instantaneamente. Utiliza infraestruturas ModelOps para estabelecer alertas automáticos de limiares de desempenho e validação contínua.

Com requisitos regulatórios globais mais rigorosos, como a Lei de IA da UE, a implementação do AI TRiSM evoluiu para um requisito de conformidade. Estudos mostram que empresas com capacidades de monitoramento e auditoria em tempo real

para IA têm 40% menos probabilidade de encontrar eventos significativos ou sofrer danos à reputação.

> Em abril de 2021, a Comissão Europeia propôs a primeira lei da UE sobre inteligência artificial, estabelecendo um sistema de classificação de IA baseado no risco. Os sistemas de IA, que podem ser aplicados em diferentes domínios, são analisados e classificados de acordo com o risco que representam para os utilizadores. Os níveis de risco também não ditar a necessidade de respeitar mais ou menos critérios de conformidade de inteligência artificial. (PARLAMENTO EUROPEU, 2023).

O TRiSM preenche a lacuna de confiança que muitas vezes causa o fracasso de projetos de IA no contexto da Engenharia de Software. Ajuda a construir uma cultura de segurança desde o início (seguro por design) usando a varredura de scripts de treinamento e pipelines de inferência para detectar riscos no design antes do lançamento de uma aplicação de IA.

Ao amadurecer a governança de IA — passando de conformidade reativa para governança preditiva — as empresas poderão inovar com segurança, com um olhar mais atento para garantir que a IA seja uma entidade produtiva e não um risco não mitigado. “fornecer orientações sobre estratégias para identificar e avaliar os riscos de ferramentas habilitadas por IA e estabelecer mecanismos de governança apropriados para a instituição se adequar às políticas e normas para uso de IA” (Almeida, Nas, 2024, p. 24).

Ao mesmo tempo, precisamos examinar os dados sobre ***IA Centrada em Dados vs. IA Centrada em Modelos*** e a transformação das metodologias de desenvolvimento tecnológico da Inteligência Artificial, analisando profundamente essa relação entre a abordagem tradicional centrada em modelos e a emergente centrada em dados.

Historicamente, a abordagem de desenvolvimento de sistemas de IA era o paradigma centrado em modelos amplamente introduzidos nas instituições acadêmicas. Nesse método, o conjunto de dados é considerado uma coisa imutável e fixa e a responsabilidade do desenvolvedor é construir o melhor modelo ajustando hiperparâmetros, estruturas ou otimização de redes neurais. Boni (2024), aponta que:

> A qualidade dos dados utilizados para treinar os modelos de IA é um fator crítico para o sucesso de sua aplicação, pois os seus modelos são altamente dependentes de dados de alta qualidade para gerar previsões precisas e relevantes, onde, a ausência de dados adequados ou a presença de dados enviesados pode comprometer seriamente os resultados e levar a falhas no processo de garantia de qualidade.

Na realidade, no entanto, essa abordagem muitas vezes falha porque, na maioria dos casos, os dados não são limpos ou bem curados. A IA Centrada em Dados, no entanto, inverte essa lógica e propõe a engenharia sistemática de dados para construir melhores sistemas. Enquanto na IA centrada em modelos a ênfase está no código e nos algoritmos, na IA centrada em dados, o valor está na melhoria iterativa (programática) da qualidade do conjunto de dados. Os princípios e guias são importantes, mas só serão efetivos se combinados "com práticas de governança que ajudem a guiar o uso de ferramentas de IA em projetos de pesquisa científica, desde o design até a produção e publicação de resultados de pesquisa" (Almeida, Nas, 2024, p. 24).

A verdade é que os dados não são fixos, mas são melhorados se as características forem removidas para eliminar ruídos e discrepâncias que são fatores excessivos ou algo assim. A chave para o sucesso da IA é que a qualidade da saída é um resultado direto da qualidade da entrada. Modelos que são imprecisos, tendenciosos ou "sujos" — ou mal rotulados — não devem confiar em nenhum algoritmo.

Empresas como a OpenAI apontam que erros nos dados e rótulos são os desafios mais importantes no treinamento de modelos como o ChatGPT e o Dall-E. Neste momento, empresas como a Tesla culpam o "Data Engine", um programa que prioriza modelos com dados aprimorados em vez de truques de modelagem, como sua fonte de ponta.

A transição para um modelo centrado em dados envolve a implementação de algoritmos para diagnosticar os problemas de dados e corrigi-los. A aplicação deve ter quatro funções principais:

- *Identificação e remoção de outliers* para exemplos anômalos que dificultam o aprendizado;
- *Correção de erros de rotulagem*: Técnicas como o Aprendizado Confiante podem filtrar automaticamente dados mal classificados;
- *Aumento*: Novos exemplos são adicionados aos dados para representar conhecimento pré-existente e aumentar a diversidade.
- *Aprendizado ativo*: Escolher inteligentemente os dados mais reveladores para rotular, que são otimizados.

Para facilitar essa conversão, os dados devem ser tratados sob regras estritas de governança. Sistemas de software modernos rastreiam métricas como a Pontuação de Confiança em IA para avaliar a qualidade da diversidade, pontualidade, precisão e

segurança dos dados para determinar se é "valioso", "oportuno" e "seguro" antes de ser usado por modelos. Na perspectiva de Seyffert (2025):

> A documentação detalhada é um pilar da governança de IA porque serve para registrar datasets, versionamento de modelos e decisões automatizadas. O melhor: com auditabilidade e rastreabilidade.
> O mesmo vale para programas de treinamento contínuo para cientistas de dados, equipes de compliance e gestão. A diferença é que essa prática ajuda a fazer com que todos todos entendam os princípios de governança e práticas éticas continuamente.
> Por fim, vale reforçar: a melhoria contínua é fundamental, pois envolve ciclos regulares de governança, como revisões, auditorias, atualizações de modelos e ajustes nas políticas, criando um ciclo virtuoso de segurança, ética e eficiência.

Esse sistema de governança torna possível continuar melhorando o modelo enquanto mantém modelos e confiabilidade, permitindo que ele evolua através da repetição dessa infraestrutura de controle, fazendo com que todo o modelo viva mais tempo. Por esse motivo, essa mudança para a IA Centrada em Dados aponta para um avanço no desenvolvimento de software, onde a curadoria manual e ad hoc de algoritmos é substituída pela curadoria algorítmica – tornando a qualidade dos dados a principal vantagem competitiva e motor de desempenho em sistemas de software de IA.

O design de inteligência artificial (IA) responsável nas organizações é um esforço de colaboração entre múltiplas partes interessadas, com diferentes objetivos e responsabilidades ao longo do ciclo de vida da IA. No entanto, é evidente na literatura e na prática existentes que isso tem sido grosseiramente desproporcional: a maioria dos sistemas de governança foi desenvolvida para dar suporte a designers e desenvolvedores na fase técnica, enquanto falha em apoiar os papéis mais relevantes dos líderes organizacionais e das comunidades impactadas. A autora citada acima, reforça que: "*Além disso, a transparência nas decisões de IA pode se tornar um ativo real de negócio, uma vez que clientes, parceiros e reguladores valorizam organizações que conseguem explicar como seus sistemas funcionam e garantem imparcialidade e segurança."*

Os líderes organizacionais tomam decisões estratégicas e definem a missão, além de criar a cultura de trabalho e políticas de alto nível. Apesar do poder, muitos desses gestores não têm ferramentas disponíveis para avaliar quão preparada uma organização está para implantar a IA de forma ética. Sua falta de artefatos explícitos para sua

liderança indica que a maioria está adotando sistemas de IA generativa sem entender o potencial estratégico e os riscos.

Estruturas como o AI TRiSM visam fechar essa lacuna ao incorporar hierarquias de governança que enfatizam a aplicação de políticas e o alinhamento estratégico, em vez de apenas avaliar o código, para que a IA produza valor de negócios com integridade. Comunidades impactadas — indivíduos ou grupos afetados direta ou indiretamente pela operação de um sistema de IA — são frequentemente consideradas uma reflexão tardia no desenvolvimento de software.

A análise de ferramentas de governança responsável indica que a maioria dos recursos para esse grupo está concentrada na fase de monitoramento, ou seja, após a implantação do sistema. Em outras palavras, as ferramentas atuais só ajudam comunidades e usuários finais quando o dano — por exemplo, violações de privacidade ou decisões discriminatórias — já começou.

Ferramentas que apoiam o empoderamento precoce, permitindo que as comunidades participem da formulação de problemas e proposição de valor, para ter voz na decisão de construir ou não o sistema, são urgentemente necessárias. Para abordar como estamos atualmente presos no "quebra-cabeça" do controle organizacional, uma estrutura de governança de ponta a ponta precisa ser desenvolvida.

No geral, a governança responsável não é apenas uma questão técnica, mas também estratégica, institucional e ética. Para dar o salto para uma IA genuinamente confiável, ferramentas que traduzam ideias abstratas em implementação concreta para todos aqueles que desempenham um papel nesses sistemas devem ser desenvolvidas — para que o desenvolvimento tecnológico não ocorra às custas da justiça social e segurança. O National Institute of Standards and Technology (2023) afirma que:

> Ao gerenciar os riscos da IA, as organizações podem enfrentar decisões difíceis ao equilibrar essas características. Por exemplo, em certos cenários, podem surgir conflitos entre otimizar a interpretabilidade e garantir a privacidade. Em outros casos, as organizações podem se deparar com um conflito entre precisão preditiva e interpretabilidade. Ou, sob certas condições, como a escassez de dados, as técnicas de aprimoramento da privacidade podem resultar em perda de precisão, afetando decisões sobre equidade e outros valores em determinados domínios.
>
> Lidar com esses conflitos exige levar em consideração o contexto da tomada de decisão. Essas análises podem destacar a existência e a extensão dos conflitos entre diferentes medidas, mas não respondem a perguntas sobre como navegar por esses conflitos. Essas questões dependem dos valores em jogo no *contexto* relevante e devem ser resolvidas de maneira transparente e devidamente justificável.

Contudo, a governança responsável da IA não é apenas uma questão técnica – mas está emergindo como um requisito estratégico, institucional e ético, chave para as organizações contemporâneas. Para que esse salto em direção a uma IA verdadeiramente confiável tenha sucesso, precisa-se de ferramentas para transformar ideias abstratas e impraticáveis em ações concretas para todos os atores do sistema — desenvolvedores, líderes e comunidades — que garantam que a tecnologia não venha à custa do avanço da justiça social e da segurança.

2. QUALIDADE E PRONTIDÃO DE DADOS PARA IA

Este capítulo foca no elemento crítico para a eficácia de qualquer sistema de Inteligência Artificial, a saber, a disponibilidade e qualidade dos dados. A engenharia sistemática de conjuntos de dados é a chave para o motor de desempenho no paradigma de IA Centrada em Dados, que não é apenas a manipulação de modelos. Para transformar dados brutos em recursos úteis para a IA Generativa (GenAI), eles precisam cumprir os seis princípios que garantem confiança, eficiência e alto desempenho.

Deste modo, sua principal característica é a capacidade de usar informações já existentes para gerar novos conteúdos. Essa habilidade é alimentada por enormes conjuntos de dados e algoritmos sofisticados que permitem à IA identificar padrões, fazer associações e criar material que simula, ou até mesmo expande, a criatividade humana. Com isso, ferramentas de IA generativa têm o potencial de revolucionar diversas áreas, como marketing, educação, design, e até mesmo a produção artística, oferecendo uma nova forma de automação que desafia as fronteiras tradicionais da criatividade e inovação (WESSEL et al., 2023).

A prontidão não é automática; exige muito esforço e tempo antes que as máquinas estejam finalmente prontas para entender o ambiente de negócios. Sem essa base, os sistemas de IA são propensos a alucinações, preconceitos e resultados sem sentido. As organizações precisam de seis princípios para mitigar riscos e maximizar o retorno sobre o investimento em IA:

1. *Diversidade*: sabemos que para ajudar a mitigar o que é chamado de preconceitos algorítmicos, é importante que os modelos de treinamento não sejam colocados em silos de isolamento. Diversidade significa adicionar uma ampla variedade

de fontes — desde sistemas legados em mainframes e SAP até arquivos não estruturados e aplicativos SaaS — para garantir que a IA aprenda diferentes padrões e pontos de vista para tomar decisões mais justas e equitativas.

2. *Atualidade*: a eficácia de um modelo de IA é determinada em uma relação linear com a atualidade das informações consumidas. Dados desatualizados causam previsões inúteis. Para torná-los oportunos, as organizações precisam implementar pipelines de dados em tempo real e de baixa latência, aproveitando a Captura de Dados de Mudança (CDC) para bancos de dados relacionais e captura de fluxo para dispositivos IoT, por exemplo.

3. *Precisão*: a precisão pode ser alcançada através de três estratégias de governança: perfilagem de dados (análise exploratória de dados) para compreensão, distribuição e redundância das informações, automação de regras de qualidade e análise de linhagem de dados. Esta última é importante para cientistas de dados que desejam rastrear a origem e o fluxo dos dados de origem para que mudanças acidentais sejam evitadas e o modelo não seja contaminado.

4. *Segurança*: sistemas de IA lidando com informações sensíveis, como Informações Pessoais Identificáveis (PII) e registros financeiros. Para a prontidão, esses sistemas devem ser classificados automaticamente com base na sensibilidade e ter implementadas medidas de proteção como mascaramento, tokenização e criptografia, juntamente com controles de acesso rigorosos para proteger a privacidade e a reputação organizacional.

5. *Identificabilidade*: Os dados preparados devem ser facilmente pesquisáveis. Fazemos isso com metadados poderosos e tipagem semântica para adicionar significado adicional para modelos automatizados. O glossário de negócios une definições técnicas a ideias de negócios, e catálogos de metadados indexam informações, para que permaneçam pesquisáveis e sejam úteis para o problema de IA.

6. *Consumibilidade pela Máquina*: A IA precisa de certos formatos para o processo de ingestão, conforme a plataforma Qlik (2026):

> De fato, tecnologias de IA, como o aprendizado de máquina (ML), já auxiliam os profissionais de dados a produzir previsões matemáticas, gerar insights e melhorar a tomada de decisões. Além disso, tecnologias emergentes de IA, como a IA generativa (GenAI), podem criar conteúdo extremamente realista, com potencial de aumentar a produtividade em praticamente todos os aspectos do negócio.

•Para ML (Machine Learning): As informações devem ser organizadas e tabulares, idealmente estruturadas em infraestruturas de data lakehouse usando SQL.

•Para IA Generativa (LLMs): O método de Geração Aumentada por Recuperação (RAG) é aplicado, no qual dados não estruturados (PDFs, e-mails e transcrições) são divididos em fragmentos e convertidos em representações numéricas, chamadas de embeddings, que são armazenadas em bancos de dados vetoriais. A fonte citada acima (2026), explica que:

> [...] As aplicações de ML e IA generativa são ferramentas poderosas, mas o potencial delas reside na capacidade de consumir os dados prontamente. Diferentemente dos humanos, que conseguem ler anotações escritas à mão e planilhas caóticas, estas tecnologias exigem que as informações sejam representadas em formatos específicos. É como cozinhar para uma pessoa exigente — se ela não comer o que você preparou, vai ficar com fome. De modo similar, as iniciativas de IA não serão bem-sucedidas se os dados não estiverem nos formatos certos para os experimentos de ML ou para as aplicações de LLM.
>
> Tornar os dados facilmente consumíveis ajuda a desbloquear o potencial destes sistemas de IA, permitindo que eles ingiram as informações tranquilamente e as traduzam em ações inteligentes para gerar resultados criativos. Consequentemente, tornar os dados facilmente consumíveis é o último princípio dos dados prontos para IA.

O Índice de Confiança em IA é proposto como uma ferramenta de monitoramento para permitir a compreensão desses princípios, convertendo-os em métricas significativas. Este indicador agrega o desempenho dos seis princípios em uma pontuação composta, permitindo que líderes e desenvolvedores determinem facilmente se os dados corporativos estão maduros o suficiente para a implantação de GenAI ou se ainda não passaram pelas intervenções iniciais de limpeza e governança.

4. APLICAÇÕES DE IA GENERATIVA NO CICLO DE VIDA DO SOFTWARE (SDLC)

A integração da GenAI no Ciclo de Vida do Desenvolvimento de Software (SDLC) é uma das transformações mais significativas e em rápida evolução na engenharia de software. Ao passar de iterações lineares como o Waterfall para processos iterativos como Agile e DevOps, o campo transitou rapidamente para um paradigma

focado em IA que pode reconhecer padrões, prever resultados e também automatizar tarefas complexas.

Recentemente, o uso de Modelos de Linguagem de Grande Escala (LLMs) está abrindo novos caminhos para produtividade e qualidade. Quando comparados aos métodos tradicionais, essas ferramentas de codificação assistida permitem que os desenvolvedores concluam tarefas cerca de 55% mais rápido, aumentando a taxa de sucesso em seus casos de teste na primeira tentativa. De acordo com Garofalo (2025, p. 26):

> Apesar do potencial transformador da IA Generativa, a relação entre os ganhos de produtividade em nível de tarefa e o desempenho global da equipe não é linear nem garantida. Fatores como a qualidade do código, a dívida técnica, a segurança e as dinâmicas sociotécnicas da equipe atuam como variáveis mediaDORAs críticas que ainda são pouco compreendidas em contextos industriais reais. Portanto, conclui-se que existe uma lacuna teórica e empírica proeminente no entendimento de como a adoção de LLMs por equipes de desenvolvimento de software impacta concretamente as métricas de fluxo de valor, como lead time e cycle time.

Mas, embora o impacto no processo de escrita de código seja o mais visível, o potencial da GenAI se estende por todo o espectro do SDLC — desde a extração inteligente de requisitos e design arquitetônico até a garantia de qualidade (QA) e manutenção preditiva. No entanto, embora a promessa por trás da GenAI esteja presente, a aplicação sistemática ainda é frequentemente fragmentada, onde as ferramentas existem de forma holística em silos e não compartilham descobertas entre as fases.

A introdução da IA no SDLC não é uma mudança incremental; ao contrário, é uma mudança de paradigma, que passa de tarefas manuais e propensas a erros para automação inteligente. A fase de engenharia de requisitos, por muitos anos caracterizada por um trabalho humano pesado que pode falhar na interpretação, está mudando com a fusão do Processamento de Linguagem Natural (NLP)- No que diz respeito ao processamento de linguagem natural (PLN) e à visão computacional, o aprendizado de máquina tem possibilitado o desenvolvimento de aplicações avançadas, como assistentes virtuais, serviços de tradução e sistemas de reconhecimento de imagem. (PALO ALTO, 2026) - e da Inteligência de Documentos.

A mudança de procedimentos tradicionais para um fluxo de trabalho assistido por IA abre a possibilidade de automatizar a extração de especificações de grandes quantidades de dados não estruturados (por exemplo, e-mails, transcrições de reuniões,

feedback de usuários, relatórios técnicos), transformando informações heterogêneas em conhecimento bem estruturado e validado. A Palo Alto[4] (2026), afirma que:

> Um LLM (Large Language Model, ou Modelo de Linguagem Grande) é um tipo específico de modelo de aprendizado de máquina, geralmente baseado em técnicas de aprendizado profundo, projetado para tarefas de processamento de linguagem natural. LLMs, como o GPT-3 ou o BERT, são pré-treinados com grandes quantidades de dados textuais e podem gerar textos semelhantes aos escritos por humanos ou compreender padrões linguísticos complexos. Os LLMs são uma subcategoria dentro do escopo mais amplo do aprendizado de máquina, com foco na compreensão e geração de linguagem natural.

A aplicação de modelos baseados em transformadores, como BERT e GPT, aumentou a precisão do entendimento semântico, permitindo que os sistemas não apenas identifiquem requisitos funcionais e não funcionais, mas também detectem ambiguidades e inconsistências em um estágio inicial. Essa capacidade de diagnóstico proativo é crítica para mitigar falhas interpretativas, cuja propagação, se levada para as etapas de codificação e teste, pode causar incursões com custos de remediação exponencialmente mais altos, além de dívida técnica.

A IA generativa também se estende além da extração, com a capacidade de criar histórias de usuários de forma a apoiar os objetivos de negócios — usando modelos treinados em dados passados para modelar quanto trabalho e esforço de codificação podem se tornar viáveis. A classificação desses requisitos agora é um processo analítico, em vez de baseado em critérios relevantes para as partes interessadas, complexidade de desenvolvimento e impacto projetado. Mas o sucesso de tal automação depende de uma infraestrutura de governança responsável. Precisamos de "Human-in-the-loop".
Marimoto (2025), explica que:

> ***Human in the Loop*** **(HITL) é uma arquitetura de integração entre sistemas de inteligência artificial e humanos**. A arquitetura de integração é a maneira como diferentes partes de um sistema (humanos, algoritmos, sensores, bancos de dados, interfaces) se conectam e trocam informações para funcionar juntas.
>
> No caso do HITL, essa arquitetura define em que momento e de que forma o ser humano entra no fluxo da IA, ou seja, determina como o ser humano

[4] A Palo Alto Networks é uma das maiores e mais conceituadas empresas globais de cibersegurança corporativa

> interage com o modelo de inteligência artificial. Dessa forma, há uma participação ativa do indivíduo nas etapas de aprendizado, análise e tomada de decisão. A IA processa dados, identifica padrões e propõe ações; o especialista avalia essas propostas, aplica ajustes com base em conhecimento de campo e orienta o sistema para que ele aprenda de acordo com a realidade operacional.

Para garantir que as decisões da IA sejam explicáveis e possam ser justas, para que os preconceitos contidos no conhecimento historicamente registrado não conduzam a novas instruções.

Apesar do domínio do kit de ferramentas orientado para desenvolvedores, a importância de envolver a liderança organizacional e as comunidades já nesta fase inicial de 'proposta de valor' é essencial — podemos permitir que as ferramentas tecnológicas realmente atendam aos requisitos humanos além dos números que entregam na métrica de produtividade técnica. A fase de design e arquitetura — tipicamente baseada em experiência heurística ou julgamento humano — está sendo redefinida por ferramentas de IA que fornecem um nível de automação cognitiva e avaliação preditiva em um momento de inovação e automação sem precedentes.

Essa evolução oferece a oportunidade de se afastar de designs estáticos em direção a modelos de sistema inteligentes e adaptáveis. A IA ajuda os desenvolvedores a

identificar automaticamente o que implementar como parte de um padrão de design (Singleton, Factory, Observer, etc.) e analisá-los em relação aos requisitos de seu sistema e sua base de código existente.

Sistemas construídos para processar e, em seguida, analisar esses dados por meio de algoritmos de aprendizado de máquina e Redes Neurais de Grafos (GNNs) derivam padrões de repositórios complexos que otimizam a eficiência, escalabilidade e manutenibilidade do software. Além de fornecer uma sugestão inicial de implementação, a IA identifica antipadrões ou decisões arquitetônicas subótimas e faz recomendações de refatoração que minimizam a dívida técnica e melhoram a arquitetura geral do sistema.

A modelagem de sistemas é aprimorada com ferramentas que extraem representações abstratas, como diagramas UML, diretamente das especificações de requisitos ou do código legado. Por meio da análise de grafos e reconhecimento de padrões de dados, a IA não apenas encurta significativamente o processo de modelagem, mas também mantém a documentação consistente com os detalhes reais de implementação.

A colaboração entre Engenharia Dirigida por Modelos (MDE) e IA também permite a geração de código executável a partir desses modelos, ajudando a alcançar o equilíbrio ideal entre o paradigma de design e implementação. Ferramentas de IA generativa são assistentes inteligentes que combinam raciocínio arquitetônico com conhecimento de domínio, propondo arquiteturas que reconciliam demandas conflitantes, como desempenho e segurança, e oferecem sugestões feitas de dentro do modelo.

Sistemas como esses também permitem simulações para reproduzir o comportamento do software em vários cenários e analisar modelos preditivos antecipadamente, para que recursos como modelos avançados possam ser examinados com antecedência, permitindo verificar falhas de design cedo, sem a necessidade de código de software.

Esse processo é crucial para todos os sistemas de nível industrial. A confiabilidade, conformidade ética e regulatória dependem disso, e tais padrões de qualidade ou conformidade são obrigatórios e devem ser observados, no mínimo. *A confiabilidade* é definida na mesma norma como a "capacidade de um item de funcionar conforme o esperado, sem falhas, durante um determinado intervalo de tempo, sob determinadas condições" (Fonte: ISO/IEC TS 5723:2022). A visão futura da arquitetura de software é de sistemas Self (auto evolutivos, auto reparadores, autoconscientes).

O sistema, com seus agentes autônomos de IA à disposição, pode se adaptar às mudanças nas necessidades dos usuários e se autoajustar em tempo de execução usando sua arquitetura e código. Ao mesmo tempo, essa autonomia exige que a segurança seja integrada ao processo de design (seguro por design), no qual pipelines de geração identificam vulnerabilidades e dependências inseguras já na fase de definição arquitetônica.

A codificação, que no passado era a etapa central do desenvolvimento de software, está passando por uma transformação profunda à medida que a popularização de assistentes baseados em grandes modelos de linguagem (LLMs) continua. Além da melhoria de velocidade, a incorporação dessas IAs também ajudou a elevar a qualidade básica do software, aumentando a taxa de sucesso inicial dos casos de teste na primeira tentativa. Essa automação está forçando uma mudança nos papéis profissionais dentro das equipes de engenharia. O papel dos desenvolvedores mudou de escrita ativa para mais um papel de supervisão e orquestração.

A maioria dos engenheiros de software agora dedicam maior esforço à engenharia de prompts, avaliando criticamente as sugestões geradas por IA e dividindo problemas complexos em seus componentes que a IA pode então digerir de forma mais eficiente. Esta é uma mudança que exige habilidade técnica e expertise no domínio de negócios, características geralmente associadas a desenvolvedores de nível superior.

Mas o uso de assistentes de codificação está expondo grandes lacunas, que precisam ser mantidas seguras com mecanismos de governança, conforme declarado pelo framework AI TRiSM:

- *Segurança*: A IA pode recomendar bibliotecas desatualizadas em seu código ou código com sérias vulnerabilidades de segurança (injeção de SQL, falta de validação de entrada). Onde ferramentas como o Snyk Code ajudam neste exercício é ao escanear scripts de treinamento e pipelines de inferência para risco de envenenamento de dados e problemas em tempo real.
- *Qualidade*: O código gerado por IA é potencialmente deficiente em termos de documentação, bem como em termos de desvio do usual, tornando a manutenção a longo prazo uma luta e introduzindo dívida técnica adicional.
- *Perda de Habilidade*: O problema moral e institucional sobre isso é que desenvolvedores novatos nesta indústria podem se tornar excessivamente dependentes da IA e não serem capazes de realizar análises profundas e criativas.

Em conclusão, a codificação apoiada por IA oferece grandes benefícios de eficiência, mas para que sejam sustentáveis, pipelines de validação contínua são cruciais e uma organização precisa estabelecer uma cultura onde a intervenção humana supere a automação direta e o software resultante não seja apenas rápido, mas também seguro e facilmente mantido.

O Controle de Qualidade (QA) está historicamente preso no meio do SDLC, pois é tedioso e exige ampla cobertura de casos extremos. Nesta fase, a combinação de modelos de Machine Learning (ML) e IA Generativa (GenAI) permite a automação da geração de scripts de teste e geração de casos de uso através de análise profunda de mudanças de código e previsão de pontos críticos de falha.

Modelos treinados em grandes conjuntos de dados históricos de bugs permitem a previsão de defeitos, que consente às equipes focarem os esforços de teste em componentes mais vulneráveis. A IA é, portanto, capaz de lidar com todos esses problemas de forma eficaz.

Além disso, a IA pode gerar espontaneamente testes unitários e scripts de integração, reduzindo os esforços manuais e acelerando a descoberta de casos de defeito desde o início. No entanto, um problema mais recente é que a natureza probabilística da IA significa que não é tão fácil configurar oráculos de teste confiáveis; pequenas diferenças nos dados de entrada produziriam grandes lacunas nos resultados.

O sucesso de qualquer sistema de IA está firmemente ligado à qualidade dos dados de treinamento. Em outras palavras, dados distorcidos, não confiáveis ou tendenciosos levam a previsões não confiáveis. Alucinações de IA representam riscos substanciais também, pois o modelo gera código que parece correto, pois soa sintaticamente correto, mas o código está na verdade usando funções que não estão presentes ou bibliotecas obsoletas ou funções com o software, aumentando a dívida técnica para o sistema que o software pode assumir e, como resultado, limitando sua manutenção a longo prazo.

GenAI amplia a superfície de ataque e introduz ameaças (por exemplo, envenenamento de dados, que é quando um atacante manipula dados de treinamento para influenciar o comportamento do modelo, e ataques adversários que enganam o sistema com entradas maliciosas) em um grau maior.

Da mesma forma, pode haver uma preocupação crítica em relação à ameaça de vazamentos de Propriedade Intelectual e Informações Pessoais Identificáveis (PII) por meio de prompts ou saídas de modelos que não foram devidamente sanitizados. A dependência excessiva de assistentes de IA é prejudicial para habilidades críticas, especialmente para desenvolvedores juniores que podem perder o aprendizado de conceitos fundamentais ao aceitar recomendações de máquinas sem questionar.

As responsabilidades do engenheiro de software podem mudar de codificação manual para orquestração de IA e curadoria de dados, o que exigirá novas competências em engenharia de prompts, supervisão ética e análise crítica de arquiteturas geradas por máquinas.

A GenAI na engenharia de software não precisa evoluir apenas por sua eficiência, mas também pela sustentabilidade ecológica e autonomia sistêmica. O gasto energético para executar esses LLMs em larga escala é enorme. Contudo, pode-se averiguar que “Diferentemente do Google Tradutor e similares, as LLMs parecem ser capazes de entregar expressões mais fidedignas, mais similares às nativas e especialmente mais próximas do tom e estrutura de materiais acadêmicos”. (Sampaio et al., 2024a, p. 12).

A sustentabilidade deve ser considerada um requisito não funcional no SDLC, não apenas como um requisito secundário ao criar modelos para medir a pegada de carbono por meio de fluxos de trabalho habilitados por IA. A direção tecnológica é em direção a sistemas auto evolutivos e autorreparáveis.

Tais sistemas seriam capazes de verificar diretamente seu desempenho em tempo de execução e modificar autonomamente seu código (ou arquitetura) para resolver falhas observadas ou atender a novas necessidades dos usuários sem intervenção humana. A transição é comparada ao desenvolvimento de sistemas de condução totalmente autônomos a partir de sistemas de assistência ao motorista, nos quais a IA deixa de ser um dispositivo de suporte e se torna a principal força motriz do desenvolvimento de software e do ciclo de vida.

5. CONSIDERAÇÕES FINAIS

A transição para sistemas movidos por IA vai além de meros desafios técnicos; é uma questão estratégica, organizacional e ética. O sucesso de tudo isso está ligado a garantir que as máquinas não prejudiquem a justiça e a segurança, com supervisão humana e responsabilidade ética em jogo para que isso aconteça.

Estruturas de IA como o AI TRiSM (Gestão de Confiança, Risco e Segurança) ilustram como a governança de IA não é mais apenas uma decisão técnica, mas uma questão de conformidade e uma nova oportunidade competitiva para as empresas. Empresas com monitoramento extensivo diminuem o risco de incidentes reputacionais graves em até 40%, mostrando claramente que ética e confiança são ativos estratégicos centrais.

Não há dúvida de que o sucesso de qualquer projeto de IA generativa ou aprendizado de máquina depende diretamente de uma base de dados sólida. Sabe-se que no mundo real os dados, muitas vezes, não são de alta qualidade e nem estruturados, e melhorar sistematicamente o conjunto de dados é a condição para construir um modelo preciso — estamos iniciando um paradigma — IA Centrada em Dados. Uma das coisas mais importantes é a qualidade dos dados de entrada, para que se tenha boa qualidade de dados de saída, especialmente em IA: um modelo construído sobre uma base ruim tem uma saída imprecisa ou perigosa.

Para que a inovação técnica se transforme em relevância para os negócios, a governança é necessária, a fim de fornecer dados que cumpram os princípios de

diversidade, pontualidade, precisão, segurança, identificabilidade e consumibilidade. Dados bem governados protegem a privacidade dos usuários, a propriedade intelectual e o alinhamento entre o desempenho do modelo e a estratégia da empresa, para que a IA possa se tornar um parceiro colaborativo, capaz de evoluir com o fluxo de trabalho em constante mudança.

No final, a engenharia de software na era da IA precisa de profissionais e gestão organizacional para desenvolver uma cultura de aprendizado constante e mentalidade ágil. A prontidão dos dados e a capacidade de governança abrangente para atender às necessidades de toda a comunidade - a verdadeira integração de desenvolvedores, gerentes e comunidades afetadas é a base para a evolução da IA generativa para soluções de software sustentáveis, seguras, confiáveis e verdadeiramente inovadoras a longo prazo. O futuro verá uma coevolução humano-máquina ao lado das máquinas, de modo que a tecnologia possa fomentar o pensamento criativo com as pessoas, enquanto é guiada por um sistema de garantia ética e de qualidade que é responsável e controlado.

REFERÊNCIAS

BONI, N. **Inteligência artificial para testes de software**. 2024.Trabalho de Conclusão de Curso, Monografia. Universidade Federal de São Carlos, Curso de Ciências da Computação, 2024. Disponível em: https://repositorio.ufscar.br/handle/ufscar/19407. Acesso em: 10 nov. 2025.

CISCO. **Visão geral da segurança de computação da Cisco** - Documento técnico. Disponível em: https://www.cisco.com/c/en/us/products/collateral/servers-unifiedcomputing/ucs-manager/compute-security-overview-wp.html#:~:text=Gest%C3%A3o%20unificada%20de%20redes,Meraki%20e%20do%20Catalyst%20Center. Acewssado em: 14 jan. 2026.

GAROFALO, Í. M. A. D. **Impacto da IA Generativa na Produtividade em Equipes de Desenvolvimento de Software**. Disponível em: https://repositorio.ufsc.br/bitstream/handle/123456789/270933/TCC_Monografia_Italo_2025.pdf?sequence=1&isAllowed=y. Acessado em: 04 jan. 2026.

GLOSSÁRIO GARTNER, AI TRiSM. **Relatório da Gartner®**: Radar de Impacto de Tecnologias Emergentes: IA Generativa. Disponível em: https://www.gartner.com/en/information-technology/glossary/ai-trism. Acessado em: 13 dez. 2025.

ISSO/IET. **Confiabilidade**. ISO/IEC TS 5723:2022. Disponível em: https://www.iso.org/standard/81608 .html . Acessado em: 11 jan. 2026. (1ª edição, 2022).

ISSO/IET. **Sistemas de gestão da qualidade** — Fundamentos e vocabulário. Disponível em: https://www.iso.org/standard/45481.html. Acessado em: 11 jan. 2026. (4ª edição, 2015).

MICROSOFT LEARN. **CheckBox**. Disponível em: https://learn.microsoft.com/ptbr/dotnet/maui/user-interface/controls/checkbox?view=net-maui-10.0. Acessado em: 13 jan. 2026.

MORIMOTO, T. **Inteligência artificial e ser humano se complementam:** o conceito de Human in the Loop (HITL) aplicado à manutenção industrial. (2025). Acessado em: 09 jan. 2026. Disponível em: https://blog.futago.ai/ia-industria/inteligencia-artificial-eser-humano-se-complementam-o-conceito-de-human-in-the-loop-hitl-aplicado-amanutencao-industrial.

NIST. **Riscos e Confiabilidade da IA**. National Institute of Standards and Technology (NIST). Disponível em: https://airc.nist.gov/airmf-resources/airmf/3-seccharacteristics/. Acessado em: 05 jan. 2026.

PALO ALTO. **O que é Aprendizado de Máquina (ML)?** https://www.paloaltonetworks.com/cyberpedia/machine-learning-ml#:~:text=Os%20modelos%20de%20aprendizado%20de%20m%C3%A1quina%20podem%20processar%20grandes%20quantidades,o%20desempenho%20geral%20dos%20neg%C3%B3cios. Acessado em: 17 jan. 2026.

PARLAMENTO EUROPEU. **Lei da UE sobre IA: primeira regulamentação de inteligência artificia**. (2023). Disponível em: https://www.europarl.europa.eu/topics/pt/article/20230601STO93804/lei-da-ue-sobre-a-primeira-regulamentacaodeinteligencia-artificial#:~:text=e%20mais%20sustent%C3%A1vel.-,Regular%20a%20intelig%C3%AAncia%20artificial%20na%20Europa%3A%20o%20primeiro%20quadro%20abrangente,de%20IA%20baseado%20no%20risco. Acessado em: 06 jan. 2026.

QLIK. **Os seis princípios dos dados prontos para IA criando uma base confiável de dados para IA.** Disponível em: https://assets.qlik.com/image/upload/v1744376151/qlik/docs/resource-library/whitepapers/resource-wp-the-6-principles-of-aready-data-br_i5flfx.pdf#:~:text=Dessa%20forma%2C%20a%20precis%C3%A3o%20dos%20dados%20%C3%A9,suas%20caracter%C3%ADsticas%2C%20integridade%2C%20distribui%C3%A7%C3%A3o%2C%20redund%C3%A2ncia%20e%20forma. Acessado em: 17 jan. 2026.

SAMPAIO, R. C. et al. **Uma revisão de escopo assistida por Inteligência Artificial (IA) sobre usos emergentes de IA na pesquisa qualitativa e suas considerações éticas**. Revista Pesquisa Qualitativa, v. 12, p. 01-28, 2024b. DOI: https://doi.org/10.33361/RPQ.2024.v.12.n.30.729.

SEYFFERT, P. **Governança de IA: fundamentos, práticas e liderança para grandes empresas.** (2025). Disponível em: https://www.softdesign.com.br/blog/governanca-deia/#h-o-que-e-governanca-de-ia-nbsp-nbsp-nbsp-nbsp-nbsp-nbsp-nbsp-nbsp. Acessado em: 16 jan. 2026.

SILVA, G. B. R. **Arquitetura serverless**: Economia e Automatização no Armazenamento e Processamento de Dados.2024. 81f. Trabalho de Conclusão de Curso (Graduação em Engenharia de Computação) – Universidade Federal do Pampa, Campus Bagé, Bagé, 2024.

WESSEL, M. et al. **Call for papers to the special issue**: Generative ai and its transformative value for digital platforms. Journal of Management Information Systems, 2023. Disponível em: https://www.jmis-web.org/cfps/JMIS_SI_CfP_Generative_AI.pdf. Acessado em: 11 dez. 2025.