Framework de Auditoria de Inteligência Artificial

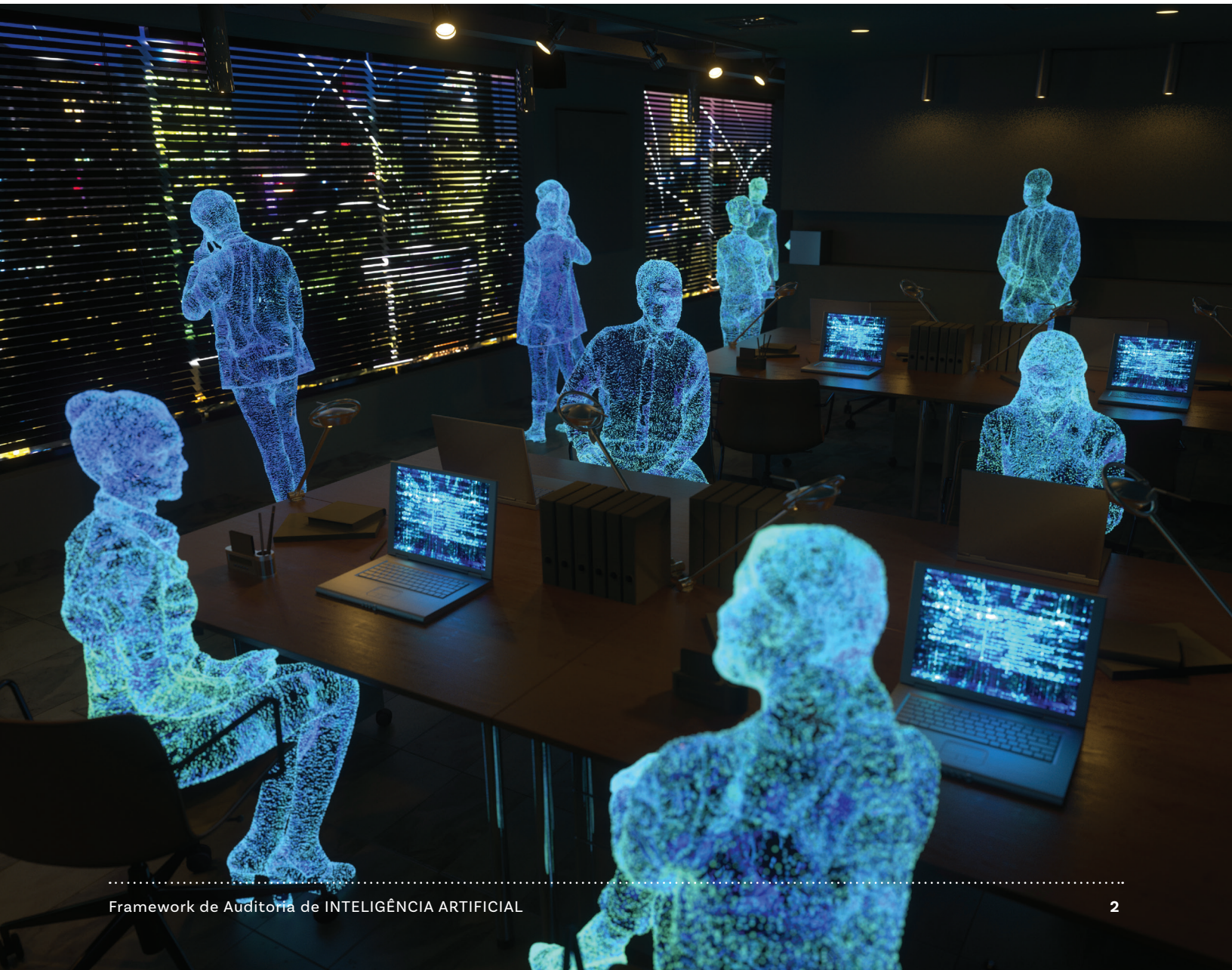
DO THE IIA

TRADUZIDO POR • INSTITUTO DOS AUDITORES INTERNOS DO BRASIL •

IIA Brasil

The Institute of Internal Auditors

Índice



Introdução

Inteligência artificial (IA) é um termo amplo que cresceu para abranger uma grande variedade de tecnologias existentes e emergentes. Embora não haja uma definição única e consensual, a IA geralmente se refere a "sistemas dotados de processos intelectuais característicos dos seres humanos, como a capacidade de raciocinar, descobrir significados, generalizar ou aprender com experiências passadas". O atual boom de aplicativos de IA demonstra maneiras aparentemente infinitas pelas quais as organizações podem alavancar tecnologias orientadas por IA para aprimorar a forma como trabalhamos, porém com riscos numerosos e significativos que são inerentes, dada a natureza da tecnologia.

A IA pode ser um tópico assustador para um auditor interno, especialmente porque a adoção e o uso da IA pelas organizações continuam crescendo. Agora, mais do que nunca, as organizações estão procurando a auditoria interna para obter maior orientação sobre IA. Seja como consultor sobre riscos e controles relacionados à IA ou em um papel de avaliação de riscos em processos que usam ou dependem da IA, é vital que os auditores internos aumentem seu conhecimento sobre o assunto da IA.

Espera-se que os auditores internos prestem atividades de avaliação em torno de processos que podem variar de simples transações comerciais a procedimentos altamente complexos que exigem profundo entendimento. A variedade e a profundidade da alfabetização em IA necessárias para apoiar as atividades de avaliação criam desafios contínuos para os auditores internos, que devem desenvolver continuamente seu conhecimento sobre IA para entender plenamente seus riscos e sua função para prestar consultoria e avaliação informadas.

A IA como assunto de auditoria apresenta seus próprios desafios únicos e sua evolução significa que os auditores internos devem reavaliar os riscos e sua mitigação no ambiente de IA. Dito isso, os auditores internos já possuem importantes habilidades fundamentais, como raciocínio crítico, mapeamento de processos, avaliação de riscos, avaliação de controles de tecnologia da informação, compreensão das estratégias organizacionais e prestação de avaliação independente para a função de governança.

A intenção do Framework de Auditoria de IA do The Institute of Internal Auditors (IIA) é ajudar os auditores internos a entender o risco e identificar as melhores práticas e controles internos para a IA. O framework ajudará os auditores internos a desenvolver um conhecimento básico. Ele é apresentado em quatro partes:

1. ***Visão geral – História e usos da IA.***
2. ***Primeiros passos – Entendendo como uma organização usa a IA.***
3. ***Framework de Auditoria de IA – Governança, Gestão e Auditoria Interna.***
4. ***Guia e Glossário do Profissional.***

Esse framework, que alavanca aspectos do Modelo das Três Linhas do The IIA,[1] incluirá referências ao Framework Internacional de Práticas Profissionais (IPPF) do The IIA, que fornece uma base de requisitos obrigatórios e princípios orientadores para a profissão de auditoria interna. As Normas aplicáveis deveriam ser revisadas para obter informações adicionais. Orientações relacionadas do The IIA, como os *Global Technology Audit Guides* (GTAGs), são citadas para fornecer conteúdos sobre tópicos específicos. Outros frameworks relevantes, como o *NIST Artificial Intelligence Risk Management Framework* (AI RMF 1.0), estão listados como recursos adicionais para os profissionais de auditoria interna.

PARTE 1
Visão Geral

História e Evolução

Como uma visão geral do tópico, é importante que os auditores internos entendam o desenvolvimento histórico da IA, a forma como a IA é usada atualmente em diversas indústrias e quais tendências emergentes de IA os auditores internos deveriam considerar.

A ideia da IA remonta a 1950, quando o matemático britânico Alan Turing fez a pergunta: "As máquinas podem pensar?" em seu artigo, "*Computing Machines and Intelligence.*"[2] Ele é considerado um dos fundadores da IA, ao sugerir que as máquinas futuramente seriam capazes de ter inteligência semelhante à humana. Dois anos depois, Arthur Lee Samuel, um cientista da computação americano da IBM, desenvolveu um programa capaz de jogar o jogo de tabuleiro de damas usando valores programados para identificar a melhor jogada.[3] The projeto *Summer Research Project on Artificial Intelligence* de Dartmouth em 1956 marcou um dos primeiros usos do termo IA, creditado a John McCarthy, um cientista americano cognitivo e de computação.[4]

A década de 1960 registrou muitos avanços em IA, inclusive o uso da robótica, programas de resolução de problemas e o primeiro programa de computador interativo (também conhecido como programa de compreensão de linguagem natural ou PLN) chamado ELIZA, criado por Joseph Weizenbaum, cientista de computação e professor germano-americano. ELIZA poderia ser considerado o primeiro "chatbot", criado para simular uma conversa com um usuário humano.[5]

O desenvolvimento da IA na década de 1970 incluiu o primeiro robô inteligente chamado WABOT, desenvolvido pela Escola de Ciências e Engenharia da Universidade de Waseda, em Tóquio, bem como o trabalho contínuo em PLNs pelo cientista da computação indiano-americano Raj Reddy.[6,7] Os avanços na década de 1980 incluíram a criação de uma van Mercedes Benz sem motorista em 1986, sob a supervisão de Ernst Dickmanns, líder alemão da tecnologia de direção autônoma.[8]

A década de 1990 registrou avanços nas tecnologias relacionadas à IA, incluindo o software de reconhecimento de fala no Windows da Microsoft. A IBM desenvolveu IAs altamente eficazes, como a "Deep Blue", que virou manchete em 1997 quando derrotou o grande mestre do xadrez Garry Kasparov.[9]

Na década de 2000, a IA passou a fazer parte do nosso cotidiano, incluindo aplicativos como Alexa da Amazon, Siri da Apple e Google Assistant. O ano de 2023 marcou o ano do aumento da adoção de modelos grandes de linguagem (LLMs), como o ChatGPT, que elevaram ainda mais os recursos da IA, desde a simples previsão de resultados até uma variedade de criação de conteúdo.

Níveis de Adoção

De acordo com o Índice Global de Adoção de IA de 2023 da IBM, 42% das empresas entrevistadas relataram usar IA em seus negócios e outros 40% relataram que estão explorando a IA.[10] A expansão

contínua da IA destaca por que os auditores internos devem se certificar de que estão incorporando os riscos relacionados à IA no planejamento de suas auditorias. Além disso, os auditores internos deveriam desenvolver continuamente seu conhecimento sobre IA. A Parte 2, Primeiros Passos, aprofundará as considerações que um auditor interno pode usar para identificar o uso da IA em suas organizações.

Assim como o cenário da IA continua evoluindo, o mesmo acontece com as formas de se categorizar a IA. Embora haja diferentes perspectivas sobre como agrupar as diversas formas de IA, a seção a seguir fornece um breve resumo das formas comuns de IA, seja em uso atualmente (Máquina Reativa e Memória Limitada) ou estritamente teórica (Teoria da Mente e Autoconsciência).

Do ponto de vista de funcionalidade, a IBM[11] categoriza a IA em quatro tipos:

1. ***IA de máquina reativa***
2. ***IA de memória limitada***
3. ***IA de teoria da mente***
4. ***IA autoconsciente***

Uso da IA

- ***AI Business Survey* (EUA) de 2022 da PwC — A tomada de decisões apoiada por IA está sendo utilizada por 74% dos líderes de tecnologia entrevistados, 62% dos líderes de operações e manutenção, 61% dos líderes de experiência do cliente e 60% dos líderes de estratégia.**
- ***Global CEO Outlook Pulse Survey* de 2023 da EY — 99% dos CEOs entrevistados estão fazendo ou planejando grandes investimentos em IA generativa.**
- **State of AI de 2023 da McKinsey — 79% de todos os entrevistados globais relatam alguma exposição à IA generativa e 22% disseram que a usam regularmente.**

1. IA de Máquina Reativa

- A IA de Máquina Reativa é um tipo de IA sem memória, projetada para executar tarefas com base apenas no input ou treinamento humano. Às vezes chamados de IA Estreita ou Fraca, esses sistemas dependem do "humano no circuito" para a programação gerada pelo humano, que instrui a máquina a operar por conta própria. Essa programação é comumente chamada de "algoritmo", ou seja, um conjunto de cálculos que inclui aspectos da ciência da computação e da matemática ou estatística.

EXEMPLOS:

- Deep Blue da IBM.
- Alguns aplicativos de aprendizado de máquina são classificados como IA de máquina reativa. O aprendizado de máquina geralmente se baseia em modelos estatísticos que analisam dados e produzem resultados preditivos a partir dos dados inseridos. Por exemplo, um varejista on-line poderia usar a IA em seu aplicativo móvel ou site que sugere produtos com base no histórico de compras de um consumidor. O histórico de compras de um usuário individual é o conjunto de dados que gera o resultado, que é personalizado para esse usuário.

2. IA de Memória Limitada

- A IA de Memória Limitada é menos dependente da interação humana para produzir resultados, o que lhe dá a capacidade exclusiva de aprender e melhorar com base no treinamento de conjuntos de dados maiores. Enquanto a IA de Máquina Reativa só pode utilizar dados disponíveis no momento, a IA de Memória Limitada pode incorporar dados passados e atuais para melhorar o desempenho em uma série de novas ferramentas de IA.
- Outros exemplos de aprendizado de máquina acessam conjuntos de dados muito maiores e realizam análises mais complexas. Isso é chamado de "*deep learning*" (aprendizado profundo), que é um subconjunto do aprendizado de máquina e da IA de Memória Limitada. Diferencia-se por ser menos dependente da interação humana para produzir resultados. A IA generativa se enquadra nessa categoria e pode ser utilizada para criar conteúdo com base nos algoritmos programados.

EXEMPLOS DE IA GENERATIVA INCLUEM:

- ChatGPT, LLaMA e Bard são exemplos de chatbots LLM que dependem do *deep learning* e podem produzir conteúdo de texto, enquanto outras formas de IA generativa podem produzir conteúdo como música (MusicLM), arte (DALL-E) e até mesmo código de programação de computador (OpenAI Codex).
- Chatbots e assistentes virtuais são uma forma comumente usada de IA de memória limitada que utiliza o processamento de linguagem natural (PLN) e o aprendizado por reforço para se envolver em conversas semelhantes às humanas com os usuários finais. São frequentemente usadas por empresas como instituições financeiras, permitindo que o usuário solucione suas próprias demandas de suporte mesmo fora do horário comercial.

Além disso, o aprendizado de máquina é frequentemente subdividido em quatro categorias:[12]

1. ***Aprendizado supervisionado*** – o aprendizado passado é aplicado a dados estruturados com resultados predeterminados.
2. ***Aprendizado não supervisionado*** – não há resultados "corretos" predeterminados; em vez disso, procura padrões em dados não estruturados.
3. ***Aprendizado semissupervisionado*** – contém elementos do aprendizado supervisionado e do não supervisionado.
4. ***Aprendizado reforçado*** – programação dinâmica em que os algoritmos são treinados por meio de um sistema de recompensas e punições; ele aprende sem interação humana.

Outros tipos de IA:

Sistemas especializados simulam o julgamento ou o comportamento humano. Incorporam o conhecimento de várias pessoas na solução de problemas e, em teoria, fornecem soluções mais eficazes. São usados em pesquisas químicas, para analisar e prever a estrutura molecular, e na medicina, para identificar bactérias nocivas.

A tecnologia de visão computacional, combinada com o deep learning, permite que as máquinas analisem imagens. Atualmente, ela está sendo utilizada na área da saúde, para detectar e diagnosticar anomalias em pacientes com base em radiografias, ressonâncias magnéticas ou tomografias computadorizadas. O reconhecimento facial é outra forma de visão computacional, que tem vários usos, incluindo a autenticação ao tentar acessar uma conta bancária ou restringir o acesso físico a edifícios que abrigam dados confidenciais.

Embora a robótica e a IA sejam campos distintos, os dois são frequentemente combinados para criar ferramentas emergentes que podem ser usadas no mundo real. Por exemplo, os robôs físicos que usam sensores visuais e processamento de imagens usam a IA para aprender a navegar em seu ambiente. A manufatura, agricultura e bens de consumo embalados são setores que também dependem da robótica combinada com a IA de memória limitada para criar eficiência e aumentar a produtividade em suas operações. O uso de cirurgias robóticas e assistidas por IA no setor de saúde promove maior precisão, podendo resultar em uma recuperação mais rápida do paciente.

3. IA de Teoria da Mente

Embora a IA de Teoria da Mente ainda não exista, as pesquisas atuais visam desenvolver sistemas de IA que compreendam e interajam com fatores diferenciados, como emoções e motivações, de forma semelhante à humana. O trabalho em andamento nessa área inclui esforços para desenvolver sistemas que possam analisar e interagir com humanos com base em informações de suas vozes, expressões faciais e sentimentos em tempo real e responder de forma semelhante à humana.

4. IA Autoconsciente

A IA Autoconsciente, assim como a IA da Teoria da Mente, é atualmente teórica e não existe na prática. Mais intimamente relacionada às conversas recentes sobre a possibilidade de "AGI" (artificial general intelligence) ou inteligência artificial geral, essa versão hipotética de IA seria especialmente consciente de si mesma, com o que muitos imaginam como uma consciência interna rica, que iguala ou excede o que os humanos são capazes de fazer. Embora seja um ponto de discussão popular nos últimos meses, há um debate contínuo sobre a viabilidade desse nível de funcionalidade da IA.

PARTE 2
Primeiros Passos

Conforme as organizações continuam implantando a IA de diversas formas, os auditores internos devem ser proativos e colaborar estreitamente com a gestão para entender a estratégia geral da organização para a IA, como a IA está sendo usada atualmente e qual uso futuro está planejado. Especialmente durante o processo de planejamento, os auditores internos deveriam começar pesquisando e reunindo informações relevantes sobre o uso potencial da IA, sob revisão por várias fontes internas e externas.

Informações internas importantes podem incluir:

- Políticas e procedimentos que fazem referência à IA, que podem ser coletados e revisados para entender melhor os processos organizacionais.
- Iniciativas estratégicas documentadas de uma organização ou o plano estratégico, incluindo aspectos de IA.
- Relatórios recentes do conselho, contendo a visão e informações sobre como a liderança e o conselho estão discutindo tópicos como o uso da IA e preocupações relacionadas a riscos.
- Informações obtidas em reuniões de avaliação de riscos em andamento com os stakeholders.

Recursos externos podem fornecer referências adicionais quando os auditores internos começarem a analisar a estratégia de IA da organização. Recursos externos valiosos podem incluir:

- *Global Perspectives & Insights* de três partes do The IIA: *The Artificial Intelligence Revolution*.[13]
- Série *Artificial Intelligence 101* do The IIA.
- *Analytics, Automation, and AI Virtual Conference* do The IIA.
- Recursos fundamentais de auditoria de TI e cibersegurança, como os Programas de Certificação do The IIA *Auditing the Cybersecurity Program* e *IT General Controls*.
- Guias Práticos e *Global Technology Audit Guides* (GTAGs) do The IIA.
- *NIST AI Risk Management Framework* (AI RMF 1.0).
- *Guidelines for Secure AI System Development*, do National Cybersecurity Centre.[14]
- Ordem executiva de IA da Casa Branca de outubro de 2023.[15]
- eBook de governança de IA da IBM.[16]

Ambiente de Controle no Nível da Entidade: Execução e Estratégia

Uma vez que os auditores tenham se equipado com esses recursos, perguntar "Como a IA está sendo usada?" é uma pergunta inicial simples e eficaz para coletar informações. A resposta a essa pergunta provavelmente implicará em perguntar a vários indivíduos ou departamentos, pois muitas organizações não têm gestão centralizada de IA, nem políticas estabelecidas (incluindo a definição do que é IA), procedimentos ou uma estratégia relativa ao uso aceitável de IA.

Para organizações em que a IA foi desenvolvida e implantada, um auditor interno deveria ter uma discussão com a equipe de IA/ciência de dados. Essa discussão deveria incluir pedir que explicassem quais IA/algoritmos foram implantados, incluindo sua função, fontes de dados utilizados, uso, limitações, riscos e implicações éticas. Os auditores internos também deveriam começar a entender quais controles existentes estão em vigor para ajudar a gerenciar os riscos apresentados pela IA — ou se a gestão implementou novos controles relacionados ao seu uso e implantação de sistemas de IA. Obter uma compreensão preliminar da criação dos controles usados para gerenciar os riscos relacionados à IA é um passo importante que pode ser realizado em conjunto com essas discussões iniciais.

Para organizações em que não está claro se ou como a IA está sendo utilizada (formal ou informalmente), a função de TI da organização é um bom ponto de partida, porque, como observado na seção Níveis de Adoção na Parte 1, os líderes de tecnologia parecem ter uma tendência maior de experimentar e utilizar a IA em seu departamento. Se a TI confirmar que a IA está sendo usada, ou se as investigações iniciais determinarem que a IA está sendo usada na organização, a próxima investigação lógica é determinar até que ponto a IA é utilizada.

Embora uma conversa inicial com a equipe de IA/ciência de dados ou com a gestão de TI seja um bom primeiro passo, a discussão não deveria se limitar a esses grupos. A partir dessas discussões iniciais, os auditores internos podem descobrir que outros departamentos ou usuários individuais estão usando a IA para sua função específica, o que exigiria conversas adicionais. É orientado trabalhar com a gestão para revisar ou colaborar na criação de um inventário de quais departamentos estão usando IA atualmente e atualizar essa lista com frequência. O inventário deveria incluir outros aspectos principais, como a meta ou o objetivo da IA, quem a utiliza, quem a gerencia, as ferramentas específicas de IA em uso, considerações de risco e quem a supervisiona. O processo de revisão ou colaboração com a gestão para desenvolver um inventário de IA também poderia ser realizado durante o processo anual de avaliação de riscos.

A maioria dos auditores internos trabalha em estreita colaboração com seu CFO em relação ao teste dos controles internos sobre o reporte financeiro, ou com outros executivos, como o CISO, CIO, etc., portanto, ter esse relacionamento profissional com membros da gestão executiva deveria proporcionar outra oportunidade para conversas iniciais sobre IA. Perguntas importantes que os auditores internos podem fazer a seus executivos incluem:

- "Foi definida uma estratégia de IA e, se sim, quais são os detalhes dessa estratégia (incluindo aspectos como o uso da IA para maximizar a eficiência das operações ou IA para reduzir custos)?"
- "A gestão executiva determinou quem é responsável pelo gerenciamento de riscos relacionados à IA?"
- "Que papel a gestão executiva desempenha em envolver o Conselho de Administração (ou equivalente) em considerações de governança de IA?"

Nesse ponto, os auditores internos terão:

- Pesquisado sobre a IA em sua organização e revisado os recursos externos.
- Conduzido conversas iniciais sobre IA com os gestores, incluindo sua equipe de IA/ciência de dados ou gestão de TI (ou ambos) e a equipe de liderança executiva (CFO, CISO, CIO, etc.).
- Colaborado com a gestão na revisão ou desenvolvimento de um inventário para registrar como a IA está sendo utilizada (ou planejada para uso futuro).
- Iniciado o processo de compreensão da governança de IA em vigor.

Realizar essas quatro tarefas indicaria que a auditoria interna deu os primeiros passos para estabelecer um conhecimento básico da IA na organização. Também ofereceria uma oportunidade para a auditoria interna enfatizar quaisquer observações imediatas que deveriam ser comunicadas à gestão de forma tempestiva.

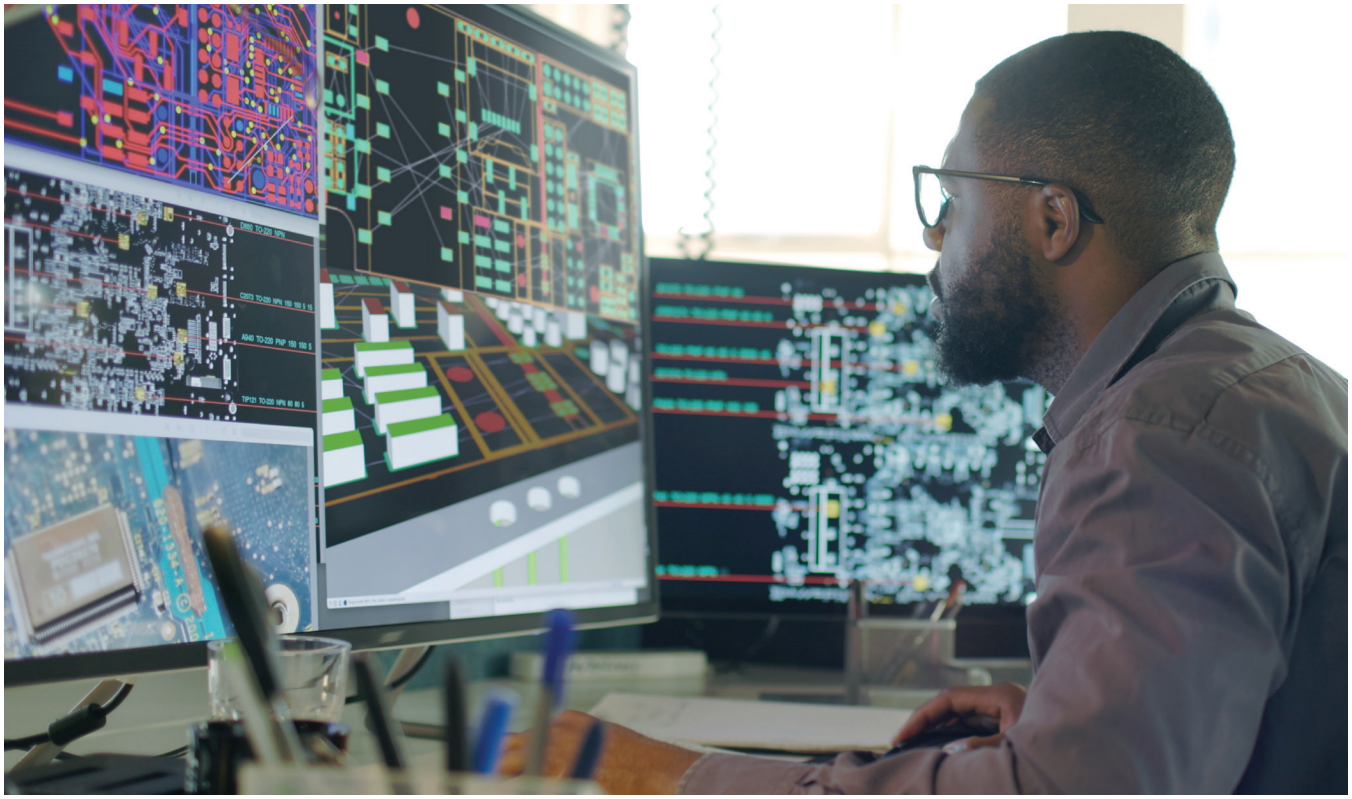
Dados

Depois que os auditores internos tiverem um entendimento fundamental de como a IA está sendo usada, eles deveriam desenvolver um conhecimento mais robusto do uso da IA dentro da organização. Como os algoritmos usados para alimentar a IA dependem de grandes volumes de dados (também chamados de "Big Data"), é fundamental determinar quais dados organizacionais estão sendo usados em qualquer aplicativo de IA e como esses dados são gerenciados. Um algoritmo é um conjunto de regras a serem seguidas pela IA e é o que permite que uma máquina processe rapidamente grandes quantidades de dados que um ser humano não pode razoavelmente processar com a mesma facilidade ou velocidade. Dada a capacidade da IA de processar e responder rapidamente a grandes quantidades de conjuntos de dados diversos, a arquitetura, o desempenho e a precisão dos algoritmos envolvidos são muito importantes.

Os algoritmos são inicialmente desenvolvidos por humanos, portanto, erros e vieses humanos (intencionais e não intencionais) poderiam afetar o desempenho do algoritmo. A Parte 3 deste framework fornecerá mais detalhes sobre os riscos relacionados a erros e vieses dos algoritmos.

Fora da IA, muitas organizações já desenvolveram uma estratégia para coletar, armazenar, usar, gerenciar e proteger dados. A IA é como outros aplicativos orientados por dados, pois os mesmos aspectos importantes sobre os dados são relevantes e deveriam ser considerados, incluindo integridade, privacidade, confidencialidade, validade, precisão e completude.

Big Data significa mais do que apenas grandes quantidades de dados — Big Data refere-se a dados que atingem volume, variedade, velocidade e variabilidade tão altos que as organizações investem em arquiteturas de sistemas, ferramentas e práticas especificamente projetadas para gerenciar os dados. Grande parte desses dados pode ser gerada pela própria organização, enquanto outros dados podem estar disponíveis publicamente ou ser adquiridos de fontes externas. Para obter orientações abrangentes sobre a compreensão e a auditoria do Big Data, incluindo uma discussão sobre oportunidades e riscos e um exemplo de programa de trabalho, consulte o *"GTAG: Understanding and Auditing Big Data"* do The IIA.



Outro aspecto crítico do uso de dados e aplicativos de IA relacionados é se os dados são hospedados ou processados por uma parte externa à organização. Os auditores internos devem sempre considerar os riscos relativos às transações de terceiros (e quartos), porque os ambientes de controle interno dos fornecedores podem não ser tão abrangentes quanto o ambiente da organização (ou o ambiente de controle desejado do fornecedor). O Guia Prático do The IIA *"Auditing Third-party Risk Management"* fornece aos auditores internos uma abordagem mais detalhada sobre os riscos relativos ao uso de fornecedores externos.

Outro aspecto importante dos dados é o acesso do usuário. Entender quem pode editar ou fazer alterações nos dados é fundamental, pois a manipulação de conjuntos de dados, do ponto de vista da entrada de dados, pode certamente afetar o resultado posterior da IA. Também é fundamental entender e documentar o acesso do usuário administrador aos dados que dependem da IA. O *"GTAG: Auditing Identity and Access Management"* do The IIA oferece uma visão mais detalhada das considerações de auditoria interna relativas a como a organização assegura que os usuários tenham acesso apropriado aos recursos de TI.

Cibersegurança

A cibersegurança também deve ser considerada no que se refere a restringir o acesso de usuários não autorizados aos dados e assegurar a privacidade, confidencialidade e proteção dos dados. A adoção e a evolução da IA estão forçando as organizações a enfatizar novamente seus recursos de resiliência cibernética. Conforme a IA se torna mais poderosa e mais decisões são entregues a algoritmos novos, complicados e opacos usando enormes conjuntos de dados, proteger esses sistemas de forças externas e danosas é fundamental para o sucesso organizacional. A resiliência cibernética é vital para qualquer organização que utilize IA.

Os auditores internos estão tipicamente envolvidos em testar a eficácia dos controles internos de TI. Essa familiaridade de como a organização implementou controles internos relativos à cibersegurança pode ajudar os auditores internos a validar se esses mesmos controles estão sendo usados para proteger os dados relacionados à IA. Exemplos de controles de cibersegurança incluem:

- Uso de criptografia.
- Presença de software antivírus.
- Uso de sistemas de prevenção/detecção de intrusão.
- Registro de eventos de segurança, tanto de solicitações quanto de respostas.
- Assegurar que um teste de penetração seja realizado periodicamente, para buscar vulnerabilidades de forma proativa.
- Treinamento de funcionários sobre práticas recomendadas para detectar e evitar *phishing*, *smishing* ou outros mecanismos de engenharia social.

Para obter mais detalhes, consulte o *"GTAG: Auditing Cybersecurity Operations: Prevention and Detection"*.

Os auditores internos precisam determinar onde os dados dependentes de IA são armazenados (internamente, externamente ou ambos) e considerar quais controles de cibersegurança estão em vigor. Para dados armazenados externamente, um relatório SOC *(service organization company)* deveria ser obtido, para conhecer o ambiente de controle do fornecedor. A gestão deveria estar consciente de quaisquer deficiências de controle encontradas no relatório SOC e assegurar que essas deficiências não coloquem em risco os dados dependentes de IA. Os contratos de nível de serviço *(service-level agreements* — SLAs) com os fornecedores deveriam incluir uma cláusula de "direito de auditoria".

PARTE 3

Framework de Auditoria de IA

A primeira versão do Framework de Auditoria de IA do The IIA foi emitida em 2017. Ofereceu aos profissionais de auditoria interna uma abordagem para a execução de serviços de avaliação e consultoria de IA de forma sistemática e disciplinada. Essa versão atualizada do framework moderniza o conteúdo com exemplos do ambiente atual de IA, ao mesmo tempo em que fornece detalhes adicionais para auxiliar os auditores internos como consultores e prestadores de avaliação. O framework tem três domínios:

Framework de Auditoria de IA do IIA

- ***Governança***
- ***Gestão***
- ***Auditoria Interna***

O framework está vinculado ao Modelo das Três Linhas do The IIA: o órgão de governança (Governança) supervisiona a gestão (Primeira e Segunda Linhas), enquanto o papel da auditoria interna é abordado no terceiro domínio, que inclui tanto atividades de avaliação independente (Terceira Linha) quanto de consultoria.

O Framework de Auditoria de IA do The IIA destina-se ao uso por auditores internos. Entretanto, os domínios de Governança e Gestão do framework descrevem as atividades e funções fora da auditoria interna necessárias para gerenciar a IA em uma organização. O principal objetivo do framework é equipar os auditores internos com o conhecimento básico essencial de IA para servir à sua organização como 1) conselheiro da gestão, para consultoria sobre a abordagem geral de como a IA é gerenciada, executada e monitorada e/ou como 2) prestador de avaliação, para auditar os processos e controles que a gestão estabeleceu para gerenciar, executar e monitorar a IA.

A maturidade organizacional do uso da IA contribui para a forma como a auditoria interna será aproveitada. Por exemplo, organizações menos maduras em IA podem precisar que a auditoria interna assuma um papel de conselheira na exploração inicial da IA, enquanto uma organização mais madura em IA provavelmente trabalharia com a auditoria interna para prestar atividades de avaliação, como a análise dos processos estabelecidos e dos controles internos quanto à eficácia operacional. Para desempenhar ambos os papéis com sucesso, a auditoria interna precisa de um sólido entendimento de como a IA deveria ser gerenciada e como a organização a está gerenciando atualmente.

Governança — o primeiro domínio do framework — baseia-se na abordagem da organização ao planejamento estratégico da IA e na supervisão e monitoramento de como a IA é planejada, gerenciada e executada pela gestão. O órgão de governança baseia-se nas informações que lhe são fornecidas pela função de auditoria interna. Os auditores

internos deveriam se esforçar para desenvolver um relacionamento de conselheiro confiável com os órgãos de governança, como o comitê de auditoria, conselho ou órgão de governança equivalente, e esse relacionamento deveria incluir tópicos emergentes, como IA, que apresenta novos desafios de supervisão.

O domínio **Gestão** do framework descreve uma abordagem que a organização usaria ao planejar e executar a IA dentro da organização. O ambiente de controle interno que envolve a IA é estabelecido pela gestão na "Primeira Linha". Também inclui aspectos estratégicos, como definição de metas e objetivos relacionados ao plano estratégico geral de IA. A auditoria interna deve garantir que entenda a direção estratégica da IA para a organização e a abordagem da gestão para gerenciar a IA.

O domínio de Gestão do framework também contém aspectos de monitoramento de "Segunda Linha" da IA, como, por exemplo, quais aspectos o gerenciamento de riscos corporativos deveria considerar ao monitorar a "Primeira Linha". Espera-se frequentemente que a auditoria interna participe do processo de avaliação de riscos de uma organização. Esse domínio será relevante para a auditoria interna, visto que ela detém conhecimento dos riscos relacionados à IA.

O terceiro domínio do framework, **Auditoria Interna,** inclui aspectos tanto das atividades de consultoria para a gestão quanto da prestação de serviços de avaliação em uma capacidade de auditoria ("Terceira Linha"). A auditoria interna pode usar o framework como ponto de partida em ambas as funções, quando encarregada de participar de atribuições de IA.

Como a IA está evoluindo rapidamente, o framework exigirá atualizações periódicas. Essa evolução, combinada com a natureza complexa da IA, significa que a auditoria interna provavelmente será capaz de prestar apenas avaliação limitada. O framework, por si só, pode não abranger todos os aspectos da IA, mas fornece uma base sólida para os auditores internos, conforme desenvolvem o conhecimento fundamental da IA como tópico de auditoria.

Governança

A governança da IA refere-se às estruturas, processos e procedimentos implementados para dirigir, gerenciar e monitorar as atividades de IA da organização. A governança inclui ajudar a garantir que as atividades, decisões e ações de IA sejam consistentes com os valores da organização, bem como com suas responsabilidades éticas, sociais e legais. Inclui também a supervisão para garantir que os funcionários com responsabilidades de IA tenham as habilidades e os conhecimentos necessários.

Conforme refletido no Modelo das Três Linhas, a auditoria interna funciona como a "Terceira Linha", prestando avaliação independente e objetiva sobre a validação dos controles internos usados pela organização para o gerenciamento de riscos, incluindo todos os aspectos da IA. A auditoria interna pode prestar serviços de consultoria relacionados à IA para a organização, mas, do ponto de vista da governança, o órgão de governança depende muito das atividades de avaliação prestadas pela auditoria interna para entender melhor a eficácia operacional da organização.

A governança da IA é vital. Dois dos papéis mais importantes que a governança desempenha são a avaliação de quão bem a organização está gerenciando as operações de IA e se as metas e objetivos estratégicos de IA da organização estão sendo concretizados de forma consistente com os valores estabelecidos. Conforme apresentado nas seções anteriores, há uma série de riscos específicos à IA; no entanto, uma das principais considerações é supervisionar se a IA está sendo usada de uma forma que não cause danos.

Estratégia

Um plano estratégico permite que a organização esclareça e comunique a direção e a visão necessárias para concretizar suas metas; o mesmo ocorre com uma estratégia de IA. A estratégia de IA de cada organização deveria ser única, com base em sua abordagem para capitalizar as oportunidades que a IA oferece e, ao mesmo tempo, estar atenta às circunstâncias específicas de uma organização, como os detalhes dos serviços de tecnologia atuais ou iniciativas atuais de governança de dados. Uma abordagem estratégica de IA cuidadosa e metódica apoiará a capacidade da organização de concentrar seus recursos e promover o alinhamento entre todos os funcionários, ao mesmo tempo em que mitiga os riscos potenciais.

Dois pontos importantes para se ter em mente:

1. O planejamento de uma estratégia de IA não é um evento único; é um processo iterativo que deveria ser realizado de forma periódica. A auditoria interna deveria trabalhar com a gestão para determinar um cronograma para revisões da estratégia de IA.

2. Uma estratégia de IA não deveria ser planejada isoladamente; dada a gama de potenciais fontes de dados e casos de uso, as estratégias organizacionais de IA deveriam ser multifuncionais. Dada a importância crítica da IA, é provável que haja envolvimento e supervisão do conselho, pois a IA tem o potencial de alterar ou modificar drasticamente as estratégias de negócios.

Abordar esses pontos ajudará a garantir que as iniciativas de IA apoiem os objetivos gerais da organização e se alinhem aos valores organizacionais declarados. A formulação de metas para a IA permite que as organizações enquadrem considerações estratégicas importantes, incluindo a resposta a perguntas básicas como: "Por que estamos usando IA?" e "O que estamos tentando concretizar?" As metas de IA deveriam ser desenvolvidas como outras metas organizacionais "SMART" — específicas, mensuráveis, concretizáveis, relevantes e baseadas no tempo —, para evitar a adoção de ferramentas e serviços de IA sem um escopo claro da razão da organização para fazê-lo.[17]

Os atributos desejados para a IA deveriam ser incluídos na definição de metas, objetivos e expectativas. As expectativas ou objetivos organizacionais podem incluir os seguintes atributos desejáveis para a inteligência artificial:

Atributos Desejáveis para a Inteligência Artificial

- **Eficaz**
- **Válida**
- **Confiável**
- **Segura**
- **Imparcial**
- **Transparente**
- **Ética**
- **Explicável**
- **Privada**
- **Obediente às leis**
- **Justa**
- **Confidencial**
- **Responsável**
- **Precisa**
- **Eficiente**
- **Responsabilizável**

A atitude e a abordagem gerais da organização em relação ao risco e ao gerenciamento de riscos deveriam ser uma consideração primária ao desenvolver ou atualizar o plano estratégico e as metas de IA. Ter um apetite a risco maior na busca das metas de IA pode não ser apropriado para uma organização que seja aversa ao risco em outros aspectos, ao passo que organizações com tolerância a risco historicamente alta podem estar mais dispostas a aceitar riscos relacionados à IA. Independentemente da tolerância a riscos de uma organização, é essencial reconhecer e mapear os riscos de IA durante o planejamento estratégico de IA.[18]

Gestão – Primeira e Segunda Linhas

Ao desenvolver a estratégia de IA, a gestão é responsável por assegurar que os controles internos tenham sido devidamente projetados e estejam funcionando com eficácia para mitigar os riscos. Conforme descrito nas seções anteriores, controles internos eficazes são um requisito crítico da IA. Muitas organizações testam e reportam os resultados dos controles de TI trimestral e/ou anualmente. A gestão deveria estar ciente de quaisquer questões de controle interno que também possam ter impacto sobre o uso da IA, especialmente em relação às áreas do ambiente de controle interno que já estão sendo avaliadas, tais como:

- Integridade e governança de dados.
- Acesso do usuário.
- Cibersegurança.
- Ciclo de vida de desenvolvimento de sistemas.
- Gestão de mudanças.
- Controles de backup/recuperação.

O COBIT e o COSO são exemplos de frameworks de controle interno que podem ser usados pelas organizações para auxiliar na abordagem e na avaliação do ambiente de controle interno.[19,20]

Gestão de Primeira Linha

Liderança

Definir papéis e responsabilidades relacionados às iniciativas baseadas em IA apoiará a organização a determinar quais recursos são necessários para operar com eficácia. A identificação da propriedade executiva, ao mesmo tempo em que incorpora a contribuição dos outros membros da gestão executiva, ajudará a garantir a prestação de contas.

Uma Equipe de Liderança de IA com membros multifuncionais é outra forma de as organizações monitorarem e comunicarem as iniciativas de IA e apoiarem a prestação de contas. Essa equipe deveria incluir:

- Gestores de IA e/ou ciência de dados.
- O CISO da organização.
- Pessoal essencial de TI.
- Jurídico (para fornecer orientação sobre considerações regulatórias).
- Finanças/contabilidade (para rastrear os custos e o ROI dos projetos de IA).
- Gerenciamento de riscos.
- Compliance.

A auditoria interna, com sua amplitude de conhecimento sobre a organização, está posicionada de forma única para servir como consultora para apoiar as iniciativas de IA e deveria ser considerada como membro da Equipe de Liderança de IA. A participação da auditoria interna deveria ser estruturada para garantir que sua independência como prestadora de avaliação não seja comprometida.

Um processo de planejamento bem pensado apoiará a organização na execução de projetos de IA. Os funcionários envolvidos na execução dos projetos precisam estar cientes dos riscos mais críticos, incluindo resultados indesejados. É importante destacar e assegurar que a execução diária dos projetos inclua a conscientização sobre os aspectos sociais, éticos, ambientais e econômicos. Além disso, encorajar um ambiente que incentive os funcionários a discutir abertamente ideias e preocupações relacionadas às iniciativas de IA pode ajudar a criar uma cultura de transparência, conscientização e responsabilidade mútua para apoiar projetos ambiciosos de IA.

Políticas e Procedimentos – Uso Interno e Aplicativos de Negócios

Definir, adotar e disseminar políticas e procedimentos organizacionais robustos sobre o uso da IA dentro da organização é outro aspecto importante da estratégia de IA de uma organização. Políticas e procedimentos claros fornecem orientação aos funcionários diretamente envolvidos nas iniciativas de IA e aos funcionários que podem usar a IA como parte de suas responsabilidades diárias de trabalho. Criar uma política de uso aceitável de IA deveria ser uma das principais prioridades da organização. Ela deveria incluir aspectos de práticas recomendadas de cibersegurança, propriedade intelectual/considerações legais e os riscos associados a diversas ferramentas de IA. A política deveria ser complementada por um processo documentado que os usuários devem seguir ao solicitar o uso de IA. O uso de um processo de aprovação formal para o uso de IA também apoiará os esforços da organização para manter um inventário de usuários ou departamentos que utilizam IA.

Políticas e procedimentos que esclarecem as diretrizes e expectativas usadas para desenvolver, implementar e monitorar iniciativas de IA formalizam o processo. Fornecem uma linha de base para validar se os projetos estão sendo realizados de forma consistente com as políticas aprovadas pela organização, a ética e a cultura geral de riscos organizacionais. Os auditores internos estão em uma posição única para fornecer feedback imediato sobre esse tópico, dado seu conhecimento e experiência prestando avaliação sobre as principais políticas e procedimentos. Em muitos casos, como ponto de partida, as políticas e procedimentos existentes podem fornecer medidas razoavelmente eficazes para mitigar os riscos apresentados pelo desenvolvimento da IA. Por exemplo, os sistemas de IA que estão sendo desenvolvidos podem estar sujeitos ao Ciclo de Vida de Desenvolvimento de Sistemas (SDLC) existente ou a processos de controle de gestão de mudanças. Com o tempo, conforme as organizações evoluem e elevam os casos de uso de IA, controles mais maduros ou novos certamente precisarão ser considerados.

Assim, políticas e procedimentos que esclareçam as expectativas e diretrizes para terceiros envolvidos em iniciativas de IA também são importantes. A coordenação entre as equipes que gerenciam a IA e o grupo da organização que gerencia os relacionamentos com terceiros (como o jurídico) promoverá relacionamentos consistentes com os fornecedores de IA. Como os terceiros são uma extensão dos processos da organização, é fundamental manter um bom entendimento dos ambientes de controle dos fornecedores. Quando disponível, a gestão deveria obter os relatórios SOC dos fornecedores de IA, para entender seus processos de controle e estar ciente de quaisquer preocupações, como constatações de auditoria. O uso de terceiros, no que se refere ao desenvolvimento de recursos de IA ou ao apoio contínuo de iniciativas de IA, deveria ser claramente definido e monitorado, incluindo SLAs que contenham o direito de auditoria.

Depois que essas políticas e procedimentos forem delineados, as organizações poderão promover adesão multifuncional às políticas e procedimentos de IA, compartilhando a documentação da política organizacional elaborada, como a política de uso

aceitável, com toda a equipe e solicitando feedback durante um período aberto para comentários. As organizações também deveriam planejar os recursos necessários para treinar a equipe sobre as novas políticas, para garantir que os funcionários estejam prontos para adotar e aderir aos papéis, controles e responsabilidades recém-definidos relativos ao uso da IA.[21]

Recursos de TI para Apoiar a IA

A otimização eficaz dos recursos de TI é necessária para apoiar as iniciativas de IA e deveria ser orçada pela gestão tendo em vista esse fato. O uso da IA exige um desempenho intensivo dos ativos de computador para sustentar um processamento confiável. Exemplos de capacidades de recursos de TI usados para apoiar as iniciativas de IA de uma organização incluem:

- ***Unidades centrais de processamento (CPUs)*** – os "cérebros" do computador; processadores que executam comandos ou instruções.[22]
- ***Unidades de processamento gráfico (GPUs)*** – cérebros mais capazes, que podem processar muitos dados simultaneamente com recursos matemáticos adicionais; capazes de produzir gráficos, imagens e são mais predominantes na IA de produção criativa.[23]
- ***Armazenamento*** – local dos dados necessários para processamento pela IA. O armazenamento é comumente mensurado em terabytes (1.000 gigabytes) ou petabytes (1.000 terabytes); como referência, um vídeo de alta definição de 5 a 10 minutos mede aproximadamente um gigabyte (1 bilhão de bytes); servidores hospedados no local ou soluções baseadas em nuvem são exemplos de onde os dados podem ser armazenados.
- ***Memória*** – também chamada de RAM (memória de acesso aleatório); local onde são armazenados os dados de curto prazo que ficam disponíveis mais rapidamente do que os dados de armazenamento; medida em gigabytes, sendo que as estações de trabalho de computadores individuais têm de 8 a 48 gigabytes de RAM; quanto mais complexa for a IA em execução, mais RAM será necessária.
- ***Supercomputadores*** – computadores de processamento mais rápido, que são usados para computação de alto desempenho e contêm várias CPUs.
- ***Estações de trabalho*** – incluem desktops e laptops com especificações técnicas que apoiam os requisitos da IA que está sendo utilizada.
- ***Software*** – plataformas, programas e aplicativos usados para desenvolver, implantar e gerenciar a IA; software de desenvolvimento. Exemplos incluem Microsoft Azure AI, IBM Watsonx.ai e Google Cloud AI Platform; software de implantação, usado para integrar a IA aos aplicativos existentes; exemplos incluem Docker e MLflow.
- ***Conectividade de rede*** – essa é uma categoria ampla que inclui hardware, software e os serviços que permitem que os usuários compartilhem recursos digitais e troquem informações; exemplos incluem servidores de arquivos e roteadores.

Embora não se espere que os auditores internos conheçam todas as especificações técnicas e detalhes dos requisitos de IA, eles deveriam ter um conhecimento básico dos recursos de TI.

Estruturação de Equipe e Treinamento

A estruturação devida da equipe é um elemento importante da estratégia de IA de uma organização. Os recursos humanos deveriam colaborar com a gestão, para garantir que funcionários com a experiência necessária em IA sejam recrutados em toda a organização. A experiência em IA deveria ser priorizada não apenas para funcionários encarregados de gerenciar os aspectos cotidianos da IA, mas também para a liderança que dirigirá as iniciativas de IA.

Como a IA está se desenvolvendo muito rapidamente, é importante que os funcionários da organização estejam cientes dos avanços e dos riscos correspondentes. As organizações deveriam assegurar que o treinamento geral de conscientização sobre IA seja fornecido a todos os funcionários e que mais oportunidades de treinamento técnico, como seminários, treinamento on-line ou cursos educacionais, estejam disponíveis para os funcionários que se concentram nas iniciativas de IA.

Conforme mencionado acima, na seção Políticas e Procedimentos, a implementação de treinamento sobre a política formal de uso aceitável da IA e a inclusão da IA no manual do funcionário e na orientação para novos contratados são boas formas de aumentar a conscientização organizacional sobre a IA e seus possíveis riscos. Ao integrar iniciativas de treinamento com foco em IA e alfabetização digital, políticas e procedimentos organizacionais e oportunidades de aprimoramento de habilidades, as organizações podem apoiar iniciativas de IA por meio de investimento direto nos funcionários atuais e futuros. A implementação e os resultados dessas iniciativas deveriam ser monitorados pela auditoria interna como parte dos controles de IA da organização.

Execução

Gerenciamento de Riscos pela Primeira e Segunda Linhas

A Parte 2 discutiu a importância de identificar os riscos de IA relacionados à segurança, integridade, privacidade e confidencialidade dos dados, e a abordagem dessas preocupações deveria ser um foco, conforme a organização executa projetos de IA. Os algoritmos de IA dependem de dados precisos e confiáveis, e as equipes de projeto deveriam monitorar de perto os dados inseridos. As organizações têm várias formas de validar a completude dos dados usados em projetos de IA, inclusive assegurando que os totais registrados correspondam e analisando o reporte de erros quando os dados são transferidos entre sistemas. A gestão deveria criar e monitorar controles internos que detectem anomalias na qualidade ou completude dos dados.

Outras considerações importantes sobre os dados incluem restringir o acesso do usuário apenas aos funcionários que estão trabalhando em um projeto de IA, o que inclui o acesso do administrador. Determinar as funções dos usuários e garantir a devida segregação de funções também é fundamental. Por exemplo, os administradores de banco de dados supervisionam os dados inseridos subjacentes e não deveriam ter acesso para modificar os algoritmos que processam esses dados; uma tarefa que tradicionalmente é responsabilidade do desenvolvedor.

Quando um projeto de IA está sendo implementado, é importante que a organização garanta que o projeto seja transparente, explicável, responsável e auditável:

- ***Transparência*** – capacidade de entender facilmente, em termos simples, o propósito da IA ou do algoritmo.
- ***Explicabilidade*** – capacidade de explicar a mecânica, cálculos ou resultados processados pela IA ou pelo algoritmo.
- ***Responsabilidade*** – uso da IA ou dos algoritmos de forma ética, segura, justa e confiável.
- ***Auditabilidade*** – como os aplicativos de IA podem começar a substituir ou aumentar certos processos importantes de conformidade ou outros processos de negócios importantes, manter a rastreabilidade, por meio de registros de auditoria eficazes ou informações relacionadas, será um componente importante para o desenvolvimento da IA, porque a avaliação desses processos provavelmente será necessária para muitos dos possíveis casos de uso.

A gestão de projetos de IA deveria definir os seguintes aspectos para cada iniciativa:

- ***Objetivos, funções e prazos*** – o que a iniciativa pretende concretizar, quem participa e quando ela ocorre.
- ***Requisitos de recursos*** – quais recursos tecnológicos e/ou de pessoal são necessários para obter sucesso.

- ***Requisitos de dados*** – quais entradas de dados são necessárias para a IA ou algoritmo(s).
- ***Requisitos de privacidade, legais e regulatórios*** – quais são os requisitos de conformidade relacionados.
- ***Avaliação de riscos*** – quais são os riscos relevantes que ameaçam o atingimento dos objetivos do projeto ou resultados indesejados, como viés, tratamento antiético ou uso indevido.
- ***Métricas de sucesso ou principais indicadores de desempenho (KPIs)*** – como o sucesso do projeto é monitorado e quantificado.
- ***Requisitos de teste*** – em um determinado momento, como validar se a IA ou algoritmo está funcionando conforme projetado e quais alterações são necessárias; isso incluirá tanto os usuários finais (que acabarão usando a IA) quanto os profissionais de IA/ciência de dados; identificar e comunicar questões será fundamental nessa etapa. Analisar como os desenvolvedores terceirizados testam e confirmam a eficácia de seus algoritmos é uma consideração importante.
- ***Requisitos de teste*** – de um ponto de vista contínuo, como os resultados da IA, por sua própria natureza, mudam devido aos dados inseridos, deveria ser considerado o teste contínuo ou a avaliação da qualidade do modelo – dependendo do caso de uso, esse conceito pode precisar ser incorporado aos requisitos de negócios ou ao projeto da IA.

O monitoramento contínuo dos projetos de IA deveria ser realizado pela gestão, para garantir que a iniciativa esteja ocorrendo conforme o planejado e para identificar quaisquer questões ou preocupações que tenham ocorrido. Conforme indicado no Modelo das Três Linhas, a gestão desempenha um papel vital no ambiente de controle interno, fornecendo o primeiro nível de ações para mitigar os riscos. O monitoramento no nível do projeto é importante, porque é onde as questões são inicialmente detectadas. O reporte da gestão à liderança executiva (gestão executiva) e ao conselho deveria fazer parte desse processo. É importante não apenas monitorar o progresso geral do projeto, mas também identificar e reportar quaisquer resultados negativos, como preocupações éticas ou violação de informações confidenciais. Também é importante incluir uma avaliação de terceiros, para garantir que estejam cumprindo suas responsabilidades no projeto de IA.

O monitoramento e reporte também deveriam incluir a divulgação de quaisquer questões de controle interno específicas do projeto ou a análise de questões de controle interno de outras áreas da organização que possam afetar o projeto de IA. O gerenciamento de riscos corporativos e/ou compliance também deveriam fazer parte do processo de monitoramento de questões de controle, a partir de uma perspectiva de "Segunda Linha".

Apoio da Segunda Linha no Gerenciamento de Riscos

Os principais objetivos do processo de gerenciamento de riscos corporativos de uma organização são entender como os riscos podem ameaçar o atingimento dos objetivos e, em seguida, tomar medidas para mitigar esses riscos. As categorias de risco incluem estratégico, financeiro, ambiental, de mercado, social, ético, tecnológico, econômico, político, legal e regulatório. A IA é um tópico geralmente considerado um risco tecnológico; no entanto, é importante reconhecer que o risco de IA pode se enquadrar em qualquer uma das categorias mencionadas anteriormente, exigindo um processo robusto de gerenciamento de riscos para projetos de IA, que considere preocupações tecnológicas e não tecnológicas.

O Framework de Auditoria de IA do The IIA fornece considerações de gerenciamento de riscos para apoiar os projetos organizacionais de IA, seguindo as melhores práticas de gerenciamento de riscos de IA. Quando apropriado, outros frameworks existentes deveriam ser considerados, em especial, o Artificial Intelligence Risk Management Framework do NIST. Os auditores internos frequentemente colaboram com os profissionais de gerenciamento de riscos em atividades como o processo anual de avaliação de riscos da organização; portanto, é vital que os auditores internos entendam os riscos relacionados à IA e continuem aumentando sua base de conhecimento. Além disso, os auditores internos deveriam considerar os riscos relacionados à IA no nível do trabalho de auditoria, ou seja, ao auditar processos que incluam algum aspecto da IA.

Identificação

Identificar os riscos relativos à IA pode ser uma tarefa nova para muitas organizações. Idealmente, o gerenciamento de riscos corporativos (com a auditoria interna, compliance e jurídico) participará das discussões iniciais de todas as iniciativas de IA, para ajudar a enquadrar os riscos relativos ao projeto de IA. Como mencionado na seção Estratégia, uma Equipe de Liderança de IA multifuncional é uma forma eficaz de identificar proativamente potenciais riscos ou ameaças antes que se concretizem e, ao mesmo tempo, assegurar que controles e técnicas de mitigação de riscos estejam em vigor em toda a organização.

As organizações que já estabeleceram um processo eficaz de avaliação de riscos em toda a empresa deveriam considerar a realização de uma avaliação inicial de riscos com foco em IA. Se uma avaliação de risco de IA separada não for viável, as organizações deveriam assegurar, no mínimo, que a IA seja incluída durante o processo geral de avaliação de riscos.

Por exemplo, as organizações trabalham periodicamente com a equipe de liderança executiva para identificar riscos em diversas áreas. Em muitos casos, as discussões ou pesquisas incluem perguntas específicas, como: "Quais são os maiores riscos estratégicos da organização?" Para trazer a IA para a vanguarda do processo de avaliação de riscos, as organizações deveriam destacar a IA como uma área de risco emergente, compartilhar o feedback contínuo da Equipe de Liderança de IA e/ou da equipe interna e, de acordo com isso, coletar as contribuições dos executivos. A etapa de identificação de riscos no processo de gerenciamento de riscos é importante, porque pode destacar riscos que não foram identificados anteriormente.

As organizações que estabeleceram uma estratégia clara de IA, com objetivos e metas definidos, estão fornecendo o contexto de que o gerenciamento de riscos corporativos precisa para auxiliar na identificação dos riscos de IA. Esse contexto permite que o gerenciamento de riscos corporativos desenvolva um inventário dos riscos que ameaçam o atingimento dos objetivos e metas, permitindo que as organizações incorporem em seus planos estratégicos salvaguardas contra os potenciais danos do uso da IA. É importante que as organizações estejam cientes de que o cenário de risco em torno da IA continua evoluindo rapidamente, causando consequências indesejadas e negativas de riscos não contabilizados que podem incluir:[24]

- Resultados enviesados ou discriminatórios, que podem afetar injustamente segmentos específicos da população.
- Prejuízos à privacidade ou confidencialidade.
- Falta de prestação de contas.
- Falta de transparência.
- Falta de explicação.
- Danos financeiros ou desigualdade econômica.
- Danos ambientais.
- Desinformação ou manipulação.
- Violação de direitos autorais.

A "Caixa Preta"

Embora grande parte dos processos de identificação, avaliação e mitigação de riscos para projetos de IA siga as práticas recomendadas existentes, é importante observar que a "caixa preta" da IA representa um risco distinto. O termo refere-se à falta de transparência dos sistemas de IA e às formas de tomar decisões. Os modelos de deep learning, em especial, podem ser difíceis de entender, dado o processamento complexo realizado pelos algoritmos, em conjunto com uma visibilidade ou compreensão potencialmente limitada ou inexistente de como um resultado foi produzido. Isso pode representar desafios específicos à medida que o gerenciamento de riscos corporativos (e os auditores internos) tentam capturar a documentação necessária para apoiar o ciclo de gerenciamento de riscos definido anteriormente. Os profissionais de risco e os auditores internos podem lidar diretamente com a "caixa preta":

Identificando e comunicando claramente onde informações possam estar ausentes ou incompletas em um projeto de IA.

- ***Exemplo:*** Se uma organização estiver usando um fornecedor terceirizado de IA que não forneça informações detalhadas sobre os dados de treinamento de um algoritmo, isso deveria ser documentado e divulgado como um risco potencial.

Avaliando e atualizando continuamente o conselho sobre os impactos potenciais relacionados às lacunas de informações identificadas.

- ***Exemplo:*** Uma vez que a falta de documentação do fornecedor sobre o conjunto de treinamento tenha sido documentada, os auditores internos deveriam atualizar o conselho diante de uma consequência relacionada ao risco (tais como diversos exemplos de resultados de IA enviesados que sugerem problemas com os dados de treinamento).

Apresentando orientações sobre como mitigar os riscos associados às lacunas de conhecimento da "caixa preta" previamente documentadas e avaliadas.

- ***Exemplo:*** Com base na avaliação e consequências apresentadas, a organização toma a decisão de migrar para um novo fornecedor de IA com uma documentação de dados mais transparente.

Avaliação

A avaliação e a análise dos riscos identificados relacionados à IA deveriam seguir um processo semelhante ao que a organização usa para revisar outros riscos – o impacto e a probabilidade deveriam ser considerados primeiro. O impacto dos riscos relacionados à IA pode ser difícil de quantificar devido às inúmeras considerações, como ramificações legais, regulatórias, sociais, financeiras, ambientais e éticas. Os danos à reputação da marca são outra consideração sobre o impacto.

A combinação de impacto e probabilidade resulta em risco inerente, que é uma métrica do risco que existe sem a consideração de controles internos. Após a avaliação do risco inerente, o risco residual deveria ser a próxima determinação, que inclui a consideração de quão bem os riscos são mitigados.

Como exemplo de avaliação da segurança da IA como um objetivo, as ameaças cibernéticas podem ser identificadas como um risco significativo. Para lidar com os riscos cibernéticos, as organizações implementam controles de cibersegurança, com o objetivo de reduzir o risco inerente a um nível aceitável de risco residual. Se o gerenciamento de riscos corporativos avaliar que o risco residual não foi reduzido a um nível aceitável, a organização deverá determinar como proceder.

A priorização de riscos é o processo que uma organização usa para classificar os riscos em ordem de importância, ou seja, os riscos de maior impacto são tratados primeiro. As organizações têm recursos limitados, mas enfrentam riscos ilimitados, portanto, é importante assegurar que os riscos relacionados à IA sejam priorizados na análise de risco mais ampla de toda a entidade. A classificação dos riscos relacionados à IA nas organizações individuais varia de acordo com o processo de avaliação de riscos, o quanto elas utilizam a IA e o nível de maturidade de seu ambiente de controle interno. Simplificando, não há uma abordagem “de tamanho único” para a avaliação de riscos relacionados à IA.

Mitigação

A mitigação de riscos é uma ação (ou ações) que a gestão toma para reduzir os riscos a um nível mais aceitável. Em muitos casos, as organizações optam por tratar os riscos relacionados à IA por meio de ações de mitigação, como a adição de controles internos; no entanto, há outras possíveis respostas aos riscos, conforme descrito na tabela abaixo.[25]

Diversos fatores influenciam a forma como uma organização determina como responder aos riscos relacionados à IA. Portanto, é fundamental ter um processo definido e repetível de resposta a riscos. Os riscos relacionados à IA podem mudar durante um projeto, portanto, a organização deveria revisar continuamente como responde aos riscos e os mitiga.

Respostas Básicas de Riscos

RESPOSTA	CARACTERÍSTICAS	DEFINIÇÃO
Ameaça	Reduzir, Mitigar, Melhorar, Explorar, Alavancar, Otimizar	Aplicar controles para reduzir o risco inerente a um nível residual aceitável ou aplicar outras medidas para maximizar e tirar proveito das possíveis variações potenciais nos resultados.
Tolerar	Aceitar, Buscar	Determinar se os benefícios potenciais justificam a tomada de riscos, tendo estabelecido as medidas consideradas necessárias para mitigar ou alavancar a probabilidade e/ou o impacto.
Transferir	Compartilhar, Espalhar	Espalhar o risco, transferindo parte ou sua totalidade a terceiros (p. ex., por meio de seguro ou terceirização) ou aplicando recursos de múltiplas equipes para se proteger contra possíveis perdas.
Terminar	Evitar	Terminar ou evitar o risco, abandonando a ação planejada ou eliminando a meta por completo, priorizando outras metas de preferência.

Auditoria interna – Atividades de Consultoria e Avaliação

Depois de descrever como uma organização deveria abordar a IA nos dois domínios anteriores da estrutura, há um domínio restante – a Auditoria Interna.

Os dois primeiros domínios fornecem uma linha de base de como a auditoria interna pode prestar serviços de consultoria e auditoria à organização. Os domínios de Governança e Gestão contêm os detalhes que um auditor interno deveria usar para orientar a organização a se mover em direção a essas práticas, ou para formar uma base para avaliar como a organização está abordando, utilizando, gerenciando e monitorando a IA.

"Garantia razoável" é um termo frequentemente mencionado na profissão de auditoria interna. Do ponto de vista do controle interno, garantia razoável significa que há uma alta probabilidade de os controles mitigarem os riscos, mas não é absoluta. A mesma lógica deveria ser considerada para os auditores internos que têm a tarefa de prestar avaliação em relação à IA.

Desafios

Diversos aspectos da IA tornam as atividades de avaliação difíceis para os auditores internos, incluindo:

- A IA (ou, mais especificamente, os algoritmos) é inerentemente altamente complexa – um problema de "caixa preta" mais difícil.
- As capacidades e os riscos da IA estão se multiplicando em um ritmo acelerado.
- A IA como tópico de auditoria está evoluindo com ferramentas limitadas ou abordagens amplamente adotadas.
- Há oportunidades limitadas de treinamento disponíveis para aprimorar os conjuntos de habilidades de auditoria de IA.

A IA como tópico de auditoria pode parecer avassaladora, mas o foco nas considerações a seguir ajudará os auditores internos a desenvolver uma mentalidade positiva e confiante:

- Não se espera que os auditores internos sejam especialistas em todos os tópicos de auditoria; em vez disso, ter uma abordagem disciplinada e metódica com foco no raciocínio crítico e na identificação de riscos deveria ser o objetivo

de todas as auditorias, não apenas da IA. A familiaridade e o conhecimento prático da IA são vitais; entretanto, não é provável que se conheçam todos os aspectos técnicos da IA. Pode ser necessário contratar recursos técnicos externos para auxiliar nos aspectos mais técnicos, como decifrar algoritmos.

- Como a IA é altamente complexa e mutável, é improvável que os auditores internos algum dia dominem o conhecimento sobre o tópico; pense na auditoria de IA como uma progressão, não como um destino – aumente a compreensão da IA ao longo do tempo.
- Esteja disposto a fazer perguntas relevantes sobre IA dentro da organização:
    - Como a IA nos ajuda a atingir nossos objetivos estratégicos?
    - Quais são os riscos envolvidos e como os estamos mitigando?
    - Há controles internos adequados envolvidos nos processos relacionados à IA?
    - Os dados que serão usados para IA são completos, precisos e confiáveis?
    - Como a IA é testada antes da implementação, para garantir que não haja viés?
    - Como a IA é testada após a implementação, para garantir que não haja viés?
    - Como a IA é governada?
    - Como a organização garante a existência de treinamento e conscientização adequados sobre IA?

Como descrito na Parte 2, a compreensão do uso organizacional da IA começa com pesquisa e discussão. É vital que os auditores internos alavanquem os relacionamentos profissionais que desenvolveram. É importante ser transparente tanto com a gestão quanto com o órgão de governança. Explique em termos simples como está pensando sobre a IA como tópico e como planeja envolver a organização para aprender mais sobre ela.

As auditorias internas de IA são uma responsabilidade relativamente nova para muitas organizações. Embora, como prestadores de avaliação, não se espere que os auditores internos sejam especialistas no tópico de IA, eles devem identificar oportunidades para aumentar seu conhecimento e conscientização sobre o assunto. Obter uma melhor compreensão dos aspectos mais técnicos da IA, como algoritmos, será importante para a educação profissional futura.

Embora a IA certamente tenha elementos complexos, é importante lembrar que ela produz alguma forma de resultado a partir da entrada que recebe. Do ponto de vista da avaliação, os auditores internos talvez nunca tenham conhecimento absoluto de todo o funcionamento interno da IA; entretanto, ajudar uma organização a 1) avaliar o que está fazendo para garantir que os dados inseridos sejam os mais precisos possíveis e, em seguida, 2) entender como esse resultado é examinado deveriam ser os principais objetivos dos profissionais. Os auditores internos aplicam esses conceitos, atualmente, ao realizar auditorias de TI de aplicativos de negócios. O ponto em comum é a noção de rastreabilidade – garantindo que os dados e resultados estejam alinhados com os objetivos de negócios e requisitos do caso de uso de IA.

Referências

1. Modelo das Três Linhas do The IIA.
2. A.M. Turing, "Computing Machinery."
3. A.L. Samuel, "Checkers."
4. McCarthy, Dartmouth.
5. Weizenbaum, "ELIZA."
6. Humanoid Robot, "Waseda University."
7. Hsu, "Raj Reddy."
8. "Prometheus Project."
9. Britannica, "Deep Blue."
10. Índice *AI Adoption Index* da IBM.
11. IBM, "Different types of AI."
12. Certes, "Types of AI."
13. The IIA, Global Perspectives & Insights.
14. National Cybersecurity Centre.
15. Folheto Informativo da Casa Branca.
16. IBM, e-Book *AI Governance*.
17. Doran, "Smart Way."
18. Moyer, ISACA, "Quantitative Approach."
19. COBIT, "Framework."
20. COSO, "Guidance."
21. Deloitte, "3 Lines."
22. "Gartner Glossary."
23. Intel, "GPU."
24. Forbes, 15 Biggest Risks.
25. The IIA, CRMA.

PARTE 4

Guia do Profissional e Glossário

O guia do profissional é um simples checklist que os auditores internos podem usar para iniciar sua avaliação de como a organização aborda, usa, gerencia e reporta sobre a IA. Os auditores internos podem usar os pontos principais descritos nas seções Governança, Gestão e Auditoria Interna da Parte 3 para desenvolver seu plano de auditoria ou como considerações em uma atuação consultiva. Muitos dos aspectos ou considerações na seção de avaliação abaixo estão intimamente ligados a itens previamente listados nos outros domínios.

Este checklist destina-se a fornecer um guia de início rápido, mas deveria ser personalizado com base em considerações organizacionais, como a extensão em que a IA já está sendo usada e se foram estabelecidos planejamento estratégico, políticas, procedimentos, processos e reporte formais para a IA.

Aspectos ou Considerações	Status/ Resultados
Criar uma visão, estratégia e priorização para a IA e atualizá-la com frequência.	
Vincular a iniciativa de IA aos objetivos estratégicos da organização. (Pode incluir casos de uso que aumentem a receita ou aplicativos internos para reduzir custos ou melhorar a eficiência).	
Garantir que os aspectos éticos, de viés, sociais e legais sejam incluídos na estratégia.	
Determinar como mensurar o sucesso das iniciativas de IA, incluindo metas e ROI.	
Garantir que o plano estratégico de IA seja consistente com a cultura de risco da organização.	
Garantir que o plano estratégico de IA seja consistente com os valores da organização.	
Garantir que o plano estratégico de IA seja formalmente comunicado ao conselho.	
Garantir que o plano estratégico inclua a otimização dos recursos de IA.	

Aspectos ou Considerações	Status/ Resultados
Garantir que o ambiente de controle interno seja propício para apoiar a IA. Considerar quais mudanças imediatas nas políticas são necessárias para apoiar o crescimento da IA – adicionar uma pergunta sobre o uso da IA na política de gestão de fornecedores terceiros, por exemplo.	
Definir a gestão executiva responsável pela supervisão das iniciativas de IA.	
Estabelecer Equipe de Liderança de IA multifuncional para monitorar todas as iniciativas de IA.	
Garantir que as equipes jurídica e de conformidade monitorem todos os requisitos regulatórios atuais e emergentes.	
Definir o papel da auditoria interna como consultora e/ou prestadora de avaliação.	
Garantir que o Modelo das Três Linhas esteja em vigor e inclua a IA.	
Garantir que o CISO (ou equivalente) esteja envolvido em todas as iniciativas de IA.	
Garantir que funções de terceiros nas iniciativas de IA sejam claramente definidas e monitoradas.	
Garantir que o financeiro/contábil monitore o ROI das iniciativas de IA.	
Desenvolver uma política de uso aceitável de IA que seja obrigatória para todos os funcionários.	
Desenvolver políticas e procedimentos para executar e manter iniciativas de IA.	
Desenvolver políticas e procedimentos para iniciativas de IA que utilizem terceiros.	
Garantir que os recursos de TI sejam suficientes para apoiar as iniciativas e controles de IA.	
Garantir que os níveis de estruturação de equipe sejam suficientes para apoiar as iniciativas e controles de IA.	
Garantir que o recrutamento de RH tenha foco em práticas de contratação de profissionais com experiência em IA.	
A liderança de IA mantém o conhecimento necessário de gestão de IA.	
Os funcionários operacionais de IA mantêm o conhecimento técnico necessário de IA.	
Todos os funcionários concluem o treinamento sobre o uso aceitável e os riscos da IA.	
Incluir o tema da IA no manual do funcionário e na orientação para novos contratados.	
Garantir que aspectos sociais, ambientais e econômicos justos sejam considerados em todos os projetos relacionados à IA.	
Garantir que os dados relacionados à IA sejam seguros, privados e confidenciais.	
Garantir que os dados relacionados à IA sejam transparentes, explicáveis e responsáveis.	
Definir objetivos, metas, prazos e requisitos de recursos para projetos de IA.	
Definir responsabilidades operacionais para todos os funcionários relevantes em projetos de IA.	
Garantir que o acesso do usuário à IA seja proporcional aos deveres do cargo.	
Definir requisitos de dados e considerações de privacidade para projetos de IA.	

Aspectos ou Considerações	Status/ Resultados
Definir os requisitos legais e regulatórios aplicáveis aos projetos de IA.	
Realizar a avaliação de riscos do projeto de IA, para identificar possíveis ameaças ao sucesso.	
Definir possíveis vieses, incluindo considerações éticas e sociais para projetos de IA.	
Definir métricas de sucesso ou principais indicadores de desempenho para projetos de IA.	
Estabelecer parâmetros de reporte, como frequência, conteúdo e marcos para projetos de IA.	
Estabelecer uma abordagem de teste para validar se a IA está funcionando como pretendido antes e depois de entrar em operação.	
Reportar sobre a concretização das métricas/KPIs à liderança executiva e ao conselho.	
Garantir que o reporte inclua a divulgação de vieses, preocupações éticas ou sociais.	
Garantir que o reporte inclua a conformidade com os requisitos legais e regulatórios.	
Garantir que o reporte inclua a divulgação de quaisquer resultados não intencionais ou negativos.	
Garantir que o reporte inclua a divulgação de possíveis perdas de dados ou violações de privacidade.	
Garantir que os controles internos relacionados sejam avaliados e reportados periodicamente.	
Incluir a IA como parte do processo de gerenciamento de riscos corporativos (ERM).	
Identificar riscos que ameacem as metas e os objetivos estratégicos de IA.	
Identificar riscos que possam ter consequências éticas, sociais, ambientais ou financeiras.	
Identificar riscos relacionados ao uso de terceiros para a IA.	
Garantir a existência de um processo para registrar riscos novos ou emergentes.	
Garantir que os funcionários com responsabilidades de gerenciamento de riscos de IA sejam devidamente treinados.	
Realizar uma avaliação de riscos baseada em IA e atualizá-la periodicamente.	
Priorizar riscos relativos à IA com base na pontuação de gravidade (impacto e probabilidade).	
Garantir que haja um processo para selecionar respostas apropriadas aos riscos, incluindo o monitoramento do progresso das respostas.	
Garantir que a organização envolva o conselho quanto à estratégia, metas e objetivos de IA.	
Garantir que a organização forneça atualizações periódicas ao conselho com relação à IA de uma forma clara e facilmente compreensível.	
Garantir que a organização envolva o conselho quanto à abordagem de gerenciamento de riscos de IA.	
Realizar pesquisas iniciais internas e externas sobre IA.	
Determinar se foi desenvolvida uma estratégia formal de IA.	
Conduzir discussões iniciais com relacionamentos organizacionais estabelecidos (como TI e CFO), para entender como a IA está sendo usada e gerenciada atualmente.	

Aspectos ou Considerações	Status/ Resultados
Conduzir discussões iniciais com a equipe de IA/ciência de dados (se aplicável) e/ou gestão de TI.	
Criar um inventário dos usos atuais e planejados da IA.	
Para os usos atuais da IA, desenvolver a compreensão de como ela está sendo usada, as metas e os objetivos.	
Para os usos planejados da IA, desenvolver a compreensão da abordagem, como os riscos são avaliados e planejar os testes antes da implementação.	
Desenvolver a compreensão dos seguintes aspectos de dados inseridos relativos à IA: • Governança. • Arquitetura. • Acesso do usuário. • Controles de cibersegurança. • Controles de processamento (integridade, precisão, completude). • Considerações de terceiros (relatórios SOC).	
Verificar como a IA é testada e revisada, para garantir que ela concretize seus objetivos e esteja livre de vieses, tanto antes quanto depois da implantação.	
Verificar se as iniciativas de IA têm objetivos e metas claros e se os projetos são gerenciados por um nível apropriado da liderança.	
Verificar se a gestão realiza o reporte periódico ao órgão de governança.	
Verificar se a IA é considerada como parte do processo de gerenciamento de riscos corporativos e se inclui riscos relacionados a: • Ética. • Considerações sociais e econômicas. • Aspectos ambientais. • Consequências financeiras. • Violações legais e regulatórias.	
Verificar se foram desenvolvidas políticas e procedimentos que delineiem como a IA deveria ser usada e gerenciada pela organização, incluindo uma política de uso aceitável da IA.	
Desenvolver uma compreensão de como a organização apoia o aprendizado e o treinamento sobre IA, para aumentar o conhecimento e a conscientização de todos os funcionários.	

Normas do IIA Relacionadas

6.1 Mandato da Auditoria Interna

6.2 Estatuto da Auditoria Interna

6.3 Apoio do Conselho e da Alta Administração

7.1 Independência Organizacional

7.2 Qualificações do Chefe Executivo de Auditoria

8.1 Interação com o Conselho

8.2 Recursos

8.3 Qualidade

8.4 Avaliação Externa de Qualidade

Glossário

As definições dos termos marcados com um asterisco foram extraídas do Glossário do International Professional Practices Framework® do IIA, edição de 2017. Outras definições são definidas para os propósitos deste documento ou derivadas das seguintes fontes:

IBM. "Explainers." *IBM*. https://www.ibm.com/topics.

Institute of Risk Management. "Risk Culture." *Institute of Risk Management*, 2023. https://www.theirm.org/what-we-say/thought-leadership/risk-culture/.

Anderson, Urton; Michael J. Head; Steve Mar; Sridhar Ramamoorti; Chris Riddle; Mark Salamasick; Paul J. Sobel. *Internal Auditing: Assurance & Advisory Services, 5ª Edição*. (Lake Mary, FL: The Internal Audit Foundation, 2022.) https://www.theiia.org/en/products/bookstore/internal-auditing-assurance-and-advisory-services-5th-edition/.

ISACA. "Glossary," *ISACA*. 2022. https://www.isaca.org/resources/glossary.

NIST Computer Security Resource Center. "Glossary." Gaithersburg, MD.: *NIST*. https://csrc.nist.gov/glossary.

Joint Task Force. *NIST SP 800-53: Security and Privacy Controls for Information Systems and Organizations, Revisão 5, Anexo A: Glossary*. Gaithersburg, MD: NIST, September 2020. https://doi.org/10.6028/NIST.SP.800-53r5.

Grassi, Paul; Michael E. Garcia; James L. Fenton. *NIST SP 800-63-3: Digital Identity Guidelines, Anexo A: Definitions and Abbreviations*. Gaithersburg, MD: *NIST*, junho de 2017. https://doi.org/10.6028/NIST.SP.800-63-3.

Sawyer's Internal Auditing: Enhancing and Protecting Organizational Value, 7ª Edição. (Lake Mary, FL: The Internal Audit Foundation, 2019.) https://www.theiia.org/en/products/bookstore/sawyers-internal-auditing-enhancing-and-protecting-organizational-value-7th-edition/.

Techopedia.com. "TechDictionary." https://www.techopedia.com/dictionary.

Algoritmo – Um processo matemático claramente específico à computação; um conjunto de regras que, se seguidas, darão um resultado prescrito. (Glossário do NIST).

Apetite a risco* – O nível de risco que uma organização está disposta a aceitar.

Avaliação – O processo de identificação de riscos para as operações organizacionais (incluindo missão, funções, imagem, reputação), ativos organizacionais, indivíduos, outras organizações e a Nação, resultante da operação de um sistema de informações. Parte do gerenciamento de riscos incorpora análises de ameaças e vulnerabilidades e considera as mitigações fornecidas pelos controles de segurança planejados ou em vigor. Sinônimo de análise de risco. (Glossário do NIST).

Backup – Arquivos, equipamentos, dados e procedimentos disponíveis para uso em caso de falha ou perda, se os originais forem destruídos ou estiverem fora de serviço. (ISACA).

Backup e recuperação – Refere-se ao processo de fazer backup dos dados em caso de perda e de configurar sistemas que permitam a recuperação dos dados devido à perda. O backup de dados requer a cópia e o arquivamento dos dados do computador, de modo que

fiquem acessíveis em caso de exclusão ou corrupção de dados. Os dados de um momento anterior só podem ser recuperados se tiverem sido salvos em um backup. (Techopedia)

Big data – Termo usado para se referir à grande quantidade de informações digitais em fluxo constante, ao aumento maciço da capacidade de armazenar grandes quantidades de dados e à quantidade de poder de processamento de dados necessária para gerenciar, interpretar e analisar os grandes volumes de informações digitais. (Internal Auditing, 5ª edição)

Black box testing (teste de caixa preta) – Uma abordagem de teste que se concentra na funcionalidade do aplicativo ou produto e não requer conhecimento dos intervalos de código. (ISACA)

Chatbot – Um chatbot é um programa de inteligência artificial que simula uma conversa humana interativa, usando frases-chave pré-calculadas do usuário e sinais auditivos ou baseados em texto. Os chatbots são frequentemente usados por organizações para fornecer serviços de gestão de relacionamento com o cliente (CRM) 24 horas por dia. Esse tipo de bot de software também pode ser usado como um assistente virtual inteligente. (Technopedia)

Cibersegurança – Refere-se a qualquer tecnologia, medida ou prática para prevenir ciberataques ou mitigar seu impacto. A cibersegurança tem como objetivo proteger os sistemas, aplicativos, dispositivos de computação, dados confidenciais e ativos financeiros de indivíduos e organizações contra vírus de computador simples e irritantes, ataques de ransomware sofisticados e caros e todos os demais. (IBM)

Ciência da computação – É o estudo do design de hardware e software de computadores. Ela abrange o estudo de algoritmos teóricos e os problemas práticos envolvidos em sua implementação por meio de hardware e software de computador. O estudo da ciência da computação tem muitos campos, incluindo inteligência artificial, engenharia de software, programação e computação gráfica. (Technopedia)

Comitê de Auditoria – Um comitê do conselho encarregado de recomendar ao conselho a aprovação dos auditores e dos relatórios financeiros. (Sawyer's).

Conselho* – O corpo administrativo de mais alto nível (p. ex.: um conselho de administração, conselho supervisor ou um conselho de gestores ou curadores) que detém a responsabilidade de dirigir e/ou supervisionar as atividades da organização e de cobrar prestação de contas por parte da alta administração. Embora os sistemas de governança variem entre jurisdições e setores, o conselho normalmente inclui membros que não fazem parte da gestão. Se não houver um conselho, a palavra "conselho" nas Normas se refere a um grupo ou pessoa responsável pela governança da organização. Além disso, "conselho" nas Normas pode se referir a um comitê ou outro órgão ao qual o corpo administrativo tenha delegado certas funções (p. ex.: um comitê de auditoria)

Controle interno – Um mecanismo abrangente que uma empresa usa para concretizar e monitorar os objetivos corporativos. (Glossário do NIST).

Cultura de risco – É um termo que descreve os valores, crenças, conhecimentos, atitudes e compreensão sobre o risco compartilhados por um grupo de pessoas com um propósito comum. Isso se aplica a todas as organizações, incluindo empresas privadas, órgãos públicos, governos e organizações sem fins lucrativos. (Institute of Risk Management)

Deep learning (aprendizado profundo) – É uma abordagem iterativa da inteligência artificial (IA) que empilha algoritmos de aprendizado de máquina em uma hierarquia de complexidade e abstração crescentes. Cada nível de aprendizado profundo é criado a partir do conhecimento obtido na camada anterior da hierarquia. (Technopedia)

Governança* – Combinação de processos e estruturas implantadas pelo conselho para informar, dirigir, gerenciar e monitorar as atividades da organização, com o intuito de alcançar os seus objetivos.

Governança de dados – Refere-se ao processo de gestão da qualidade dos dados em uma organização, para garantir que, em todos os momentos de seu ciclo de vida, os dados sejam precisos, disponíveis, consistentes, seguros e utilizáveis. Os analistas de negócios e cientistas de dados buscam informações em toda a empresa, para obter insights e compreensão dessas informações, apoiando as necessidades dos negócios. (IBM)

Integridade dos dados – A propriedade de que os dados não tenham sido alterados de forma não autorizada. A integridade dos dados abrange dados armazenados, em processamento e em trânsito. (Glossário do NIST)

Inteligência artificial – Um sistema avançado de computador que pode simular capacidades humanas, como análise, com base em um conjunto predeterminado de regras. (ISACA).

Large language Model (LLM) – Um modelo grande de linguagem é um tipo de modelo de aprendizado de máquina que pode realizar uma variedade de tarefas de processamento de linguagem natural *(natural language processing* – NLP), como gerar e classificar texto, responder a perguntas de forma conversacional e traduzir texto de um idioma para outro. O rótulo "grande" refere-se ao número de valores (parâmetros) que o modelo de linguagem pode alterar de forma autônoma conforme aprende. Alguns dos LLMs mais bem-sucedidos têm centenas de bilhões de parâmetros. (Technopedia)

Machine learning (ML) – *Machine learning* (aprendizado de máquina) é a subcategoria da inteligência artificial (IA) que cria modelos algorítmicos para identificar padrões e relacionamentos nos dados. Nesse contexto, a palavra máquina é sinônimo de programa de computador e a palavra aprendizado descreve como os algoritmos de ML se tornam mais precisos conforme recebem dados adicionais. (Technopedia)

NIST – National Institute of Standards and Technology.

NIST Artificial Intelligence Risk Management Framework – Conforme orientado pela lei *National Artificial Intelligence Initiative Act* de 2020 (P.L. 116-283), o objetivo do AI RMF é oferecer um recurso às organizações que criam, desenvolvem, implantam ou usam sistemas de IA para ajudar a gerenciar os muitos riscos da IA e promover o desenvolvimento e o uso confiáveis e responsáveis de sistemas de IA. O Framework pretende ser voluntário, preservador de direitos, não específico a um setor e agnóstico em relação a casos de uso, fornecendo flexibilidade a organizações de todos os tamanhos e em todos os setores e em toda a sociedade para implementar as abordagens do Framework. O Framework foi elaborado para equipar organizações e indivíduos - aqui referidos como usuários da IA - com abordagens que aumentam a confiabilidade dos sistemas de IA e para ajudar a promover a criação, desenvolvimento, implantação e uso responsáveis dos sistemas de IA ao longo do tempo.

Reconhecimento de fala – Também conhecido como reconhecimento automático de fala (ASR), reconhecimento de fala por computador ou fala para texto, é um recurso que permite que um programa processe a fala humana em um formato escrito. Embora seja comumente confundido com o reconhecimento de voz, o reconhecimento de fala concentra-se na tradução da fala de um formato verbal para um formato de texto, enquanto o reconhecimento de voz busca apenas identificar a voz de um usuário individual. (IBM)

Reconhecimento facial – É um tipo de tecnologia biométrica que usa dados para verificar a presença do rosto de um ser humano em uma captura digital. Há dois usos principais para o software de reconhecimento facial: reconhecimento e autenticação. (Technopedia)

Relatório SOC (Service Organization Company) – Relatório de auditoria, completado por um avaliador independente, que avalia o ambiente de controle interno de uma organização; pode ser fornecido por fornecedores aos clientes para fins de avaliação de que seus controles internos estão operando com eficácia.

Risco – Algo que ameaça a concretização de um objetivo.

Robótica – É a engenharia e a operação de máquinas que podem executar tarefas físicas de forma autônoma ou semiautônoma em nome de um ser humano. Tipicamente, os robôs realizam tarefas que são altamente repetitivas ou perigosas demais para serem executadas com segurança por um ser humano. (Technopedia)

Referências

Ambrozi, Austin. "11 Challenges Of Adopting AI In Business (And How To Address Them Head-On)," *Forbes*, 24 de outubro de 2023. https://www.forbes.com/sites/forbesbusinesscouncil/2023/10/24/11-challenges-of-adopting-ai-in-business-and-how-to-address-them-head-on/?sh=6710c8474bfe.

Ankers, Damon. "Types of Artificial Intelligence: A Detailed Guide." *Certes IT Service Solutions*. https://certes.co.uk/types-of-artificial-intelligence-a-detailed-guide.

Appel, Gil; Juliana Neelbauer; David A. Schweidel. "Generative AI Has An Intellectual Property Problem." *Harvard Business Review*, 7 de abril de 2023. https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem.

Billington, James. "The Prometheus Project: The Story Behind One of AV's Greatest Developments." *ADAS & Autonomous Vehicle International*, 22 de agosto de 2018. https://www.autonomousvehicleinternational.com/features/the-prometheus-project.html.

Britannica. "Artificial Intelligence." *Encyclopedia Britannica*. 2023. https://www.britannica.com/technology/artificial-intelligence.

Britannica. "Deep Blue Computer Chess-Playing System." *Encyclopedia Britannica*. 2023. https://www.britannica.com/topic/Deep-Blue.

COSO. "Guidance, Internal Control Integrated Framework." *COSO*. 2023. https://www.coso.org/guidance-on-ic.

Deloitte. "Modernizing the three lines of defense model: An internal audit perspective." Deloitte, 2023. https://www2.deloitte.com/us/en/pages/advisory/articles/modernizing-the-three-lines-of-defense-model.html.

Doran, George T. "There's a SMART Way to Write Management's Goals and Objectives." *Management Review*, 70, novembro de 1981, 35-36. https://community.mis.temple.edu/mis0855002fall2015/files/2015/10/S.M.A.R.T-Way-Management-Review.pdf.

EY. "The CEO Outlook Pulse – October 2023," EY, 2023. https://www.ey.com/en_us/ceo/ceo-outlook-global-report#:~:text=The%20CEO%20Outlook%20Pulse%20

Gartner. "Gartner Glossary." *Gartner*. 2023. https://www.gartner.com/en/information-technology/glossary/cpu-central-processing-unit.

Hsu, Hansen. "Meet 2021 CHM Fellow Honoree Raj Reddy." *Computer History Museum*. https://computerhistory.org/blog/meet-2021-chm-fellow-honoree-raj-reddy/.

Humanoid Robotics Institute. "History of Humanoid Robot in Waseda University." Waseda University. https://www.humanoid.waseda.ac.jp/history.html.

IBM. "eBook: Build responsible AI workflows with AI governance." *IBM*. https://www.ibm.com/account/reg/us-en/signup?formid=urx-51898.

IBM. "IBM Global AI Adoption Index." *IBM*, 2023. https://www.ibm.com/watson/resources/ai-adoption.

IBM. "Understanding the Different Types of Artificial Intelligence." *IBM*, outubro de 2023. https://www.ibm.com/blog/understanding-the-different-types-of-artificial-intelligence/.

The Institute of Internal Auditors. *CRMA Study Guide and Practice Questions, 3ª Edição*. The IIA. 2023. https://www.theiia.org/en/products/bookstore/crma-study-guide-and-practice-questions-3rd-edition/.

The Institute of Internal Auditors. *Global Perspectives & Insights: the Artificial Intelligence Revolution*. The IIA. 2023. https://www.theiia.org/en/content/articles/global-perspectives-and-insights/2023/global-perspectives-insights-the-artificial-intelligence-revolution/.

The Institute of Internal Auditors. *The IIA's Three Lines Model: An update of the Three Lines of Defense*. The IIA. 2020. https://www.theiia.org/en/content/position-papers/2020/the-iias-three-lines-model-an-update-of-the-three-lines-of-defense/?msclkid=f2923355c01e11ecb401fe1dc46cbc38.

The Institute of Internal Auditors. *International Professional Practices Framework. Edição de 2017.* (Lake Mary, FL: The Institute of Internal Auditors, 2017). https://www.theiia.org/en/products/bookstore/international-professional -practices-framework---ippf---2017-edition/.

Intel. "What is a GPU?" Intel. https://www.intel.com/content/www/us/en/products/docs/processors/what-is-a-gpu.html.

ISACA. "COBIT, An ISACA Framework." *ISACA*. 2023. https://www.isaca.org/resources/cobit.

ISACA. "Glossary." *ISACA*. 2022. https://www.isaca.org/resources/glossary.

Marr, Bernard. "The 15 Biggest Risks of Artificial Intelligence." *Forbes*. 2 de junho de 2023. https://www.forbes.com/sites/bernardmarr/2023/06/02/the-15-biggest-risks-of-artificial-intelligence/?sh=16756d8b2706.

McCarthy, J; M.L. Minsky; N. Rochester; C.E. Shannon. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." *BibSonomy*. https://www.bibsonomy.org/bibtex/24550126962fd8014daa80db1ffae4df2/mhwombat.

Moyer, Steven; Gunter Brunhart; Richard Dubs, Thomas Erickson, Robert Skalamera, Rob Kepner, Marty Meyer, "A (Kind of) Quantitative Approach to Organizational Risk Tolerance." *ISACA*, 8 de julho de 2021. https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2021/a-kind-of-quantitative-approach-to-organizational-risk-tolerance.

National Cyber Security Centre. "Guidelines for secure AI system development." *National Cyber Security Centre*. 2023. https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development.

NIST. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, Gaithersburg, Md.: NIST, 2023. https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf.

NIST Computer Security Resource Center. "Glossary." Gaithersburg, Md.: NIST. https://csrc.nist.gov/glossary.

PwC. "PwC 2022 AI Business Survey (U.S.)." *PwC*. https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-business-survey.html.

QuantumBlack AI by McKinsey. "The State of AI in 2023: Generative AI's Breakout Year." *McKinsey*. 2023. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-AIs-breakout-year.

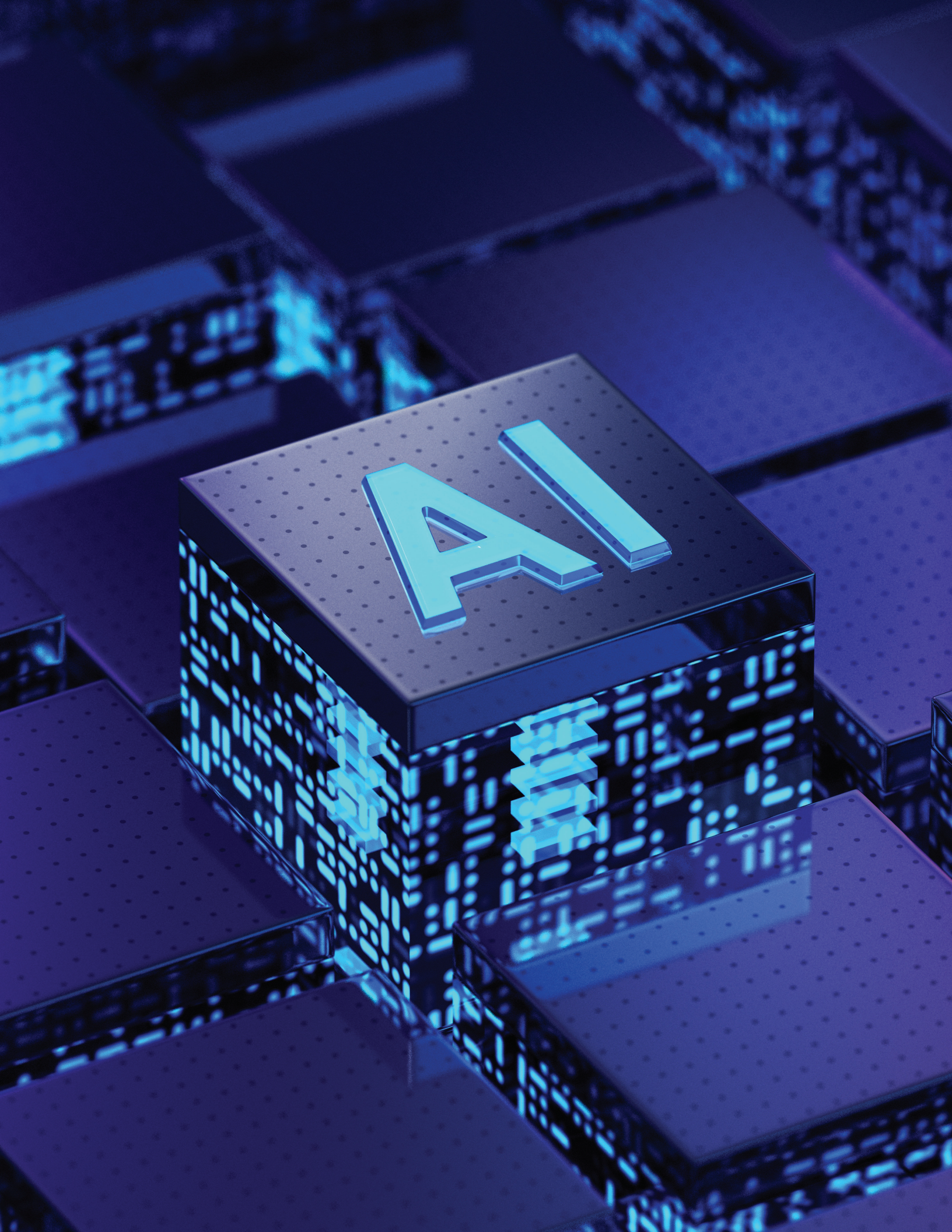
Samuel, A.L. "Some Studies in Machine Learning Using the Game of Checkers." *IBM Journal of Research and Development*. Julho de 1959. https://ieeexplore.ieee.org/document/5392560.

Techopedia.com. "TechDictionary." https://www.techopedia.com/dictionary.

Turing, A.M. "Computing Machinery and Intelligence." *Mind*, Volume LIX, Edição 236, outubro de 1950, 433–460. https://academic.oup.com/mind/article/LIX/236/433/986238.

Weizenbaum, Joseph. "ELIZA—A Computer Program For the Study of Natural Language Communication Between Man and Machine." *Communications of the Association for Computing Machinery*. Janeiro de 1966. https://dl.acm.org/doi/10.1145/365153.365168.

Casa Branca. "Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence." *The White House*. 30 de outubro de 2023. https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.


AI

Equipe de Desenvolvimento do Framework
George Barham, Allison Banzon, Anne Mercer, Kat Seeuws, Geoff Nordhoff.

Colaboradores
Andrew Cook, Pam Stroebel Powers, Robert Perez, Jim Enstrom, Scott Moore.

Sobre o The IIA
The Institute of Internal Auditors (IIA) é o mais reconhecido defensor, educador e fornecedor de normas, orientações e certificações da profissão de auditoria interna. Fundado em 1941, o The IIA atende hoje a mais de 245.000 membros de mais de 170 países e territórios. A sede global da associação fica em Lake Mary, Flórida, EUA. Para mais informações, acesse **theiia.org**.

Isenção de Responsabilidade
O The IIA publica este documento para fins informativos e educacionais. Este material não tem a intenção de fornecer respostas definitivas para circunstâncias individuais específicas e, portanto, deve ser usado apenas como guia. O The IIA recomenda a busca de assessoria especializada independente relacionada diretamente a qualquer situação específica. O The IIA não se responsabiliza por qualquer pessoa que confie exclusivamente neste material.

Sede Global do The IIA
1035 Greenwood Blvd., Suíte 401
Lake Mary, FL 32746 EUA

The Institute of
Internal Auditors