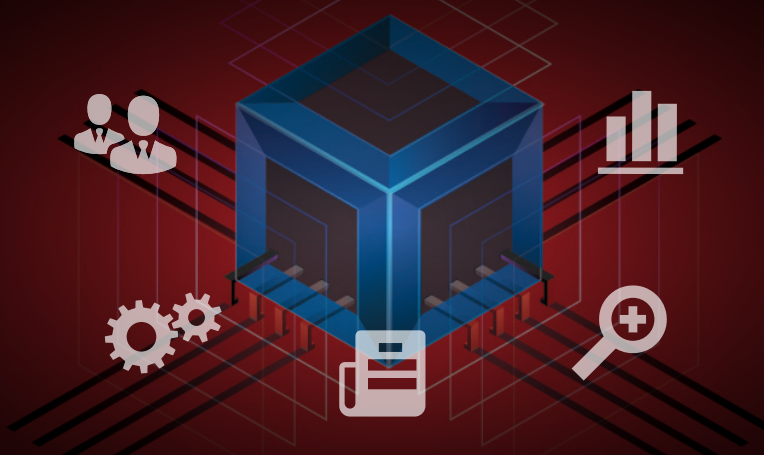




COSO

Committee of Sponsoring Organizations of the Treadway Commission

Gerenciamento de Riscos Corporativos



CONCRETIZE TODO O POTENCIAL DA INTELIGÊNCIA ARTIFICIAL

**APLICANDO O FRAMEWORK E OS PRINCÍPIOS DO COSO PARA AJUDAR
A IMPLANTAR E DIMENSIONAR A INTELIGÊNCIA ARTIFICIAL**

Patrocinado por

Deloitte.

Keri Calagna | Brian Cassidy | Amy Park

Setembro de 2021

As informações aqui contidas são de natureza geral e baseadas em autoridades sujeitas a alterações. A aplicabilidade das informações a situações específicas deve ser determinada por meio de consulta com seu assessor profissional, e este documento não deve ser considerado substituto para os serviços de tais assessores, nem deve ser usado como base para qualquer decisão ou ação que possa afetar sua organização.

Autores



Keri Calagna
Risk & Financial Advisory Principal
Deloitte & Touche LLP



Brian Cassidy
Audit & Assurance Partner
Deloitte & Touche LLP



Amy Park
Audit & Assurance Partner
Deloitte & Touche LLP

Agradecimentos

Gostaríamos de reconhecer e agradecer a John Fogarty, Gerente Sênior da Deloitte & Touche LLP, Hemant Dhengane, Gerente da Deloitte & Touche LLP, Mary Schmidlin, Gerente Sênior da Deloitte & Touche LLP, e Edward Bowen, Diretor Administrativo da Deloitte & Touche LLP por suas contribuições e conselhos técnicos.

O Conselho do COSO gostaria de agradecer à Deloitte & Touche LLP por seu apoio.

Membros do Conselho do COSO

Paul J. Sobel
Presidente do COSO

Daniel C. Murdock
Financial Executives International

Douglas F. Prawitt
American Accounting Association

Jeffrey C. Thomson
Institute of Management Accountants

Jennifer Burns
American Institute of CPAs (AICPA)

Patty K. Miller
The Institute of Internal Auditors

Prefácio

Este projeto foi encomendado pelo *Committee of Sponsoring Organizations of the Treadway Commission* (COSO), que se dedica a ajudar as organizações a melhorar seu desempenho, desenvolvendo liderança criativa que aprimora o controle interno, gerenciamento de riscos, governança e dissuasão de fraudes.

O COSO é uma iniciativa do setor privado, patrocinada e financiada em conjunto pelas seguintes organizações:



American Accounting Association (AAA)



American Institute of CPAs (AICPA)



Financial Executives International (FEI)



The Institute of Management Accountants (IMA)



The Institute of Internal Auditors (IIA)

COSO

Committee of Sponsoring Organizations
of the Treadway Commission

coso.org

Gerenciamento de Riscos Corporativos



CONCRETIZE TODO O POTENCIAL DA INTELIGÊNCIA ARTIFICIAL

APLICANDO O FRAMEWORK E OS PRINCÍPIOS DO COSO PARA AJUDAR
A IMPLANTAR E DIMENSIONAR A INTELIGÊNCIA ARTIFICIAL

Pesquisa Encomendada por



Committee of Sponsoring Organizations of the Treadway Commission

Setembro de 2021

Copyright © 2021, Committee of Sponsoring Organizations of the Treadway Commission (COSO).
1234567890 PIP 198765432

As imagens do COSO são do *COSO Enterprise Risk Management - Integrating with Strategy and Performance* ©2017, *American Institute of Certified Public Accountants* em nome do *Committee of Sponsoring Organizations of the Treadway Commission* (COSO). COSO é uma marca registrada do *Committee of Sponsoring Organizations of the Treadway Commission*.

Todos os direitos reservados. Nenhuma parte desta publicação pode ser reproduzida, redistribuída, transmitida ou exibida de qualquer forma ou por qualquer meio sem permissão por escrito. Para obter informações sobre licenças e permissões para reimpressão, entre em contato com o *American Institute of Certified Public Accountants*, que lida com licenciamento e permissões para materiais protegidos por direitos autorais do COSO. Direcione todas as perguntas para copyright-permissions@aicpa-cima.com ou AICPA, com destinatário "Attn: Manager; Licensing & Rights", 220 Leigh Farm Road, Durham, NC 27707 USA. Perguntas por telefone podem ser direcionadas para 888-777-7077.

Design e produção: Sergio Analco.

Conteúdos	Pág.
Introdução	1
A Revolução da IA: Transformando Negócios e a Inovação	3
O Framework de ERM do COSO: Abordando os Riscos de IA em Alinhamento com sua Estratégia Geral de Negócios e TI	7
Governança & Cultura	9
Definição de Estratégia e de Objetivos	11
Desempenho	13
Análise e Revisão	17
Informação, Comunicação e Reporte	19
Observações de Conclusão	21
Sobre os Autores	23
Sobre o COSO	24
Sobre a Deloitte	24





INTRODUÇÃO

A inteligência artificial (IA) tem transformado e continuará a transformar as estratégias, soluções e operações de negócios. Os riscos relacionados à IA precisam vir em primeiro lugar e ter uma prioridade fundamental, para que as organizações adotem e escalem as aplicações da IA e concretizem todo o potencial da IA. A aplicação dos princípios do gerenciamento de riscos corporativos (ERM) a iniciativas de IA pode ajudar as organizações a fornecer uma governança integrada de IA, gerenciar riscos e impulsionar o desempenho para maximizar o atingimento das metas estratégicas. O Framework de ERM do COSO, com seus cinco componentes e vinte princípios, é um framework amplo e abrangente, pode alinhar o gerenciamento de riscos com a estratégia e o desempenho da IA, para ajudar a concretizar o potencial da IA.

Figura 1. **COSO Enterprise Risk Management – Integrating with Strategy and Performance Framework**



2017 COSO Enterprise Risk Management – Integrating with Strategy and Performance



A REVOLUÇÃO DA IA: TRANSFORMANDO NEGÓCIOS E A INOVAÇÃO

Conforme a IA se expande em quase todos os aspectos da vida moderna, ela está se tornando uma capacidade de negócios necessária. Seja gerenciando relacionamentos com clientes, identificando e respondendo a ameaças cibernéticas ou ajudando a orientar decisões médicas, a IA está abordando uma ampla gama de problemas de negócios. A rápida adoção da IA está fornecendo informações sobre os dados das organizações que, por sua vez, fornecem inteligência para apoiar a tomada de decisões. Isso levou as organizações a investir em iniciativas de IA em grande escala. Prevê-se que os gastos com IA dobrem até 2024, passando de US\$ 50,1 bilhões em 2020 para mais de US\$ 110 bilhões em 2024. A taxa de crescimento anual composta (CAGR) prevista para este período é de aproximadamente 20%.¹ Além disso, as receitas mundiais para o mercado de IA, incluindo software, hardware e serviços, devem crescer para US\$ 327,5 bilhões em 2021 e atingir US\$ 554,3 bilhões até 2024, com uma CAGR de cinco anos de 17,5%.²

O que está fomentando a revolução? As organizações estão aplicando a IA por seu potencial transformador: automatizar processos, tarefas e ações de negócios para reduzir custos, aumentar a eficiência e melhorar a previsibilidade dos resultados. Com a IA, elas estão vendo melhores insights de dados, levando a decisões de negócios mais informadas, resultados comerciais e operacionais positivos e maior inovação.

Como empresas usam a IA para gerar valor

REDUÇÃO DE CUSTOS

Aplicam a IA para automatizar de forma inteligente processos, tarefas e interações de negócios para reduzir custos, aumentar a eficiência e melhorar a previsibilidade.

VELOCIDADE À EXECUÇÃO

Aplicam a IA para acelerar o tempo de resultados operacionais e de negócios, minimizando a latência.

ANÁLISE PREDITIVA

Aplicam a IA para oferecer insights sobre dados de uma empresa e melhorar a compreensão e a tomada de decisões, decifrando padrões, conectando pontos e prevendo resultados de fontes de dados cada vez mais complexas.

ENGAJAMENTO DIGITAL

Aplicam a IA para mudar a forma como humanos interagem com sistemas inteligentes, expandindo as formas de engajamento por meio de voz, visão, texto e toque.

INOVAÇÃO FOMENTADA

Aplicam a IA para gerar insights para novos produtos, oportunidades de mercado e modelos de negócios.

Estudos recentes indicam que as organizações estão agindo para aproveitar esses benefícios com investimentos de curto prazo em IA:

- 75% dos entrevistados esperam mudar da IA piloto para a IA operacionalizada até o fim de 2024.³
- 75% dos usuários de IA entrevistados esperam uma transformação organizacional dentro de três anos.⁴
- 61% dos usuários de IA entrevistados estão antecipando a transformação do setor no mesmo prazo.⁵
- Os usuários de IA entrevistados estão investindo significativamente, com 53% gastando mais de US\$ 20 milhões em 2020 em tecnologia e talentos relacionados à IA.⁶
- 71% dos usuários de IA entrevistados esperam aumentar o investimento no próximo ano fiscal, em uma média de 26%.⁷

¹ International Data Corporation (IDC), "Worldwide Spending on Artificial Intelligence is Expected to Double in Four Years, Reaching \$110 Billion in 2024, According to New IDC Spending Guide," 25 de agosto de 2020. <https://www.idc.com/getdoc.jsp?containerId=prUS46794720>

² International Data Corporation (IDC), "IDC Forecasts Improved Growth for Global AI Market in 2021," 23 de fevereiro de 2021. <https://www.idc.com/getdoc.jsp?containerId=prUS47482321>

³ Gartner, Accelerating AI Deployments – Paths of Least Resistance, julho de 2020.

⁴ Deloitte, State of AI in the Enterprise, 3rd Edition, 2020. Figura 2, página 7.

⁵ Ibid., Figura 2, página 7.

⁶ Ibid., página 6.

⁷ Ibid., página 6.

Para colocar a transformação organizacional e da indústria em perspectiva, muitas empresas estão investindo em recursos de IA para impulsionar sua estratégia de negócios. Em alguns casos, a IA sustenta modelos de negócios, como o caso de algumas empresas de tecnologia financeira que se afastam das pontuações FICO tradicionais e usam vários

parâmetros e modelos com inteligência artificial para informar as decisões de crédito. O processo é automatizado, tornando o esforço mais eficiente, e alerta os usuários quando os casos precisam de revisão adicional. Pode melhorar a tomada de decisões e aprimorar os serviços e a experiência existentes para os clientes.

IA e Aprendizado de Máquina: Uma introdução prática

Um entendimento dos algoritmos associados à IA e como são criados é fundamental para identificar e gerenciar devidamente os riscos relacionados à IA. Na prática, a IA é desenvolvida por humanos através do uso de programação de software (código). Semelhante à necessidade de governança e controles no reporte financeiro ou no desenvolvimento de software, devido ao elemento humano, as organizações precisam de governança e controles para IA também. Mas os conselhos e executivos não podem ajudar a monitorar os controles com eficácia sem um entendimento básico do que a IA faz e como ela é construída.

O que os algoritmos fazem

Há três classes comuns de algoritmos de aprendizado de máquina: aprendizado não profundo, aprendizado profundo e aprendizado por reforço. O objetivo desses modelos de IA é criar uma classificação, previsão ou a geração de novos dados.

- **O aprendizado não profundo classifica, encontra padrões e prevê resultados.** Modelos comuns incluem regressões, clustering, árvores de decisão e máquinas de vetor de suporte. Podem ajudar com muitos problemas úteis e comuns, como previsão de demanda, propensão a vendas cruzadas e classificação de riscos.
- **Os algoritmos de aprendizado profundo mudaram o jogo.** Esses métodos de classificação e previsão impulsionaram a revolução da IA da última década. Imagens, processamento de linguagem natural e detecção de anomalias alcançaram resultados de última geração usando redes neurais profundas. Os bots de conversação que estão ajudando as pessoas a navegar no atendimento ao cliente em um site vêm dessa tecnologia de IA. Uma automação simples pode ser aplicada de forma mais ampla, como voz para texto em um telefone celular, ou pode ser usada para reconhecer e traduzir a escrita à mão, utilizando os dados para auxiliar no esforço.
- **Modelos de aprendizagem por reforço** examinam um ambiente e desenvolvem a capacidade de tomar uma sequência de decisões para encontrar o melhor caminho positivo a seguir. Esses modelos podem aprender a ganhar torneios de Xadrez e de Go contra grandes mestres humanos. As aplicações práticas incluem otimização de rotas, otimização de fábrica e testes de vulnerabilidade cibernética.

Como os algoritmos são construídos

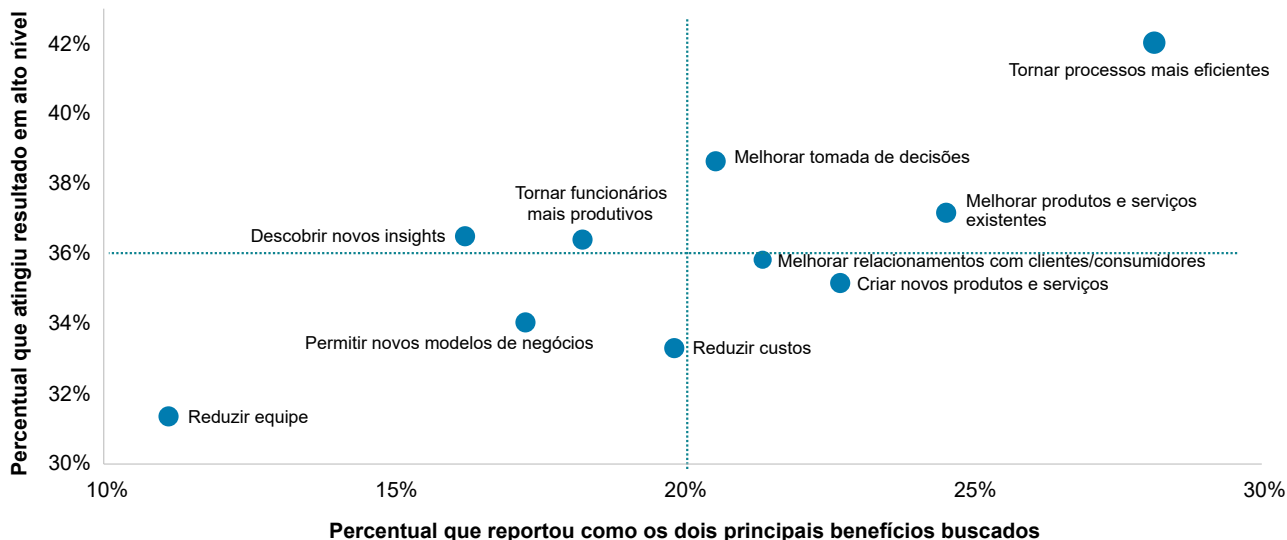
Todo algoritmo deve estar vinculado à estratégia de negócios. Os algoritmos são criados por humanos para contribuir para a tomada de decisão informada, que cria o valor comercial pretendido. Há seis etapas principais para criar um modelo de aprendizado de máquina:

1. **Definição do problema** - Considerar um problema de negócios e como o aprendizado de máquina pode resolvê-lo.
2. **Perfilação de dados** - Identificar as fontes de dados necessárias para resolver o problema e quais dados adicionais são necessários. Uma tendência emergente na IA é o desenvolvimento de novos sensores e coleta de dados com o único propósito de melhorar o desempenho da IA. As organizações precisam garantir que os dados sejam justos e equilibrados em todas as dimensões éticas e de desempenho.
3. **Preparação de dados** - Determinar o que é necessário para transformar, normalizar e limpar os dados e criar uma abordagem de teste e validação.
4. **Avaliação de algoritmos** - Alavancar as principais práticas para selecionar os algoritmos necessários para resolver o problema. Muitas vezes, as equipes de ciência de dados desenvolvem vários algoritmos em paralelo, para determinar o modelo com melhor desempenho. É importante estabelecer os critérios corretos de avaliação de desempenho.
5. **Desenvolvimento do modelo** - Treinar, testar e validar todos os algoritmos identificados com os dados e implantar abordagens, como regularização.
6. **Implantação, monitoramento e manutenção de modelos** - Incorporar operações de aprendizado de máquina (MLOps) e estruturas de monitoramento, além de processos para lidar com desvios de modelos. O desempenho do modelo pode diminuir, se as atividades no ambiente mudarem ao longo do tempo (por exemplo, modelos que preveem o consumo de eletricidade precisam ser atualizados ao longo do tempo, conforme os painéis solares ganham força com os consumidores).

A IA serve a um mundo maravilhoso... até que haja um resultado infeliz

Conforme a implantação da IA e do aprendizado de máquina aumentou, os dois principais benefícios da implantação citados pelos usuários pesquisados são o aumento da eficiência do processo e o aprimoramento dos produtos e serviços existentes. (Veja a Figura 2) Além disso, uma pesquisa realizada pelo Gartner indica que os dois principais motivos para as organizações investirem em recursos de IA são o desejo de aumentar a receita ou reduzir os custos, e abordar vulnerabilidades de concorrentes e startups.⁸

Figura 2. Eficiência do Processo Lidera a Lista de Benefícios Atingidos com a IA



As linhas azuis pontilhadas representam a média das respectivas dimensões.

Fonte: State of AI in the Enterprise, 3rd Edition, Deloitte. Copyright © 2020 Deloitte Development LLC. Todos os direitos reservados.

A IA impulsiona a eficiência por meio de algoritmos de computador que usam dados para criar previsões ou recomendações prescritivas, gerar classificações e inventar novas construções. Muitos casos de uso de IA implantados hoje estão fazendo coisas que os humanos podem fazer, mas com muito mais rapidez e eficiência. Nos próximos dez anos, a ênfase provavelmente evoluirá para a implantação da IA para fazer coisas que os humanos não podem fazer, porque os humanos são incapazes de ver as sutilezas e nuances que a IA pode detectar. Por exemplo, as empresas farmacêuticas podem usar a IA para interpretar nuances em imagens microscópicas que os cientistas humanos não conseguem detectar. Esse perfil de célula baseado em imagem em grande escala está verificando rapidamente as diferenças entre grandes conjuntos de dados de células saudáveis e doentes, a fim de projetar novos compostos de drogas altamente específicos para tratar doenças. Em teoria, os pesquisadores poderiam fazer as comparações a olho nu; no entanto, comparar milhares de células com diferenças pequenas, mas consistentes, seria muito difícil sem o uso de IA. Em essência, a IA está impulsionando a inovação transformadora. Essas tendências podem acelerar ou evoluir ainda mais no futuro. Embora a IA pareça uma panaceia para a transformação dos

negócios, a tecnologia e a aplicação da tecnologia não estão isentas de riscos que possam resultar em sérios problemas para uma organização. Esses riscos podem ser mitigados pela consideração cuidadosa e preventiva do Framework de ERM do COSO. Mas primeiro, vamos falar sobre os riscos. Há um amplo espectro de riscos relacionados à IA que incluem, mas não se limitam ao seguinte:

- Viés e falhas de confiabilidade, devido a falhas inadequadas ou dados não representativos
- Incapacidade de entender ou explicar as saídas do modelo de IA
- Uso inadequado de dados
- Vulnerabilidades a ataques adversários para obter dados ou manipular o modelo de IA
- Estresses sociais devido à rápida aplicação e transformação de tecnologias de IA

⁸ 2019 Gartner, AI in Organizations Survey. 735439_C.

As consequências potenciais desses riscos podem incluir danos à reputação, destruição do valor do acionista, multas regulatórias e ações judiciais. Devido a esses riscos emergentes, 56% dos usuários da IA entrevistados dizem que sua organização está diminuindo a adoção de tecnologias de IA.⁹ No entanto, isso pode não ser viável por muito tempo, se as organizações permanecerem competitivas. Em vez de pisar no freio, uma estratégia mais prudente pode ser gerenciar melhor os riscos associados. As organizações não podem ignorar os riscos ou consequências não intencionais da IA.

A pesquisa “State of AI in the Enterprise” da Deloitte ilustra que os implantadores e usuários da IA têm sérias preocupações sobre o uso de IA, envolvendo uma variedade de áreas de risco além do viés. (Veja a Figura 3) Além disso, os entrevistados da pesquisa indicam que há lacunas significativas nas habilidades atuais de suas organizações para lidar com essas preocupações. Os resultados de uma pesquisa separada conduzida pela Gartner citaram as principais barreiras à implantação da IA como preocupações com segurança ou privacidade, e a complexidade da integração de soluções de IA com a infraestrutura existente.¹⁰

Figura 3. Comparação de preocupação versus preparação para riscos relacionados à IA



Impacto da incerteza regulatória

Os requisitos regulatórios são outra consideração importante e aderir à conformidade regulatória significa não apenas seguir a legislação atual, mas também demonstrar compromisso com práticas seguras de IA que podem se tornar necessárias no futuro. As organizações devem considerar a extensão aplicável dos requisitos regulatórios pendentes, ao avaliar sua estrutura de governança sobre a IA e dados relacionados.

Figura 4. Conformidade regulatória

Exemplo de Envolvidos	Exemplo de Normas, Políticas e Leis	
	Exemplo de Regulamento	O que significa para seu negócio
The World Economic Forum's Council on the Future of AI and Robotics	Leis de responsabilidade do produto se aplicam a indivíduos feridos ao usar produto orientado por IA	As empresas devem monitorar a IA da mesma forma que monitoram funcionários humanos (digital não é igual a infalível)
AI Now Initiative	A Restatement of Torts refere-se a defeitos de design e fabricação de IA e falha em avisar	Cuidado extra deve ser tomado ao desenvolver rótulos de advertência para produtos orientados por IA, p. ex., "Este produto foi auditado com IA"
The Stanford One Hundred Year Study on AI	Fair Credit Reporting Act a aplicação da FTC contra o conluio através da IA	Conclusão poderia ser não intencional sem transparência nos métodos de IA, o que significa que empresas precisam de forte controle em vigor
MIT Media Lab, AI, Ethics and Governance Project	Regulamento geral de proteção de dados da UE que afeta empresas dos EUA que operam na UE	Empresas precisam criar políticas sobre IA que atendam às expectativas, até mesmo nos mercados mais regulamentados
The Partnership on AI	Leis de comunicação aplicadas à comunicação entre bots e pessoas	Controles extras devem ser implantados em casos de uso de IA conversacional para incorporar leis de comunicação
Data & Society's Intelligence and Autonomy Initiative	Bot Disclosure and Accountability Act 2018, para regular bots de notícias	Bots de mídia social já exigem a divulgação de que estão operando com IA; regulamentação futura pode ir além dos bots sociais

Copyright © 2020 Deloitte Development LLC. Todos os direitos reservados.

⁹ Ibid., página 13.

¹⁰ 2019 Gartner, AI in Organizations Survey. 729419_C.

O FRAMEWORK DE ERM DO COSO: ABORDANDO OS RISCOS DE IA EM ALINHAMENTO COM SUA ESTRATÉGIA GERAL DE NEGÓCIOS E TI

Conforme a IA se torna mais difundida nos negócios e em nossas vidas cotidianas, as organizações provavelmente não terão mais a opção de ignorar ou evitar os riscos exclusivos que acompanham a adoção da IA. Em vez disso, devem aprender a identificar e gerenciar esses riscos de forma eficaz. Para agravar o problema, há o fato de que a IA geralmente não está isolada em uma função específica, como TI, mas afeta várias funções em uma organização. As organizações precisam criar e implantar estratégias e estruturas de governança, gerenciamento de riscos e controle para concretizar o potencial dos humanos em colaboração com a IA. Felizmente, a IA é como outros componentes tecnológicos de uma organização e, portanto, pode ser governada com sucesso por um ERM eficaz.

Desde 1985, o comitê voluntário do setor privado Committee of Sponsoring Organizations of the Treadway Commission (COSO) tem se concentrado em ajudar as organizações a melhorarem

seu desempenho, desenvolvendo uma liderança criativa que aprimora o controle interno, gerenciamento de riscos, governança e dissuasão de fraudes. A atualização mais recente do Framework de ERM do COSO – adotada em 2017 – destaca a importância de incorporá-lo em toda a organização em cinco componentes críticos:






-  **Governança e Cultura**
-  **Definição de Estratégia e Objetivos**
-  **Desempenho**
-  **Análise e Revisão**
-  **Informação, Comunicação e Reporte**

Figura 5. **COSO Enterprise Risk Management - Integrating with Strategy and Performance Framework**



2017 COSO Enterprise Risk Management - Integrating with Strategy and Performance

Ao alavancar o Framework de ERM do COSO, as organizações podem identificar e gerenciar riscos específicos à IA e estabelecer práticas para otimizar os resultados, enquanto gerenciam a exposição a riscos como o viés não intencional e a falta de transparência. A implantação pode ajudar a melhorar a confiança entre os stakeholders dentro e fora da organização e a abordar proativamente os riscos emergentes relacionados à IA.





GOVERNANÇA E CULTURA

Governança e cultura, juntas, formam a base para todos os componentes do gerenciamento de riscos. A governança reforça a importância do ERM e a cultura se reflete na tomada de decisões em todos os níveis de uma organização. De acordo com o Framework de ERM do COSO, esses componentes devem incorporar o compromisso de uma organização com sua visão, missão e valores centrais. Os valores fundamentais fornecem uma base importante para a supervisão adequada das iniciativas de IA e modelos de IA, para ajudar a atingir a estratégia e os objetivos de negócios da organização. O componente de Governança e Cultura e os seguintes princípios do Framework de ERM do COSO servem como base para esta seção do documento:

- 1 **Exercita Supervisão de Riscos do Conselho**
- 2 **Estabelece Estruturas Operacionais**
- 3 **Define a Cultura Desejada**
- 4 **Demonstra Compromisso com os Valores Essenciais**
- 5 **Atrai, Desenvolve e Retém Indivíduos Capazes**

O conselho de uma organização, geralmente, não está envolvido em iniciativas de IA, ou pode não estar totalmente ciente delas para fazer as devidas perguntas de gestão relacionadas a riscos. Quando executivos de alto nível e membros do conselho entendem a IA e suas consequências e estão ativamente engajados, eles definem o tom no topo sobre a importância do gerenciamento de riscos. Tal engajamento é imperativo.

Apenas cerca de 26% dos usuários de IA entrevistados têm um único executivo responsável pelo gerenciamento de riscos relacionados à IA.¹¹ Assim como outros elementos fundamentais de um negócio, os membros do conselho precisam entender o framework da organização para avaliar o risco associado às iniciativas de IA e determinar o limiar de risco que requer supervisão da alta administração. Algumas iniciativas podem ser limitadas a um pequeno número de modelos simples de IA e ter um perfil de risco menor. Outras iniciativas podem ter um grande número de modelos complexos de IA ou envolver atividades críticas de negócios, como oferecer assistência médica ao paciente, garantir a segurança do cliente ou controlar atividades de fabricação, e ter um perfil de risco mais alto. As iniciativas de IA de alto risco exigem supervisão de perto por um executivo sênior, em colaboração com um diretor de risco ou líder de risco equivalente. As organizações podem precisar de pessoal com experiência em desenvolvimento de IA e análise de

dados, para supervisionar devidamente suas iniciativas de IA, ou buscar assessores externos com experiência relevante, se o conjunto de habilidades necessário estiver ausente na organização. Esses indivíduos podem assessorar os membros do conselho, oferecer insights sobre riscos/recompensas e promover a tomada de decisões informadas sobre os riscos. Esse envolvimento é fundamental para a adoção e implantação eficazes da IA e para a prevenção de eventos de crise organizacional.

A Importância da Governança

Conforme a IA é implantada em uma escala mais ampla dentro das organizações, a governança tem um papel fundamental na supervisão adequada das iniciativas de IA e modelos relacionados. As organizações estão enfrentando um maior escrutínio de vários stakeholders (p. ex., reguladores, clientes, usuários, etc.), devido, em parte, à supervisão inadequada percebida da IA.

A governança desempenha um papel fundamental nas seguintes áreas principais:

1. Para apoiar o desenvolvimento e a operação de modelos de IA, as organizações estão coletando quantidades sem precedentes de dados. Os participantes têm preocupações, incluindo, mas não limitadas a, como seus dados estão sendo usados e quem mais tem acesso aos seus dados. As organizações precisam ter regras claras sobre o uso de dados, coleta de dados, retenção de dados e acesso de dados e aplicar consistentemente essas regras em toda a organização, como parte de sua resposta a essas preocupações. A falha em abordar devidamente essas questões pode prejudicar as pessoas e causar danos à reputação corporativa e ao valor para os acionistas.
2. As organizações estão aplicando cada vez mais a IA a situações que exigem mais julgamento e podem ter um impacto significativo sobre os envolvidos. Modelos de IA que realizam ou informam julgamentos significativos (p. ex., decisões de subscrição, elegibilidade a vários benefícios, diagnóstico médico e tratamento recomendado, etc.), que têm um impacto significativo nos envolvidos, podem apresentar preocupações éticas. Como parte de sua resposta, as organizações precisam avaliar quando, onde e como a IA é ou será usada, e se tal uso é consistente com os valores e o design da organização, e como as estruturas de supervisão da organização se relacionam com as maiores preocupações sociais, se aplicável.

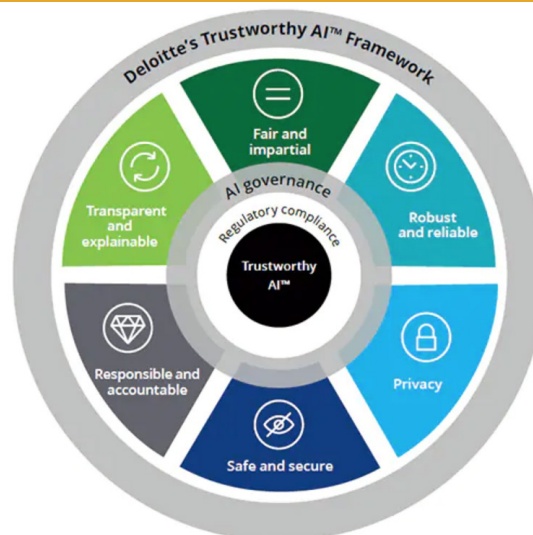
¹¹ Ibid., based on average from Figure 9 on page 15.

Além disso, os líderes precisam entender como definem o sucesso ao desenvolver, implantar, monitorar e manter a IA, e como ela se relaciona com o objetivo da empresa. Aspectos importantes da definição de sucesso incluem determinar quais medidas ou métricas são mais aplicáveis, além de como a organização identifica e avalia custos versus benefícios. Esses aspectos estão intimamente relacionados à gestão vincular as iniciativas de IA ao compromisso mais amplo da organização com seus valores fundamentais, fornecendo a base para aplicar prestação de contas pelas ações e alinhando comportamentos conscientes de risco e tomada de decisão com desempenho. Como tal, as organizações precisam de um processo rigoroso e controlado para documentar o propósito do algoritmo, bem como as necessidades e objetivos da organização. Isso deve ser incluído no documento de arquitetura de IA de uma organização e nos processos de desenvolvimento de software relacionados.

Juntamente com a visibilidade clara para principais executivos e membros do conselho, a governança dos dados subjacentes é fundamental para um framework eficaz de ERM. Para uma implantação bem-sucedida, as organizações devem avaliar quais dados são necessários para desenvolver a IA. Os algoritmos de IA usam dados para treinar e criar um novo modelo. Os modelos preveem resultados futuros conforme recebem novos dados. As considerações de governança de dados necessárias, baseadas em valores fundamentais, podem incluir 1) representação da população apropriada para o caso de uso de IA e redução de viés; 2) regras claras para uso e divulgação de dados, incluindo privacidade na coleta de dados, bem como divulgação de uso e descarte; e 3) formas de proteger ativos de dados.

A IA e os modelos que a fazem funcionar também precisam ser monitorados de perto em toda a organização. Ao projetar e implantar a IA, seis dimensões principais podem ajudar a proteger a ética e construir uma estratégia de IA confiável para a empresa que as pessoas possam adotar. Embora, atualmente, não exista um framework oficial de ética para a IA, o *Trustworthy AI™ Framework* da Deloitte pode servir como meio para entender e avaliar os riscos e as considerações éticas específicas da IA, e pode ser uma lente valiosa para complementar o Framework de ERM do COSO, especialmente no que se refere à governança e desempenho. As organizações podem usá-lo para ajudar a determinar e monitorar os riscos contínuos.

Figura 6. Deloitte's Trustworthy AI™ Framework



Copyright © 2020 Deloitte Development LLC. All rights reserved.

O *Trustworthy AI™ Framework* da Deloitte (veja a Figura 6) inclui:

- **Justiça e imparcialidade** – Avalie se os sistemas de IA incluem verificações internas e externas, para ajudar a permitir uma aplicação equitativa entre todos os participantes.
- **Transparência e explicabilidade** – Ajude os participantes a entender como seus dados podem ser usados e como os sistemas de IA tomam decisões. Algoritmos, atributos e correlações estão abertos à inspeção.
- **Responsabilidade e responsabilização** – Implante uma estrutura organizacional e políticas que possam ajudar a determinar claramente quem é responsável pelo produto das decisões do sistema de IA.
- **Robustez e confiabilidade** – Confirme se os sistemas de IA têm a capacidade de aprender com humanos e outros sistemas, para produzir resultados consistentes e confiáveis.
- **Privacidade** – Respeite a privacidade dos dados e evite usar a IA para alavancar os dados do cliente além do uso pretendido e declarado. Permita que os clientes ativem ou desativem o compartilhamento de seus dados.
- **Segurança e proteção** – Proteja os sistemas de IA contra riscos potenciais (incluindo riscos cibernéticos) que possam causar danos físicos e digitais.

Pontos a Considerar

- A organização tem um programa integrado de governança de IA?
- Como as considerações éticas são abordadas na implantação da IA? Deve haver um diretor de ética para controlar o monitoramento contínuo da IA?
- A organização tem um diretor de riscos, diretor de dados ou líder de riscos equivalente para ajudar com os riscos associados às iniciativas de IA em toda a empresa?
- O conselho tem um membro que seja especialista em tecnologia ou IA?
- Quais aprovações ou consultas no nível do conselho ocorrem quanto à implantação e às mudanças pós-implantação da IA?



DEFINIÇÃO DE ESTRATÉGIA E OBJETIVOS

Toda organização tem uma estratégia para concretizar sua missão e visão e gerar valor. As organizações devem integrar o ERM com a definição da estratégia, para obter informações sobre o perfil de risco associado à estratégia e aos objetivos de negócios da organização. O componente Definição de Estratégia e de Objetivos e os seguintes princípios do Framework de ERM do COSO servem como base para esta seção do documento:

- 6 **Analisa o Contexto de Negócios**
- 7 **Define o Apetite a Risco**
- 8 **Avalia Estratégias Alternativas**
- 9 **Formula Objetivos de Negócios**

As organizações devem estabelecer uma estratégia e objetivos de negócios quanto à IA. De acordo com a pesquisa *Chief Strategy Officer Survey* de 2020 da Deloitte, 51% dos entrevistados indicaram que a IA era importante para a estratégia de sua organização, mas 17% achavam que sua organização tinha recursos para executar estratégias relacionadas à IA.¹² Com um entendimento do contexto de negócios e estratégico, os líderes da organização podem entender os fatores internos e externos que afetam os riscos de suas iniciativas de IA. Componentes importantes são a classificação dos usos atuais ou potenciais da IA e dados relacionados, e a avaliação da exposição potencial ao uso da IA.

Usar modelos de IA que não se alinham aos valores da organização pode prejudicar os objetivos estratégicos. Há vários exemplos em que as organizações usaram modelos de IA que produziram resultados que levaram a um tratamento diferente ou desigual dos participantes com base em características protegidas (p. ex., gênero, raça, etc.). Esses incidentes sugerem um foco inadequado na identificação e abordagem de questões relacionadas à imparcialidade e transparência durante o desenvolvimento e monitoramento contínuo da IA.

Definir o apetite a risco permite que as organizações alinhem à estratégia de negócios a identificação, avaliação e resposta a risco. Para uma discussão mais aprofundada sobre a definição do apetite a risco, consulte "*Risk Appetite – Critical to Success: Using Risk Appetite to Thrive in a Changing World*", do COSO.¹³ Uma consideração adicional sobre desenvolver o apetite a risco é fazer a comparação com colegas da indústria. O apetite a risco de uma organização também é uma consideração importante para facilitar a tomada de decisões informadas sobre riscos no que se refere à IA. Não há como eliminar totalmente o risco, portanto, a organização deve determinar seu apetite a risco e avaliar quanto risco é aceitável ao identificar e avaliar investimentos na IA. Há uma compensação entre risco e recompensa a ser considerada. As organizações que adotam a IA estão fazendo investimentos significativos em desenvolvimento e implantação, e devem alinhar seu gerenciamento de riscos de IA com esforços mais amplos de gerenciamento de riscos. De acordo com o último relatório *State of AI in the Enterprise* da Deloitte, dos usuários experientes de IA entrevistados estão buscando esse alinhamento.¹⁴

Atingindo Resultados com Menor Risco

A IA pode criar eficiência e benefícios significativos para uma organização. As organizações estão usando a IA para monitorar componentes de uma variedade de processos de fabricação. Por exemplo, um fabricante pode usar a IA para prever quando as esteiras transportadoras provavelmente falharão. Em vez de usar dados atuais de falhas das correias, o fabricante pode usar a IA por meio de medições de temperatura, feeds de câmeras de vídeo e outras variáveis inovadoras para identificar pontos de falha, criando assim novos dados para construir um modelo útil. Os dados de treinamento e os dados dentro do próprio modelo podem ajudar a aumentar a eficiência dos processos do fabricante. Este exemplo de IA ilustra um caso de uso de IA relativamente comum com risco reduzido.

¹² Deloitte, 2020 *Chief Strategy Officer Survey*, a Monitor Deloitte and Kellogg School of Management study.

¹³ Committee of Sponsoring Organizations of the Treadway Commission (COSO), "Risk Appetite – Critical to Success: Using Risk Appetite to Thrive in a Changing World," maio de 2020. <https://www.coso.org/Documents/COSO-Guidance-Risk-Appetite-Critical-to-Success.pdf>.

¹⁴ Ibid., Figura 2 na página 7.

A interação entre o desenvolvimento da estratégia e o apetite a risco é uma entrada fundamental para a avaliação de riscos de uma organização. Informada por sua avaliação de riscos, uma organização determina suas respostas aos riscos identificados. A resposta de uma organização deve incluir a criação de atividades de controle (p. ex., inventário, benchmarking e análise de tendências) que gerenciem os riscos identificados. Após a implantação, é importante mensurar os resultados, para determinar se os objetivos de negócios foram alcançados com menor risco. Apenas cerca de 34% dos usuários de IA entrevistados mantêm um inventário formal de todas as implantações da IA.¹⁵ Sem manter esse inventário, é difícil monitorar e avaliar a exposição potencial de casos de uso de IA.

Partir de uma definição da tolerância a risco, que é uma das principais partes do Framework de ERM do COSO, ajuda a estabelecer os principais indicadores de desempenho e risco quanto à IA, para monitorar o desempenho dos algoritmos ao longo do tempo. A configuração de principais indicadores de desempenho e riscos e de níveis de tolerância enquanto o algoritmo está sendo desenvolvido ajuda a criar uma linha de base de desempenho para articular a confiança. Reportar tais métricas traz transparência entre os stakeholders, o que pode ajudar a melhorar o desempenho do algoritmo e a integridade dos dados de entrada subjacentes.

¹⁵ Ibid., com base na média da Figura 9 na página 15.

Pontos a Considerar

- A organização usa técnicas estratégicas de avaliação de riscos, como planejamento de cenário e teste de premissas para programas de IA?
- Os recursos de IA são usados para identificar riscos emergentes e buscar feedback dos stakeholders sobre produtos, serviços e a marca em si?
- As iniciativas de IA oferecem suporte à análise de riscos para monitorar os riscos?
- As avaliações de riscos de IA consideram os *riscos* e as *recompensas* associados a cada caso de uso de IA, e consideram essas compensações nas decisões sobre ir em frente ou não, bem como na criação e na finalidade dos modelos de IA relevantes?





Identificar, avaliar e responder a riscos são atividades principais que as organizações devem realizar para apoiar a execução da estratégia e o atingimento dos objetivos de negócios da organização. O risco, especialmente o risco relacionado à IA, emana de várias fontes, e as organizações precisam adotar uma variedade de respostas em toda a organização e em todos os níveis. O componente Desempenho e os seguintes princípios do Framework de ERM do COSO servem como base para esta seção do documento:

- 10 **Identifica Riscos**
- 11 **Avalia Gravidade dos Riscos**
- 12 **Prioriza os Riscos**
- 13 **Implanta Respostas aos Riscos**
- 14 **Desenvolve uma Visão de Portfólio**

As organizações não devem implantar aplicações da IA sem abordar sua confiabilidade. Para liberar todo o valor potencial, os modelos de IA devem ser criados com uma IA confiável em mente e devem incluir considerações de desempenho que ajudem a tornar a IA robusta, confiável, segura e protegida, mantendo a privacidade.

Nem todos os modelos de IA têm o mesmo perfil de risco. As organizações precisarão realizar avaliações de riscos para solidificar cada caso de negócios. A identificação de riscos relacionados às iniciativas de IA também é necessária, para avaliar a exposição e identificar oportunidades para uma maior adoção da IA para criação de valor. As organizações também precisam priorizar os riscos, avaliando os modelos de IA e determinando o nível de precisão, confiabilidade e transparência necessários para os casos de uso relacionados. Modelos de IA que exijam um alto nível de precisão, confiabilidade ou transparência para alcançar o sucesso provavelmente têm um perfil de risco mais alto. Além disso, um modelo de IA que esteja sendo usado para fornecer uma sugestão para uma decisão de baixo impacto (p. ex., qual música tocar em seguida) terá um perfil de risco menor do que um modelo de IA que esteja sendo usado para automatizar decisões anteriormente tomadas por humanos (p. ex., decidir sobre os termos de subscrição de uma apólice de seguro).

As organizações devem considerar a gravidade e a prioridade do risco, bem como o contexto de negócios, objetivos de negócios e metas de desempenho do modelo de IA ao selecionar e implantar uma resposta a risco. As respostas a risco relacionadas a modelos de IA geralmente se enquadram nas seguintes categorias:

- **Aceitar:** Nenhuma ação é tomada para alterar a gravidade do risco. Essa resposta é apropriada quando o risco à estratégia e aos objetivos de negócios já está dentro do apetite a risco. O risco que está fora do apetite a risco da organização e que a gestão procura aceitar geralmente exigirá a aprovação do conselho ou de outros órgãos de supervisão.
- **Evitar:** Ações são tomadas para remover o risco, o que pode significar não usar o modelo de IA, limitar o escopo de uso do modelo de IA ou modificar a funcionalidade do modelo de IA para limitar a complexidade.
- **Buscar:** São tomadas ações que aceitam o aumento do risco para conseguir melhor desempenho. Isso pode envolver a expansão do escopo de uso dos modelos de IA, ou a modificação da funcionalidade do modelo de IA para aumentar a complexidade. Ao optar por buscar o risco, a gestão entende a natureza e a extensão de quaisquer mudanças necessárias para atingir o desempenho desejado, sem exceder os limites de tolerância aceitável a risco.
- **Reduzir:** Ações são tomadas para reduzir a gravidade do risco. Isso envolve estabelecer processos e controles de negócios que reduzam o risco residual a um nível aceitável, alinhado com o perfil de risco e o apetite da organização. (As ações que as organizações podem tomar para reduzir o risco associado aos modelos de IA estão descritas a seguir.)
- **Compartilhar:** Ações são tomadas para reduzir a gravidade do risco, transferindo ou compartilhando parte do risco. Um exemplo comum é a terceirização do desenvolvimento, implantação ou monitoramento de modelos de IA para prestadores de serviços especializados.

Embora não seja possível evitar completamente o risco de IA, existem ações que as organizações podem tomar para reduzir o risco. Uma é desenvolver um regime de teste para soluções de IA desenvolvidas ou implantadas, e aplicar o regime de teste em todo o ciclo de vida das soluções de IA. Aproximadamente 40% dos usuários de IA entrevistados atualmente realizam auditorias internas e testes de suas implantações de IA.¹⁶

A Inteligência Artificial Às Vezes Tem Consequências Não Intencionais

O desempenho dos algoritmos deve ser avaliado de forma abrangente quanto à imparcialidade, transparência e robustez. Eles têm o potencial de desviar da intenção estratégica original, conforme ingerem mais dados.

- **Justiça e imparcialidade:** Há viés quanto a determinados grupos, tratamento diferenciado justificado de grupos ou uma representação justa de populações relevantes?
- **Transparência e explicabilidade:** Quais são os principais contribuintes que influenciam a saída do modelo e como cada fator de entrada influencia o resultado?
- **Robustez e confiabilidade:** O modelo permanecerá estável no futuro e generalizará bem para dados não vistos, ou há o risco de viés futuro, conforme o modelo recebe novos dados?

As principais ações na avaliação do desempenho dos modelos de IA incluem, mas não se limitam a:

- **Revisão de riscos** ajuda a identificar fatores de risco, incluindo cibersegurança, riscos de dados, preconceito e ética, que podem impedir ou prejudicar os objetivos da implantação bem-sucedida da IA. Uma visão de portfólio dos riscos associados a todos os projetos de IA deve ser revisada com a alta administração e com o conselho. Um aspecto principal desta revisão é a implantação de respostas a risco, em que cada resposta e o nível residual de risco devem ser cuidadosamente avaliados em relação à definição de apetite a risco.
- **Revisão de dados** ajuda a avaliar a qualidade e a integridade dos dados e seu impacto sobre os modelos de IA e seus resultados. A revisão de dados também ajuda a identificar correlações entre variáveis. Por exemplo, a idade e/ou o índice de massa corporal se correlacionam com o câncer? As organizações podem realizar análises multivariadas de dados subjacentes, para identificar fontes históricas de viés que possam ser usadas como entrada para os algoritmos.

- **Revisão do Modelo** testa os resultados usando as seguintes ações:
 1. Análise da forma funcional e dos parâmetros do algoritmo, para entender possíveis problemas no processo de tomada de decisão.
 2. Avaliação do desempenho do algoritmo com dados reais, para testar vieses ocultos resultantes de correlações complexas ou outras fontes inesperadas de erro do mundo real. A correlação é importante, porque ajuda a identificar a presença de uma associação entre uma variável protegida (p. ex., gênero, raça, etc.) e variáveis que possam servir como proxies potenciais para uma variável protegida usada no modelo. Se tal relação existir, o modelo pode conter viés. A significância estatística indica que a relação entre essas variáveis não é causada por acaso.
- **Revisão de implantação** ajuda a garantir que um algoritmo de IA esteja funcionando corretamente. Essa revisão ajuda a avaliar se o algoritmo continuará robusto, eficaz e justo no futuro e identifica riscos potenciais.
- **Revisão pós-implantação** examina os algoritmos repetidamente. É necessário avaliar periodicamente o desempenho e a imparcialidade do modelo após a implantação. Essa avaliação provavelmente requer um mecanismo de monitoramento que teste continuamente os dados subjacentes e a funcionalidade do modelo.

Complicações Quando Modelos de IA Funcionam Fora do Ambiente de Teste

Testar o desempenho e resultados de um modelo de IA inclui considerar dados/comportamentos inesperados ou alterações nos dados, para avaliar a confiabilidade dos resultados do modelo de IA. Dependendo de como o modelo de IA é criado, a introdução de dados/comportamentos inesperados ou alterações nos dados podem fazer com que o modelo de IA produza saídas incorretas/prejudiciais ou não funcione mais. Pode haver consequências significativas para organizações que implantem modelos de IA que não sejam robustos e confiáveis.

¹⁶ Ibid., based on average from Figure 9 on page 15.

Tenha em mente que programas de IA podem ser hackeados como qualquer outra fonte de dados ou empresa. A Deloitte descobriu que 62% dos entrevistados têm preocupações significativas sobre vulnerabilidades de cibersegurança, mas apenas 39% estão abordando esses riscos.¹⁷ Para manter aplicações de IA e dados relacionados seguros e protegidos — um pilar confiável da IA —, as organizações devem implantar e manter uma metodologia de controle de versão de modelo, incluindo a manutenção de uma versão de linha de base do modelo de IA, e o rastreamento de cada versão subsequente e as alterações feitas nas mesmas, para permitir auditabilidade, transparência e reprodutibilidade do modelo de IA. A metodologia de controle de versão de dados fornece a base. As organizações devem estabelecer controles incrementais preventivos, de detecção e monitoramento em torno do modelo, bem como os dados usados para treinar os algoritmos subjacentes no modelo, para evitar e detectar alterações não autorizadas ou maliciosas. Devido ao poder de computação necessário para aplicar muitos desses modelos, o processamento ocorre na nuvem, o que também introduz preocupações de confiabilidade e privacidade de terceiros.

Além disso, são necessárias políticas que abordem a retenção segura de dados pessoais (criptografia, anonimização, etc.), descarte de dados e comunicação do que é obtido, como é usado e como é mantido. A Deloitte descobriu que 57% dos entrevistados têm preocupações significativas sobre as consequências do uso de dados pessoais sem consentimento, mas apenas 37% estão abordando esses riscos.¹⁸ A privacidade é um pilar importante para alcançar a IA confiável.

Devem ser estabelecidas regras sobre quando uma revisão adicional é necessária. As organizações devem definir deficiências, métricas de desempenho e limites que exijam investigação adicional ou revisão escalonada. Além do Desempenho, essas regras apoiam os componentes de Governança e Cultura, e Definição de Estratégia e de Objetivos do Framework de ERM do COSO. As principais

¹⁷ Ibid., Figura 8, página 14.

¹⁸ Ibid., Figura 8, página 14.

entradas incluem, mas não estão limitadas aos seguintes itens:

- A definição de sucesso da organização (não apenas financeiro ou operacional) para iniciativas de IA e modelos de IA relacionados
- Riscos identificados ao atingimento desse sucesso
- Controles criados e implantados para gerenciar esses riscos

Como parte da responsabilidade e responsabilização, um dos pilares da IA confiável, as organizações precisam definir e executar processos para monitorar o sucesso contínuo. Também devem definir e executar a correção, quando o sucesso não for alcançado. Pessoas precisam ser especificamente responsáveis por essas atividades. Para ajudar os responsáveis, a arquitetura necessária para dar suporte ao monitoramento e ao escalonamento pode ser incorporada à plataforma de IA. A automação pode ajudar a facilitar o monitoramento e encaminhar as revisões para as pessoas designadas em tempo real.

Pontos a Considerar

- As revisões de desempenho do modelo de IA incluem avaliação e gerenciamento de riscos para melhorar os resultados?
- Os principais indicadores de riscos e desempenho para aplicações de IA são monitorados por meio de painéis executivos e reportados a usuários de dados autorizados?
- Qual é o grau de confiança de que a aplicação de IA e controles relacionados estão operando conforme o esperado, e gerando as informações corretas para a tomada de decisões?







ANÁLISE E REVISÃO

Em um ambiente de negócios em constante mudança, a estratégia ou objetivos de negócios de uma organização e as práticas e recursos de ERM podem mudar ao longo do tempo. Em específico quanto à IA, as mudanças contínuas nos recursos e a expansão dos usos exigem que a organização avalie continuamente suas práticas e recursos de ERM e os revise, se necessário. O componente Análise e Revisão e os seguintes princípios do Framework de ERM do COSO servem como base para esta seção do documento:

- 15 **Avalia Mudanças Substanciais**
- 16 **Revisa os Riscos e o Desempenho**
- 17 **Busca Melhoria no Gerenciamento de Riscos Corporativos**

Conforme mencionado anteriormente neste documento, as organizações estão adotando cada vez mais a IA e antecipando a transformação organizacional e de toda a indústria a partir de seus investimentos em IA. Além disso, agências reguladoras e governos promulgaram e estão deliberando sobre regulamentações adicionais relacionadas ao uso de IA e dados relacionados. Esses acontecimentos podem levar a mudanças substanciais, incluindo a funcionalidade dos modelos de IA, que podem resultar em riscos novos ou alterados. Esse acontecimento também pode afetar o ERM, bem como a execução da estratégia e o atingimento dos objetivos de negócios. Um processo iterativo que pode afetar vários componentes do ERM envolve identificar mudanças substanciais e seus efeitos, e responder a essas mudanças.

A revisão das práticas e recursos de ERM, juntamente com o desempenho da organização em relação às suas metas, ajuda a permitir que as organizações monitorem como suas aplicações de IA aumentam o valor e continuarão a impulsioná-lo. A gestão precisa testar e monitorar aplicações de IA e aprendizado de máquina, para ajudar a garantir que funcionem conforme o esperado. O monitoramento contínuo do desempenho e dos riscos ajuda a avaliar se a IA está cumprindo com seus objetivos pretendidos e estabelece um ciclo de tomada de decisão informada sobre riscos.

Uma taxonomia de riscos focada no modelo de IA e na iniciativa relacionada deve ser desenvolvida, para abordar o universo de riscos de IA. As equipes de gerenciamento de riscos devem ajudar a desenvolver a taxonomia que orientará os esforços de identificação e avaliação de riscos. As organizações podem usar o Framework de ERM do COSO e outras orientações, para ajudar a identificar, avaliar, priorizar e monitorar os riscos relacionados à IA. Avaliar o atingimento dos objetivos do modelo de IA demonstra o valor do gerenciamento de riscos e destaca as oportunidades de melhoria.

É importante manter os principais indicadores de desempenho e de riscos a longo prazo, porque os algoritmos mudam conforme aprendem e podem produzir consequências não intencionais no futuro. Além disso, mesmo os algoritmos mais bem intencionados estão sujeitos a vieses ou problemas relacionados à confiabilidade. A simples omissão de informações de identificação pessoal (PII), como raça e sexo, pode não ser suficiente. O monitoramento e teste contínuos de algoritmos são necessários, especialmente porque os dados usados por algoritmos e tendências dentro dos dados mudam ao longo do tempo.

O modelo das três linhas pode ser usado, de acordo com o qual cada stakeholder pode desempenhar um papel na análise e revisão das aplicações de IA e de seu desempenho. A primeira linha, guiada pelo ERM, pode identificar e abordar proativamente os fatores de riscos da IA, enquanto o ERM (a segunda linha) pode colaborar com a primeira linha e tornar as avaliações de riscos eficazes, dinâmicas e aplicáveis. A ERM também pode colaborar com os stakeholders da primeira linha, para apresentar relatórios de riscos e recomendações perspicazes à liderança. A auditoria interna, usando uma abordagem baseada em riscos, pode desempenhar um papel de revisor independente e avaliar criticamente as aplicações de IA quanto ao desempenho de negócios e às metas de gerenciamento de riscos.

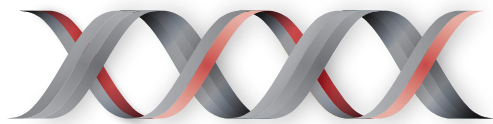
Problemas Sérios Podem Surgir Quando o Desempenho Não É Revisado e Monitorado

Por exemplo, os modelos de IA estão sendo cada vez mais usados na área da saúde, para auxiliar no diagnóstico de doenças e oferecer aconselhamento médico. Se organizações ou profissionais médicos não monitorarem devidamente o desempenho desses modelos, eles podem não identificar e corrigir casos em que os modelos de IA forneçam diagnósticos ou conselhos médicos imprecisos. A falha em identificar e corrigir resultados imprecisos pode levar a danos médicos, preocupações do paciente e perguntas sobre o processo de construção dos modelos de IA relacionados.

Pontos a Considerar

- A organização realiza uma revisão de portfólio de todos os programas de IA, para entender as sinergias e os riscos em um nível agregado?
- Um diretor de riscos participa de análises de desempenho de IA, para compartilhar perspectivas de gerenciamento de riscos?
- As descobertas, tanto positivas quanto negativas, são compartilhadas com membros da alta administração e do conselho nessas revisões?
- A alta administração toma as devidas medidas corretivas para abordar quaisquer descobertas negativas?
- Você tem uma equipe multidisciplinar de gerenciamento de riscos que possa ajudar no planejamento da mitigação do risco do modelo de IA?





INFORMAÇÃO, COMUNICAÇÃO E REPORTE

As organizações são continuamente desafiadas a usar a enorme quantidade de dados gerados, juntamente com as crescentes preocupações com privacidade e segurança de dados e transparência dos modelos de IA relacionados. Nesse ambiente, é importante que as organizações forneçam as informações certas, na forma certa, no nível certo de detalhes, para as pessoas certas, em tempo hábil. O componente de Informação, Comunicação e Reporte e os seguintes princípios do Framework de ERM do COSO servem como base para esta seção do documento:

- 18 Alavanca Informações e Tecnologia**
- 19 Comunica as Informações de Riscos**
- 20 Reporta sobre Riscos, Cultura e Desempenho**

Relatórios sobre riscos, cultura e desempenho usam sistemas de TI para capturar, processar e gerenciar dados e informações. A gestão usa essas informações para informar e apoiar o gerenciamento de riscos, incluindo o gerenciamento de riscos relacionados a modelos de IA. Um processo de reporte é necessário para informar os stakeholders internos e externos sobre o desempenho, benefícios e riscos potenciais dos modelos de IA. O processo de reporte também considera como, quando e com que frequência os stakeholders receberão as informações. Ao desenvolver a resiliência de uma organização, é necessário entender o cenário de riscos e um relatório unificado de

riscos de IA deve ser compilado para a gestão executiva e os membros do conselho, para auxiliar em seus esforços de supervisão. Este relatório pode incluir atualizações sobre as principais métricas de desempenho e indicadores de riscos para o desempenho dos modelos de IA da organização, bem como os resultados dos principais processos de supervisão e monitoramento. A comunicação tempestiva dos resultados, incluindo descobertas inesperadas, é vital para a identificação e resolução de problemas, antes que se transformem em problemas maiores.

Para evitar crises, gerenciar problemas e se preparar para os piores cenários que podem surgir a partir do desempenho indesejado ou de incidentes relacionados a iniciativas de IA, um framework e protocolos de resposta às comunicações de crise devem funcionar como guia. (Veja a Figura 7) Esse manual de comunicação de crise explicará como a organização deve responder para controlar o impacto e a exposição de quaisquer incidentes, enquanto mantém os negócios funcionando. Também deve incluir etapas para ajudar na recuperação.

Dados sobre as reações dos stakeholders são um componente importante para se reconstruir e emergir mais forte após uma crise. Essas respostas ajudarão a informar a estratégia e a implantação da IA, e ajudarão a organização a atender às expectativas de transparência.

Figura 7. Desenvolvendo Resiliência



Uso de IA sob os Holofotes

A IA está se tornando, cada vez mais, uma grande parte das operações de negócios da organização. Por conta do crescente interesse dos investidores no uso de IA, várias grandes empresas de base tecnológica incluíram divulgações em seus reportes 10-K que descrevem como os modelos de IA atualmente impactam as operações de negócios e seu potencial impacto no futuro.

Pontos a Considerar

- Existe um plano implantado de resposta a crises?
- Qual reporte do desempenho do programa de IA é divulgados aos stakeholders e ao público?
- Os executivos e órgãos de supervisão dentro da organização recebem informações relevantes de desempenho sobre os programas de IA?



OBSERVAÇÕES DE CONCLUSÃO

Para perceber o valor da IA e alavancar seu potencial, as organizações devem alinhar o gerenciamento de riscos com sua estratégia e sua execução das iniciativas de IA. O Framework de ERM do COSO pode ajudar as organizações a desenvolver a governança integrada da IA, gerenciar riscos e impulsionar o desempenho para atingir metas estratégicas. Ao implantar a governança integrada da IA, as organizações podem ter melhores informações sobre os riscos relevantes. Isso pode apoiar um aumento na gama de oportunidades e na flexibilidade para assumir riscos estratégicos calculados e tornar-se mais ágil e adaptável no planejamento e execução de suas iniciativas de IA. Embora não oficial, o Deloitte Trustworthy AI™ Framework pode ajudar as organizações a pensar nos riscos ao implantar o Framework de ERM do COSO para a IA.

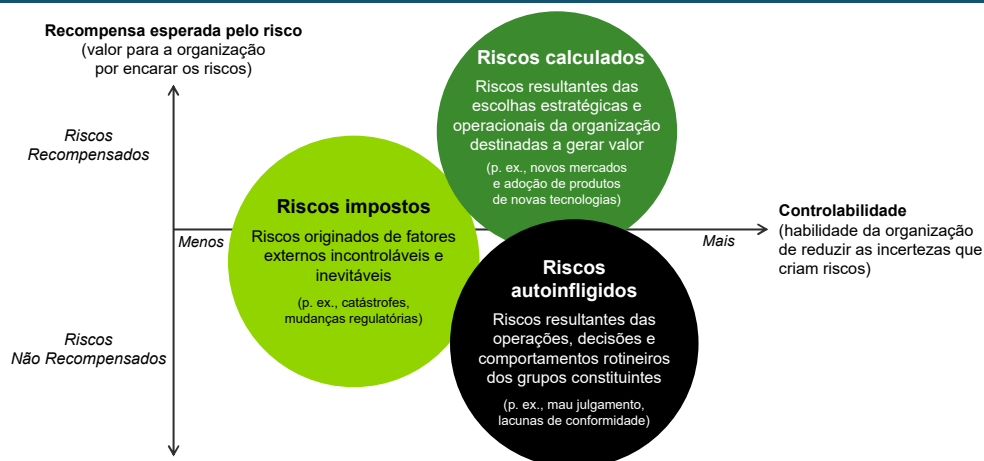
Figura 8. **COSO Enterprise Risk Management - Integrating with Strategy and Performance Framework**



Por meio do ERM informado pelo Framework de ERM do COSO, as organizações podem reduzir a variabilidade do desempenho e melhorar a probabilidade de sucesso de suas iniciativas de IA. Ao identificar sinais para corrigir o curso antecipadamente, as organizações podem aumentar os resultados positivos, reduzir surpresas negativas e melhorar a resiliência ao risco. A alocação de recursos informada sobre o risco também pode ser melhorada e, ao compreender seu risco, a organização pode estar mais bem equipada para entregar retorno sobre o investimento e atender às expectativas dos stakeholders. Além disso, ao implantar o ERM, as organizações podem refinar e adaptar suas iniciativas de inovação, para apoiar suas estratégias em um ambiente de negócios em rápida mudança.

O gerenciamento de riscos implantado devidamente pode ajudar as organizações a tirar vantagem de riscos calculados com altas recompensas, gerenciar os riscos inerentes e ajudar a diminuir significativamente os riscos autoinfligidos. (Ver Figura 9).

Figura 9. **Um programa de ERM ajuda organizações a alcançar o sucesso quanto às suas iniciativas de IA**



Copyright © 2020 Deloitte Development LLC. Todos os direitos reservados.

As soluções de IA precisam ser confiáveis, testadas e verdadeiras. Confiáveis – porque o ERM é transparente por natureza e ajuda a manter uma organização a par de seus riscos e oportunidades. Testadas – no sentido de que os modelos sejam continuamente testados e verificados, para verificar se estão funcionando conforme o esperado. E verdadeiras – os regimes de governança, gerenciamento de riscos, testes e monitoramento ajudam os modelos a operar de maneira que reflitam os valores da organização e protejam sua reputação. O Framework de ERM do COSO, quando devidamente considerado, pode resultar em uma IA confiável, testada e verdadeira.

Chamada à Ação: 5 Próximos Passos a Considerar com Base no Framework de ERM do COSO

Use o framework do COSO e seus componentes e princípios para estabelecer um programa de IA confiável. Veja como começar:

- 1. Estabeleça a estrutura de governança para o programa de IA.** Determine quando e como a organização usará a IA e defina o propósito e os objetivos das iniciativas de IA propostas. Isso inclui avaliar as considerações éticas aplicáveis. Traga várias iniciativas de IA para toda a sua organização sob um programa geral de IA e uma estrutura de governança que forneça visibilidade à alta administração e ao conselho. Identifique um executivo sênior para liderar seu programa de IA e fornecer supervisão de riscos e desempenho.
- 2. Monte uma estratégia de risco de IA.** Colabore com os stakeholders para elaborar uma estratégia em toda a organização para gerenciar os riscos estratégicos, técnicos, regulatórios e operacionais da IA. Certifique-se de que sua organização tenha a experiência técnica de IA para executar a estratégia de risco de IA. A estratégia deve definir funções, responsabilidades, controles e procedimentos de mitigação.
- 3. Tome a iniciativa de realizar uma avaliação de riscos de IA.** Para cada modelo de IA que sua organização usa, avalie o impacto potencial de resultados estratégicos abaixo do ideal, falhas operacionais ou vies. Além disso, avalie como o algoritmo gerencia e usa os dados e se introduz algum vies não intencional. Para processos de negócios que se integram à IA, busque vulnerabilidades e veja a probabilidade de ocorrência, depois registre os riscos conhecidos e os controles correspondentes.
- 4. Desenvolva uma visão de portfólio de riscos e oportunidades para iniciativas de IA.** O Chief Risk Officer e o líder de IA podem trabalhar juntos para revisar proativamente os modelos de IA quanto a riscos relacionados a vies, adulteração e mau funcionamento do modelo. Eles devem reportar uma visão de portfólio dos riscos de IA aos executivos seniores e ao conselho, para conscientização e suporte à tomada de decisões.
- 5. Estabeleça uma abordagem para gerenciar os riscos de IA e reporte aos stakeholders para transparência.** Isso inclui avaliar a relação risco-recompensa para iniciativas de IA e alocação de recursos. Considere montar uma equipe de especialistas no risco de modelo de IA, para oferecer as principais práticas, objetividade e metodologias de resposta a riscos. Estabeleça as principais métricas de desempenho e de riscos, para mensurar metas como eficácia, imparcialidade e transparência de cada modelo. Para cada métrica, defina limites que acionariam revisões adicionais do modelo e ações corretivas. Desenvolva painéis de reporte para executivos e conselhos de administração, e divulgue o desempenho da IA e as ações de gerenciamento de riscos aos stakeholders externos para conscientização.

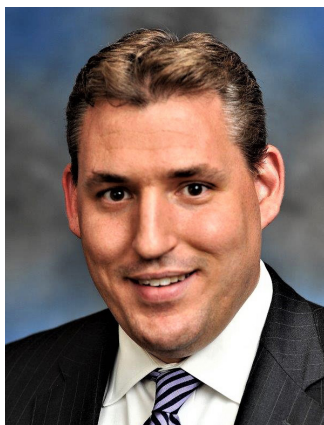
SOBRE OS AUTORES



Keri Calagna, Risk & Financial Advisory Principle, Deloitte & Touche LLP

Keri é líder do departamento de Riscos Cibernéticos e Estratégicos da Deloitte. Com mais de 25 anos de experiência em riscos, a Keri ajuda as organizações a desenvolver sua cultura, recursos e processos para criar programas de riscos integrados que ajudam a expandir os negócios, acelerar o desempenho, melhorar a resiliência e atingir metas estratégicas. Ao longo de sua carreira, ela ajudou empresas a avaliar, gerenciar e monitorar uma ampla gama de riscos, incluindo riscos financeiros, operacionais, reputacionais, regulatórios, empresariais, estratégicos e tecnológicos. Keri frequentemente assessora conselhos e equipes de liderança executiva na criação e implantação de governança, monitoramento e reporte de riscos no nível corporativo e ajuda os executivos a alinhar e mobilizar os principais riscos à sua organização.

Keri tem um Bacharelado e um MBA em Empreendedorismo pela *Lally School of Management & Technology* do *Rensselaer Polytechnic Institute*.



Brian Cassidy, Audit & Assurance Partner, Deloitte & Touche LLP

Brian é o líder de Inteligência Artificial/Algoritmos de Auditoria e Avaliação dos EUA com experiência diversificada na prestação de serviços de auditoria e assessoria para empresas da Fortune 500. Com fortes habilidades técnicas, de gerenciamento de riscos, comunicação e organização, concentra-se na prestação de serviços de auditoria, contabilidade e assessoria para empresas públicas e privadas no setor de serviços financeiros. A experiência de Brian atravessa uma ampla gama de indústrias no setor de serviços financeiros, incluindo bancos (corretores/representantes), empresas de investimento, empresas de desenvolvimento de negócios e fundos alternativos, incluindo private equity, hedge e imobiliário. Também lidera os esforços da Deloitte na área de avaliação de Algoritmos/IA, conforme as tecnologias emergentes continuam a impactar os clientes e o mercado.

Brian é membro do *American Institute of Certified Public Accountants (AICPA)*, do *Pennsylvania Institute of Certified Public Accountants (PICPA)*, e da *New York State Society of CPAs (NYSSCPA)*. Brian é bacharel em contabilidade e em administração de empresas pela *Villanova University*.



Amy Park, Audit & Assurance Partner, Deloitte & Touche LLP

Amy é a líder de Ideação do departamento de Serviços de Assessoria e Transformação em Contabilidade da Deloitte. Nesta função, Amy lidera o desenvolvimento de ideias em potenciais novas ofertas de serviços que podem aumentar o valor que a Deloitte pode trazer para o mercado, incluindo áreas de tecnologias emergentes e avaliação expandida, como IA, algoritmos, blockchain e ativos digitais. Também é partner no Departamento Nacional de Serviços de Contabilidade e Reporte da Deloitte e é especializada em assuntos técnicos de contabilidade, consolidação, instrumentos financeiros e contabilidade de transações de ativos digitais.

Amy é membro do *American Institute of Certified Public Accountants* e atua na Força-Tarefa de Ativos Digitais da AICPA, com foco em questões contábeis relacionadas a ativos digitais. Tem mais de 17 anos de experiência em contabilidade pública, incluindo um *practice fellowship* no *Financial Accounting Standards Board*, e atendeu empresas públicas e privadas nos setores bancário e de valores mobiliários e ativos digitais.

SOBRE O COSO

Originalmente formado em 1985, o COSO é uma iniciativa conjunta de cinco organizações do setor privado e se dedica a ajudar as organizações a melhorar o desempenho, desenvolvendo liderança criativa que aprimora o controle interno, o gerenciamento de riscos, a governança e a dissuasão de fraudes. As organizações de apoio ao COSO são a *American Accounting Association* (AAA), *American Institute of Certified Public Accountants* (AICPA), *Financial Executives International* (FEI), *Institute of Management Accountants* (IMA) e *The Institute of Internal Auditors* (IIA).



The Association of
Accountants and
Financial Professionals
in Business



Esta publicação contém apenas informações gerais e nenhuma parte do COSO, nenhuma de suas organizações constituintes ou nenhum dos autores desta publicação está, por meio desta publicação, prestando assessoria ou serviços contábeis, comerciais, financeiros, de investimento, jurídicos, fiscais ou outros. As informações aqui contidas não substituem tais conselhos ou serviços profissionais, nem devem ser usadas como base para qualquer decisão ou ação que possa afetar seus negócios. As visões, opiniões ou interpretações aqui expressas podem diferir daquelas de reguladores relevantes, organizações autorreguladoras ou outras autoridades, e podem refletir leis, regulamentos ou práticas sujeitas a mudanças ao longo do tempo. A avaliação das informações aqui contidas é de responsabilidade exclusiva do usuário. Antes de tomar qualquer decisão ou ação que possa afetar seus negócios com relação aos assuntos descritos neste documento, você deve consultar assessores profissionais qualificados relevantes. O COSO, suas organizações constituintes e os autores se isentam expressamente de qualquer responsabilidade por qualquer erro, omissão ou imprecisão contida neste documento ou qualquer perda sofrida por qualquer pessoa que confie nesta publicação.

SOBRE A DELOITTE

Esta publicação contém apenas informações gerais e a Deloitte não está, por meio desta publicação, prestando assessoria ou serviços contábeis, comerciais, financeiros, de investimento, jurídicos, fiscais ou outros. Esta publicação não substitui tais assessorias ou serviços profissionais, nem deve ser usada como base para qualquer decisão ou ação que possa afetar seus negócios. Antes de tomar qualquer decisão ou ação que possa afetar seu negócio, você deve consultar um assessor profissional qualificado.

A Deloitte não se responsabiliza por qualquer perda sofrida por qualquer pessoa que confie nesta publicação.

Deloitte refere-se a uma ou mais partes da Deloitte Touche Tohmatsu Limited, uma empresa privada do Reino Unido limitada por garantia (DTTL), sua rede de firmas-membro e suas entidades relacionadas. A DTTL e cada uma de suas firmas-membro são entidades juridicamente separadas e independentes. A DTTL (também referida como “Deloitte Global”) não presta serviços a clientes. Nos Estados Unidos, Deloitte refere-se a uma ou mais firmas-membro da DTTL nos Estados Unidos, suas entidades relacionadas que operam usando o nome “Deloitte” nos Estados Unidos e suas respectivas afiliadas. Certos serviços podem não estar disponíveis para atestar clientes sob as regras e regulamentos da contabilidade pública. Acesse [Deloitte.com/about](https://www.deloitte.com/about) para saber mais sobre nossa rede global de firmas-membro.

Deloitte.

Gerenciamento de Riscos Corporativos



COSO

Committee of Sponsoring Organizations
of the Treadway Commission

coso.org



CONCRETIZE TODO O POTENCIAL DA INTELIGÊNCIA ARTIFICIAL

APLICANDO O FRAMEWORK E OS PRINCÍPIOS DO COSO
PARA AJUDAR A IMPLANTAR E DIMENSIONAR A AI

COSO

Committee of Sponsoring Organizations of the Treadway Commission

coso.org

