

GOVERNANÇA E ÉTICA

Da Inteligência Artificial nas Companhias Abertas

Relatório de Pesquisa

Governança e Ética de Inteligência Artificial nas Companhias Abertas

Outubro, 2024

Coordenação Geral

Dora Kaufman

Alexandre Zavaglia Coelho

Realização



Coordenação Executiva

Rafael Sasso

Parceria Estratégica



Apoiadores



Apoio Institucional



Coordenação Geral

Dora Kaufman
Alexandre Zavaglia Coelho

Coordenação Executiva

Rafael Sasso

Equipe Accenture

Daniel Lázaro
Victor Lima
Olivia Ferreira
Ralph Lagnado
Iberê Borges
Caroline Mutus
Thayana Almeida
Rodrigo Acosta
Luana Nagai
Vitória Oliveira

Design e diagramação

Rafael Garcia Posada

Como citar este relatório

KAUFMAN, D.; ZAVAGLIA, A. C. 2024. Relatório final da Pesquisa Governança e Ética de IA nas Companhias Abertas. Realizada pela Associação Brasileira das Companhias Abertas – ABRASCA, sob a coordenação dos autores.

Apresentação

O avanço da inteligência artificial (IA), particularmente o hype protagonizado pelo ChatGPT e demais soluções de IA generativa, mudou o comportamento do mercado e de seus profissionais em relação à tecnologia, acelerando a percepção de sua relevância estratégica para o futuro das companhias e, conseqüentemente, o ritmo da adoção da inteligência artificial em diferentes tarefas e setores. Junto aos extraordinários benefícios da IA emergem incertezas, riscos e potenciais danos — não apenas aos usuários e clientes, e à sociedade em geral, como igualmente à operação e à reputação das companhias. Como reação, mundo afora estão em curso iniciativas para identificar boas práticas e o uso responsável dessa tecnologia, ao mesmo tempo que avança a agenda para regulamentar o desenvolvimento e o uso da IA, inclusive no Congresso Nacional. Independentemente das obrigações a serem definidas em um futuro Marco Regulatório da Inteligência Artificial no Brasil, é recomendável construir um conjunto sólido de diretrizes de governança de IA.

Diante desse cenário, a Associação Brasileira das Companhias Abertas (ABRASCA) assumiu a missão de realizar uma tomada de subsídios, por meio desta pesquisa, com informações estruturadas e parâmetros atualizados sobre como as companhias de capital aberto no Brasil estão lidando com a inteligência artificial. O objetivo é oferecer às associadas, e ao mercado em geral, diretrizes e caminhos para o desenvolvimento de um Programa de Governança e Gestão de Riscos de IA contemplando a constituição de um Comitê de Ética — ou a expansão das competências dos comitês de ética ou outras estruturas orgânicas já existentes — e de Códigos de Conduta adaptados para essa finalidade.

Como ponto de partida, empreendemos um diagnóstico de como a inteligência artificial está sendo adotada pelas companhias de capital aberto, buscando identificar o grau de maturidade no uso e na gestão de risco dessa tecnologia. Particularmente, o tema da transparência é sensível às companhias de capital aberto, sujeitas às regras pré-estabelecidas nas relações com seus investidores e com os órgãos que controlam e gerenciam o mercado de capitais.

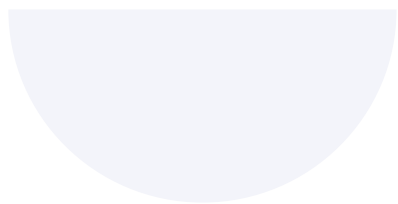
Para coordenar o projeto, convocamos dois especialistas, Dora Kaufman¹ e Alexandre Zavaglia², e contamos com o apoio de cinco empresas líderes em seus respectivos setores: Accenture, Vivo, Itaú, Cielo e Comgás. A Accenture, além de apoiadora máster, foi uma parceira estratégica ao disponibilizar seus profissionais para somar esforços junto à coordenação da pesquisa, formando um time diversificado cognitivamente, com diferentes campos de conhecimento e experiências. Contamos, igualmente, com o apoio institucional do Instituto Brasileiro de Governança Corporativa (IBGC).

Somos gratos aos especialistas à frente da pesquisa, aos apoiadores e à inestimável colaboração das companhias e de seus executivos, que cederam tempo valioso para nos municiar com um quadro fidedigno da relação atual de suas organizações com os modelos e/ou soluções de inteligência artificial e os aspectos de governança, e nos dando feedback sobre as conclusões preliminares.

O relatório final é apenas o início de uma longa jornada. Pretendemos fomentar o progresso do desenvolvimento e do uso da IA no Brasil, buscando identificar e enfrentar os novos desafios. Agradecemos a todos pela leitura e pelas contribuições, que nos permitiram e ainda permitem incorporar novas evidências a este documento.

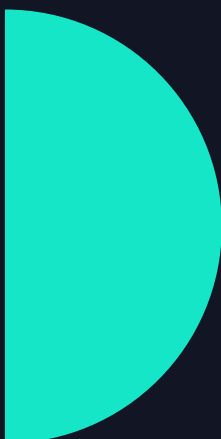
1 Professora do Programa de Pós-graduação em Tecnologias da Inteligência e Design Digital (TIDD) da PUC-SP e doutora em mídias digitais pela ECA-USP, com estágio sanduíche na Université Paris Sorbonne IV, na França. Tem dois pós-doutorados, na COPPE-UFRJ e no TIDD/PUC-SP. É autora de vários livros, entre eles "A inteligência artificial irá suplantará a inteligência humana?" e "Desmistificando a Inteligência Artificial", colunista da revista Época Negócios e colaboradora dos jornais O Globo e Valor Econômico.

2 Doutorando no TIDD/PUC-SP com pesquisa voltada para governança e ética de IA e cursos de direito digital na FGV-SP, Strategic Data Storytelling na Universidade de Chicago, e Management of Technology: Roadmap & Development no MIT, os dois últimos nos Estados Unidos. É professor da pós-graduação e do mestrado (LLM) AI&LAW da Universidade de Lisboa, em Portugal; membro do RECSAI.org.; líder de pesquisas na área de regulação e governança de IA no CEPI-FGV Direito SP; e membro do grupo de trabalho do CNJ sobre regulamentação do uso de IA no Poder Judiciário.



SUMÁRIO

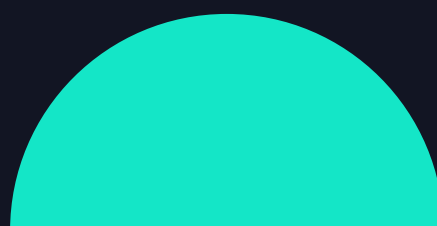
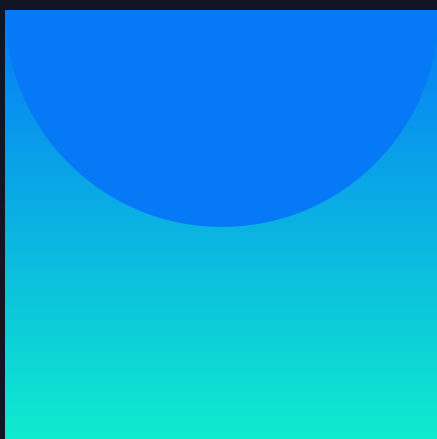
SUMÁRIO EXECUTIVO	7
1. Sobre este relatório	9
2. Destaques do relatório	11
3. Metodologia	15
CONTEXTUALIZAÇÃO	17
II. Contextualização	18
1. Avanços e desafios da IA	21
2. Princípios éticos gerais, regulação e governança de IA	33
2.1. Princípios éticos gerais	33
2.2. Regulação de IA	34
2.3. Governança de IA	35
RELATÓRIO	39
III. Relatório	40
1. Resultados da pesquisa	43
2. Recomendações	64
2.1. Programas de Governança e Gestão de Riscos de IA	65
2.2. Conselho de Administração	71
2.3. Chief Artificial Intelligence Officer (CAIO)	71
2.4. Parcerias estratégicas	72
2.5. Comitê de Ética de IA	72
2.6. Capacitação dos colaboradores	73
CONSIDERAÇÕES FINAIS	74
IV. Considerações finais	75
GLOSSÁRIO	79
V. Glossário	80

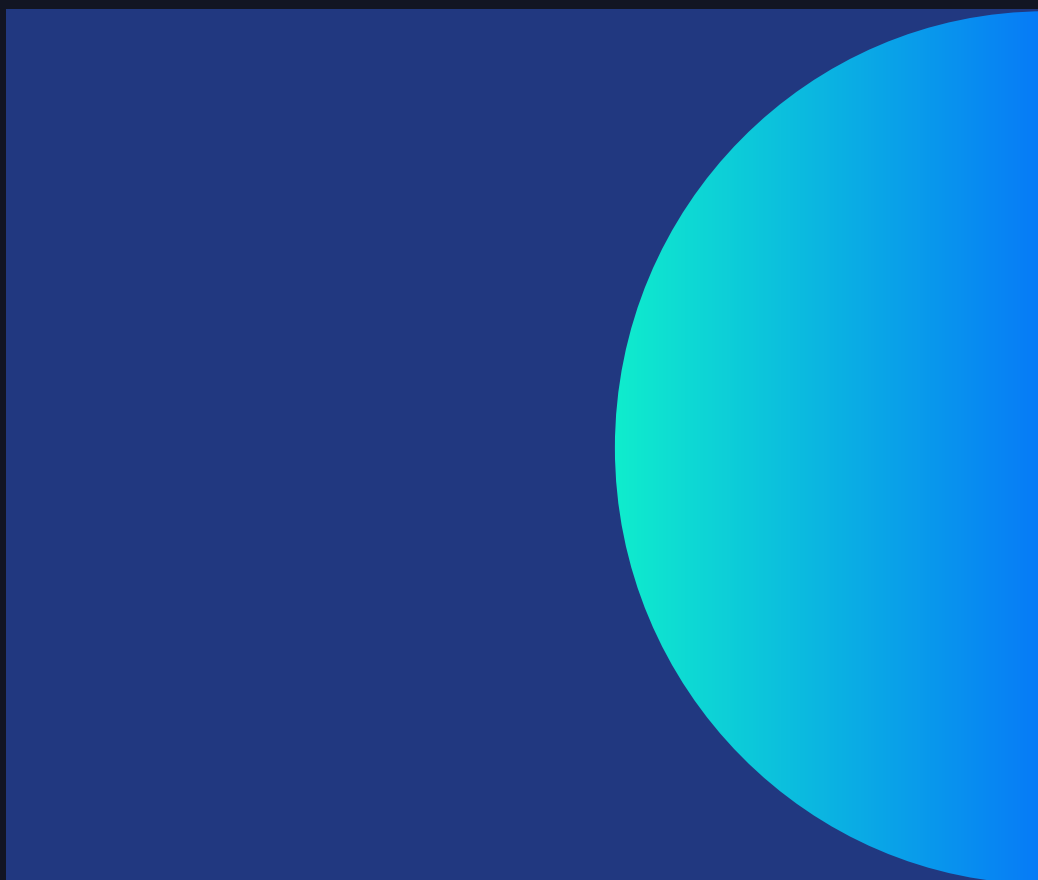


Parte 1

SUMÁRIO EXECUTIVO

Clique para ir até o tópico de interesse





II . Sumário Executivo

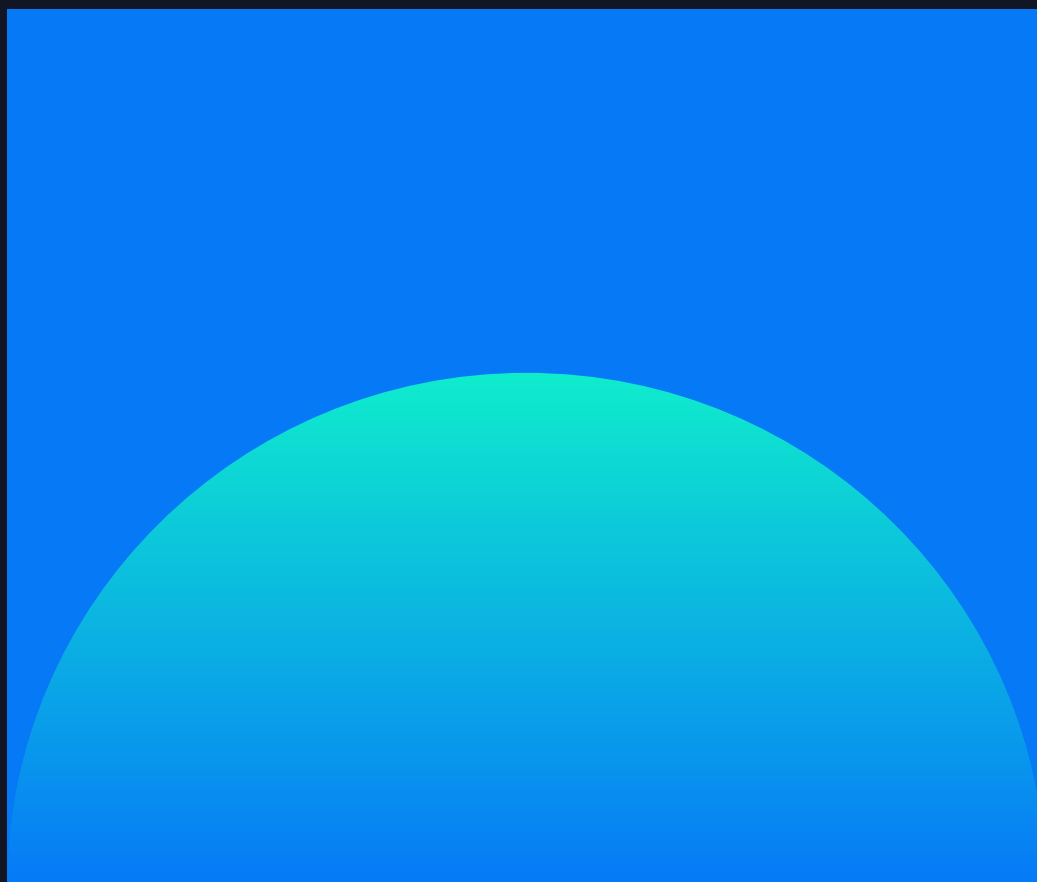
1. SOBRE ESTE RELATÓRIO





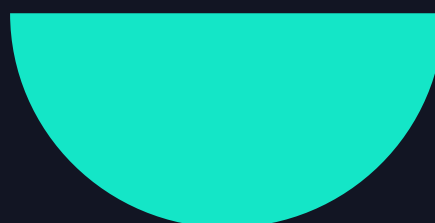
1. Sobre este relatório

- Esta é a publicação, sujeita a revisões, de uma investigação inédita sobre como a inteligência artificial está sendo adotada pelas companhias de capital aberto no Brasil, com foco nos aspectos de governança e gestão de riscos, dividida em duas etapas: qualitativa e quantitativa. Primeiramente, foram realizadas entrevistas qualitativas com companhias associadas à ABRASCA, previamente selecionadas, envolvendo em momentos distintos um membro do conselho, o CEO e o head de tecnologia. Em um segundo momento foram distribuídos questionários on-line para os associados da ABRASCA e do IBGC.
- Em um período de progresso sem precedentes no avanço da IA, a pesquisa não se restringiu a um tipo específico de IA, mas buscou contemplar as diferentes técnicas e soluções desenvolvidas e adotadas pelos associados da ABRASCA e do IBGC.
- Apesar do foco ter sido as companhias de capital aberto, ao agregar percentual significativo do PIB brasileiro, acreditamos que as evidências identificadas na pesquisa possam ser de grande valia para o mercado como um todo.
- O projeto é uma contribuição para melhorar o entendimento coletivo do amplo espectro e dos desafios associados ao desenvolvimento e à adoção da inteligência artificial, cooperando para capacitar os gestores na identificação de oportunidades de negócio e no monitoramento de forma eficaz dos riscos, possibilitando que as companhias experimentem e usufruam os benefícios da IA de forma sustentável.
- Liderados pelos coordenadores, em conjunto com os especialistas que auxiliaram na construção deste relatório, os autores tiveram total poder discricionário sobre o seu conteúdo.



II . Sumário Executivo

**2. DESTAQUES
DO RELATÓRIO**





2. Destaques do relatório

- A adoção de sistemas de IA pela maioria das companhias de capital aberto do Brasil é recente, sendo que poucas utilizaram a inteligência artificial em suas operações na última década — as que o fizeram, geralmente, foi com foco em resolver problemas pontuais ou de negócio específicos.
- Especialmente após o advento dos LLMs (large language models), em particular o ChatGPT, que teve significativo efeito cultural, observa-se um quase consenso entre os conselhos e os C-Level sobre o caráter estratégico da IA para impulsionar o desenvolvimento de produtos e serviços, a experiência do cliente (UX), a expansão para novos mercados e negócios, e a sustentabilidade da própria companhia.
- Não identificamos nas companhias participantes da pesquisa, em ambas as etapas, estratégia que ligue de forma consistente e estruturada a adoção da IA ao planejamento estratégico, ou que demonstre o envolvimento direto da alta gestão e dos conselhos nesta agenda. Em geral, a adoção de modelos de inteligência artificial tem sido pontual, para equacionar desafios específicos, em sua maioria por iniciativas de profissionais das áreas de TI e, em menor escala, por iniciativa de profissionais das áreas de negócio.
- A IA generativa é a categoria com maior potencial de rápida expansão, oferecendo contribuições significativas na geração de valor e, conseqüentemente, para atingir os benefícios desejados e estabelecer um novo patamar de competitividade. Esses modelos têm maior potencial de impacto na experiência do cliente (UX), na otimização dos processos de backoffice, e no entendimento de padrões a partir de grandes bases de dados para aprimorar ou criar novos produtos e serviços.
- Predomina a adoção de modelos ajustados para um caso de uso específico, a fim de resolver um problema de negócio de uma área ou tipo de atividade — o ajuste fino desses modelos requer menos dados, com custo e tempo de conclusão menores. A IA generativa é percebida como um ponto de inflexão com potencial de servir de interface de linguagem natural, integrada com



sistemas que se utilizam de outras técnicas, para automatizar tarefas rotineiras, possibilitando migrar de projetos-piloto periféricos para o core do modelo de negócio da companhia.

- Os sistemas de inteligência artificial ofertados no mercado são cada vez mais intensivos em dados, logo demandam grande poder computacional. Como resultado, as companhias estão migrando para plataformas em nuvem, permitindo consolidar os dados em um único ambiente e gerenciar seus gastos de forma mais eficiente (“as a service”) e segura (cybersecurity), escalando sua operação.
- A maior parte das companhias ainda se encontra na fase de estruturação de dados e formação de data lakes que sejam capazes de servir como fontes de treinamento e produção de modelos de IA adaptados às suas especificidades.
- Predomina a adoção de soluções de inteligência artificial ofertadas no mercado, acessadas por meio de APIs em plataformas de tecnologia ou por contrato direto com startups e empresas de tecnologia. Um número reduzido de companhias desenvolve soluções customizadas, geralmente compostas por algoritmos de IA ofertados por plataformas de tecnologia, acessados por meio de APIs, ou disponíveis em comunidades de código aberto.
- Em ambos os casos, as companhias supõem que os sistemas e/ou algoritmos de inteligência artificial foram previamente validados em relação aos riscos pelas plataformas de tecnologia e/ou pelas comunidades de código aberto, pressupostos não comprovados por evidências.
- Não foram identificadas diretrizes formalizadas de governança de IA, apenas iniciativas de formação de grupos de trabalho ou comissões para liderar o processo. Nas companhias que utilizavam sistemas de inteligência artificial em período anterior ao atual hype identificamos procedimentos a serem observados pelos desenvolvedores e usuários internos, especialmente associados à proteção de dados e à segurança da informação; e em um pequeno grupo de companhias observamos iniciativas de governança em áreas ou projetos específicos, e as primeiras políticas com diretrizes para o uso de LLMs acessados diretamente pelos colaboradores. Em todos os casos, distantes



da abrangência e da precisão de um Programa de Governança e Gestão de Riscos de IA ou da adoção de um framework ou protocolo de governança de inteligência artificial para toda a companhia.

- Consolida-se a percepção de que os riscos no desenvolvimento e na adoção de soluções de IA não apenas podem provocar efeitos danosos sobre as pessoas afetadas e para toda a sociedade, como riscos significativos de reputação, regulatórios e éticos às companhias, o que tem influenciado as primeiras iniciativas para estruturar a governança adequada para essa finalidade. Os potenciais riscos são sempre significativos e variados, porque os sistemas de inteligência artificial operam em larga escala, além da complexidade e da transversalidade da tecnologia.
- Observa-se uma assimetria de conhecimento sobre a inteligência artificial entre os gestores, tomadores de decisão, e os profissionais de TI especializados em IA. Foram identificados raros esforços de capacitação dos colaboradores das áreas técnicas, principalmente nas questões relacionadas a governança e ética. Praticamente não existem programas estruturados de capacitação para os colaboradores das demais áreas das companhias.



II . Sumário Executivo

3. METODOLOGIA





3. Metodologia

- Entrevistas qualitativas com companhias de capital aberto associadas à ABRASCA, previamente selecionadas entre as que efetivamente utilizam sistemas de inteligência artificial. Foram entrevistados, em momentos distintos, sempre por um grupo de especialistas, um membro do Conselho de Administração, o CEO ou membro da alta gestão, e o responsável pela área de tecnologia (TI). Em alguns casos foram entrevistados os responsáveis pelo departamento jurídico e/ou pelo departamento de compliance. As entrevistas foram gravadas, com autorização prévia dos entrevistados; transcritas por um sistema de IA; agregadas e analisadas, respeitando o sigilo e a confidencialidade das informações.
- Questionário on-line para os associados da ABRASCA e do IBGC, em torno dos seguintes tópicos:
 - a. Perfil, setor e faixa de receita da companhia;
 - b. Se utilizam soluções de inteligência artificial, em que setores ou áreas, e se o desenvolvimento é in house, em parceria ou ofertado por terceiros;
 - c. Se a IA é pauta do Conselho de Administração, se é percebida como estratégica para a sustentabilidade do negócio, e se está inserida no planejamento estratégico da companhia;
 - d. Indicadores mais relevantes para avaliar os desafios e o sucesso da implementação das soluções de inteligência artificial;
 - e. Ações relacionadas à transparência e à confiança na adoção de IA na interação com os stakeholders;
 - f. Iniciativas ou programas de gestão de riscos de IA;
 - g. Programas de transformação cultural e capacitação em IA.

Os resultados da pesquisa quantitativa (questionários on-line) foram compilados automaticamente, sob supervisão dos especialistas, e as respostas foram agregadas por temática.



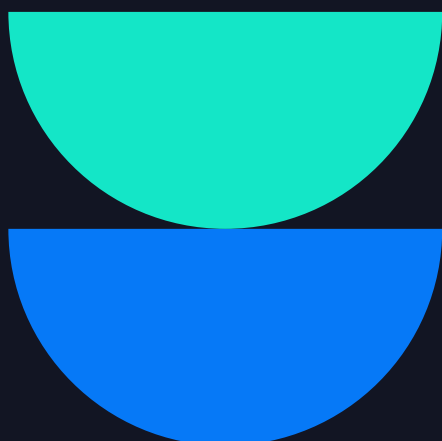
Observamos um descompasso entre os resultados apurados nas pesquisas qualitativa e quantitativa em todos os temas e níveis, que será abordado no capítulo “Resultados da pesquisa”.



Parte 2

CONTEXTUALIZAÇÃO

Clique para ir até o tópico de interesse





II. Contextualização

A pesquisa confirmou a hipótese inicial de que estamos em um momento de transição — denominado por Ajaw Agrawal, Joshua Gans e Avi Goldfarb de “*Between Times*”³ —, após o testemunho do poder da inteligência artificial e anterior à sua adoção generalizada, predominando casos de uso pontuais. Em parte, substituindo os modelos preditivos tradicionais por modelos preditivos habilitados por IA, com foco na automação e na melhoria de processos existentes e, conseqüentemente, com impacto limitado na produtividade. Ainda são poucas as companhias que implementaram a inteligência artificial para redesenhar produto ou serviço — ou, mais radicalmente, mudar seu modelo de negócio —, aproveitando plenamente os benefícios, inclusive justificando o relativo alto custo de desenvolvimento e/ou de adoção da tecnologia.

Previsões com relativa assertividade, ou altas taxas de acurácia, desempenham papel crítico no aprimoramento da tomada de decisão. A técnica de redes neurais profundas (deep learning), modelo estatístico de probabilidade capaz de lidar com grandes conjuntos de dados (big data), tem o potencial de mudar o processo de tomada de decisão, desde que ajustado aos procedimentos e à cultura da companhia, garantindo que as pessoas certas tomem as decisões no momento certo.

Com o pressuposto de que qualquer previsão visa reduzir a incerteza, pode-se considerar que uma previsão é valiosa se melhorar a decisão existente ou se promover novas decisões, ou seja, quando gera uma tomada de decisão aprimorada por meio do incremento das habilidades do gestor ao transformar insights em decisões fundamentadas. Se os sistemas de inteligência artificial forem implementados adequadamente, as companhias têm uma oportunidade de gerar decisões não somente com maiores taxas de acurácia, mas com mais transparência para seus stakeholders. Os sistemas de IA são mais suscetíveis ao escrutínio a partir de evidências, comparativamente às decisões humanas permeadas por subjetividades. Essa inédita interação homem-máquina é capaz de aprimorar e/ou amplificar as habilidades humanas, construindo cenários para escalar a tomada de decisão com mais probabilidade

3 AGRAWAL, A.; GANS, J.; GOLDFARB, A. Power and Prediction: The Disruptive Economic of Artificial Intelligence. Boston: Harvard Business Review Press, 2022.



de acerto e a um custo menor.

O desafio das companhias está na capacidade de traduzir seus principais problemas de negócios em dados solucionáveis e problemas de modelagem, permitindo estabelecer estratégias de inteligência artificial alinhadas com suas principais metas de curto e longo prazo. O valor agregado do investimento é função de estruturas de gestão ágeis para responder eficazmente às previsões dos modelos de IA.

Os impactos da inteligência artificial não são uniformes nem simultâneos. Embora em níveis e intensidades distintos, a IA deve afetar todos os setores, organizações e regiões geográficas. E, se adequadamente incorporada, vai promover oportunidades substanciais às companhias na criação de valor. O vetor final, contudo, depende da capacidade de cada companhia transformar sua cultura e seus processos, e identificar os projetos candidatos a causarem maior impacto nos resultados.



II . Contextualização

1. AVANÇOS E DESAFIOS DA IA





1. Avanços e desafios da IA

As tecnologias não são todas iguais, algumas adicionam valor incremental à sociedade e outras são disruptivas. Ao reconfigurar a lógica de funcionamento da economia e da sociedade e aportar inéditos modelos de negócios, as tecnologias disruptivas provocam períodos de reorganização no que Joseph Schumpeter, no livro “Capitalismo, socialismo e democracia” (1942)⁴, denominou de “destruição criativa”, derrubando os sistemas existentes e construindo novos. As tecnologias de propósito geral (general purpose technologies – GPT) estão no último bloco. São tecnologias-chave, moldam toda uma era e reorientam as inovações nos setores de aplicação, como a máquina a vapor, a eletricidade e a computação. A inteligência artificial é a tecnologia de propósito geral do século XXI e tende a impactar cada vez mais a economia e a vida em sociedade. Com a IA, migramos de um mundo de máquinas programadas para um mundo de máquinas probabilísticas, implicando lógicas e riscos diferenciados (identificação e monitoramento).

Diferentemente de tecnologias de propósito geral anteriores — como o carvão, a eletricidade e a computação —, a adoção de sistemas de inteligência artificial está se disseminando aceleradamente na sociedade. A eletricidade, por exemplo, demorou 40 anos para ser adotada em larga escala. Em 1879, Thomas Edison demonstrou a famosa lâmpada elétrica. Duas décadas depois, apenas 3% dos lares americanos tinham eletricidade, que em quatro décadas atingiu 50% da população americana.

A técnica de inteligência artificial que permeia a maior parte das implementações atuais, as redes neurais profundas (deep learning), é um modelo estatístico de probabilidade que identifica padrões em grandes volumes de dados. Essa técnica tem duas categorias: IA preditiva e IA generativa. O que distingue as categorias é a “arquitetura” — forma como se combina os elementos que compõem a técnica, fundamentalmente os neurônios artificiais, a conexão entre eles e o número de camadas.

Os modelos de inteligência artificial que servem como base para diversos tipos de aplicação têm sido denominados

4 SCHUMPETER, J.A. 1961. Capitalismo, socialismo e democracia. Editado por George Allen e Unwin Ltd., traduzido por Ruy Jungmann). Rio de Janeiro: Editora Fundo de Cultura.



de “Foundation Models”. Em analogia, seriam o “alicerce de uma casa” sobre o qual serão construídas soluções de IA para distintos propósitos e tarefas. O desenvolvimento desses modelos concentra-se nos Estados Unidos (Figura 1) e nas big techs americanas (Figura 2).

Figura 1

Número de modelos fundacionais por área geográfica

Fonte: Bommasani et al., 2023 | Chart: 2024 AI Index report.

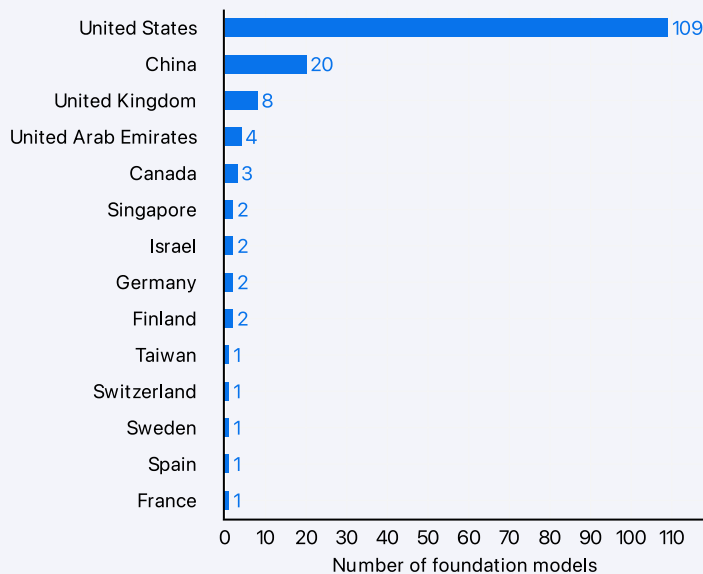
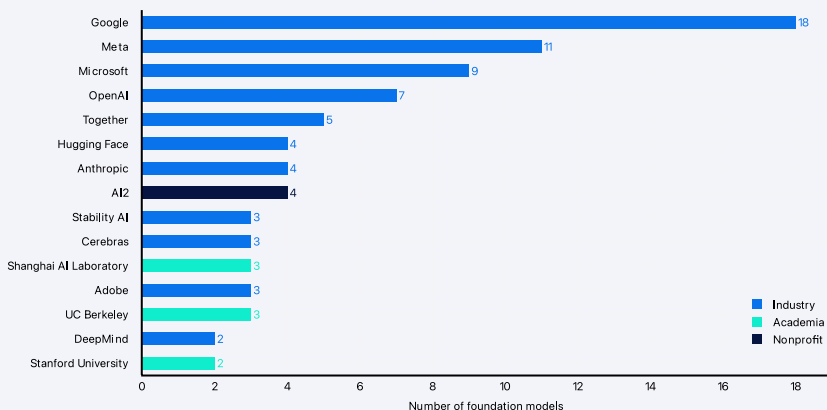


Figura 2

Número de modelos fundacionais por organização

Fonte: Bommasani et al., 2023 | Chart: 2024 AI Index report.



A IA generativa (GenAI) democratizou o acesso à tecnologia com uma interface com o usuário final por meio do diálogo (linguagem natural) e é uma “parceira” dos profissionais, com potencial de gerar impacto positivo sobre a produtividade do trabalho — menor tempo de realização e maior assertividade — e o gerenciamento de forma proativa. As soluções de GenAI, por exemplo, podem criar gráficos personalizados com interação em linguagem natural, se interligar com outras soluções para elaborar cenários preditivos identificando



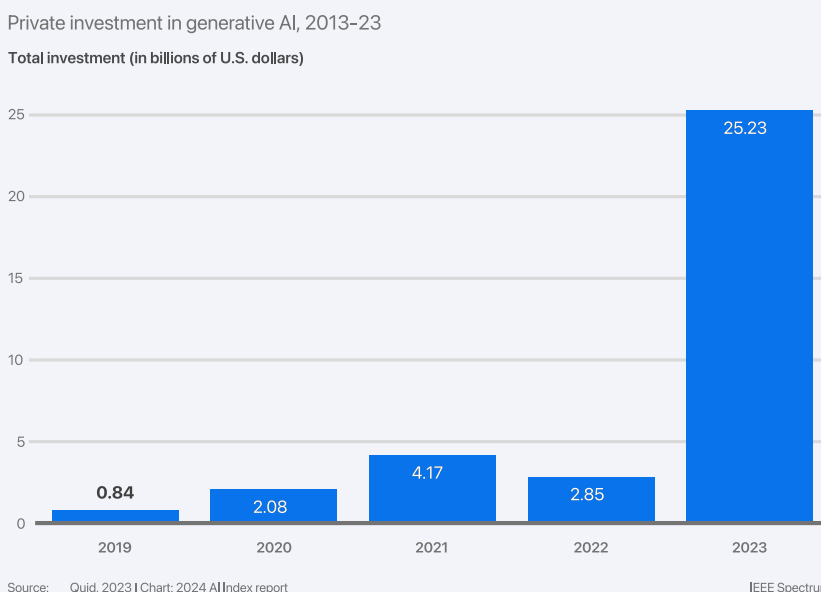
variáveis-chave, avaliar o histórico de risco de um cliente, ou produzir as primeiras versões de documentos e relatórios e extrair insights de regulamentos e processos.

O estudo “Walking the Walk of AI Ethics in Technology Companies”⁵ do Stanford University Human-Centered Artificial Intelligence (HAI), de dezembro de 2023, apurou que o número global de novos modelos de LLM (large language models) dobrou entre 2022 e 2023, sendo dois terços de código aberto. Contudo, os de melhor desempenho são sistemas fechados desenvolvidos pelos líderes da indústria. Segundo o IEEE Spectrum, que analisou os pontos principais da publicação do HAI de 2024 (Figura 3), o investimento no ano passado em IA generativa foi em torno de 25 bilhões de dólares.

Figura 3

Investimento em IA generativa

Fonte: IEEE Spectrum, 2024.



5 STRICKLAND, E. 15 Graphs that explain the state of the art of AI in 2024. IEEE Spectrum, 15 abr. 2024. Disponível em: <https://spectrum.ieee.org/ai-index-2024/15-ai-makes-people-nervous>. Acesso em: 15 de agosto de 2024. Embora o investimento empresarial tenha diminuído globalmente no ano passado, o investimento em IA generativa disparou. Nestor Maslej, editor-chefe do relatório de 2024, declarou ao Spectrum que o boom é indicativo de uma tendência de massivos investimentos. Outro gráfico do relatório mostra que a maior parte do investimento privado em GenAI está nos Estados Unidos.



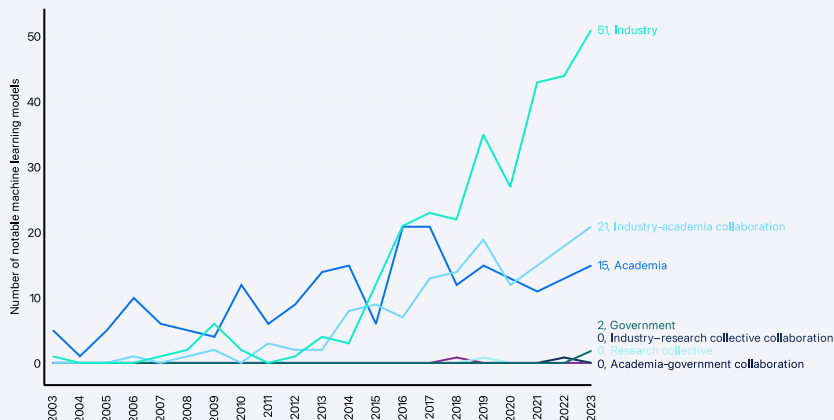
O estudo do HAI indicou dois cenários futuros: a melhoria contínua da tecnologia, com conseqüente expansão do uso de sistemas de inteligência artificial pelas organizações, com efeitos na produtividade e no emprego; e a adoção restrita da IA pelas organizações, função da não superação das limitações da tecnologia. Apesar do ambiente em constante mudança e incertezas, algumas descobertas do estudo:

- a. A inteligência artificial supera os humanos em algumas tarefas, mas não em todas. A tecnologia já superou o desempenho humano, por exemplo, em classificação de imagens, raciocínio visual e compreensão do idioma inglês, mas ficou atrás em tarefas mais complexas, como matemática de nível competitivo, raciocínio visual de bom senso e planejamento;
- b. A indústria continua a dominar a pesquisa de ponta em IA, tendo produzido em 2023 51 modelos notáveis de aprendizado de máquina (machine learning – ML) contra 15 produzidos pela academia – 21 modelos foram produzidos em parceria entre indústria e academia (Figura 4);

Figura 4

Número de modelos notáveis de machine learning por setor

Fonte: Epoch, 2023 | Chart: 2024 AI Index report.



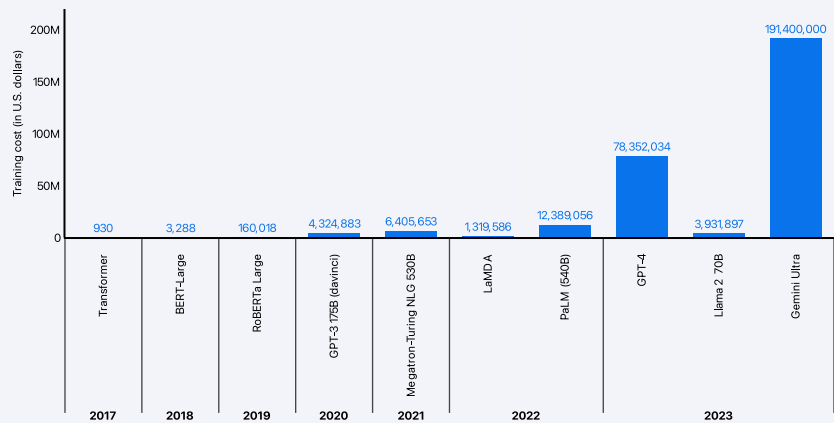


- c. Os custos dos modelos de IA de última geração atingiram níveis sem precedentes. O GPT-4, da OpenAI, por exemplo, gastou cerca de 78 milhões de dólares em computação para treinar, enquanto o Gemini Ultra, do Google, custou 191 milhões de dólares em computação (Figura 5);

Figura 5

Custos estimados para o treinamento de modelos seccionados de IA

Fonte: Epoch, 2023 | Chart: 2024 AI Index report.



- d. Os Estados Unidos lideram como a principal fonte dos modelos de inteligência artificial. Em 2023, 61 modelos notáveis de IA foram desenvolvidos por organizações sediadas nos EUA, ultrapassando em muito os 21 da União Europeia e os 15 da China (Figuras 6 e 7);

Figura 6

Número de modelos notáveis de machine learning por área geográfica

Fonte: Epoch, 2023 | Chart: AI Index report.

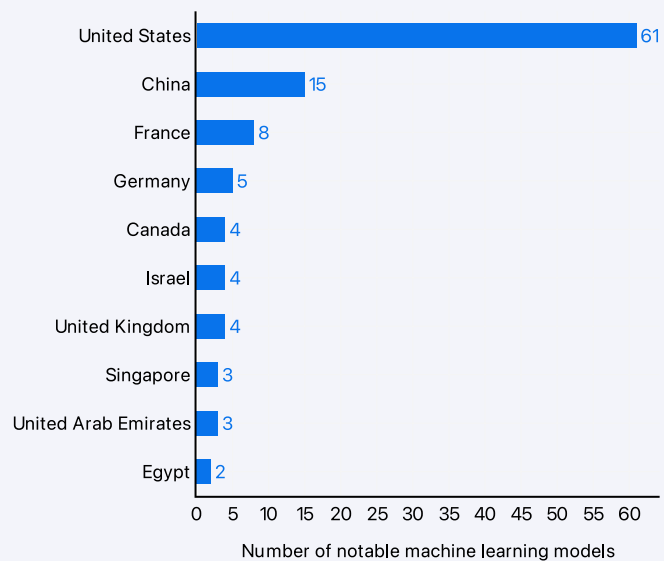
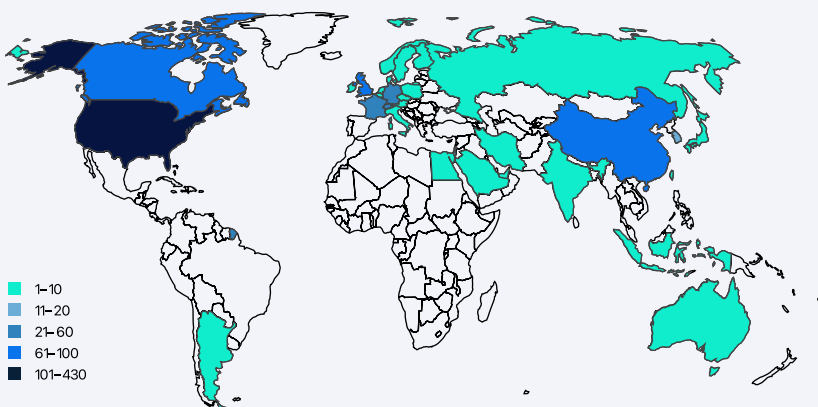




Figura 7

Número de modelos notáveis de machine learning por área geográfica

Fonte: Epoch, 2023 | Chart: AI Index report.



- e. China domina o conjunto de patentes de inteligência artificial: em 2022, o país liderou o registro de patentes de IA com 61,1%, ultrapassando significativamente os Estados Unidos, com 20,9%. Desde 2010, a participação dos EUA nas patentes de IA diminuiu 54,1%;
- f. Faltam avaliações robustas e padronizadas nos relatórios de “IA responsável”. Os principais desenvolvedores — OpenAI, Google e Anthropic — testam seus modelos em distintos benchmarks, dificultando comparações sistemáticas dos riscos;
- g. O investimento global em IA generativa disparou em 2023, duplicando o volume de 2022 e atingindo 25,2 bilhões de dólares. OpenAI, Anthropic, Hugging Face e Inflection relataram rodadas substanciais de captação de fundos (Figuras 8 e 9);

Figura 8

Investimentos privados em IA por área geográfica

Fonte: Quid, 2023 | Chart: 2024 AI Index report.

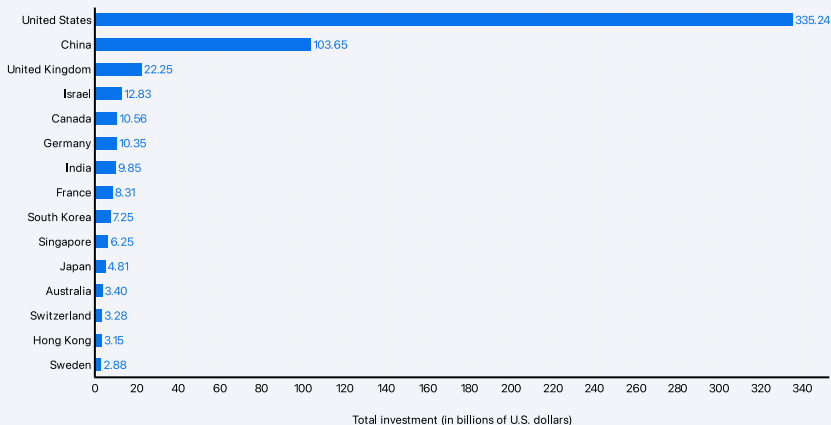
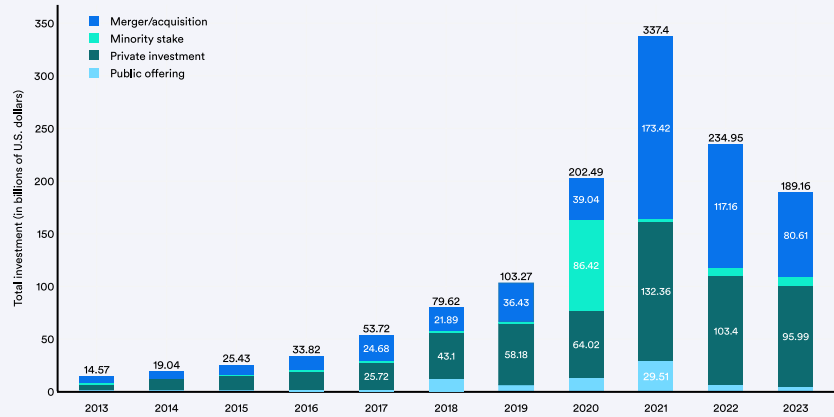




Figura 9

Investimentos corporativos globais em IA por modalidade de investimento

Fonte: Quid, 2023 | Chart: 2024 AI Index report.

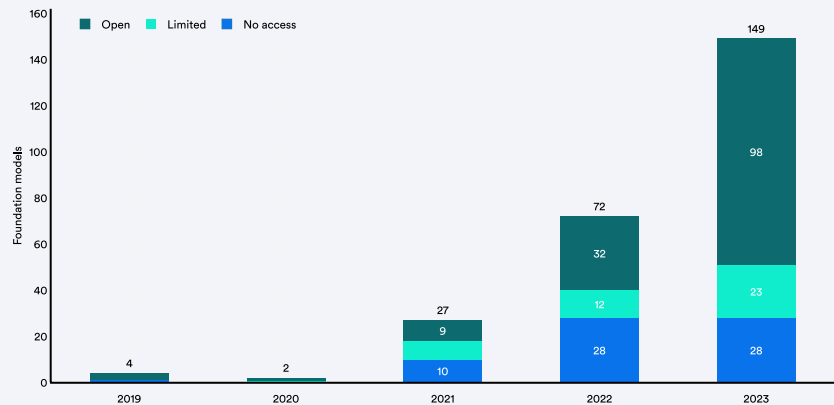


h. A pesquisa de inteligência artificial em código aberto apresenta aumento consistente desde 2011: de 845 projetos para aproximadamente 1,8 milhão em 2023, ano em que o crescimento foi de 59,3% (Figura 10);

Figura 10

Modelos fundacionais pela forma de acesso

Fonte: Bommasani et al., 2023 | Chart: 2024 AI Index report.



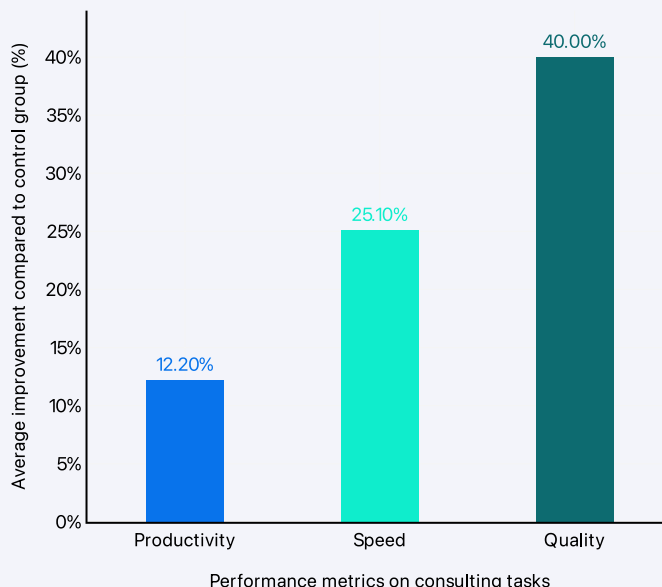
i. A IA tem o potencial de tornar os trabalhadores mais produtivos, executando tarefas em menor tempo e com resultados de melhor qualidade, reduzindo a disparidade de competências entre trabalhadores com baixa e alta qualificação. A adoção de inteligência artificial sem supervisão e os devidos cuidados de governança e riscos, contudo, pode comprometer o desempenho (Figura 11);



Figura 11

Efeito do uso de GPT-4 em um grupo de consultores

Fonte: Dell'Acqua et al., 2023 | Chart: 2024 AI Index report.

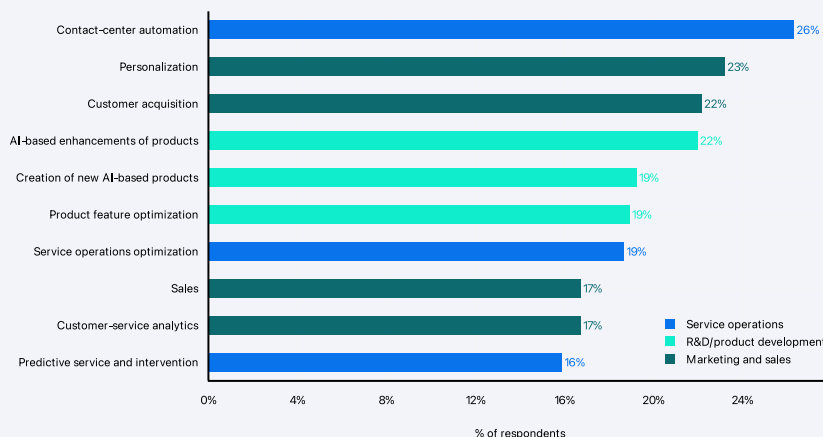


j) A adoção de sistemas de IA tende a reduzir custos operacionais e aumentar as receitas das companhias. Em comparação com 2022, 2023 teve um aumento de 10% no número de companhias que relataram redução de custos, sugerindo geração de ganhos significativos de eficiência nos negócios em parte pela diversidade de aplicações (Figura 12);

Figura 12

Casos de uso de IA mais comuns por funcionalidade

Fonte: McKinsey; Company, 2023 | Chart: 2024 AI Index report.



k) A inteligência artificial, particularmente a IA generativa, está na pauta das 500 maiores companhias dos EUA. Em 2023, a IA foi mencionada em 394 teleconferências de resultados, um aumento notável em relação às 266 menções em 2022. Desde 2018, as menções à inteligência artificial nas teleconferências de resultados das empresas "Fortune 500" quase duplicaram;



- l) O número de publicações sobre inteligência artificial entre 2010 e 2022 quase triplicou, passando de aproximadamente 88.000, em 2010, para mais de 240.000 em 2022;
- m) O número de regulamentações sobre inteligência artificial nos Estados Unidos aumentou 56,3% em 2023, com 25 regulamentações relacionadas à IA contra apenas uma em 2016 (Figuras 13 e 14);

Figura 13

Número de projetos de lei relacionados à IA aprovados em países selecionados (2023)

Fonte: AI Index 2024 | Chart: 2024 AI Index report.

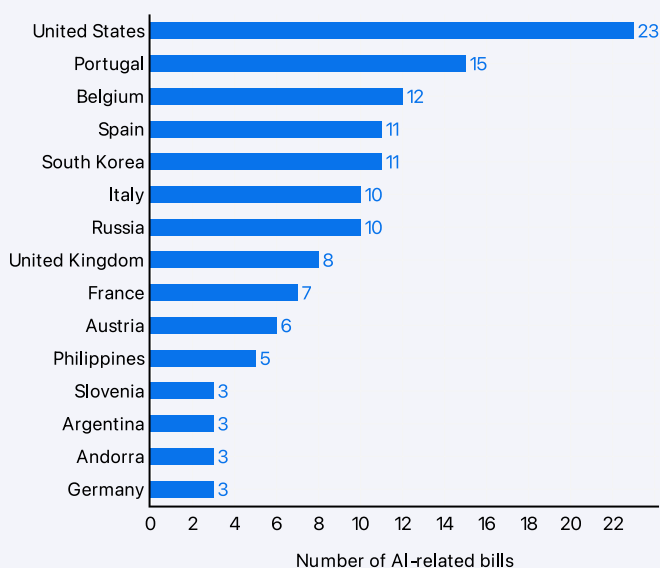
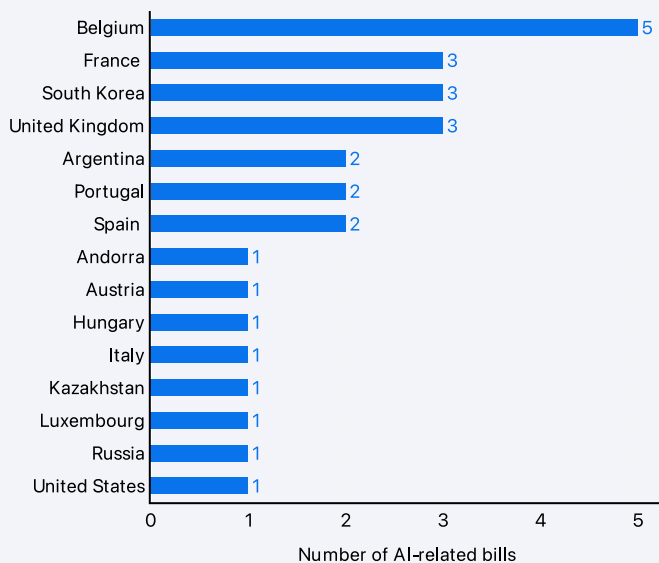


Figura 14

Número de projetos de lei relacionados à IA aprovados em países selecionados (2016-2022)

Fonte: AI Index 2024 | Chart: 2024 AI Index report.



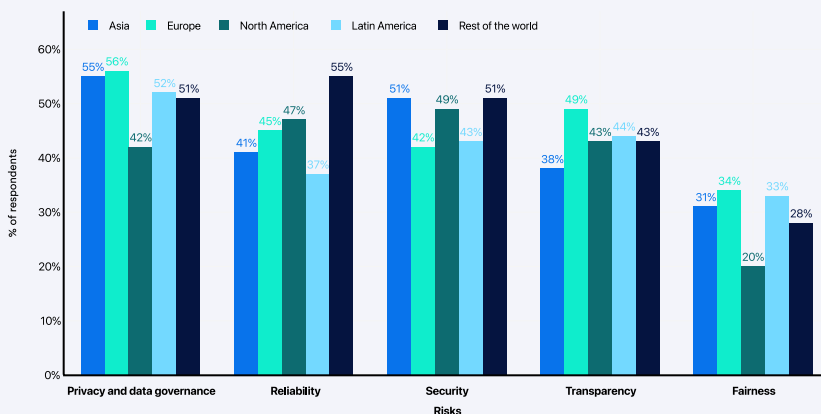


- n) Mundo afora, as pessoas estão mais conscientes — e mais apreensivas — sobre o impacto potencial da inteligência artificial. Pesquisa da Ipsos mostra que, em 2023, a proporção daqueles que pensam que a IA irá afetar dramaticamente as suas vidas nos próximos três a cinco anos aumentou de 60% para 66%. E 52% dos entrevistados expressaram nervosismo em relação aos produtos e serviços de IA, um aumento de 13 pontos percentuais em relação a 2022. Nos Estados Unidos, os dados do Pew sinalizam que 52% dos americanos relatam se sentir mais preocupados do que entusiasmados com a inteligência artificial (eram 37% em 2022)⁶;
- o) Foram identificadas vulnerabilidades mais complexas nos modelos de LLMs, requerendo estratégias menos óbvias para mitigar os comportamentos prejudiciais;
- p) As organizações estão mais preocupadas com os riscos dos sistemas de IA, particularmente privacidade e segurança. Contudo, neutralizaram apenas uma pequena parcela dos riscos (Figuras 15 e 16).

Figura 15

Relevância do risco relacionado à Responsible AI para a organização por região

Fonte: Global State of Responsible AI report, 2024 | Chart: 2024 AI Index report.



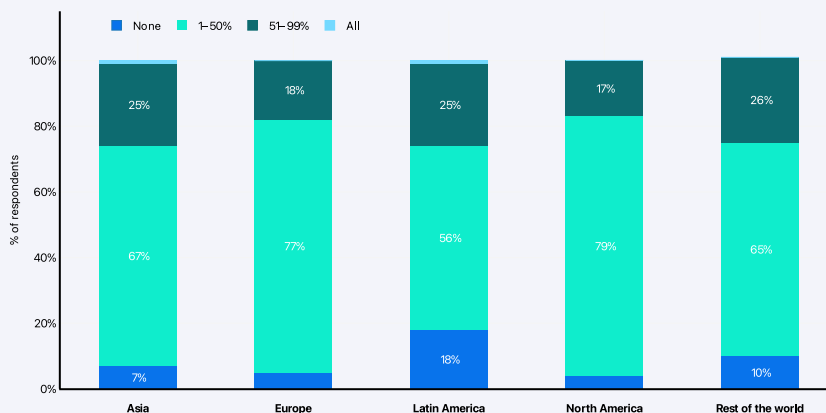
6 A Genial/Quaest, de 2 a 6 de maio 2024, investigou a relação dos brasileiros com a IA: 21% dos entrevistados declararam já ter usado alguma solução de IA, com uma distribuição quase equitativa por região — 20% Nordeste, 23% Sudeste e Sul, e 18% Centro-Oeste/Norte — e discrepância em relação à faixa etária — 33% de 16 a 34 anos, 19% de 35 a 59 anos, e 6% acima de 60 anos — e à escolaridade — 9% até o ensino fundamental, 21% ensino médio completo/incompleto e 42% ensino superior ou mais. Quanto ao tipo de solução, a concentração, com quase 50%, recai sobre os recursos de geração de texto automático, atingindo 59% entre os adolescentes e jovens adultos.



Figura 16

Adoção global de Responsible AI por região

Fonte: Global State of Responsible AI report, 2024 | Chart: 2024 AI Index report



O estudo do HAI identificou uma ampla gama de obstáculos que impedem as organizações do setor de tecnologia de incorporar padrões éticos e de governança no desenvolvimento de sistemas de IA. Em geral, são iniciativas isoladas sem apoio institucional, logo sem autoridade para enfrentar a prevalência dos objetivos de inovação. Os gestores de produtos tendem a considerar que as atividades dos eticistas (especialistas em ética) atrasam o lançamento de produtos e serviços, comprometendo os resultados financeiros.

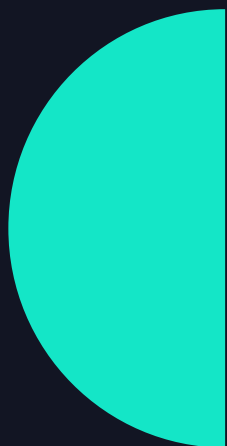
Geralmente, as variáveis críticas a balizar as prioridades são as métricas de desempenho, e não as considerações de sustentabilidade. As variáveis relacionadas à governança, à gestão de riscos e à ética precisam ter evidências quantitativas irrefutáveis para serem aceitas. Alterar a prática e a cultura do mercado passa por regulamentação, com regras e padrões técnicos a serem cumpridos, sob pena de sanção; e sensibilização dos colaboradores de que a conformidade regulatória e ética não se restringe a um posicionamento legal ou moral, mas tem impacto efetivo na reputação e na responsabilidade da organização, assim como na credibilidade junto aos stakeholders e à sociedade — e, por conseguinte, nos resultados financeiros.



II . Contextualização

2. PRINCÍPIOS ÉTICOS GERAIS, REGULAÇÃO E GOVERNANÇA DE IA

Clique para ir até o tópico de interesse





2. Princípios éticos gerais, regulação e governança de IA

2.1. Princípios éticos gerais

A reação inicial à disseminação da inteligência artificial, com seus potenciais riscos e danos, foi a elaboração de um conjunto de princípios gerais. A primeira publicação sobre princípios relacionados ao uso de IA — originados na *Conference on Beneficial AI*, evento organizado pelo instituto Future of Life em 2017, na Califórnia— ficou conhecida como “*Asilomar Principles*”⁷ e teve um papel importante na sensibilização do mercado e da sociedade sobre o tema. No entanto, mostrou-se de aplicabilidade restrita pelos princípios não serem universais nem representáveis em termos matemáticos, pré-condição para inseri-los em modelos estatísticos.

Em 2018, a Comissão Europeia constituiu o “High-Level Expert Group on Artificial Intelligence”, grupo de trabalho responsável pela elaboração do documento “Ethics Guidelines for Trustworthy AI”⁸. Segundo essas diretrizes, o desenvolvimento de projetos de IA devem ser: legais, em conformidade com todas as leis e regulamentos aplicáveis; éticos, garantindo a adesão a princípios e valores éticos; e robustos. Elas têm como base os direitos humanos, a liberdade dos indivíduos, o respeito pela democracia e pela justiça, a não discriminação, a solidariedade e os direitos dos cidadãos.

Com foco no papel das organizações no uso responsável e sustentável da inteligência artificial, em 2019 a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) publicou uma lista de princípios⁹: a) crescimento inclusivo, desenvolvimento sustentável e bem-estar; b) valores centrados no ser humano e na justiça; c) transparência e explicabilidade; d) robustez, segurança e proteção; e e) responsabilidade.

As Nações Unidas (ONU), por meio da UNESCO e de sua plataforma “AI for Good” — liderada pela International

7 FUTURE OF LIFE INSTITUTE. Asilomar AI principles. Futureoflife.org, 11 ago. 2017. Disponível em: <https://futureoflife.org/open-letter/ai-principles/>. Acesso em: 22 jun. 2024.

8 COMISSÃO EUROPEIA. Ethics guidelines for trustworthy. Commission.europa.eu, 8 abr. 2019. Disponível em: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>. Acesso em: 22 jun. 2024.

9 OCDE (Organização para a Cooperação e Desenvolvimento Econômico). OCDE AI principles. OECD.AI, maio 2019. Disponível em: <https://oecd.ai/en/ai-principles>. Acesso em: 22 de jun. 2024.



Telecommunication Union (ITU), agência da ONU especializada em tecnologias de informação e comunicação —, têm promovido diversas iniciativas associadas aos objetivos de desenvolvimento sustentável (ODS), obtendo a adesão de 193 países no primeiro acordo global sobre ética de IA¹⁰. Em paralelo, emergem iniciativas de autorregulamentação pela indústria de tecnologia, soft law baseada em códigos de conduta e padrões éticos, sem necessidade de controles externos ou imposições regulatórias.

Embora relevantes, essas iniciativas não são suficientes para proteger a sociedade, requerendo que o poder público estabeleça arcabouços regulatórios para o desenvolvimento e o uso da inteligência artificial.

2.2. Regulação de IA

A União Europeia, em uma abordagem conceituada como horizontal, promulgou em 1º de agosto de 2024 a primeira lei de inteligência artificial, conhecida como Artificial Intelligence Act (AI Act), com um cronograma de vigências parciais entre fevereiro de 2025 e agosto de 2027¹¹. Sua abordagem é baseada em riscos, classificando-os como “inaceitáveis”, com veto a determinadas soluções; e “alto risco”, com parâmetros de governança, incluindo a obrigatoriedade de realizar análises de impacto, documentar o ciclo de vida de cada solução, respeitar as leis de proteção de dados, e empreender testes de performance e análise de saídas, entre outras disposições. O propósito da lei é garantir os direitos fundamentais, a saúde e a segurança das pessoas, dos grupos e de toda a sociedade, por meio de transparência, supervisão humana, prevenção de riscos e danos, e promoção de seu uso confiável e responsável.

Nos Estados Unidos, com uma abordagem conceituada como vertical, a “Executive Order” do presidente Biden, em 30 de outubro de

10 NAÇÕES UNIDAS. 193 countries adopt first-ever global agreement on the Ethics of Artificial Intelligence. UN News, 25 nov. 2021. Disponível em: <https://news.un.org/en/story/2021/11/1106612>. Acesso em: 8 jun. 2024.

11 Fevereiro-2025: as disposições dos capítulos I e II começam a ser aplicadas. Agosto-2025: a maior parte das obrigações constantes nos capítulos III, V, VII, XII e artigo 78 serão aplicadas a partir desta data, exceto o artigo 101, ou seja, obrigações relacionadas a modelos de uso geral, governança, confidencialidade e sanções. Agosto-2026: aplicação da lei completa, com exceção do artigo 6(1) e obrigações correlatas. Agosto-2027: aplicação do artigo 6(1) e obrigações correspondentes.



2023¹², convocou as agências reguladoras e os demais organismos setoriais a uma tomada de subsídios em seus respectivos domínios sobre as aplicações em uso e os riscos envolvidos da inteligência artificial, e eventuais normas setoriais ou precedentes judiciais com potencial de ser aplicados ou adaptados. Em paralelo, foram definidas diretrizes gerais de cumprimento voluntário (sem força de lei).

No Brasil, tramita no Senado o Projeto de Lei (PL) 2.338/2023, visando constituir o “Marco Regulatório da IA”. A versão atual do PL 2.338 adota igualmente a abordagem baseada em riscos, com algumas diferenças em relação à Lei de IA europeia. O projeto brasileiro contempla uma extensa relação de princípios e direitos, ao mesmo tempo que atribui um protagonismo para agências e órgãos setoriais, tanto na definição de parâmetros técnicos para o cumprimento de suas disposições como na harmonização do teor da futura lei, com as especificidades de cada domínio.

O ritmo acelerado de adoção da inteligência artificial pelas companhias e o provável estabelecimento de um marco regulatório da IA no Brasil sinalizam a premência do debate sobre a governança de IA.

2.3. Governança de IA

A governança e a ética organizacional significam um compromisso de atuação de forma responsável e sustentável. Construir bases sólidas de governança e princípios éticos, alinhados com os valores da organização, fornece uma base de confiança entre as partes interessadas, reputação de marca e posicionamento competitivo no mercado.

Em todo o mundo já foram publicados mais de 60 frameworks de governança e ética de IA, com o propósito de oferecer uma metodologia para organizar, documentar e gerir o ciclo de vida das soluções de inteligência artificial; incentivar a colaboração entre as áreas das organizações; proporcionar transparência, explicabilidade e auditabilidade; e possibilitar ajustes visando mitigar, de forma

12 THE WHITE HOUSE. Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Whitehouse.gov, 30 out. 2023. Disponível em: <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>. Acesso em: 19 jun. 2024.



recorrente, os impactos negativos. Um grupo de pesquisadores do TIDD da PUC-SP¹³ analisou alguns desses frameworks, destacando três pela consistência do conteúdo e pela relevância das instituições promotoras:

- a. AI Risk Management Framework (AIRMF), do National Institute of Standards and Technology (NIST);
- b. Information technology. Artificial intelligence. Guidance on risk management - ISO/IEC 23894, 2023;
- c. Ethically Aligned Design-for-Business do Institute of Electrical and Electronics Engineers (IEEE).

As Tabelas 1 a 3 descrevem cada um dos três frameworks em termos de objetivos, princípios, profundidade, processos e uso indicado:

Tabela 1

NIST

Elaboração:
Grupo de pesquisa TIDD/PUC-SP

Objetivos	Princípios	Principais conceitos abordados	Processos	Profundidade da discussão do tema	Para quem é indicado
<p>O principal objetivo do NIST AI RMF é oferecer às organizações uma abordagem estruturada para gerenciar riscos associados à IA de forma eficaz. O framework busca aprimorar a confiabilidade dos sistemas de IA e promover o desenvolvimento e a implementação responsável dessas tecnologias. Isso envolve a compreensão, avaliação e gestão de potenciais riscos e impactos, visando maximizar os benefícios e minimizar as consequências negativas para indivíduos, comunidades, organizações e para a sociedade como um todo.</p>	<p>O framework é fundamentado em princípios de confiabilidade, que incluem validade e confiabilidade, segurança, proteção e resiliência, responsabilidade e transparência, explicabilidade e interpretabilidade, aprimoramento da privacidade e equidade com vieses prejudiciais gerenciados. Esses princípios têm o objetivo de orientar o desenvolvimento e o uso de sistemas de IA para garantir que eles estejam alinhados com os valores sociais e os padrões éticos.</p>	<ul style="list-style-type: none"> • Gestão de Riscos: Processo sistemático para identificar, avaliar e gerenciar riscos ao longo de todo o ciclo de vida da IA. • Confiabilidade: Características como validade, confiabilidade, segurança, proteção, transparência e equidade. • Abordagem ao Ciclo de Vida: Gestão contínua de riscos desde o design até a implementação e além. • Centrado no Humano: Ênfase na supervisão e envolvimento humano nos sistemas de IA. • Responsabilidade Social e Ética: Abordagem dos impactos sociais e considerações éticas no desenvolvimento de IA. 	<p>O NIST AI RMF delinea um processo estruturado dividido em quatro funções principais:</p> <p>Governar: Estabelecer estruturas e processos de governança para supervisionar a gestão de riscos de IA.</p> <p>Mapear: Identificar e compreender o contexto e os riscos específicos associados aos sistemas de IA.</p> <p>Mensurar: Empregar métodos e métricas para avaliar e monitorar os riscos de IA.</p> <p>Gerenciar: Implementar ações para mitigar riscos e garantir conformidade e melhoria contínua.</p>	<p>O framework oferece uma discussão aprofundada sobre vários aspectos da gestão de riscos em IA, incluindo a identificação de riscos, os desafios de mensuração, a priorização de riscos e a integração da gestão de riscos nas práticas organizacionais. Ele enfatiza a importância de uma abordagem diversa e multidisciplinar, a monitoração contínua e a adaptabilidade do framework às tecnologias e riscos de IA em evolução.</p>	<ul style="list-style-type: none"> • Liderança Executiva: Responsável por definir o tom e a cultura de gestão de riscos. • Equipes de Gestão de Riscos: Responsáveis por desenvolver e implementar processos de gestão de riscos. • Desenvolvedores e Engenheiros de IA: Responsáveis por integrar a gestão de riscos no design e desenvolvimento de sistemas de IA. • Departamentos Jurídico e de Conformidade: Responsáveis por garantir a adesão aos padrões regulatórios e éticos. • Equipes Operacionais: Responsáveis por monitorar e gerenciar os sistemas de IA em operação. • Departamentos de RH e Treinamento: Responsáveis por promover uma cultura de conscientização sobre riscos e fornecer o treinamento necessário.

¹³ Estudo em curso: "AI Governance and Ethical Guiding Principles for Brazilian Market", com fomento do Digital Futures Fund (DFF), colaboração da TIDES Foundation e do Google. Realizado pelos pesquisadores do TIDD/PUC-SP Alexandre Silveira, Alexandre Zavaglia, André Magno, Leilane Borges e Priscila Reis, sob a coordenação da Prof. Dora Kaufman. Colabora Gabriel Marin Rosário.

Tabela 2

ISO

 Elaboração:
Grupo de pesquisa TIDD/PUC-SP

Objetivos	Princípios	Principais conceitos abordados	Processos	Profundidade da discussão do tema	Para quem é indicado
<p>Este documento é destinado a ser utilizado em conjunto com a ISO 31000:2018. Fornece orientações específicas sobre a gestão de riscos relacionados à IA. Ajuda as organizações a integrarem a gestão de riscos em suas atividades relacionadas à IA. Descreve processos para a implementação e integração eficazes da gestão de riscos em IA.</p>	<p>A gestão de riscos deve atender às necessidades da organização utilizando uma abordagem integrada, estruturada e abrangente. Os princípios devem ser aplicados em todos os níveis organizacionais e em todos os objetivos.</p>	<p>Integrada: A gestão de riscos é parte de todas as atividades organizacionais. Estruturada e Abrangente: A abordagem leva a resultados consistentes e comparáveis. Personalizada: Adaptada ao contexto da organização. Inclusiva: Envolve os diferentes stakeholders para incorporar seus conhecimentos e perspectivas. Dinâmica: Adapta-se a riscos e oportunidades em mudança. Melhor Informação Disponível: Utiliza expectativas históricas, atuais e futuras. Fatores Humanos e Culturais: Considera a influência do comportamento humano e da cultura. Melhoria Contínua: Aperfeiçoamento contínuo por meio de aprendizado e experiências.</p>	<p>Comunicação e Consulta: Diálogo contínuo com as partes interessadas. Definição de Escopo: Definição clara da extensão da gestão de riscos. Avaliação de Riscos: Identificação, análise e avaliação dos riscos. Tratamento de Riscos: Seleção e implementação de medidas para mitigar os riscos. Monitoramento e Revisão: Verificação contínua e melhoria do processo de gestão de riscos. Registro e Relatório: Documentação e relato das atividades e resultados da gestão de riscos.</p>	<p>Discussão detalhada sobre os vários aspectos da gestão de riscos em IA, incluindo novos e emergentes riscos, o papel das partes interessadas e a necessidade de melhoria contínua. Diretrizes específicas sobre como adaptar a gestão de riscos ao contexto da organização e à natureza dinâmica das tecnologias de IA.</p>	<ul style="list-style-type: none"> • Liderança Organizacional: Supervisão e garantia do compromisso com a gestão de riscos. • Gestores de Riscos: Implementação e supervisão do processo de gestão de riscos. • Cientistas de Dados: Avaliação de riscos técnicos e qualidade dos dados. • Consultores Jurídicos: Garantia de conformidade com as regulamentações pertinentes. • Equipes de Engajamento de Partes Interessadas: Facilitação da gestão de riscos inclusiva através do engajamento com grupos diversos.

Tabela 3

IEEE

 Elaboração:
Grupo de pesquisa TIDD/PUC-SP

Objetivos	Princípios	Principais conceitos abordados	Processos	Profundidade da discussão do tema	Para quem é indicado
<p>O objetivo é gerar consciência e iniciar uma conversa.</p> <p>É uma introdução às necessidades, conceitos gerais e primeiros passos para o desenvolvimento e implementação de Ética em AI, além do compliance, trazendo a responsabilidade para todas as pessoas que estão trabalhando no mundo da IA.</p>	<p>A responsabilidade por uma IA ética é de todos os envolvidos e não apenas de alguns especialistas. O valor aportado quando se cria sistemas e práticas éticas em IA, vai além do compliance. O approach deve ser a "Ética by Design" e preocupação com o usuário. Adotar e praticar competências éticas em IA, as empresas tornam-se mais conscientes do ponto de vista ético, aumentam a transparência, mitigam vieses na sua IA, ao mesmo tempo que desenvolvem tecnologias que trarão maior benefício para a humanidade.</p>	<ul style="list-style-type: none"> • Ética by Design • Sistemas Autônomos e Inteligentes (A/IS) • Pesquisa e design sustentados por normas éticas e legais • Prontidão para Ética em IA • Bem-estar humano no centro do desenvolvimento de A/IS 	<p>Não introduz um processo, mas apresenta o framework "AI Ethics Readness" para se entender o momento da empresa em relação ao desenvolvimento da Ética e Ética de AI e as necessidades para desenvolver os skills e escoar o pensamento de "ética by design"</p>	<p>A discussão maior é sobre Ética, responsabilidades e responsabilização. Se aprofunda no tema na medida que precisa ser pensado "by design" por todos os envolvidos. Para isso, se explica um pouco mais os skills necessários para desenvolver ou contratar pessoas.</p>	<ul style="list-style-type: none"> • Empresas em estágios iniciais ou com nenhum desenvolvimento de IA • Iniciar a conversa com outros stakeholders não envolvidos diretamente no projeto • Gerar conscientização no C-Level • Evangelistas da Ética • Treinamento da Equipe

O grupo elaborou um Mapa de Calor, classificando cada item dos frameworks por graus de superficialidade e de profundidade.

Figura 17

Mapa de calor dos frameworks

Elaboração: Grupo de pesquisa TIDD/PUC-SP



Os frameworks, de adoção voluntária, não têm o objetivo de criar certificações de adequação, mas proporcionar um processo flexível, estruturado e mensurável para abordar os riscos da IA ao longo do ciclo de vida dos sistemas. O objetivo do NIST, por exemplo, é “oferecer uma orientação para o desenvolvimento e utilização de uma IA confiável e responsável, melhorando a compreensão das organizações, indústrias e sociedade para determinar seus níveis aceitáveis de risco, para gerenciar os impactos empresariais e sociais relacionados ao desenvolvimento, implantação e uso dessa tecnologia”¹⁴.

Apesar da Lei de IA europeia (AI Act), das diretrizes das agências reguladoras americanas, do escopo do PL2338/23, dos frameworks de governança, e dos debates nos fóruns especializados, ainda não há clareza nem consenso sobre, por exemplo, o conceito de explicabilidade; o teor das análises de impacto e da documentação associada; os procedimentos técnicos para auditar e manter em compliance com as questões ética, bem como em relação a normas e regulamentos, os sistemas habilitados por inteligência artificial que são probabilísticos e em contínua mutação; e a fronteira entre transparência e segredo de negócio. Nesse sentido, para além da norma, é preciso avançar no estabelecimento de padrões técnicos tanto para satisfazer as demandas de governança, especialmente no alto risco, como para garantir sua efetividade e segurança jurídica das organizações desenvolvedoras ou aplicadoras dessa tecnologia.

14 AI Risk Management Framework (AIRMF), do National Institute of Standards and Technology (NIST). Disponível em: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>. Acesso em 17 de ago 2024.



Parte 3

RELATÓRIO

Clique para ir até o tópico de interesse





III. Relatório

Diante da recente visibilidade dos sistemas de inteligência artificial, os gestores indagam em quais atividades deve-se aplicar IA para automação ou otimização de processos, criação de novos produtos ou serviços. Quais atividades e/ou tarefas devem ser priorizadas? Qual o custo envolvido na implementação, considerando inclusive a capacitação dos colaboradores? Como apurar os benefícios e mensurar os resultados futuros do investimento em IA? Como selecionar fornecedores? Recomenda-se contratar especialistas ou formar ecossistemas de parceiros? Qual o grau de conhecimento exigido para as equipes, particularmente os gestores e times técnicos diretamente envolvidos e/ou beneficiados pelos sistemas de IA? É imprescindível ter um Chief Artificial Intelligence Officer (CAIO)? Como a IA se relaciona com os sistemas legados? Como construir uma cultura sólida de dados e de IA na companhia? Como estruturar um Programa de Governança e Gestão de Riscos para essa finalidade?

Parte fundamental do sucesso da adoção da inteligência artificial depende de talentos, tanto tecnólogos com competência para entender as especificidades dos negócios e traduzi-las para soluções de IA, quanto profissionais de negócios familiarizados com a lógica e os fundamentos da tecnologia.

Em geral, as atividades e os processos digitalizados têm potencial de gerar significativos volumes de dados. O setor de logística, por exemplo, lida com milhões de pacotes por dia, sendo que cada pacote é escaneado, em média, entre 15 e 20 vezes. A inteligência artificial permite extrair insights dos dados gerados na operação para prever rotas de entrega, rastrear pacotes, otimizar a logística, prever demandas de estoque e de vendas, e tornar mais eficiente o relacionamento com os fornecedores.

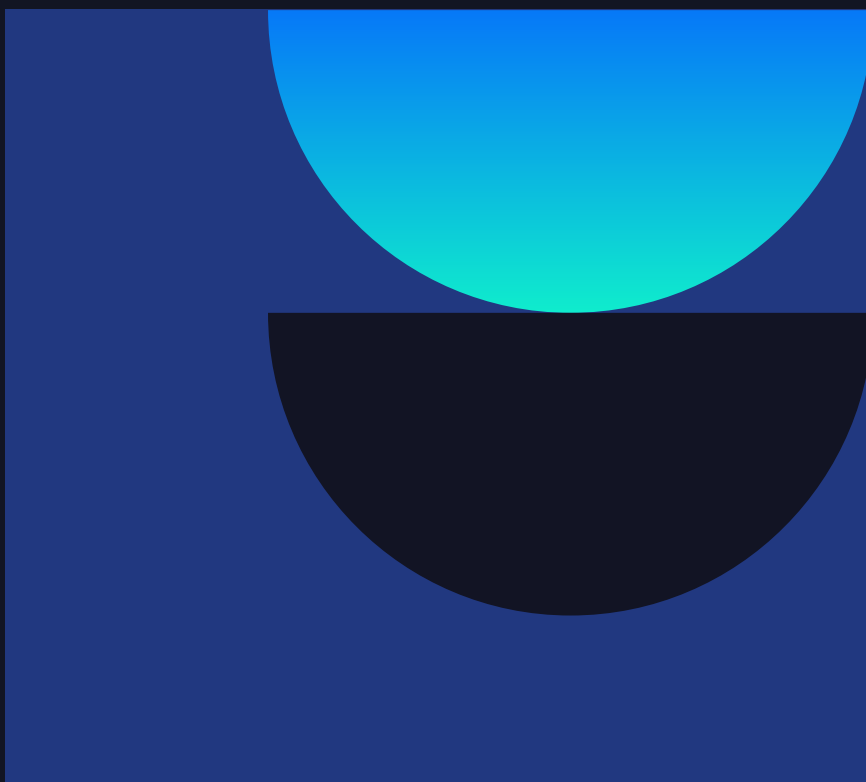
No setor bancário, a IA promove a otimização de serviços de backoffice; avalia o risco de crédito; identifica a melhor opção de investimento adequado ao perfil do cliente-investidor; sumariza e classifica documentos, como as atas do Comitê de Política Monetária (Copom) do Banco Central; detecta e previne fraudes; acompanha os processos judiciais; planeja as operações de logística, como transportar dinheiro aos caixas eletrônicos; e atende digitalmente o cliente, gerando mais autonomia e reduzindo o tempo de resposta



estimado — consequentemente, com ganhos na experiência do cliente (UX).

Companhias de energia adotam sistemas de visão computacional para monitorar as redes de transmissão, e as companhias de óleo e gás para identificar o ponto ideal de perfuração nas plataformas de extração de petróleo. As varejistas adotam a IA para automatizar os armazéns, otimizando o tempo de empacotamento e entrega; compatibilizar as aquisições de suprimentos com a previsão de vendas; e melhorar a experiência do cliente (UX). Em todos os setores, a inteligência artificial generativa engendra novos casos de uso, como distribuição horizontal da informação em reuniões — respeitando os segredos de negócio e os protocolos de consentimento de dados pessoais previstos na LGPD (Lei Geral de Proteção de Dados) —; leitura e análise de grandes quantidades de documentos; geração de textos e imagens; e aperfeiçoamento das tarefas cotidianas.

A prática de mercado tem sido começar por projetos pontuais, permitindo validar conceitos, experimentar a tecnologia, criar cultura de IA nas companhias, medir o retorno do investimento e os riscos envolvidos. A maioria das companhias investigadas se encontra nessa fase, com o uso de soluções de inteligência artificial preferencialmente nas áreas de apoio e suporte. Um grupo restrito de companhias de setores altamente regulados, como energia e varejo financeiro, utilizam soluções de IA há mais tempo e em atividades do core business.



III . Relatório

**1. RESULTADOS
DA PESQUISA**





1. Resultados da pesquisa

Classificamos as companhias em quatro perfis, segundo o grau de adoção de sistemas de IA — fora do core business e no core business — e o grau de maturidade de governança de IA, sintetizados no diagrama abaixo:



Diagrama 1 Perfil das Companhias



A partir do perfil de cada quadrante da matriz, analisamos os resultados quantitativos (pesquisa on-line) com base nos resultados qualitativos (entrevistas qualificadas). Os conflitos atestam, em parte, a relativa limitação de pesquisas realizadas somente por meio de questionários on-line, e a relevância da metodologia escolhida ao promover o cruzamento das respostas quantitativas com as evidências coletadas nas entrevistas qualitativas.

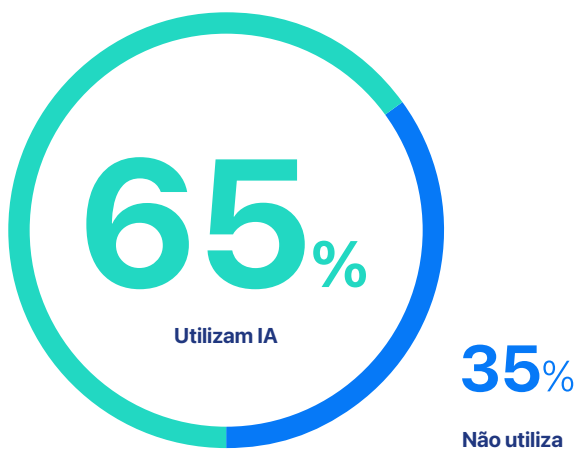


A

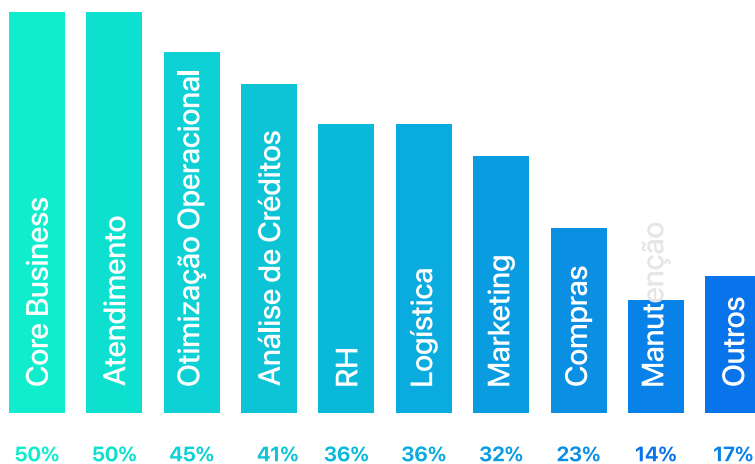
Primeiro bloco de resultados da pesquisa quantitativa

Adoção de IA e tipo de utilização

I. Utiliza IA?

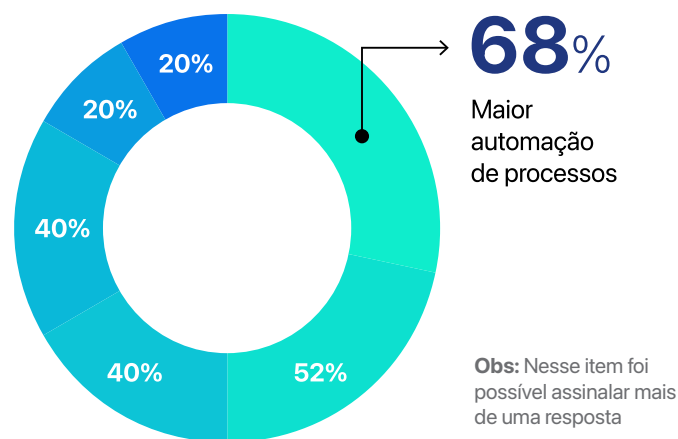


II. Entre as que utilizam: Para quê?





III. Que mudanças a IA pode trazer para a organização nos próximos anos?



- 68% Maior automação de processos
- 40% Transformação da experiência do cliente
- 52% Ganhos de Eficiência
- 20% Mudanças na estrutura organizacional
- 40% Novos modelos de negócio
- 20% Novos processos internos

Comentários com base na pesquisa qualitativa

Nos quatro grupos predomina a adoção de inteligência artificial para otimizar atividades e/ou tarefas operacionais e de suporte; o uso de deep learning para mapear padrões nos dados e gerar insights para otimizar a produtividade e os resultados nas mais diversas áreas, ou de aplicativos de IA generativa ofertados pelo mercado, como o Copilot, para melhorar o desempenho dos colaboradores; e de chatbots para apoiar os canais de relacionamento com os clientes. É cauteloso o uso de chatbots habilitados por IA na interação direta com os clientes e/ou consumidores, em geral os sistemas estão em teste.

Observa-se que a maioria das companhias está estruturando seus bancos de dados, formando data lakes para adequar as soluções de mercado às especificidades de sua operação, processo conhecido como *fine tuning*¹⁵.

¹⁵ Fine Tuning (ajuste fino) é o retraining de modelos pré-treinados em novos dados, no caso gerados na operação da companhia, adequando os modelos às suas especificidades.



Os grupos **1** e **3** apresentam um índice maior de projetos relacionados ao core business, impulsionando novos modelos de negócio, produtos e serviços, embora concentrados em iniciativas pontuais. A conformidade dos grupos 1 e 3 em relação à natureza dos projetos adotados não se mantém em relação à governança: o Grupo 1 possui estruturas e procedimentos em cibersegurança e proteção de dados pessoais (LGPD), com algumas iniciativas relacionadas a criação de normas internas e diretrizes para o uso de IA Generativa, e o Grupo 3, embora sensibilizado, ainda não tem ações concretas.

Os grupos **2** e **4** adotam sistemas de inteligência artificial nas áreas de suporte — financeiro, recursos humanos, marketing, jurídico, manutenção e logística —, com foco na automação dos processos para reduzir custos e ganhar eficiência.

Exemplos ilustrativos

Grupo 1

Companhia do setor de telecomunicações está em processo de habilitar seus chatbots com IA generativa para aprimorar a interação com o cliente por meio de linguagem natural e hiper personalização por perfil (região, frequência, comportamento e hábitos). Ainda na fase de teste, dedicada a identificar e mitigar os riscos inerentes aos sistemas de inteligência artificial generativa, o projeto restringe-se a municiar de informações precisas e mais rápidas as equipes de atendimento ao cliente.

Companhia do setor de energia, ao usar sistemas de visão computacional e identificação de padrões para as atividades de monitoramento e manutenção de suas redes, transformou a experiência exitosa em um serviço a ser comercializado, inclusive para os concorrentes — spin-off de um projeto interno para um serviço oferecido no mercado, inaugurando um novo centro de resultado.



Grupo 3

Companhias do varejo usam dados de logística e outras fontes para a predição de tendências, vendas futuras e otimização de estoques, além de gerar inputs para as áreas administrativa e comercial. A área jurídica adota modelos de IA para classificar contratos (tipos de processo, temas e predição de desfechos ou acordos).

Grupos 2 e 4

O RH adota aplicativos de inteligência artificial para triagem de currículos, seleção de candidatos (compatibilizar o perfil da vaga com o perfil dos candidatos) e avaliação de performance de funcionários. O marketing utiliza sistemas de IA ofertados no mercado para segmentação de campanhas de venda, criação e distribuição de conteúdo. A área de logística integra automação com inteligência artificial para otimizar todas as fases da operação. As decisões de contratação, ou escolhas das soluções de mercado, estão fundamentalmente a cargo da TI, sem escrutínio de jurídico, compliance e governança — em geral, inclusive, sem interferência das áreas de negócio.

Desafios

Avançar na estratégia organizacional e ampliar projetos associados ao core business; estreitar a assimetria de informação e conhecimento entre as áreas técnicas e de gestão; incluir a IA no planejamento estratégico; capacitar os colaboradores para as novas habilidades e competências; definir critérios para a escolha de fornecedores de IA; elaborar diretrizes de gestão de riscos, com destaque para os riscos associados na adoção de soluções de terceiros, contemplando a transparência em relação aos dados de treinamento; e comprovar os resultados e o retorno do investimento.

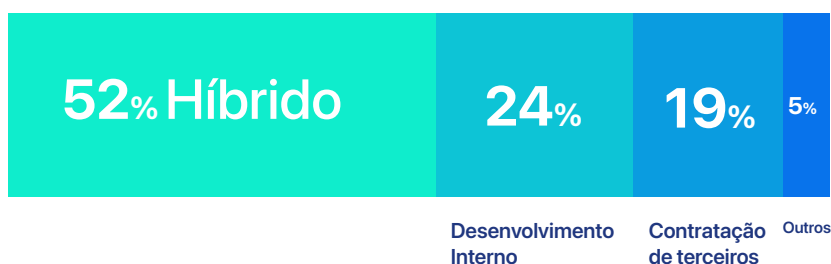


B

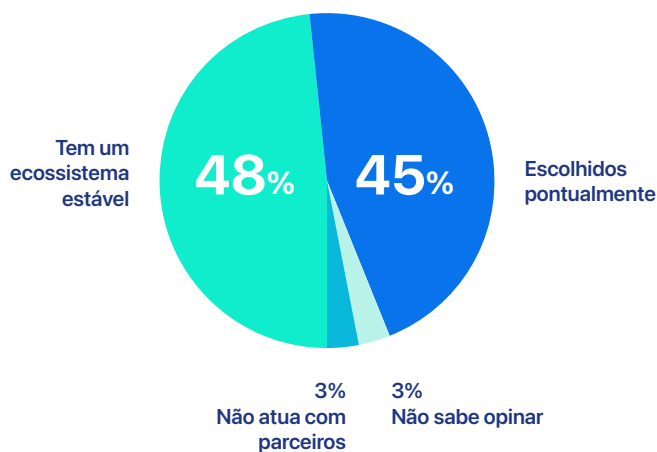
Segundo bloco de resultados da pesquisa quantitativa

Modelos de desenvolvimento de projetos de IA e parcerias

I. Quais as bases para desenvolvimento e uso de soluções de IA?



II. Tem um ecossistema de parceiros de tecnologia estável ou são escolhidos pontualmente?



Comentários com base na pesquisa qualitativa

Companhias do Grupo 1 desenvolvem soluções combinando algoritmos de IA de plataformas de tecnologia e/ou comunidades de código aberto, com relativa estabilidade nas parcerias. Não



identificamos iniciativas de auditoria prévia. O pressuposto é que os algoritmos foram validados pelo fornecedor. O foco é na eficiência, ou seja, se a solução resolve o problema sem atentar para os potenciais riscos.

No Grupo 3 prevalece o modelo híbrido, com parceiros ocasionais. Os grupos 2 e 4 contratam soluções de inteligência artificial ofertadas por terceiros na modalidade “model-as-service” (MaaS)¹⁶. Em todos os grupos, a tendência das companhias é expandir as parcerias estratégicas com os fornecedores de tecnologia. Predomina a Microsoft, seguida do Google e da AWS/Amazon. No caso de aplicativos de terceiros, os critérios de seleção mais comum são a referência de mercado e a reputação do fornecedor.

Não identificamos iniciativas de gestão de risco sólidas, evidenciando a pouca maturidade das companhias em relação à governança de IA, seja pela falta de regulação, de apoio da alta gestão, ou pela lacuna de capacitação e/ou recursos tecnológicos. Algumas companhias têm estruturas de análise de dados e segurança da informação, e um grupo ainda mais restrito possui protocolos de governança integrando as áreas técnica e de negócio nas aplicações voltadas para a criação de produtos ou serviços. Quanto ao uso de dados pessoais, a tendência é confiar que os fornecedores de soluções de inteligência artificial observaram os procedimentos previstos da LGPD e trataram os dados adequadamente.

A liderança de projetos de IA concentra-se na área de TI, com pouco ou nenhum envolvimento das áreas jurídica, de governança ou compliance. Exceto no cadastro do fornecedor e na parte contratual, as decisões são reputadas como “decisões técnicas”. Soluções de inteligência artificial pontuais frequentemente são contratadas diretamente pelas áreas administrativa e de suporte, com a TI interferindo apenas nas questões de segurança da informação e do uso de dados pessoais.

¹⁶ Esses serviços são divididos em três categorias: Software como serviço (SaaS), Plataforma como serviço (PaaS) e Infraestrutura como serviço (IaaS).



Exemplos ilustrativos

Companhias do setor financeiro adotam soluções de reconhecimento facial ofertadas pelo mercado para integrar o processo de abertura de contas digitais, sem auditoria prévia, inclusive sem avaliar a conformidade com a LGPD e os aspectos éticos relacionados.

Adoção de modelos de IA generativa para chatbots de atendimento aos clientes, sem transparência no processo de coleta de dados e de treinamento, limitando a possibilidade de prever e mitigar riscos relacionados a respostas discriminatórias, conteúdo equivocado em relação a serviços e produtos, ou que implique em responsabilização da companhia.

Adoção de soluções de RH para triagem de candidatos, sem transparência sobre o funcionamento da solução, apostando somente na declaração do fornecedor.

Desafios

Garantir a transparência dos modelos ou soluções de terceiros com a devida gestão de riscos; constituir estruturas especializadas para auditar ofertas de terceiros; fortalecer o processo de cadastro de fornecedores de IA; implementar testes para avaliar a performance das soluções de terceiros, inclusive identificando riscos por meio de evidências; implementar protocolos em relação à documentação, observando a tensão entre a explicabilidade do modelo de IA e o segredo de negócio.

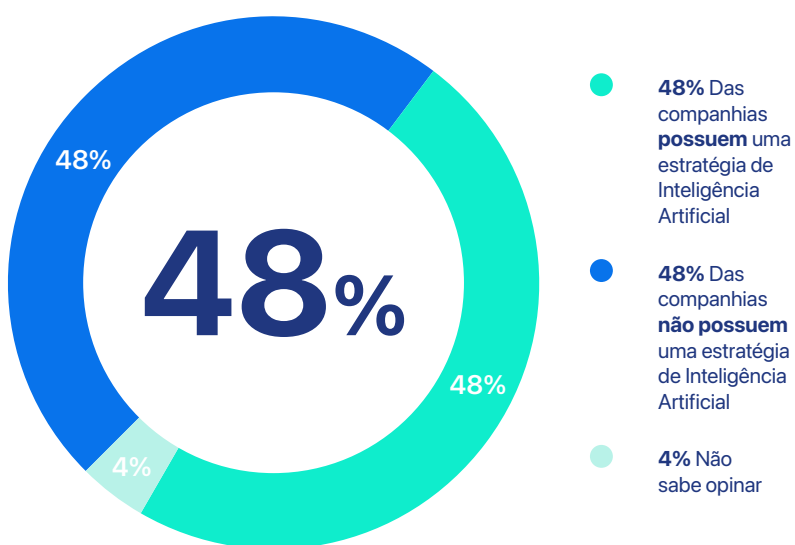


Terceiro bloco de resultados da pesquisa quantitativa

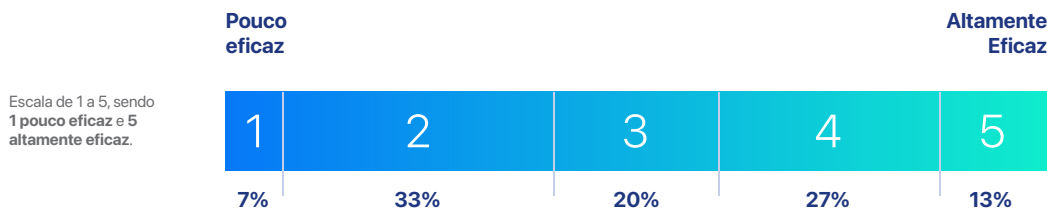


Estratégia de Inteligência Artificial?

I. A companhia tem uma Estratégia de Inteligência Artificial?



II. Quão efetivo é o plano estratégico de adoção de IA na organização?



Comentários com base na pesquisa qualitativa

Na pesquisa qualitativa, na qual foram entrevistadas as maiores companhias de capital aberto do Brasil, não identificamos estratégias de inteligência artificial — apenas, como já mencionado,



usos pontuais de soluções de IA para apoiar a criação de produtos, serviços e ações de marketing; facilitar os processos de RH; otimizar o backoffice; e apoiar as equipes de atendimento ao cliente.

Os resultados apurados nas entrevistas qualitativas conflitam frontalmente com as respostas do questionário on-line, no qual 48% das companhias respondentes afirmaram possuir uma estratégia de IA sendo que 40% atestaram positivamente sobre a efetividade do plano estratégico. Provavelmente, o conflito decorre de interpretações distintas do que seria uma Estratégia de Inteligência Artificial (harmonização conceitual possível nas entrevistas qualitativas). O que observamos, ao longo das entrevistas, foi o interesse crescente do C-Level e o aumento de casos de uso, principalmente no Grupo 1, a conscientização sobre os impactos da IA nos negócios e a necessidade de ajustes no planejamento estratégico no médio e longo prazo.

Exemplos ilustrativos

Grupo 1

Companhia no setor financeiro com uma divisão de “soluções de dados” ou “agenda de dados”, combinando modelos preditivos com modelos de IA generativa, oferece um aplicativo para o setor de seguro que calcula a probabilidade de determinada transação em curso ser fraude, gerando um relatório com sistemas de processamento de linguagem natural (LLMs) com recomendação de encaminhamento.

Desafios

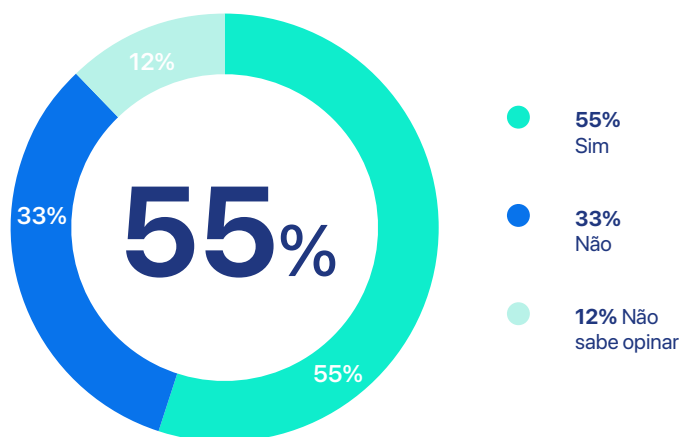
Validar os casos de uso e gerar visibilidade sobre eficiência e retorno do investimento em projetos de IA, com indicadores quantitativos; incluir a inteligência artificial no planejamento estratégico da companhia, precedido de um diagnóstico sobre onde e como o uso de inteligência artificial agrega valor não apenas na produtividade, mas no core business.



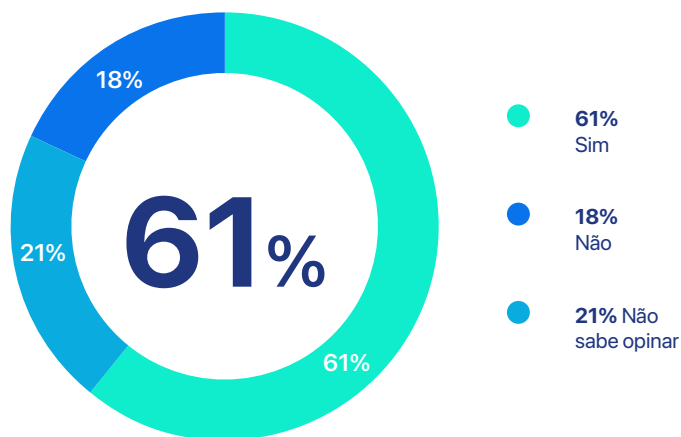
Quarto bloco de resultados da pesquisa quantitativa

D

I. A inteligência artificial é pauta do Conselho de Administração?

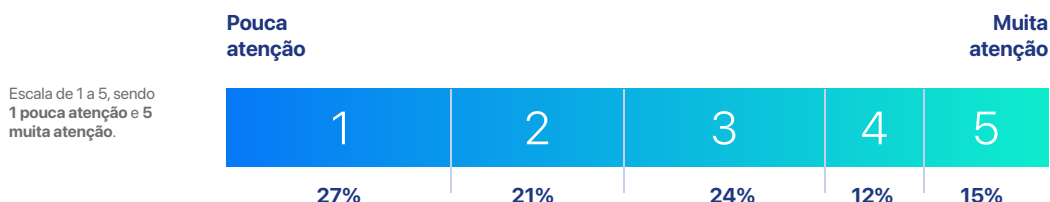


II. O Conselho de Administração reconhece a IA como estratégica para a sustentabilidade do negócio?





III. Quão atentos estão o conselho e os diretores da organização aos processos internos de governança e gestão de riscos no uso de IA?



Comentários com base na pesquisa qualitativa

Não observamos o envolvimento direto dos conselhos e dos C-Level nas iniciativas de IA das companhias. No Grupo 1 detectamos algum nível de reporte e acompanhamento dos projetos ao conselho e ao C-Level, embora tímidos e não estruturados, particularmente no âmbito do conselho. Predomina o interesse comercial, ou seja, quais oportunidades a inteligência artificial oferece para incrementar o negócio, e quase nenhuma preocupação efetiva sobre os riscos associados — menos ainda sobre a percepção de responsabilidade do conselho e do C-Level em garantir o bom desenvolvimento e uso da IA na companhia.

No Grupo 3 identificamos um tênue acompanhamento macro, sem adentrar no âmbito de cada iniciativa. Nos grupos 2 e 4 não identificamos interação estruturada sobre o uso de IA entre as áreas de negócio, de tecnologia e a alta gestão.

As respostas ao questionário on-line são intrinsecamente contraditórias: 55% dos respondentes afirmaram que a inteligência artificial é pauta do conselho; enquanto 61% dos conselhos reconhecem a IA como estratégica para a sustentabilidade do negócio. Por outro lado, 72% dos conselhos e dos C-Level não estão muito atentos aos processos de adoção e de gestão de risco de IA nas suas companhias.



Exemplos ilustrativos

Grupo 1

Em companhia do setor de energia, a formação e experiência profissional progressiva do CEO em tecnologia é um elemento chave para fomentar e disseminar a cultura de IA, assegurando a colaboração entre áreas de desenvolvimento de projetos, e interlocução com o conselho. O conselho de uma companhia com diversos produtos baseados em dados e uso de IA, desenvolve estudos sobre o potencial da tecnologia visando promover ajustes no planejamento estratégico, incluindo a previsão de investimento para os próximos anos.

Desafios

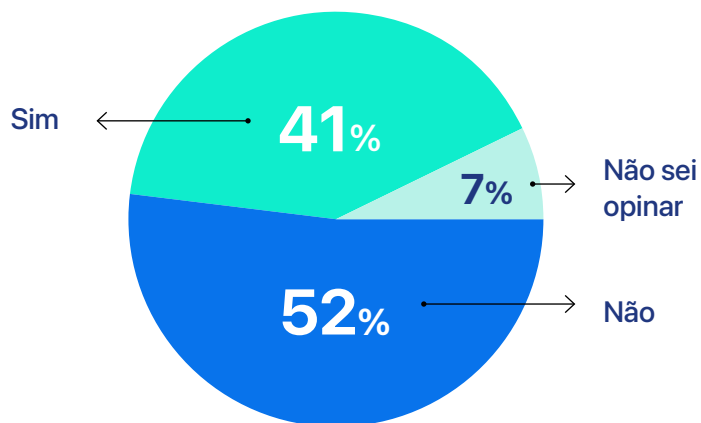
Capacitar em IA os membros do conselho e da alta administração; diversificar a composição do conselho (diversidade cognitiva); criar canais de comunicação eficientes entre o C-Level e o conselho, consolidando a liderança da alta gestão nas estratégias de inteligência artificial; e estreitar a interação com as áreas no desenvolvimento e/ou na adoção de soluções de IA.



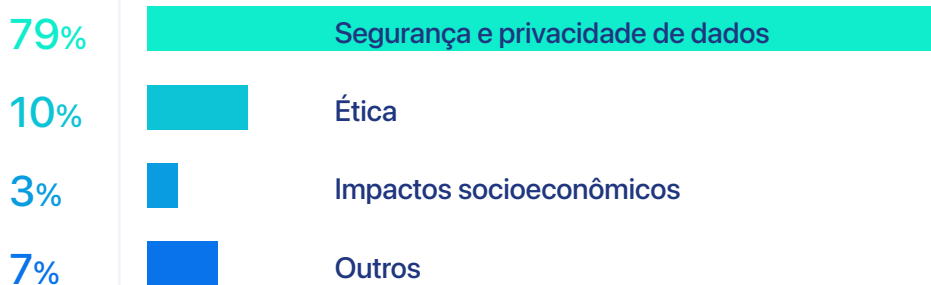
Quinto bloco de resultados da pesquisa quantitativa

E

I. Existe na organização estratégia de gestão de riscos em relação à IA?



II. Se sim, quão efetivas são, especialmente nas fases de desenvolvimento e implementação de sistemas de IA



Comentários com base na pesquisa qualitativa

Nas companhias do Grupo 1 identificamos movimentos objetivando a formação de comitês, grupos de trabalho ou fóruns para a elaboração de programas de governança e risco com formulação de políticas e códigos de conduta. A partir de casos de uso de IA, são ainda raros os protocolos envolvendo as áreas de risco, compliance e jurídico.



No Grupo 1 o monitoramento de riscos é predominantemente de responsabilidade da TI, com estruturas relativamente consolidadas de governança de segurança da informação e governança de dados. No grupo 3 percebe-se uma falta de discernimento (conhecimento) sobre os riscos associados aos modelos de IA, inclusive os impactos da inteligência artificial na segurança da informação e na proteção de dados. Nos grupos 2 e 4, não observamos estruturas para análise de impacto e de risco das soluções de IA nem o envolvimento das áreas de compliance, risco ou jurídica.

Nesse caso, o resultado da pesquisa qualitativa é coerente com o resultado da pesquisa quantitativa (questionário on-line): 41% afirmam ter estratégias de gestão de risco, mas quando perguntados sobre as preocupações, 79% dos respondentes mencionaram segurança e privacidade dos dados e apenas 10% as questões éticas.

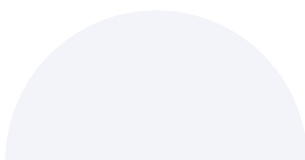
Em um cenário pré-regulação, e com as companhias na fase de exploração e validação das primeiras iniciativas, observamos o início tímido de estruturação dos programas de governança e gestão de riscos de IA, principalmente no Grupo 1.

Exemplos ilustrativos

Grupo 1

Companhia do setor de telecomunicações constituiu um grupo de trabalho, ainda não formalizado e liderado pela área de TI, para a governança e a análise de riscos de projetos pontuais, reportando diretamente para o conselho e o C-Level. Esse mesmo grupo formulou as diretrizes para o uso de LLMs pelos colaboradores, com foco principal no não compartilhamento de dados pessoais e de negócios, e na supervisão e correção dos resultados.

Companhia do setor financeiro elaborou um protocolo para o desenvolvimento de projetos para concessão e prevenção à fraude, envolvendo as áreas de negócios, jurídico e compliance, lideradas pela TI.



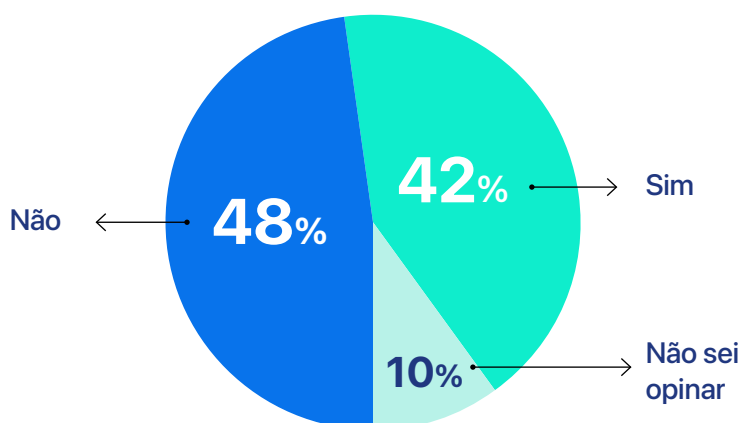
Desafios

Sensibilizar o conselho e o C-Level sobre a importância de estruturar governança e gestão de riscos de IA para a sustentabilidade do negócio. Envolver a área de risco, compliance e jurídico nas análises.

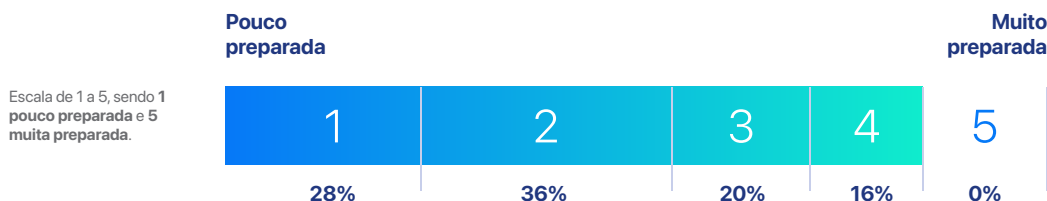
Sexto bloco de resultados da pesquisa quantitativa

F

I. Existe uma instância ou um órgão na organização para avaliar questões éticas relacionadas ao uso de IA?



II. A organização está preparada para enfrentar as discussões sobre dilemas éticos relacionados a IA?





Comentários com base na pesquisa qualitativa

Os resultados do questionário on-line conflitam não apenas com o cenário observado nas entrevistas qualitativas, como descrito ao longo do relatório, mas igualmente entre si: 42% dos respondentes afirmam possuir um órgão para avaliar as questões éticas envolvidas nos sistemas de IA, sendo 38% de estruturas próprias para essa finalidade. No entanto, 64% consideram que essas estruturas estão despreparadas para exercer suas funções. Observamos a utilização de estruturas de conformidade ética não especializadas em IA.

Exemplos ilustrativos

Grupos 1 e 3

Companhia do setor financeiro criou um fórum interno para avaliar projetos pontuais, inclusive com a possibilidade de participação de especialistas de mercado, enquanto empresas de tecnologia tem contado com o apoio de estruturas criadas por associações especializadas.

Desafios

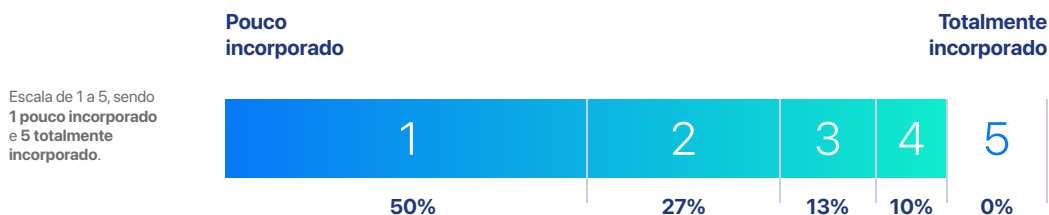
Capacitar os colaboradores sobre a lógica e o funcionamento da IA; estabelecer estruturas de gestão de riscos regulatórios e éticos diversificadas cognitivamente, e capacitá-las para o exercício da função.



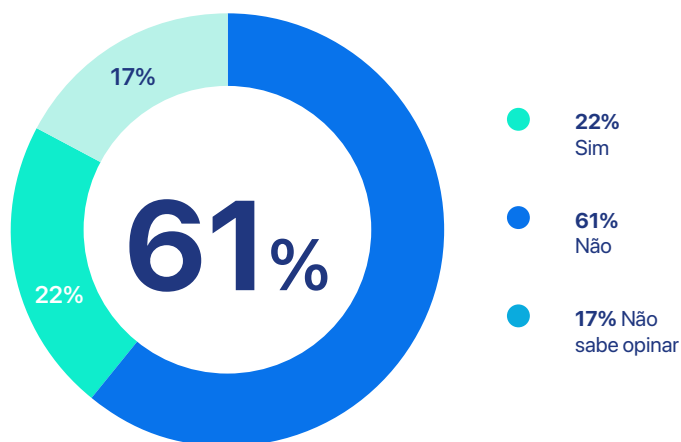
Sétimo bloco de resultados da pesquisa quantitativa



I. Em que medida a organização já incorpora os impactos da IA na agenda de ESG?

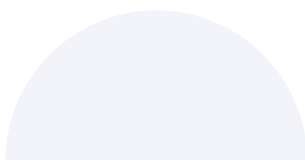


II. Isso já se reflete nos relatórios de sustentabilidade?



Comentários com base na pesquisa qualitativa

As entrevistas qualitativas sinalizaram que o debate e a adoção da agenda ESG ainda não é uma pauta consolidada, menos ainda o entendimento dos impactos da inteligência artificial sobre o meio ambiente (E - “environmental”), o social (S - “social”), e a gestão (G - “governance”). Neste sentido, alinha-se com os resultados dos questionários on-line, em que 77% dos respondentes indicaram baixa consciência da companhia.



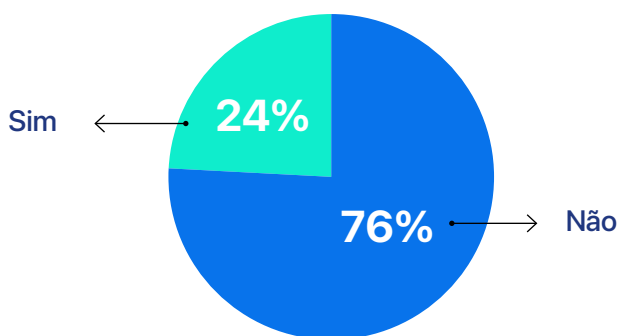
Desafios

Sensibilizar a companhia, começando pela alta gestão, sobre o potencial da inteligência artificial de impulsionar a agenda ESG; contemplar nas futuras estruturas de governança e gestão de riscos instrumentos para mapear e monitorar os efeitos ambientais e sociais dos sistemas de IA desenvolvidos e/ou adotados; e promover a inserção de indicadores associados ao uso responsável de IA nos relatórios de sustentabilidade.

Oitavo bloco de resultados da pesquisa quantitativa

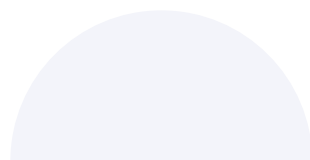


I. Existe um programa de sensibilização dos colaboradores em governança de IA?



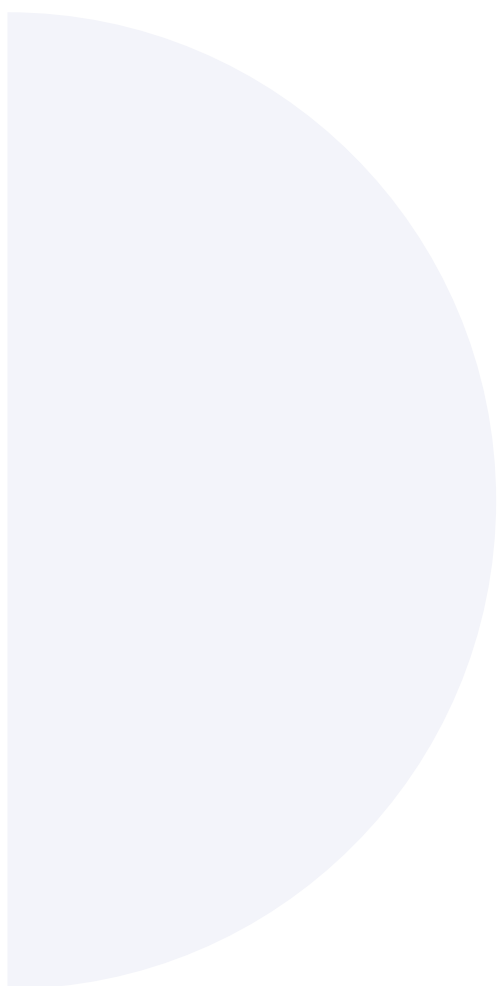
Comentários com base na pesquisa qualitativa

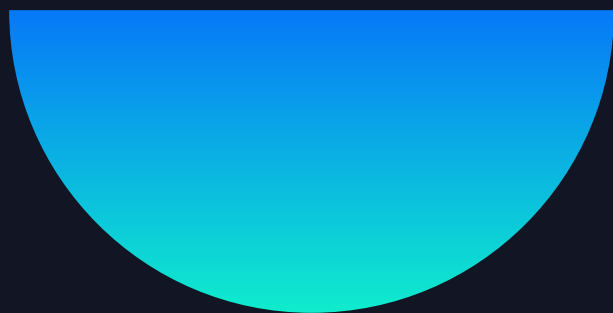
No Grupo 1 observamos iniciativas pontuais de capacitação sobre governança e ética da IA para as equipes de TI, e praticamente nenhuma iniciativa de programas de capacitação para os colaboradores das demais áreas. Pesquisa qualitativa e quantitativa convergentes.



Desafios

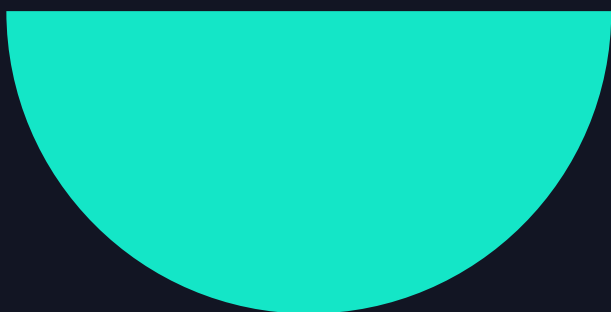
Estruturar programas de capacitação em inteligência artificial para todos os colaboradores, com conteúdos abrangentes e diversificados, contemplando aspectos técnicos e de governança, inserindo-os na prática de *longlife learning* da companhia.





III . Relatório

2. RECOMENDAÇÕES





2. Recomendações

Os modelos de inteligência artificial não são plug and play. Para aproveitar plenamente seus benefícios requer preparar previamente a companhia. Construir uma estratégia de IA envolve:

- Declarar o problema, identificando quais desafios podem ser melhor equacionados com sistemas de inteligência artificial, avaliando custo e benefício;
- Estruturar os dados gerados nas diversas etapas operacionais e nas interações com seus stakeholders, adaptando os modelos de mercado às particularidades de seu negócio, zelando pela qualidade e respeitando os requisitos da LGPD e a segurança da informação;
- Garantir suporte, manutenção e atualização contínua dos sistemas (gestão de riscos);
- Formar times mesclando profissionais de diferentes formações e áreas, incluindo TI, negócios e áreas de apoio (jurídico, compliance, administrativo, financeiro e ambiental), entre outras, para a colaboração no desenvolvimento e na adoção de sistemas de inteligência artificial de forma responsável e sustentável;
- Prover a alta direção - conselho e C-Level - de capacitação para exercer a liderança da transformação digital orientada por inteligência artificial;
- Construir indicadores de desempenho (KPIs) para apurar o retorno do investimento em IA considerando o custo de: desenvolvimento e/ou adaptação de soluções de mercado; capacitação dos colaboradores; recursos de gestão de risco; custos de energia e outros insumos no processamento dos sistemas;



- Identificar e mitigar os impactos ambientais e sociais dos sistemas;
- Estruturar programas de governança de IA em conformidade com a cultura e a ética da companhia e com os regulamentos e normas gerais e setoriais;
- Estabelecer um ecossistema de parcerias estratégicas de longo prazo.

Perguntas iniciais a serem respondidas pelo red-teaming:

- Qual a taxonomia de riscos de IA e como esses riscos se enquadram no apetite de risco geral da companhia?
- Como o sistema ou o modelo de IA em avaliação se qualifica em relação a risco segundo os parâmetros indicados do PL2338/23 em tramitação no Senado, tendo em vista que cada uma das categorias de risco tem obrigações específicas para o desenvolvedor, o distribuidor e o aplicador?
- Quais os processos técnicos e barreiras legais e éticas existentes para documentar e abordar as vulnerabilidades identificadas?

2.1. Programas de Governança e Gestão de Riscos de IA

Para o endereçamento dos desafios listados recomenda-se a formação de um ambiente adequado, uma estrutura própria ou a adaptação das estruturas existentes, considerando:

- Envolver a alta direção;
- Estabelecer diretrizes, políticas e procedimentos sobre o desenvolvimento e a adoção de IA, alinhados com a ética, a cultura, os valores, o propósito e o apetite de risco, além das normas e regulamentações gerais e setoriais;



- Supervisionar e controlar os benefícios e riscos, pela aplicação das diretrizes de governança e procedimentos correspondentes, ao longo de todo o seu ciclo de vida e de forma constante;
- Supervisionar os aspectos sociotécnicos no desenvolvimento e na implementação das soluções, especialmente em relação aos dilemas éticos, integrando nesse programa um Comitê de Ética próprio ou se valendo de outra estrutura organizacional adaptada para essa finalidade, assim como um Código de Conduta;
- Construir um ecossistema ativo de parceiros, diversificados cognitivamente, complementando os talentos internos e compartilhando boas práticas de mercado;
- Capacitar as equipes sobre a natureza, as características e as possibilidades trazidas pela adoção de IA nas diversas áreas, estruturando programas de formação e treinamento.

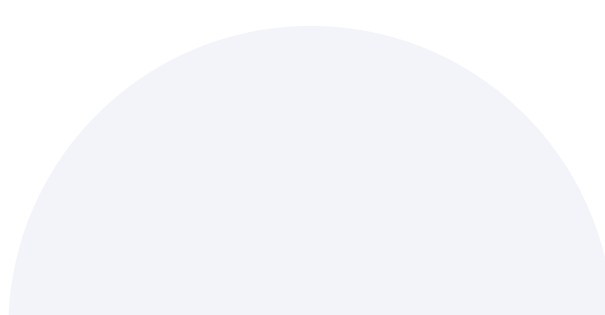
O Programa de Governança e Gestão de Riscos é um dos pilares para o desenvolvimento e uso responsável e sustentável da inteligência artificial, garantir sua conformidade legal e ética, e promover a sustentabilidade dos projetos de IA e da própria companhia. Seus objetivos são: criar, adaptar, liderar, controlar e monitorar os seus elementos principais, fazer os ajustes e mitigar os riscos envolvidos, e instaurar um Comitê de Ética para assessorar e subsidiar as decisões (trade-offs) ao longo do ciclo de vida das soluções.



Diagrama 2 Programa de Governança e Gestão de Riscos de IA



Com base no Diagrama 2, criado a partir de pontos comuns identificados entre projetos legislativos, frameworks de governança e discussões em fóruns especializados, entre os principais tópicos do programa estão a definição, a construção, a documentação e o monitoramento dos seguintes pontos:





Principais tópicos do Programa de Governança e Gestão de Riscos de IA

01

Propósito do sistema: a que se destina

Principais objetivos e funcionalidades

02

Dados: qualidade e adequação dos dados de treinamento e produção

Observância à LGPD e aos princípios de finalidade, adequação e necessidade, avaliando a representatividade e a diversidade das bases de dados, entre outras características que podem impactar em efeitos ou saídas indesejadas, como a replicação de desigualdades sociais históricas e de discriminações abusivas ou ilegais.

03

Avaliação de impacto: analisar os potenciais impactos das soluções nas pessoas, nos grupos, na sociedade e na própria companhia

Riscos relacionados aos direitos fundamentais, saúde e segurança, para identificar os efeitos negativos e promover ações de mitigação.

04

Avaliação de performance: testes de acurácia e de resultados dos modelos antes de colocar as soluções em produção

Monitorar o ciclo de vida, para identificar o que deve ser feito de forma recorrente após sua implantação e uso, para garantir a qualidade e a mitigação dos riscos, assim como as melhorias, correções e ajustes ao longo do tempo.

05

Documentação: etapas do ciclo de vida dos projetos

Documentação dos dados, tipo de tecnologia, riscos, performance dos modelos, robustez e segurança, análise das saídas e decisões (trade-offs), tanto para prover as informações necessárias para a transparência e auditabilidade dessas soluções como para a sua explicabilidade, com duas vertentes, no sentido de (a) melhorar o entendimento de funcionamento para a supervisão humana voltada para implementar melhorias e (b) permitir a verificação e a correção de problemas (relação homem-máquina) por parte dos usuários.

06

Supervisão humana: interação homem-máquina em todas as fases

Em todas as etapas do ciclo de vida, tendo em vista a natureza e limitações da IA.

07

Princípios éticos organizacionais: definição dos pilares éticos da companhia relacionados ao uso responsável e sustentável de IA

Com base nos princípios éticos gerais e organizacionais, e alinhados com os valores e o planejamento estratégico da companhia, para nortear o programa de governança, o Comitê de Ética de IA e colaborar com as diretrizes para a mitigação de riscos e efeitos indesejados, assim como a elaboração de códigos de conduta.

08

Parcerias e relacionamento com terceiros: fortalecer ciclo de parceiros e definir protocolo de gestão de terceiros

Especialmente para avaliação de riscos, governança de dados, segurança da informação e demais aspectos de conformidade na gestão de terceiros.

09

Políticas e normativos internos: criar ou adaptar políticas e normativos para o uso de IA e suas especificidades, assim como códigos de conduta

De forma integrada com diretrizes regulatórias, normas setoriais, e demais normativos ou políticas da companhia

10

Cultura e treinamento: programa de capacitação de profissionais em diferentes níveis para desenvolvimento e uso de IA

Contemplar os times envolvidos diretamente no desenvolvimento e na implementação dos projetos de IA, como gestores, áreas de negócios e áreas de suporte.

Fonte: Kaufman e Zavaglia, 2024.



A adoção de um dos frameworks de governança de IA publicados, como o NIST ou o ISO/IEC, ou a elaboração de um framework próprio, mesmo que inspirado nos existentes, maximiza a relação homem-máquina e promove a inovação com sustentabilidade. A governança de IA adequada para cada perfil ou tipo de uso proporciona diversos benefícios às companhias.

Benefícios da Governança ativa de soluções de IA para uso responsável e sustentável

Governança de IA

benefícios

- Integra as áreas no desenvolvimento de novos produtos, serviços e modelos de negócio com o uso de IA.
- Maximiza os resultados e a identificação e mitigação de riscos.
- Proporciona transparência e melhora a confiança dos investidores e clientes.
- Garante a participação e supervisão humana para maximizar resultados e permitir ajustes e correções de forma recorrente.
- Define os papéis e responsabilidades.
- Compila obrigações legais e éticas.
- Conecta com responsabilidade e sustentabilidade corporativa (ASG).

Fonte: Kaufman e Zavaglia, 2024.

O Programa de Governança e Gestão de Riscos, com as subsequentes normativas e protocolos de implantação, acompanhamento e monitoramento, traz desafios culturais e organizacionais, e necessita da liderança do conselho e do C-Level para integrá-lo ao planejamento estratégico.



Desafios para implantação Programas de Governança e Ética de IA

Governança de IA

desafios

- Integrar os projetos habilitados por IA no planejamento estratégico, como motor de transformação e sustentabilidade do negócio.
- Mitigar a assimetria de informação e conhecimento entre gestores e técnicos.
- Eliminar iniciativas isoladas sem apoio institucional.
- Definir e valorizar métricas não diretamente ligadas a desempenho, superando a suposição de que governança atrapalha o desempenho.
- Aumentar o nível de transparência dos sistemas de IA, reconhecendo o limite imposto pela opacidade da própria técnica.
- Construir evidências sobre os benefícios da governança, e das estruturas de gestão associadas.
- Formar e capacitar gestores e colaboradores com as novas habilidades e competências.
- Conectar com os princípios e diretrizes da ASG.

Fonte: Kaufman e Zavaglia, 2024.



2.2. Conselho de Administração

A inteligência artificial é estratégica para a sustentabilidade do negócio, logo deve compor a pauta de prioridades do conselho, o que requer adquirir familiaridade com os meandros da tecnologia e criar canais permanentes de atualização, particularmente em relação à concorrência (direta e indireta). Aconselhável estabelecer uma agenda regular com o C-Level sobre o progresso, as barreiras e os desafios do processo de desenvolver e incorporar a IA, assim como incluir especialistas em IA no ecossistema de parceiros estratégicos, convidando-os para apresentações no conselho com múltiplas abordagens (ética, economia, trabalho, governança, regulamentação, tecnologia, meio ambiente e social).

2.3. Chief Artificial Intelligence Officer (CAIO)

Segundo o LinkedIn, nos últimos cinco anos triplicou o número de diretores de inteligência artificial (CAIOs), tendência referendada pelo Financial Times como um esforço das companhias de capitalizar a tecnologia¹⁷. Nos EUA, esse movimento é, em parte, reação à “Ordem Executiva” da Casa Branca que recomenda às agências federais americanas designarem diretores de IA para garantir responsabilização, liderança e supervisão da tecnologia. O perfil desses profissionais combina conhecimento técnico robusto e visão de negócios e aspectos jurídicos, ou seja, perfil multidisciplinar em campos como IA, ciência da computação, estatística, análise de dados, ética, conformidade regulatória, vivência no setor, impactos sociais, além da capacidade de liderar o processo de disseminação da tecnologia na companhia. O desafio de encontrar esse perfil pode ser compensado por uma equipe multidisciplinar.

O cargo de CAIO não é uma unanimidade. Parte do mercado, ao reconhecer o papel estratégico da IA, defende que a tecnologia deve se disseminar na companhia sem centralização em um responsável. De qualquer forma, se for criado o cargo, prevalece o consenso de que o CAIO, obrigatoriamente, deve interagir com funções como CIO (Chief Information Officer) e CTO (Chief Technology Officer) e fazer parte do grupo de C-Level.

¹⁷ SHEAD, S. Rise in number of chief AI officers. LinkedIn, 15 abr. 2024. Disponível em: <https://www.linkedin.com/news/story/rise-in-number-of-chief-ai-officers-5981036/>. Acesso em: 15 de agosto de 2024.



Se constituído, cabe ao CAIO liderar a elaboração do plano estratégico para adoção e supervisão da IA, em conjunto com a alta direcção.

2.4. Parcerias estratégicas

O grau e a extensão dos desafios associados aos sistemas de inteligência artificial, em geral, extrapolam as competências internas da companhia e/ou requerem consumo intensivo de tempo e recursos. Uma alternativa são as parcerias estratégicas, sólidas e confiáveis, com fornecedores e especialistas de inteligência artificial, preservando a independência e os limites sobre a propriedade intelectual (PI). O critério é seleccionar parceiros com potencial de agregar conhecimento, acesso às inovações, velocidade de implementação, compartilhamento de dados, escalabilidade — volumes crescentes de dados, interações com usuários, expansão de canais — e interoperabilidade entre sistemas e modelos. Parceria colaborativa pressupõe planejamento compartilhado, cocriação das soluções, previsão e monitoramento dos riscos das soluções, e conformidade regulatória não apenas no uso de IA, mas também nas questões relacionadas à segurança da informação e à proteção de dados pessoais.

2.5. Comitê de Ética de IA

A diversidade das soluções de IA imputa às companhias novos desafios principiológicos, que extrapolam a conformidade legal. É crítico adequar a cultura e os valores corporativos ao inédito ambiente de negócio caracterizado por agilidade, volatilidade e tecnologias complexas. Além das questões técnicas, emergem as discussões sobre os aspectos sociotécnicos, relacionados aos reflexos negativos ou indesejáveis dessa tecnologia nas pessoas, nos grupos e na sociedade, especialmente em relação aos dilemas éticos, como a amplificação de desigualdades e discriminações sociais históricas.

A contextualidade de cada aplicação e os desafios para a identificação e a solução dessas demandas têm impulsionado a criação de Comitês de Ética de IA ou a adaptação de outras estruturas organizacionais para essa finalidade. Um comitê diversificado



cognitivamente acelera o processo de tomada de decisão — novas perspectivas, formas de pensar e de reagir às situações, e uma variedade de habilidades e repertórios.

2.6. Capacitação dos colaboradores

O letramento em inteligência artificial está previsto na Lei de IA europeia (AI Act) e no projeto de lei em tramitação no Congresso brasileiro (PL 2338/2023), ou seja, tende a se constituir em uma obrigação legal. O AI Act introduz o conceito de alfabetização em IA, definido como as habilidades, o conhecimento e a compreensão necessários para tomar decisões informadas sobre os sistemas de inteligência artificial; inclui a capacidade de avaliar as oportunidades e os riscos associados aos sistemas de IA; e ressalta os danos potenciais que esses sistemas podem causar aos usuários e à sociedade. O Artigo 4 do AI Act explicitamente exige que os fornecedores e implementadores de sistemas de IA garantam um nível suficiente de alfabetização em IA em suas equipes e outros agentes relevantes. O treinamento deve ser adaptado ao conhecimento técnico, à experiência, à educação e às funções específicas do colaborador envolvido.

No PL2338/23 a capacitação está prevista no Art. 56, II, V, que estabelece que os programas de capacitação devem assegurar uma compreensão holística da implementação da IA, abrangendo aspectos técnicos, éticos e sociais.

Independentemente das obrigações legais, equipes capacitadas para aproveitar os benefícios da tecnologia, equilibrando a conformidade com o potencial de inovação, representam impactos positivos sobre os resultados da companhia.



Parte 4

CONSIDERAÇÕES FINAIS





IV. Considerações finais

O campo da inteligência artificial, que teve seu início em 1956, levou mais de meio século para que as condições essenciais emergissem, especialmente a disponibilidade de grandes volumes de dados e poder computacional, resultando em avanços significativos na década de 2010. Levou, contudo, apenas alguns anos para a IA se tornar uma força motriz de transformações globais, capturando a atenção de investidores, governos, mercados e mídia. Com o advento das soluções de IA generativa, a tecnologia alcançou um novo patamar, apesar de ainda estar distante do que parte da indústria sugere.

A inteligência artificial agrega valor substancial aos processos de decisão ao ser capaz de identificar padrões invisíveis nos dados e usá-los para fornecer insights e análises, previsões e recomendações. A IA não é uma tecnologia, mas um conjunto de técnicas e soluções com potencial de ampliar as capacidades humanas.

Adotar a inteligência artificial é uma jornada, com tentativas e erros, experimentando novas soluções e comparando-as com as soluções tradicionais, sempre preservando a relação entre investimento e retorno. É longo o caminho a percorrer entre a etapa de experimentação e testes e a efetiva entrada em produção, o que requer transformações significativas na cultura, na linha de comando, na organização dos dados, nos processos, nos fundamentos éticos, no apetite ao risco, nos investimentos e na capacitação dos colaboradores. Definir o problema de negócio adequado é vital para evitar expectativas irrealistas e, conseqüentemente, frustrações.

A inteligência artificial não resolve todo e qualquer problema, mas pode ser solução para alguns problemas, particularmente os que têm ganhos de eficiência e escala com base em extrair insights de grandes volumes de dados, ou para automatizar processos complexos e melhorar a interação.

Como ponto de partida na construção de uma estratégia de IA é importante considerar o grau de maturidade da companhia em relação a essa tecnologia; as soluções já adotadas para enfrentar os desafios em questão; os sistemas legados; o nível de engajamento dos executivos, particularmente os C-Level; a governança e



gestão de riscos apropriadas; a aderência à cultura “data-driven”; a disponibilidade de talentos; a integração das soluções de IA aos processos de negócio; e o suporte das áreas jurídicas e de compliance, entre outros. A concorrência tende a se acirrar, pois o potencial desses modelos é um atrativo para startups e empresas inovarem com novas ofertas habilitadas por inteligência artificial.

Para maximizar os benefícios da inteligência artificial, é fundamental: a) garantir que todos os colaboradores obtenham um entendimento básico da lógica e do funcionamento da tecnologia; b) construir um repositório centralizado de conhecimento para permitir compartilhamento e colaboração constantes; c) dotar a companhia de estruturas e processos flexíveis que permitam adaptação e mudança contínua; e c) incorporar a IA nos modelos de operação e de negócio (core business). Para promover a adoção responsável da inteligência artificial é crucial diminuir a assimetria de conhecimento entre gestores e equipes de TI, capacitando os líderes e os colaboradores para lidar com um ambiente de negócio impulsionado por sistemas habilitados por IA.

É imprescindível construir estruturas de governança sólidas, sedimentadas na expertise da natureza e no funcionamento das várias técnicas e soluções de inteligência artificial. Migrar a companhia para “IA – driven” requer a liderança do conselho e dos C-Level, efetivamente inserindo a IA no planejamento estratégico da companhia.



Pressupostos de IA	
01	Sendo a IA transversal/tecnologia de propósito geral inter e multidisciplinar, requer equipes com esse perfil (ética by design).
02	Como toda tecnologia, a IA é social e humana, e seus efeitos dependem de como os seres humanos a inserem nos ambientes técnico-sociais.
03	São várias camadas de mediação da IA na comunicação e na sociabilidade (quanto mais complexo o sistema, maior o potencial dano).
04	Como modelo estatístico, a IA tem uma variável de incerteza intrínseca (migramos de um mundo de máquinas programadas para um mundo de máquinas probabilísticas).
05	Quando falamos de ética da IA estamos falando em sistemas em observância com os princípios éticos e legais da sociedade.
06	Nem tudo é IA, o desafio é identificar a especificidade da tecnologia.
07	Cada implementação tem suas particularidades, éticas e regulatórias.
08	Difícil definir o que seja IA, mais fácil definir o que não é IA.

Fonte: Kaufman e Zavaglia, 2024.



Os líderes precisam estar cientes de que as características que permitem à IA processar rapidamente grandes quantidades de dados e executar certas tarefas com mais precisão ou eficiência do que os profissionais humanos, também aumentam sua opacidade. O trade-off entre precisão e explicabilidade implica em avaliações cuidadosas para determinar quando e como valorizar um ou outro, sempre sob supervisão humana.

A inteligência artificial é estratégica porque a escala, a quantidade de dados gerados, o escopo, a complexidade e o dinamismo nos negócios atuais são tão extremos que nós humanos não conseguimos mais gerenciá-los sem os recursos dos sistemas de IA. As companhias que compreenderem o significado dessa mudança e agirem rápido, terão uma vantagem competitiva considerável.



GLOSSÁRIO





V. Glossário

Algoritmo	conjunto de instruções ou sequência de tarefas para alcançar um cálculo ou um resultado específico. Necessita ser preciso e suficientemente não ambíguo para ser executado por um computador.
Alinhamento	o processo de garantir que os objetivos e comportamentos de um sistema de IA estejam alinhados com os objetivos, valores e intenções do desenvolvedor.
Aprendizado de máquina (machine learning)	subcampo da inteligência artificial, é um sistema ou uma máquina (ou mesmo um software) que pode aprender com base em um processo de aprendizado computacional e estatístico, distinto do aprendizado humano. Com base em correlações em grandes conjuntos de dados, os algoritmos de aprendizado detectam padrões ou regras nos dados e projetam cenários futuros.
Arquitetura Transformer	arquitetura de redes neurais de aprendizado profundo (deep learning) na base de algumas soluções de IA generativa.
Auditoria de Ecossistemas (ecosystem audit)	avaliação ampla de um sistema de IA e do ecossistema circundante. As auditorias de ecossistemas podem considerar modelos de IA, seus dados de treinamento e as circunstâncias de sua implantação que envolvem a prática operacional.
Autonomia/autônomo	capaz de operar, agir ou tomar decisões sem a intenção expressa ou supervisão de um ser humano.
Avaliações	avaliações sistemáticas de desempenho, capacidade ou impacto potencial de um sistema de IA. As avaliações podem incluir benchmarking, red-teaming e auditorias.



Base de dados tendenciosa

ocorre quando os dados não representam a composição proporcional do universo em questão, refletem os preconceitos existentes na sociedade ou apresentam problemas em sua rotulagem.

Big data

uso de grandes e diversificados conjuntos de dados em análises preditivas para entender padrões, tendências e comportamentos. O big data estabelece correlações entre os dados (não causalidades).

Black-box

em analogia ao sistema de registro de voz e dados dos aviões, refere-se ao desconhecimento pelos seres humanos de como os algoritmos na técnica de redes neurais profundas estabelecem as correlações nos dados e geram os resultados; ou como transformam os dados de entrada (inputs) em dados de saída (outputs). Expressões com o mesmo significado: "questão da interpretabilidade", "opacidade do sistema" e "não explicabilidade".

Capacidades

a gama de tarefas ou funções que um sistema de IA pode executar e a proficiência com que pode executá-las.

Ciclo de vida da IA

todos os eventos e processos relacionados com a vida útil de um sistema de IA, desde a concepção até o desmantelamento, incluindo investigação, formação, desenvolvimento, implantação, integração, operação, manutenção, venda, utilização e governança.

Cluster

coleção de pontos de dados agregados devido a certas semelhanças. Os dados parecem ser "reunidos" em torno de um valor específico. Termo associado à segmentação de usuários nas redes sociais ou em qualquer outro processo de agregar usuários com perfis similares.



Database	software para armazenamento e processamento eficiente de informações representadas digitalmente.
Data Broker	empresas cujo modelo de negócio é comprar e vender dados sem o consentimento explícito dos proprietários.
Desenvolvedores de IA	organizações ou indivíduos que projetam, constroem, treinam, adaptam ou combinam modelos e aplicações de IA.
Enviesamento estatístico	quando o sistema exibe um erro sistemático no resultado.
Ética por design (ethics by design)	abordagem que visa integrar a ética na fase de design e desenvolvimento da tecnologia. Termos em inglês com o mesmo significado: “embedding values in design”, “value-sensitive design” e “ethically aligned design”.
FLOPS	Floating-point Operations Per Second” (“Operações de Ponto Flutuante por Segundo”), uma medida do poder de computação de um computador.
IA simbólica (symbolic AI)	incorporação explícita do conhecimento e do comportamento humanos nas máquinas por meio de linguagens formais e regras de inferência lógica.
IA confiável (trustworthy AI)	IA confiável pelos seres humanos. As condições para tal confiança podem se referir a princípios éticos como dignidade humana, respeito aos direitos humanos e assim por diante; e/ou a fatores sociais e técnicos que influenciam se as pessoas vão querer usar a tecnologia. O uso do termo “confiança” em relação às tecnologias é controverso.



IA restrita (narrow AI)	sistema de IA que funciona bem apenas em uma única tarefa ou em um conjunto restrito de tarefas.
Implantação (deployment)	o processo de lançamento de um sistema de IA em um ambiente do mundo real, como um sistema de IA voltado para o usuário final e/ou consumidor/cliente.
Implementadores de IA	qualquer indivíduo ou organização que utilize um sistema de IA para fornecer um produto ou serviço. A implantação pode ser “interna”, em que um sistema é usado apenas pelos desenvolvedores, ou “externa”, permitindo que o público ou outras entidades não desenvolvedoras o utilizem.
Input (para um sistema de IA)	dados ou prompts inseridos em um sistema de IA para serem processados e produzir uma saída.
Inteligência artificial – IA (artificial intelligence – AI)	existem múltiplas definições de IA, nesse relatório adotamos a definição do PL 2338/2023: “sistema baseado em máquina que, com graus diferentes de autonomia e para objetivos explícitos ou implícitos, infere, a partir de um conjunto de dados ou informações que recebe, como gerar resultados, em especial, previsão, conteúdo, recomendação ou decisão que possa influenciar o ambiente virtual, físico ou real”.
Inteligência Geral Artificial (Artificial General Intelligence – AGI)	potencial futuro sistema de IA que iguala ou supera o desempenho humano em todas ou quase todas as tarefas cognitivas. Várias empresas de IA declararam publicamente o seu objetivo de construir AGI. No entanto, o termo não tem uma definição universalmente acordada com precisão.
Interface de Programação de Aplicativos (Application Programming Interface – API)	conjunto de regras e protocolos que permite a integração e a comunicação entre sistemas de IA e outros aplicativos de software.



GPU – Graphics Processing Unit
(Unidade de Processamento Gráfico)

peça de hardware de computador montada a partir de semicondutores amplamente utilizada como fonte central de poder computacional para IA de uso geral. As GPUs foram originalmente projetadas para aplicativos de renderização gráfica.

Guardrails

restrições ou limites de segurança predefinidos estabelecidos na tentativa de garantir que um sistema de IA opere dentro dos parâmetros desejados, evitando resultados indesejados ou prejudiciais.

Large Language Model (LLM)

sistemas de inteligência artificial que podem gerar e reconhecer texto, e analisar e processar dados de linguagem natural. Eles são treinados em grandes conjuntos de dados.

Modelos de Fundação (Foundation Models)

modelos de aprendizado de máquina treinados em grandes quantidades de dados que podem ser adaptados a uma ampla gama de casos de uso.

Pesos (weights)

parâmetro aprendível de alguns modelos de aprendizado de máquina, incluindo redes neurais. Dentro de uma rede neural, os pesos controlam o sinal, ou a força de um sinal entre dois neurônios. Um peso descreve quanta influência uma entrada terá em uma saída. Os pesos podem ser otimizados durante os processos de aprendizado de máquina.

Pré-treinamento

primeira etapa do desenvolvimento de um modelo de IA, no qual os modelos aprendem com grandes quantidades de dados. O pré-treinamento é a parte do treinamento de IA que requer mais dados e recursos computacionais.



**Reconhecimento de
faces**
(face recognition)

sistema de reconhecimento da identidade de uma pessoa a partir de suas imagens de rosto capturadas por uma câmera, utilizado dentre outros por sistemas de vigilância e identificação.

**Redes neurais
profundas**
(deep learning)

técnica de aprendizado de máquina que representa o mundo como uma hierarquia de conceitos, sendo que cada conceito é definido em termos de conceitos mais simples, como representações mais abstratas computadas em termos de outras menos abstratas. É uma forma de aprendizado de máquina que usa redes neurais com várias camadas de neurônios artificiais (unidades de processamento simples interconectadas que interagem).

Riscos de IA

a combinação da probabilidade de ocorrência de danos decorrentes do desenvolvimento ou da implantação de modelos ou sistemas de IA e a gravidade desses danos.

**Sistema de inteligência
artificial de propósito
geral**
(SIAGP)

sistema de IA baseado em um modelo de IA treinado com bases de dados em grande escala, capaz de realizar uma ampla variedade de tarefas distintas e servir diferentes finalidades, incluindo aquelas para as quais não foram especificamente desenvolvidos e treinados, podendo ser integrado em diversos sistemas ou aplicações (definição do PL 2338/2023).

Superinteligência
(superintelligence)

intelecto que excede em muito o desempenho cognitivo dos seres humanos em praticamente todos os domínios de interesse; máquinas que superam a inteligência humana. Frequentemente conectada com a ideia de “explosão de inteligência” causada por máquinas inteligentes projetando máquinas ainda mais inteligentes.



Tarefas cognitivas

tarefas que envolvem uma combinação de processamento de informações, memória, recuperação de informações, planejamento, raciocínio, organização, resolução de problemas, aprendizagem e tomada de decisão orientada por objetivos previamente definidos.

**Transparência
algorítmica**

o grau em que os fatores que geram os resultados da IA, recomendações ou decisões, são conhecidos por várias partes interessadas. Tais fatores podem incluir o funcionamento interno do modelo de IA, a forma como foi treinado, em que dados foi treinado, que características do input afetaram seu resultado e que decisões teria tomado em diferentes circunstâncias.

Usuário final de IA

qualquer indivíduo ou organização intencional ou real que usa ou consome um produto ou serviço baseado em IA à medida que é implantado.

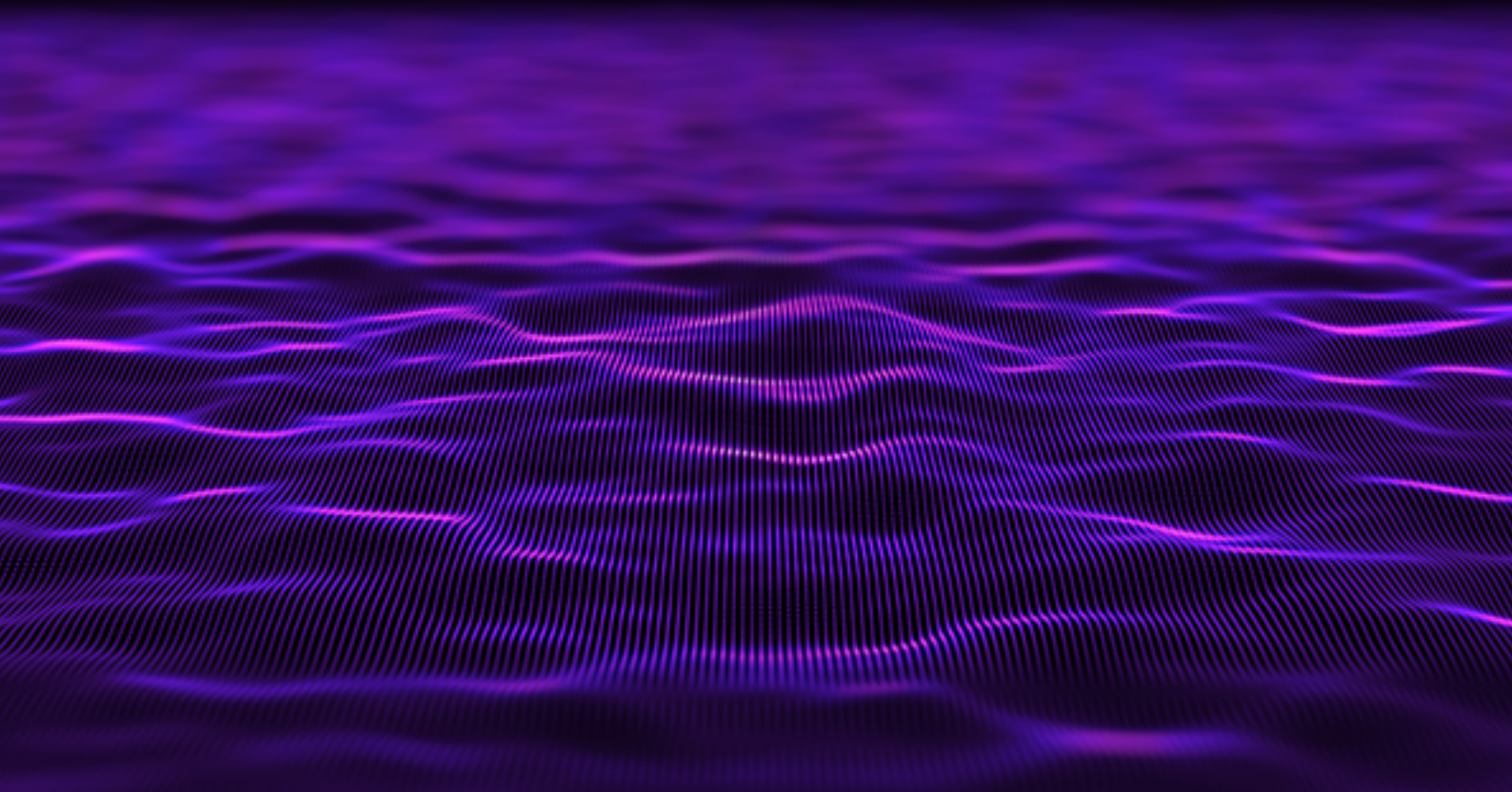
**Viés discriminatório
(bias)**

discriminação sistemática contra certos indivíduos e/ou grupos de indivíduos com base no uso inadequado de certos traços ou características (“atributos sensíveis”); discriminação contra ou a favor de determinados indivíduos ou grupos. No contexto da ética e da política, emerge a questão de saber se um determinado viés é justo ou injusto.

abrasca 

Apêndice de Futuros


accenture



Apêndice de Futuros



O estudo sobre cenários futuros de Governança de IA é de autoria exclusiva da Accenture. O conteúdo não tem relação nem faz parte do escopo e objeto da pesquisa ABRASCA.

A vertical bar with a gradient from dark purple at the bottom to light pink at the top, matching the square graphic above.

Resumo do capítulo

Aqui inicia uma jornada para possíveis futuros da ética e governança da IA no cenário brasileiro, por meio de cenários provocativos criados com a abordagem de design especulativo. Em um contexto em que as transformações tecnológicas são constantes, pensar em futuros aumenta a visibilidade e possibilita que empresas e organizações sejam resilientes às constantes transformações. A Accenture te convida para viajar para 2030 e experimentar possíveis desdobramentos Inteligência Artificial no mercado brasileiro, imaginando um novo amanhã.

Momento de reinvenção

Estamos à beira de uma transformação profunda que tem o potencial de redefinir o cenário corporativo global. Em um mundo onde a interconexão entre tecnologia, globalização e sustentabilidade se torna cada vez mais evidente, as grandes empresas enfrentam um conjunto de desafios que exige não apenas adaptação, mas uma reavaliação estratégica de suas operações e visões de futuro. Essa pesquisa, realizada em 2024, busca explorar como essas mudanças podem impactar as corporações e as maneiras como elas devem se preparar para o futuro.

O avanço tecnológico acelerado da IA está forçando organizações a modernizar rapidamente suas infraestruturas e a integrar novos sistemas com tecnologias existentes, tarefas que podem ser tanto desafiadoras quanto cruciais para manter suas vantagens competitivas.

À medida que avançamos para um amanhã em que a inclusão, a inovação e a governança serão decisivas, é imperativo que os líderes empresariais se preparem para enfrentar os desafios das diferentes possibilidades de maneira proativa, mantendo a resiliência de seus negócios. O sucesso dependerá da capacidade de não apenas responder às mudanças, mas de liderá-las com visão e integridade.

É por isso que na Accenture pensamos futuros, no plural. Por meio de nossa expertise em consultoria estratégica e inovação tecnológica, acreditamos que essas mudanças representam uma oportunidade única para redefinir o papel da tecnologia nas empresas e na sociedade como um todo.

Quando falamos em futuros, queremos falar sobre as possibilidades dentro desta unidade de tempo, considerando os desdobramentos de diferentes cenários. Este capítulo resume um estudo aprofundado acerca do futuro da governança e ética da IA, a partir de cenários construídos pela abordagem do Design Especulativo.

Essa abordagem explora possibilidades futuras e impactos sociais de tecnologias emergentes e tendências culturais e sociais em mudança, provocando novos pensamentos, debates, direções, preferências e estratégias. Ao criar cenários de futuros, experimentamos a extrapolação de tendências que começam a se desenhar atualmente, possibilitando que líderes e tomadores de decisão interajam com elas e trabalhem para criar as estratégias que os levem para o amanhã desejável.



Webb, Amy

The Signals Are Talking: Why Today's Fringe Is Tomorrow's Mainstream.

PublicAffairs
2016

“Ninguém deveria planejar para um futuro que não consegue ver. No entanto, é exatamente isso que acontece todos os dias em nossas salas de reuniões e edifícios legislativos. Frequentemente, líderes ignoram os sinais, demoram demais para agir ou planejam apenas para um cenário. Não só os primeiros a se mover criam novas estratégias, liderança de pensamento, hacks ou explorações para se alinhar à tendência, como provavelmente já estão desenvolvendo a terceira e quarta iterações.”

A partir dessas discussões, chegamos à pergunta que norteou nosso estudo e que buscamos responder: **será que as organizações estão realmente preparadas para moldar um futuro em que a IA não apenas impulsiona a inovação, mas também promove os mais altos padrões de ética e governança?**

Metodologia

O estudo de futuros contou com diferentes etapas, que se iniciaram no processo de pesquisa documental e exploratória e entrevistas em profundidade. Na pesquisa documental analisamos dados internos, como estudos da própria Accenture, POVs e a experiência adquirida nos diferentes projetos que conduzimos. Além disso, consultamos publicações externas, resultados de pesquisas, livros, notícias, relatórios e conduzimos entrevistas em profundidade com especialistas no tema de IA, ética e governança em IA e executivos do mercado. Essa abordagem nos deu acesso a diversos contextos e perspectivas, permitindo uma compreensão mais ampla dos fenômenos. Com isso, foi possível identificar elementos em diferentes áreas, integrar referências de variados campos de estudo e estimular a criatividade.

Identificar diferentes tendências e as novas manifestações de mudanças sustentadas dentro de uma indústria, setor público, sociedade ou na forma como nos comportamos uns com os outros, podem ser pontos de partida que nos ajudam a entender simultaneamente as demandas do presente enquanto planejamos o futuro (2016, Webb). Para começar a entender como elas se relacionam, as agrupamos em cinco áreas: tecnologia, social, ambiental, economia e política.

A partir disso, buscamos significado na relação entre essas diferentes tendências, em um processo que chamamos de sensemaking, nos perguntando: como esses fenômenos se relacionam e se influenciam? São esses padrões, reflexões e a criação de hipóteses que nos trazem insights.

No processo de entender em profundidade essas tendências mapeadas e a relação entre elas, surgem os impulsionadores de mudanças - que são uma coleção de sinais e planos declarados de partes interessadas influentes e outros dados que surgem por meio do processo de digitalização e agrupamentos de temáticas. São eles que influenciam em desafios, oportunidades e transformações relacionadas à governança e à ética da IA nas organizações.

Os cenários são histórias descritivas que ajudam a tangibilizar possíveis eventos futuros. Para construí-los, há um grande trabalho especulativo focado em identificar como os impulsionadores de mudança se manifestam nos futuros. Nesse processo, são

estabelecidas diretrizes, que são uma forma de visualizar como o sistema evolui à medida que interage como drivers que criam mudanças novas e inesperadas. A construção de cenários neste estudo foi baseada em uma matriz de dois eixos: o nível de maturidade de governança em IA e aplicação da IA dentro ou fora do core business. O cruzamento desses eixos gera quatro quadrantes, cada um representando um cenário projetado para o ano de 2030. Esses cenários serão apresentados e detalhados a seguir, utilizando empresas hipotéticas para criar uma narrativa envolvente.

Matriz de Cenários 2030



Cenário A

Governança madura,
IA fora do core business

Evolução acima da revolução: uma história de transformação com governança no uso de IA

Utilizar IA para converter as jornadas de clientes-pacientes em dados estruturados de alta qualidade, desde a abertura de tickets de atendimento até a recepção em suas unidades médicas ou teleconsultas, foi o que permitiu a tomada de decisões automatizadas e precisas que transformariam a experiência dos clientes da CareCura, referência no setor de saúde do país. Esse movimento, porém, não poderia ser chamado de ousado para quem acompanhou de perto toda essa transformação – contar com a parceria de consultorias e soluções IA já bem estabelecidas de big-techs, tanto para gerenciar trilhas de dados imutáveis de forma transparente, segura e rastreável, quanto para evitar ataques e fraudes usando soluções de IA generativa em cibersegurança para simular ataques e contra-ataques em seu sistema, levaram a CareCura no caminho oposto de sua concorrente, a CuraAvançada.

Enquanto o país passava por uma revolução silenciosa e um forte investimento em infraestrutura por parte de grandes corporações para viabilizar seus altos índices de adoção de IA, a CuraAvançada se movimentava de forma ousada, individual e desregulada com intuito de revolucionar a experiência de atendimento médico aos seus clientes. A falta de um comitê de governança na adoção de boas práticas de IA e uma cultura de inovação frágil, porém, foi o que fez a tão famosa empresa falhar catastróficamente: o sequestro de dados de seus clientes devido a uma falha de segurança, que expôs históricos e informações sigilosas e gerou danos e prejuízos irreparáveis.

Tal aprendizado não passaria em branco para a CareCura, que acompanhou o avanço conjunto dos consórcios e parcerias entre empresas na regulação e fiscalização do uso de dados, enquanto amadurecia sua cultura de inovação e adotava um importante comitê de governança de IA para regularizar e fiscalizar o uso dentro da organização. Esta evolução consistente foi o que permitiu também à empresa aumentar significativamente a eficiência e redução de

gastos operacionais, em equilíbrio com uma gestão de mudança também focada na experiência de seus colaboradores, capacitando-os continuamente sobre boas práticas e guias de IA responsável, aplicando agentes de inteligência artificial copilotos que permitiram a redução da jornada de trabalho para funções administrativas, além de tradutores simultâneos em chamadas de vídeo que conectaram profissionais do mundo todo, quebrando barreiras geográficas e de idiomas. As práticas sustentáveis, estimuladas pelo comitê, também foi assunto determinante para observar o presente enquanto se planejou o futuro: monitorar e analisar os consumos de suas operações com auxílio de IA permitiu a publicação de relatórios públicos periódicos alinhados à estratégia de posicionar a CareCura como uma empresa ambientalmente responsável, com emissões de carbono e pegadas hídricas zeradas e certificadas por órgãos reguladores.

Desafios e oportunidades

Desafios e oportunidades

Cultura de Adaptação e Capacitação Contínua

Integração e colaboração da gestão e áreas no desenvolvimento de novos modelos de negócio, produtos, serviços com o uso de IA.

Governança e Impacto Social

Garantir governança robusta para minimizar possíveis impactos negativos do uso de IA, como dados enviesados que aumentem a exclusão social e que não estejam alinhados à responsabilidade ética e social.

Segurança e Privacidade

Proteger dados e garantir a privacidade, usando ações preventivas e que explorem possíveis vulnerabilidades de segurança da informação, equilibrando inovação tecnológica com práticas éticas.

Inovação e Eficiência Operacional

Criar um ambiente controlado para testar novas aplicações de IA, assegurando que o avanço tecnológico esteja em harmonia com a responsabilidade ética.

Cenário B

Governança madura,
IA no core business

O desafio da integração de AI para clientes, colaboradores e sucesso do negócio nas mãos do CAIO

Adotar um Chief AI Officer (CAIO) foi fundamental para a recente transformação de tecnologia e gestão ética da StellarFine. A consultoria financeira, que oferece conselhos de investimento personalizados com base em análise de mercado em tempo real de forma automatizada, tem IA como a força motriz do seu negócio, tanto em seus produtos, mas também internamente, garantindo eficiência em processos internos. Para que essa harmoniosa integração ocorresse, a StellarFine precisou ser desafiada.

Toda essa transformação começou com o desafio da escassez de profissionais na área, que fez com que concorrentes sinalizassem a gestão de talentos como diferencial para futuros empregados. A StellarFine não aceitaria ficar para trás, mais do que isso, ela quis estar à frente. O uso de IA para contratações mais inclusivas, diminuindo vieses humanos, foi o primeiro passo; os algoritmos avançados que mapeiam continuamente habilidades e aspirações dos colaboradores e apoiam o desenvolvimento de carreira, foi o segundo e decisivo passo que apontou para uma nova direção: a do futuro.

O CAIO coordenava esforços entre diferentes departamentos, assegurando que a IA fosse utilizada de forma ética e alinhada às metas da empresa, e que estivesse em conformidade com regulamentações emergentes. Através da promoção de explainable AI (XAI), também assegurou que as decisões automatizadas fossem transparentes e justificáveis, além de possibilitar auditoria, aumentando a confiança e garantindo diferencial competitivo à StellarFine.

Otimizar operações, melhorar eficiência, garantir processos energeticamente eficientes investindo em tecnologias de IA verde, realizar parcerias e aquisições de startups, manter-se em

colaboração com associações internacionais, além de liderar um laboratório de inovação com ambiente seguro de desenvolvimento ágil para experimentações foram algumas das missões concluídas com sucesso sob a liderança do Chief AI Officer, fazendo dessa abordagem equilibrada e responsável, mais que uma fórmula para criar clientes e colaboradores satisfeitos, um modelo a ser seguido, mostrando que é possível avançar tecnologicamente sem comprometer ética e sustentabilidade.

Desafios e oportunidades

Desafios e oportunidades

Talentos e Desenvolvimento de Carreira

Atrair, reter e desenvolver talentos em um mercado competitivo, utilizando a IA para criar oportunidades de desenvolvimento de carreira e promover um ambiente de trabalho inclusivo e acessível.

Parcerias e Expansão

Fortalecer a posição da empresa por meio de parcerias estratégicas e aquisições que ampliem as capacidades de IA, impulsionando a inovação e o crescimento no mercado.

Governança Ética e Liderança

Promover o equilíbrio ético no uso da IA, garantindo conformidade com regulamentações, e adotar novas ferramentas e frameworks que possibilitem rastrear e interpretar previsões e decisões geradas por IA, como explainable AI (XAI).

Gestão de Mudança

Conduzir processos de transformação junto a um plano estruturado que considere a comunicação e capacitação da liderança e colaboradores, criando um ambiente de confiança e condições para criar um trabalho mais significativo para todos.

Cenário C

Governança imatura,
IA fora do core business

Privacidade em xeque: Transformação ou transgressão?

A inovação focada em IA, que inicialmente permitiu à VistaMartVibes expandir rapidamente sua base de clientes e aumentar sua participação no mercado, revelou-se uma faca de dois gumes. A ex-esposa de um consumidor, recém-separada, recebeu um álbum com fotos do casal geradas por IA na data de aniversário, já não mais comemorada, do seu finado casamento. Esse incidente, amplamente compartilhado nas redes sociais, levantou sérias preocupações sobre a privacidade e a transparência no uso de dados pessoais pela VistaMartVibes, gigante do varejo reconhecida por transformar todas as suas interações com clientes em ações de Inteligência Artificial Generativa e que estava testando uma novo projeto de experiência do cliente. Relatos semelhantes de outros consumidores também surgiram nas redes para questionar a invasão de privacidade e exigir maior transparência quanto ao uso de suas informações pessoais.

A falta de maturidade nas discussões de quais os limites do uso dessa tecnologia na VistaMartVibes foi o que causou o incidente. Muitos dos problemas que a empresa enfrentava eram consequências de uma implementação acelerada e sem reflexões sobre os limites no uso da tecnologia. Mas esse não era o único desafio que eles estavam enfrentando: foram feitas publicações alegando ataques coordenados por múltiplas inteligências artificiais autônomas aos sistemas da companhia, resultando no roubo de informações sensíveis e prejuízos financeiros enormes. Mesmo sendo concorrentes diretos, a MelhorCompra decidiu colaborar com a VistaMartVibes para que os danos desse ataque não fossem irreversíveis a ponto de abalar a percepção pública na confiança de toda a indústria.

Esse caso alertou todo o mercado para os riscos da adoção de IA para a transformação do negócio sem a devida governança. A VistaMartVibes, antes elogiada por suas recomendações

personalizadas de IA, não possuía um comitê de autorregulação e mapeamento de possíveis impactos da tecnologia, o que resultou em problemas de imagem e perda de confiança dos investidores. A colaboração com concorrentes para mitigar os danos sublinha a urgência de um compromisso renovado com ética, segurança e sustentabilidade na era digital.

Desafios e oportunidades

Desafios e oportunidades

Resiliência e Transformação Cultural

Promover uma cultura organizacional resiliente e adaptável, que reavalia constantemente seus valores e compromissos em torno da ética, responsabilidade e inovação sustentável diante de avanços tecnológicos.

Colaborações Estratégicas

Formar alianças e colaborações com outras empresas e instituições para co-desenvolver normas de governança de IA e reforçar a transparência no uso de dados e práticas de IA.

Confiança e Transparência

Reforçar a confiança dos stakeholders com maior transparência nas práticas de IA, comunicando claramente a finalidade, bem como os benefícios e cuidados adotados.

Segurança e privacidade

Ter um ambiente digital resiliente a novas práticas de ataques cibernéticos e uma equipe constantemente atualizada a fim de proteger os dados e privacidade dos clientes.

Cenário D

Governança imatura,
IA no core business

Aparência Frágil: *AI washing*

HorizonEdu, uma promissora empresa de tecnologia educacional, enfrentou uma grave crise de segurança após uma tentativa de invasão cibernética. A empresa acumulava grandes volumes de dados de estudantes e professores, inclusive dados sensíveis como endereço e biometria, o que a tornou um alvo atraente para hackers e uma grande preocupação para a área de TI e compliance.

Essa crise expôs os efeitos da adoção acelerada e pouco estratégica de IA pela organização, impulsionada pela pressão dos investidores e consumidores. Além da tentativa de invasão, veio a público que a HorizonEdu, que alegava usar IA para personalizar planos de estudo, fazia uma análise superficial dos dados, sem gerar insights úteis ou experiência realmente customizada. O impacto foi imediato e profundo, alarmando estudantes, pais e professores e comprometendo a credibilidade da empresa, que ao tentar se adaptar rapidamente às novas demandas do mercado sem uma estrutura estabelecida, não priorizou os esforços em governança ou gestão da mudança e fez um uso superficial e pouco produtivo da tecnologia.

Por isso, a restauração da reputação da empresa, que envolveu reestruturar práticas de governança e segurança cibernética, não endereçou a crise criada no ambiente de trabalho, que se tornou tensa, uma vez que os funcionários temiam ser responsabilizados por novos vazamentos. Como não houve um plano estruturado de adoção da tecnologia, não foi fornecido um ambiente seguro para experimentação da IA ou capacitação dos colaboradores com guidelines e boas práticas de uso, o que levou funcionários a adotarem soluções de terceiros não autorizadas na tentativa de se adequarem ao novo momento, expondo a empresa a novos riscos de vazamento de informação.

A experiência da HorizonEdu destaca a importância de práticas robustas de governança e segurança de dados no uso de

tecnologias avançadas, além da necessidade de conduzir a mudança considerando a capacitação dos colaboradores, mostrando que o avanço tecnológico deve ser acompanhado por um compromisso com a ética e a sustentabilidade.

Desafios e oportunidades

Desafios e oportunidades

Governança e Segurança

Participar de comitês de governança para monitorar o uso da IA e estabelecer um framework de governança robusto e fortalecer a segurança.

Eficácia na Implementação

Entender as alavancas de valor da adoção da IA, a fim de fazer um uso estratégico e evitar o uso superficial ou enganoso, como IA Washing.

Interoperabilidade

Assegurar que os sistemas de IA sejam interoperáveis com outras tecnologias emergentes e sistemas legados, facilitando a escala e evitando silos tecnológicos.

Inovação Contínua

Construir uma cultura que equilibre inovação tecnológica com práticas éticas e responsáveis.

Considerações Finais

Os cenários apresentados aqui são um recorte focado na importância da estruturação de governança e capacitação para uso no processo de adoção de IA. Entretanto, também foram analisadas questões relacionadas a talentos, ética, sociedade, sustentabilidade, tecnologia e economia no estudo.

Referências

WEBB, Amy. The Signals Are Talking: Why Today's Fringe Is Tomorrow's Mainstream. Edição 01 da PublicAffairs, 2016 – New York. Março 2018.

'Prediction Machines: The Simple Economics of Artificial Intelligence' by Ajay Agrawal, Joshua Gans, & Avi Goldfarb Harvard Business Review Press, 2018.

Dervin, B. (1992) "From the Mind's Eye of the User: The Sense-Making Qualitative Quantitative Methodology". In: Glazier, J.; Powell, R. "Qualitative Research in Information Management." Englewood: Libraries Unlimited, 1992. cap. 6, p. 61-84.

Disponível em:

<http://communication.sbs.ohiostate.edu/sensemaking/art/artabsdervin92powell.html>.

Accenture e Microsoft. Guia Inteligência Artificial para conselheiros de administração 2024

WORLD Economic Forum. Global Security Outlook Report 2023.

Disponível em:

https://www3.weforum.org/docs/WEF_Global_Security_Outlook_Report_2023.pdf

UMA NOVA era da IA generativa para todos. Accenture, 2023.

Disponível em:

<https://www.accenture.com/content/dam/accenture/final/accenture-com/document/Accenture-A-New--Era-of-Generative-AI-for-Everyone-PT-v2.pdf>.

IA GENERATIVA para o crescimento do cliente. Accenture, 2024.

Disponível em:

<https://www.accenture.com/br-pt/insights/song/generative-ai-customer-growth>.

Daugherty, Paul R. Humanos + Máquina: Reinventando o Trabalho na Era da IA / Paul R. Daugherty, H. James Wilson; tradução de Wendy Campos - Rio de Janeiro: Alta Books, 2019.