



Inteligência artificial no Brasil: explorando oportunidades

Sumário

Na trilha do uso responsável de IA	3
O Brasil no debate sobre IA	6
Agindo agora: parceria internacional para avançar na governança de IA	7
Impacto social	8
Engajamento além dos governos	9
Inteligência artificial no Brasil em ação	10
IA e educação	12
IA para preservar o meio ambiente	14
IA para tornar a Justiça mais eficiente	16
IA para transformar a gestão pública	18
IA para fortalecer a agricultura	19
IA para tornar as empresas mais produtivas	20
Como melhorar a governança de IA	23
Cinco pontos para a governança pública da IA	24
IA responsável: a abordagem Microsoft para a construção de sistemas de inteligência artificial que beneficiem a sociedade	26
Tornando operacional a IA responsável na Microsoft	27
Estudo de caso: aplicando nossa abordagem de IA responsável ao novo Bing	34
Avançando na IA responsável por meio da cultura empresarial	37
Fortalecendo os clientes em suas jornadas de IA responsável	42

Na trilha do uso responsável de IA



Tânia Cosentino, presidente da Microsoft Brasil

O interesse da sociedade em torno das possibilidades da inteligência artificial (IA) é cada vez maior em todo o mundo – e não poderia ser diferente no Brasil. As pessoas estão impressionadas com o poder inovador da tecnologia generativa, capaz de trazer mudanças na forma como a sociedade trabalha e se relaciona. Ainda assim, há uma legítima preocupação sobre os efeitos da IA na segurança de informações, na privacidade e no futuro do trabalho. O potencial de transformação da IA é comparado com as mudanças ocorridas com o surgimento do computador pessoal, da internet e da mobilidade e da computação em nuvem. Estamos na porta de entrada de uma quarta revolução digital.

Por tudo isso, é com satisfação que apresento o relatório **IA no Brasil: explorando oportunidades**. As próximas páginas mostram um panorama de tudo o que a Microsoft tem feito para tornar essa enorme inovação em algo verdadeiramente útil, eficiente e produtivo para todos, considerando os aspectos econômicos, sociais e até mesmo ambientais.

A Microsoft está totalmente comprometida em ajudar a garantir que a IA tenha a governança mais efetiva possível e que as pessoas e as empresas possam utilizá-la de forma responsável. No âmbito nacional, procuramos atuar em estreita

colaboração com clientes e parceiros, que incluem o governo brasileiro e empresas de diversos setores, para estimular a inovação e gerar ganhos para a sociedade como um todo. Você terá a chance de descobrir os detalhes dessas iniciativas no capítulo *Inteligência artificial no Brasil em ação*.

Mostrar como o Brasil e os brasileiros estão recebendo e interpretando os primeiros passos dessa transformação promovidas pela IA é especialmente importante em 2024, momento em que o país assume a presidência do G20, grupo formado por 19 países¹ mais União Europeia e a União Africana. Os desdobramentos da IA ganharam espaço relevante nos principais fóruns mundiais, e não será diferente no G20. As organizações internacionais buscam traçar diretrizes de comportamento, conduta e governança para determinar que os benefícios da IA possam se realizar e ser distribuídos de forma justa, prevendo e evitando eventuais problemas por mau uso da tecnologia.

¹ África do Sul, Alemanha, Arábia Saudita, Argentina, Austrália, Brasil, Canadá, China, Coreia do Sul, Estados Unidos, França, Índia, Indonésia, Itália, Japão, México, Reino Unido, Rússia e Turquia.

Na liderança desse grupo tão representativo, o Brasil está em boa posição para fortalecer seus pontos de vista e indicar os caminhos à frente. A reunião de cúpula está marcada para novembro de 2024, no Rio de Janeiro. A intenção brasileira é ir além das questões de uso ético e responsável da IA, incorporando ao debate questões como a necessidade de a tecnologia refletir a diversidade social, racial, linguística e de contextos geográficos.

Uma maneira de abordar a questão é o Brasil apontar para seus parceiros do G20 o enorme mercado potencial para desenvolvimento de soluções e aplicativos a partir de ferramentas de IA atualmente disponíveis para desenvolvedores em todo o mundo. A ideia é aproveitar as bases construídas pela Microsoft e outros fornecedores para tirar do papel ideias novas, que atendam a demandas e necessidades próprias de cada país ou região.

Ou seja: em vez de tentar “reinventar a roda”, os países podem dar preferência a utilizar o que já existe para criar aplicações práticas e resolver problemas reais – como agilizar o atendimento no setor de saúde ou tornar o transporte público mais confiável.

Muitos profissionais brasileiros podem ser convocados para o desafio, tanto no setor público quanto nas empresas privadas, graças ao alto grau de profissionalização das equipes de tecnologia da informação brasileiras.

Afinal, o Brasil tem experiência em liderar avanços digitais, como as automações e a segurança cibernética do setor bancário, que são referências em todo o mundo. Embora sejam totalmente nacionais, surgidas para pôr fim a questões específicas da nossa economia, essas soluções que revolucionaram o segmento se valem de aplicações e infraestruturas de nuvem de abrangência global, como o Microsoft Azure. Na era da inteligência artificial, esse tipo de atuação conjunta de soluções globais e locais tende a ser ainda mais proveitosa e frutífera para o futuro das nações.



A Microsoft está na vanguarda da IA – inclusive aqui no Brasil. Trabalhamos para construir uma fundação responsável sobre o tema desde 2017, com um conjunto de iniciativas em segurança digital. Essa atuação nos conectou a uma estrutura mais ampla de gerenciamento de riscos corporativos, o que foi fundamental para construir princípios, processos e sistemas de governança que nos orientassem para o uso responsável da IA.

No entanto, é preciso fazer sempre mais. Reconhecemos que a IA está em constante evolução e temos um compromisso com melhorar continuamente nossos processos, sempre ao lado de clientes, parceiros e da sociedade em geral. O mais importante é que a IA possa assistir a todos, o que abrange treinar trabalhadores, orientar estudantes e levar em conta as populações mais vulneráveis. Só assim a tecnologia terá uma efetividade global, proporcionando crescimento justo e inclusivo em todos os continentes.

No caso brasileiro, buscamos desde o primeiro momento inserir as mais diferentes abordagens e oportunidades de IA no Microsoft Mais Brasil, uma série de iniciativas coordenadas que busca combinar inclusão e crescimento econômico. Incentivamos, por exemplo, a especialização de profissionais na tecnologia generativa, condição essencial para promover a transformação digital das empresas brasileiras. Em outra linha de ação, encorajamos o uso da IA em projetos de preservação do meio ambiente.

A revolução da IA está apenas começando. É o momento de construirmos uma abordagem ética e responsável, com uma governança pública moderna e eficiente, para que as pessoas e as empresas possam aproveitar da melhor forma as oportunidades da tecnologia generativa.



O Brasil no
debate sobre IA

O Brasil no debate sobre IA nos organismos internacionais

O Brasil atua em duas frentes para buscar a adoção de critérios de adoção de governança de inteligência artificial (IA). O Ministério da Ciência e Tecnologia e Inovações está concentrado na revisão da Estratégia Brasileira de IA, apresentada originalmente em 2021, para considerar as últimas inovações em tecnologia generativa e adaptá-las às prioridades do governo.

O Congresso Nacional, por sua vez, age para formatar uma regulação legal para IA, com regras que possam garantir seu uso ético, transparente e responsável. Ambas as iniciativas, ainda que precoces, são complementares, com esforços para estabelecer um panorama legal eficaz, direcionado a fomentar a inovação em IA ao proporcionar confiança na implementação de medidas apropriadas de IA responsável.

Agindo agora: parceria internacional para avançar na governança de IA

O Brasil tem um papel central a desempenhar no diálogo global sobre IA por estar na presidência do G20, um grupo formado por 19 países mais a União Europeia e a União Africana. Trata-se de um dos mais representativos fóruns internacionais, reunindo países de perfis tão distintos e amplo impacto coletivo. Suas economias representam 85% do Produto Interno Bruto (PIB) mundial, mais de 75% do comércio do planeta e cerca de dois terços da população global. Em sua liderança, o Brasil está em posição privilegiada para fortalecer seus pontos de vista sobre o uso da inteligência artificial. Essa diversidade de perfis sociais, econômicos, ambientais, de interesses e objetivos torna o G20 especialmente relevante na questão da IA, uma tecnologia inovadora com crescente impacto global.

Construir uma abordagem global coerente para a governança de IA traz benefícios significativos para todos os stakeholders envolvidos no desenvolvimento, uso e regulamentação da IA. Esforços internacionais são úteis não apenas como soluções intermediárias importantes antes que regulamentações locais definitivas entrem em vigor, como o projeto de lei de IA

em discussão no Brasil, mas também para nos ajudar a trabalhar em um conjunto comum de princípios compartilhados, que podem fornecer pontos de referência para abordagens consistentes e interoperáveis a situações de cada país. Parcerias multilaterais são necessárias para fortalecer a governança de IA, além, e ao lado, de legislações nacionais.

A IA faz parte da agenda de inúmeros fóruns diplomáticos. Avanços significativos foram alcançados recentemente. Os integrantes do G7, por exemplo, concordaram com um código de conduta internacional para o desenvolvimento de sistemas avançados de IA.² O Reino Unido reuniu líderes internacionais em um encontro sobre segurança de IA, o que resultou na ampla adoção, incluindo o Brasil, da Declaração de Bletchley.³ No Fórum Econômico Mundial, em Davos, na Suíça, o CEO da Microsoft, Satya Nadella, reconheceu que um olhar global para a IA é “muito desejável” porque estamos nos preparando para responder a “desafios globais que exigem normas globais e padrões globais. Caso contrário, vai ser muito difícil conter, difícil de aplicar e difícil de fazer acontecer, até mesmo em relação a algumas pesquisas centrais que são necessárias”. Ele também reconheceu “que parece haver um amplo consenso no que está surgindo”.

2 Trata-se do Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI System, avaliado pelos países do G7 em outubro de 2023. Ver <https://www.mofa.go.jp/files/100573473.pdf> (em inglês).

3 A Declaração de Bletchley busca promover o desenvolvimento seguro da IA para impulsionar o crescimento econômico inclusivo e a inovação. O documento foi assinado por 27 países e divulgado após o AI Safety Summit 2023, realizado no Bletchley Park, em Londres. Ver <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> (em inglês).

O atual progresso está baseado em iniciativas mundiais anteriores e contínuas sobre a governança de IA, o que inclui a Declaração de Montreal para o Desenvolvimento Responsável de IA, os Princípios de IA da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), a recomendação da Unesco sobre ética em IA e a Parceria Global em IA (GPAI), em todas as quais o Brasil esteve engajado ou adotou.

Esse contexto posiciona bem o Brasil em sua presidência do G20, para aprofundar e estender o momento positivo surgido com o código de conduta do G7, incluindo sua aplicação aos responsáveis em implantar a tecnologia. O modelo voluntário fornece uma oportunidade para os países do G20 avançarem juntos em um conjunto de valores e princípios compartilhados. Atuar em direção a uma abordagem globalmente coerente para a implantação responsável é importante, reconhecendo que a IA, assim como muitas inovações, é e será utilizada através das fronteiras e que uma governança consistente e interoperável ajuda a permitir que todos acessem as melhores ferramentas e soluções para suas necessidades.

Impacto social

O Brasil também está especialmente bem posicionado para liderar o G20 na promoção do diálogo global e cooperação no uso da IA para enfrentar os maiores desafios da sociedade, baseando-se na Declaração de Bletchley¹ e esforços multilaterais mais amplos, incluindo os Objetivos do Desenvolvimento Sustentável das Nações Unidas. Esse foco também é consistente com a liderança do Brasil na luta contra a fome e a desigualdade, no desenvolvimento sustentável e na reforma da governança global – temas centrais da agenda brasileira do G20.



A presidência do Brasil terá a chance de encaminhar as conversas sobre IA para muitas esferas de discussão, explorando os benefícios do impacto social da inovação ao máximo. Com a IA não apenas na agenda dos grupos de trabalhos (GTs) de Economia Digital e de Pesquisa e Inovação, mas também presente nos GTs sobre Trabalho, Redução do Risco de Desastres, Saúde e Desenvolvimento, entre outros, o Brasil pode contribuir para o desenvolvimento e adoção de quadros de governança de IA que estão alinhados com valores fundamentais e inclusão social, bem como as necessidades e aspirações de uma sociedade diversa e multicultural.

O Brasil pode ainda contribuir de forma significativa compartilhando sua experiência, conhecimento e perspectiva sobre IA de um ponto de vista latino-americano, enquanto se beneficia da troca das melhores práticas, experiências e conhecimentos com outros países e regiões que estão promovendo uma governança interoperável além das fronteiras, como um facilitador no desenvolvimento e no uso da IA.

De forma notável, o G20 sob a direção brasileira está posicionado para representar um marco nas discussões de como os países abordam a IA ao colocar luzes no potencial de soluções locais de IA para abordar

questões sociais urgentes. A experiência reconhecida dos inovadores do Brasil liderando a automação e a cibersegurança no setor bancário, por exemplo, fornece um conjunto de projetos inovadores para desenvolvedores de todo o G20 construírem.

O Brasil pode ser o vetor dessa transformação, aumentando ainda mais seu protagonismo internacional e liderança global, direcionando a aplicação do IA para resolver alguns dos maiores desafios do mundo.

Engajamento além dos governos

Outra oportunidade de ação está nos grupos de engajamento do G20 Social, evento paralelo que busca estimular a participação da sociedade civil nas discussões e formulações de políticas e que vai ser ampliado durante a presidência brasileira. A IA é um tema que se relaciona, por exemplo, com as discussões do Startup20, fórum que propõe estabelecer um diálogo aberto entre as diversas partes interessadas no ecossistema de startups e tecnologia. Entre seus objetivos estão fomentar a inovação e o empreendedorismo científico e desenvolver oportunidades de negócios entre países-membros – situações em que a IA pode representar papel preponderante.



Inteligência artificial no Brasil em ação

Inteligência artificial no Brasil em ação

Os avanços recentes revelam as inúmeras possibilidades do uso da inteligência artificial (IA) e apontam como a tecnologia pode moldar o futuro dos negócios e da sociedade no Brasil. O ritmo e a escala das novidades tecnológicas são inéditos. São inovações que ajudam o país a enfrentar os desafios sociais, ambientais e produtivos.

Os números apontam que há um ambiente favorável para a IA ganhar espaço por aqui. A pesquisa [Trust in Artificial Intelligence](#) (*Confiança em inteligência artificial*, em tradução livre), realizada pela consultoria KPMG, mostrou que o grau de confiança dos brasileiros na tecnologia é de 84%, acima da média global (61%), embora os entrevistados reconheçam preocupações com segurança (61%), manipulação de informações (63%) e substituição de mão de obra humana com ferramentas de automação (57%).

As empresas, por sua vez, estão cada vez mais interessadas em incorporar soluções generativas. A consultoria Accenture fez um [levantamento com executivos seniores brasileiros](#) que decidem sobre investimentos em tecnologia. O resultado mostrou que 97% dos entrevistados têm a intenção de ampliar seus desembolsos na área, sendo que 84% citaram a inteligência artificial como prioridade.

Para atender a essas expectativas, a Microsoft atua para promover o uso da IA de forma ética e responsável, disseminando o conhecimento mais atualizado sobre o tema. Nossa intenção é oferecer recursos para que a tecnologia generativa ajude de maneira efetiva o desenvolvimento econômico e social do Brasil. Uma das linhas de ação é promover a **educação em IA**, com a disponibilização de formações profissionais especializadas para todos os tipos de público, permitindo que cada vez mais pessoas possam aproveitar as possibilidades da inovação.

Além disso, nós estamos ao lado de algumas transformações promovidas pela IA nos setores público e privado. Fornecemos softwares, infraestrutura e a experiência pioneira de nossos profissionais para projetos inovadores que estão ajudando, por exemplo, a **preservar o meio ambiente, a tornar a Justiça mais eficiente, a transformar a gestão pública, a fortalecer a agricultura e melhorar a produtividade das empresas**. A seguir, você vai conhecer algumas das iniciativas que, com a contribuição da Microsoft, estão ajudando a impulsionar a IA no Brasil.

IA e educação

A questão é matemática: quanto mais pessoas estiverem habilitadas para trabalhar com IA, maiores são as perspectivas de desenvolvimento. A lógica serve tanto para desenvolvimento de aplicações quanto para o preparo da força de trabalho com as competências digitais necessárias para os empregos que vão surgir nesta era da IA.

A Microsoft oferece oportunidades para diferentes perfis de profissionais. Em colaboração com o LinkedIn, lançamos o programa de aprendizado [Career Essentials in Generative AI](#) (Fundamentos Essenciais de Carreira em IA Generativa). O curso é gratuito, está legendado em português e estará disponível até 2025. Composto por seis módulos, a formação tem como objetivo apresentar um panorama de todo potencial da tecnologia, explicando o funcionamento dos modelos de IA generativa e destacando a importância de seus aspectos éticos e de seu uso responsável. A plataforma permite ao aluno acessar arquivos de exercícios, tirar dúvidas, ler a transcrição do conteúdo dos vídeos e fazer anotações. Quem conclui o programa recebe um certificado profissional da Microsoft.

Uma parceria produtiva é a da plataforma [Escola do Trabalhador 4.0](#), ao lado do Ministério do Trabalho e Emprego do governo brasileiro. Mais de 1 milhão de pessoas estão cadastradas na ferramenta. Seu objetivo é oferecer treinamentos em habilidades digitais para ajudar no acesso ao trabalho, emprego e renda dos brasileiros. O público-alvo são jovens e adultos trabalhadores, acima de 14 anos, que querem se reinventar e se preparar para as novas demandas do mundo de trabalho. Há desde módulos iniciantes até formação técnica específica, como IA: “Profissional de Inteligência Artificial”, com cinco módulos, e o “Conceitos Básicos de IA do Microsoft Azure: O Que é IA Generativa?”. Todos os cursos incluem certificação após a realização de um teste.



Da esquerda para a direita: Magno Lavigne, secretário de Qualificação e Fomento à Geração de Emprego e Renda do MTE; Juscelino Filho, ministro das Comunicações; Luiz Marinho, ministro do Trabalho e Emprego; Kate Behncken, vice-presidente corporativa e líder Global da Microsoft Philanthropie; Alessandra Del Debbio, vice-presidente Jurídica e de Assuntos Corporativos da Microsoft na América Latina; Lucia Rodrigues, líder de Filantropia da Microsoft Brasil; Jorge Cella, diretor da Microsoft Philanthropies para Américas; e Elías Abdala, diretor regional de Government Affairs da Microsoft na América Latina.

Tanto o treinamento em parceria com o LinkedIn quanto os da Escola do Trabalhador 4.0 (*foto*) podem ser acessados pelo Conecta+, hub que reúne todos os cursos e programas de parceria para o acesso ao trabalho, emprego e renda e permanência no mundo do trabalho oferecidos pela Microsoft Brasil.

A Microsoft também estimula a formação em tecnologia para a população de baixa renda. Um exemplo de nossa atuação foi o *hackaton* direcionado a jovens apoiados pelo Instituto Atravessando a Ponte, entidade que

oferece bolsas de estudos a quem precisa na área de tecnologia da informação. O projeto, viabilizado pelo programa de voluntariado interno GIVE, incluiu palestras de especialistas da Microsoft sobre cenários e perspectivas para o avanço na aplicação de IA no mundo dos negócios. As propostas desenvolvidas foram apresentadas a uma banca avaliadora formada por representantes do departamento de recursos humanos de empresas parceiras interessadas em preencher vagas de estágio e de jovem aprendiz.

IA para preservar o meio ambiente

Poucos temas são tão relevantes hoje quanto as mudanças climáticas. A IA, utilizada de forma inteligente, é capaz de acelerar os esforços de transição para as energias limpas e mitigar o impacto dos efeitos do aquecimento global nas pessoas e nas comunidades. O tema é ainda mais relevante no contexto brasileiro por causa da Floresta Amazônica, com seu ecossistema que desempenha um papel crucial na regulação do clima, na produção de oxigênio, no armazenamento de carbono e na conservação da biodiversidade.

A Microsoft está na linha de frente na preservação ambiental na região com a plataforma PrevisIA, uma aplicação que utiliza a IA e recursos avançados em nuvem do Microsoft Azure para detectar áreas de maior risco de desmatamento na Amazônia. Nossos aliados nessa iniciativa inovadora são o Fundo Vale e o Instituto Homem e do Meio Ambiente da Amazônia (Imazon), entidade com mais de 30 anos de atuação.

A missão da ferramenta é fornecer informações valiosas para a tomada de decisão dos setores público e privado, assim como da sociedade em geral, para a implementação de ações preventivas de combate e controle ao desflorestamento. A PrevisIA constatou que cerca de 5 mil quilômetros estavam sob risco médio, alto ou muito alto de devastação entre agosto de 2023 e julho de 2024.

O segredo está em utilizar essa nova frente das ciências de dados em favor do meio ambiente. A PrevisIA possui um algoritmo que permite monitorar por meio de imagens de satélite diversos fatores de risco de desmatamento, em especial as estradas não oficiais da região. Qualquer alteração no entorno dessas vias é identificada imediatamente, em tempo real.

Fazer do futuro
sustentável a
**nossa única
previsão.**



O monitoramento é crucial porque as vias são a principal variável na tarefa de prever desmatamentos. Estudos da Imazon apontam que 95% do desmatamento acumulado na Amazônia concentra-se em um raio de 5,5 quilômetros de suas margens. Mais: 90% das queimadas anuais ocorrem a 4 quilômetros

delas. Essas rotas são abertas na floresta para a exploração madeireira, garimpo e ocupação ilegal de terras públicas. O resultado é um mapa interativo, destacando as áreas mais ameaçadas. Com esse apoio, as autoridades ganham agilidade para proteger a floresta.

IA para tornar a Justiça mais eficiente

Uma iniciativa inovadora em ação na Advocacia-Geral da União (AGU) traz um exemplo bastante rico de como a IA pode ajudar o sistema de Justiça a atuar de maneira mais eficiente, com perspectiva de soluções mais rápidas e com menor chance de erro nos processos judiciais. A AGU, instituição responsável pela representação judicial da União, é uma das primeiras entidades nacionais do setor público a adotar o uso da IA generativa.

Ao adotar a tecnologia GPT-4, incorporada na solução Azure OpenAI Service, o órgão tem o objetivo de potencializar o trabalho de seus advogados e gerar mais eficiência. O sistema foi integrado ao Sapiens, gerenciador eletrônico de documentos da instituição. O desafio é imenso. A AGU analisa 20 milhões de processos judiciais, em uma média de 10 mil citações por dia, além de 80 mil intimações diárias. A intenção é usar a inovação para auxiliar a análise e previsão de resultados de casos judiciais com o objetivo de aprimorar a tomada de decisões e embasar a elaboração de estratégias processuais, tendo como fonte a gigantesca base de dados do Judiciário.

Nesse primeiro momento, a AGU aproveita a tecnologia para produzir os textos de processos judiciais, para os quais utiliza como base informações de petições iniciais e de

sentenças, além de avaliar o contexto do que está sendo relatado no processo e os casos similares em que o órgão participou. Com isso, a IA constrói o texto citando a legislação, a jurisprudência e argumentos que seriam usados pelo órgão, oferecendo ao servidor o material inicial para que possa revisar e complementar com informações de cada ação. Os primeiros setores a trabalhar com a novidade foram a Procuradoria-Regional da União da 1ª Região e a Procuradoria Nacional Federal de Contencioso Previdenciário.

A privacidade de dados e a conformidade são elementos essenciais nessa tarefa e, em sua execução, consideram as mais avançadas camadas de segurança presentes na nuvem Microsoft Azure. Todas as informações contidas na aplicação são armazenadas e manipuladas no domínio privado da AGU, que adotou técnicas de anonimização de dados, ocultando as informações pessoais, garantindo que os indivíduos não possam ser identificados a partir dos dados disponíveis. Tudo para que a Justiça possa ser feita com todas as demandas da lei e da ética.

Advocacia privada

A inovação também está atraindo o interesse dos escritórios de advocacia do setor privado. Com mais de 50 anos de atuação, o escritório Machado Meyer Advogados adotou o Azure OpenAI Service para desenvolver uma ferramenta segura e eficaz que utiliza a IA como complemento ao trabalho de seus profissionais. A intenção é que, inicialmente, a solução preste apoio às atividades diárias do escritório, como traduções, revisões, resumos e pesquisas.

“O projeto com a Microsoft traz mais inovação para o nosso ambiente e nos permite extrair conhecimento de forma mais rápida e fácil para nossos advogados, potencializando nossa base de informações e tornando o trabalho ainda mais eficaz”, explica Paulo Silvestre de Oliveira Junior, coordenador de Inovação e Desenvolvimento.

IA para transformar a gestão pública

O setor público tem muito a ganhar com o uso da IA, tanto para apoiar a condução de políticas públicas quanto para melhorar a administração em si, reduzindo gastos desnecessários e aumentando a transparência dos investimentos. Em parceria com a Microsoft Brasil, o Centro de Liderança Pública (CLP) desenvolveu um guia para a criação de ferramentas de tecnologia generativa na gestão pública. O CLP é uma organização suprapartidária que busca engajar a sociedade e desenvolver líderes para enfrentar os problemas mais urgentes do Brasil.

O manual, lançado em dezembro de 2023 durante o 127º Fórum Nacional de Secretários de Estado da Administração e que será apresentado pelo CLP e pela Microsoft em diferentes eventos ao longo de 2024, utiliza como base o AI 100-1 – Artificial Intelligence Risk Management Framework, um documento criado pelo National Institute of Standards and Technology (NIST), do Departamento de Comércio dos Estados Unidos. Sua intenção é ajudar a disseminar boas práticas e promover uma IA ética, confiável e inclusiva, que melhore a experiência dos brasileiros no uso dos serviços públicos. A ideia é ter um mapa para as organizações solidificarem a metodologia na gestão de riscos relacionados à IA.

“É fundamental que os servidores públicos estejam adequadamente orientados para garantir o uso ético, transparente e benéfico da inteligência artificial”, afirma Tadeu Barros, diretor-presidente da entidade.

O CLP também elaborou um segundo documento que relata boas práticas em IA na gestão pública, no qual compartilha iniciativas que estão sendo implementadas no país em níveis federal e estadual. São casos de sucesso que já estão mudando a lógica estatal. Entre os exemplos destacados está o assistente virtual de atendimento Jaque, adotado pela Secretaria do Tesouro Nacional. Ao levar a tecnologia para o portal de informações contábeis da administração pública nacional, o chatbot concluiu, nos primeiros quatro meses de operações, mais de 2.800 interações, aliviando a carga de trabalho dos servidores do órgão.

IA para fortalecer a agricultura

O agronegócio tem enorme importância para o crescimento do Brasil, com sua grande contribuição para o Produto Interno Bruto e para o desenvolvimento do interior. As técnicas agrícolas brasileiras são reconhecidas por garantir um aumento da eficiência do cultivo com sustentabilidade. O uso da IA tem tudo para impulsionar ainda mais o setor.

Uma experiência pioneira está sendo adotada pela Grão Direto, plataforma especializada em comercialização digital de grãos que atua na América Latina. A empresa atua com serviços B2B, auxiliando outras companhias do agronegócio no processo de transformação digital e em operações com trading de grãos, jornada do cliente e formalização de contratos. Com o crescimento de suas operações, viu a necessidade de consolidar os milhares de informações que obtinha em suas transações comerciais para tomar melhores decisões e realizar análises de mercado eficientes para seus clientes.

A Grão Direto passou a utilizar o serviço nuvem da Microsoft Azure e o Microsoft Power BI para gerar dados e transformá-los em inteligência de mercado. Além disso, o uso de IA está presente em diversos serviços e soluções da empresa. No Barter Digital, por exemplo, ela é capaz de “ler” contratos dos mais diversos tipos e realizar o preenchimento dos dados no sistema; basta o cliente subir o contrato e ela faz todo o restante. Ademais, a inteligência é aplicada no Clicou, Fechou em que é possível calcular a probabilidade de um usuário fechar um negócio naquele mesmo dia – a partir de dados de compor-

tamento e de mercado – e enviar notificações inteligentes a fim de estimular negócios que estão na iminência de ocorrer. Por fim, ela também está lendo certidões de penhor – documentos que reúnem informações de diversas propriedades e produtores diferentes –, estruturando os dados e organizando as informações na plataforma de Gestão de Contratos, economizando horas de trabalho de seus clientes e diminuindo a chance de erro de preenchimento.

O Microsoft Azure Databricks, software de construção de pipelines de dados e de IA, por sua vez, trouxe mais camadas de segurança, privacidade e usabilidade. A ferramenta auxiliou na previsão de preços médios para o mercado futuro, um dos serviços oferecidos pela empresa. “A IA representa um avanço significativo em nossa plataforma, permitindo maior eficiência e precisão em nossas operações. Estamos comprometidos em oferecer soluções tecnológicas que agreguem valor de verdade aos nossos clientes, sem perder a importância da confiabilidade e segurança dos dados. Em pouco tempo de desenvolvimento, já conseguimos atingir resultados impressionantes, e esse é só começo; em 2024 iremos lançar novos serviços potencializados pela inteligência artificial, tornando a experiência de nossos usuários cada vez mais fluida, assertiva e ancorada na melhor e exclusiva informação do mercado físico que a Grão Direto já possui”, afirma Fred Marques, cofundador e CTO da Grão Direto.

IA para tornar as empresas mais produtivas

A Microsoft Brasil está ao lado do setor produtivo para tornar a IA uma ferramenta mais útil e confiável. O avanço dos modelos tecnológicos de última geração contribuiu para aumentar a produtividade e a eficiência profissional e, como consequência, levando à transformação digital. É um apoio mútuo. A Microsoft transfere aos parceiros todo aprendizado e experiência com infraestrutura, softwares e uso responsável de IA. As empresas aplicam esse conhecimento em ideias práticas para impulsionar suas receitas e ajudar a sociedade dentro de suas áreas de atuação. A seguir, conheça alguns dos casos em que a IA já está fazendo a diferença, em diferentes setores.

Energia – Petrobras

A “democratização” do uso da IA já é realidade na Petrobras, a maior empresa do Brasil e uma das maiores do mundo. A companhia desenvolveu o ChatPetrobras, uma ferramenta interna de geração de conteúdos criativos e analíticos que utiliza a tecnologia de IA do Azure OpenAI Service. Um dos objetivos do projeto era oferecer uma solução que pudesse ser utilizada por toda a força de trabalho, o que de fato aconteceu. Atualmente o aplicativo está disponível por meio de um portal interno para 110 mil trabalhadores, entre empregados próprios e prestadores de serviço, auxiliando na realização de

diversas atividades, como a elaboração de relatórios, resumos, apresentações, além de responder perguntas e dúvidas dos usuários.

O ChatPetrobras foi criado com base em políticas de segurança da informação, proteção de dados, conformidade, normas e regulamentações do governo brasileiro às quais a Petrobras está sujeita como uma empresa de economia mista, evitando a exposição de informações internas e garantindo a integridade dos dados empresariais. A solução conta com as mais avançadas camadas de segurança e proteção, com todos os dados protegidos sob o domínio interno.

A solução proporciona à estatal subsídios para tomar decisões mais rápidas e precisas, gerar insights valiosos para o negócio e otimizar processos internos. Os recursos de compreensão e geração de linguagem natural, tradução e criação de código da IA beneficiam tanto as áreas de negócio, como Exploração & Produção, Reservatórios e Poços, quanto setores administrativos e corporativos (Comercialização, TIC, Engenharia, Jurídico, Recursos Humanos etc.). A solução funciona como uma arquitetura de base para incorporação de IA generativa nas novas aplicações generativas da Petrobras.

Comunicação – Globo

A busca por atenção concentrada em um ambiente altamente disperso é um dos grandes desafios da publicidade. Entendendo a importância da contextualização das mensagens para a ampliação da eficiência da publicidade, a Globo mergulhou na inteligência artificial e desenvolveu o ORA (Oportunidade Real de Agrupamento), uma plataforma de gerenciamento de solução de dados e de IA, que tem como parceiro o Azure Data & AI, da Microsoft.

A tecnologia utiliza recursos de reconhecimento de texto e imagem, impulsionados por inteligência artificial, para mapear oportunidades de veiculação de peças publicitárias, que estejam contextualizadas aos assuntos abordados pelos conteúdos da Globo, tanto na programação da TV Globo e dos canais por assinatura como no conteúdo das plataformas digitais da empresa. Com o ORA, a Globo consegue entregar para as marcas o que produz de mais valioso: o seu conteúdo de qualidade, que alcança milhões de pessoas todos os dias, como ponte para as mensagens dos anunciantes.

O ORA analisa o amplo volume de conteúdos da Globo para sugerir com agilidade aos anunciantes os melhores espaços para veiculação de campanhas publicitárias contextualizadas aos temas abordados, permitindo, assim, a integração das mensagens das marcas às histórias da Globo com máxima harmonia. Outra vantagem desse cruzamento rápido dos dados está em sua atuação ampla, já que o ORA contempla as frentes de TV aberta, canais pagos e plataformas digitais da empresa.

Dessa forma, a ferramenta traz vantagens aos anunciantes, já que a conexão entre as mensagens das marcas e o conteúdo contribui para aumentar a retenção de atenção das pessoas, tanto na TV quanto no ambiente digital. Afinal, a publicidade quando direcionada e contextualizada é capaz de ampliar a proximidade com públicos de interesse, além de aumentar atributos como favorabilidade e intenção de compra por parte da audiência. E esse é só o começo de uma nova era.

Finanças – PicPay

O ecossistema de serviços financeiros PicPay aposta no uso da IA para entregar cada vez mais valor a seus usuários. O passo mais recente de sua estratégia foi integrar o Microsoft Azure OpenAI Service ao seu Direct Message (DM), canal de atendimento que pode ser acessado diretamente do aplicativo no celular. A utilização dessa tecnologia traz benefícios, como respostas mais rápidas e contextualizadas para dúvidas do dia a dia e a possibilidade de direcionar corretamente o cliente com links para acesso direto a ações determinadas. Para o usuário, a experiência é prática e segura, preservando seus dados.

Além de atender os usuários, a IA ajuda na organização financeira. Quem paga suas contas no PicPay pode perguntar quanto gastou com energia elétrica no ano ou de quanto foi o aumento na sua conta de luz, água ou celular em um período. Tudo em um ambiente controlado e testado para assegurar a qualidade e a precisão, garantindo a privacidade dos dados individuais do cliente.

A intenção é ir bem além, transformando o DM em uma espécie de assistente virtual completo do app. O plano é que a IA apoie o cliente em transações, como direcionar recursos para investir ou fazer um Pix. Sem nunca deixar de lado os cuidados com segurança, privacidade, confiabilidade e conformidade com todas as normas, o que é ainda mais fundamental quando o assunto é dinheiro.

Tecnologia – Blip

A Blip é a principal plataforma de inteligência conversacional que cria conversas e conecta experiências entre marcas e seus clientes nos principais aplicativos de mensagem. Para cumprir sua missão, a empresa utiliza os mais novos recursos do Azure OpenAI Service integrados aos aplicativos de mensagens disponíveis no mercado, como o WhatsApp, Instagram e Messenger. Os desenvolvedores da Blip criam fluxos conversacionais cada vez mais naturais, fluidos e personalizados. Um exemplo é conseguir oferecer aos fluxos conversacionais entre marcas e consumidores sugestões de respostas com base no histórico de conversas de cada cliente.

Um aprimoramento oferecido pela empresa é a otimização de redirecionamento contínuo (conhecido como *overflow* no mundo da tecnologia) de uma conversa entre um cliente e uma marca, dentro do contato inteligente, para um atendente humano, quando necessário. Essa transição é transparente para o cliente e permite a continuação da conversa exatamente do ponto em que parou.

Por sua precisão e agilidade, a funcionalidade permite otimizar o tempo de atendimento em todas as etapas da jornada do consumidor (customer care, venda, pós-venda e marketing). Também proporciona uma melhora da experiência da marca que utiliza a plataforma da Blip para responder a seus clientes. A integração com a IA atende principalmente às grandes empresas que conversam diariamente com milhares de consumidores e entendem a importância do atendimento como estratégia de comunicação.



Pagamentos automáticos – Sem Parar

O Azure OpenAI Service também está por trás da Carol, assistente virtual de atendimento do Sem Parar, ecossistema de mobilidade com foco em veículos. A companhia integrou a IA a seu serviço de apoio pelo WhatsApp, que foi preparado para responder a mais de 15 mil hipóteses de interação.

Com recursos de linguagem natural, a Carol pode interagir de forma inteligente e contextualizada às perguntas dos clientes, em um tom mais humanizado e personalizado. A taxa de sucesso nos contatos se manteve sempre superior a 90%, chegando a 100% em determinadas interações. Essas variações ocorreram devido às diferentes abordagens dos modelos estatísticos utilizados. Para entrar na era da IA e melhorar a experiência do usuário, o Sem Parar atuou de forma interdisciplinar em times orientados pela metodologia ágil, com parceiros externos.

A expectativa de seus executivos é reduzir significativamente o tempo entre o primeiro contato e a solução da demanda e, assim, ter um aumento na percepção positiva da marca. “A inteligência artificial desempenha um papel crucial em nossa estratégia. Estamos focados no futuro e assegurando que os benefícios da transformação digital alcancem nossos clientes por meio de diversas iniciativas, aumentando a percepção de valor sobre nossos serviços”, diz Paulo Scrideli, vice-presidente de Tecnologia e Digital do Sem Parar.

Relacionamento com clientes – Atento

Líder na América Latina em prover serviços de gestão de relacionamento com clientes e terceirização de transformação de negócios, a Atento abraçou a revolução da IA generativa com a adoção do Azure OpenAI Service em soluções inovadoras, como o Atento Knowledge Assistant, que funciona como um “copiloto” para os agentes de atendimento ao cliente. A nova solução possibilita a resposta mais adequada e ágil durante qualquer tratativa. Ao longo da interação, também gerencia todas as etapas de formalização do sistema e atualizações de status, integrando-se perfeitamente com os canais de comunicação escolhidos pelo cliente.

Em uma das implantações, a Atento registrou uma melhora de 20% na satisfação dos consumidores. A colaboração da Microsoft foi essencial para calibrar a ferramenta, que segue em constante aprendizado. Além disso, o Microsoft Power Automate integrou os sistemas legados, reduzindo esforços manuais e contribuindo para a otimização do tempo de resposta. Essa mudança proporcionou um aumento de 30% de produtividade e uma redução de quase 20% nos erros de operação.

A segurança, privacidade e conformidade de dados foram outras vantagens do sistema. As informações são anonimizadas com os dados mantidos dentro do ambiente interno, aderindo a todos os padrões de confidencialidade.



Como melhorar a governança de IA

Em todo o mundo, governos e organizações estão em busca de orientações sobre como melhorar a governança de inteligência artificial (IA). Como podemos usar essa nova tecnologia para resolver nossos problemas? O que fazer para evitar ou gerenciar novas questões surgidas a partir de seu uso? Não há, é claro, uma abordagem única ou correta. No entanto, após muitos anos trabalhando para garantir que nossos próprios sistemas de IA sejam seguros, protegidos e confiáveis, oferecemos nossas ideias e sugestões para capacitar outras organizações em sua jornada para o uso responsável da IA.

Em maio de 2023, divulgamos o documento [*Governing AI: A Blueprint for the Future*](#) (*Governando a IA: Um Plano para o Futuro, em tradução livre*). A publicação apresenta um plano para abordar diversas questões atuais e emergentes de IA por meio de políticas públicas, leis e regulamentações. Esperamos que esse plano continue a guiar conversas importantes sobre uma necessária governança da IA. Aqui estão alguns de seus pontos-chave.

Cinco pontos para a governança pública da IA

1. Implementar e desenvolver sobre as novas estruturas de segurança de IA lideradas pelo governo americano. Acreditamos que uma das maneiras mais eficazes de avançar rapidamente no tema é construir sobre os recentes avanços no esforço governamental que promovem a segurança da IA. Um novo Quadro

de Gerenciamento de Riscos de IA, introduzido pelo [Instituto Nacional de Padrões e Tecnologia](#) (NIST, na sigla em inglês), do Departamento de Comércio dos Estados Unidos, é um ótimo exemplo.

2. Exigir freios de segurança eficazes para sistemas de IA que controlem infraestruturas críticas. Sugerimos novos requisitos de segurança de IA para controlar a operação de infraestrutura crítica designada, com uma abordagem abrangente à segurança do sistema que teria como prioridade a supervisão humana eficaz, a resiliência e a robustez.
3. Desenvolver um amplo arcabouço legal e regulatório baseado na arquitetura tecnológica para IA. Acreditamos que as responsabilidades regulatórias devem ser atribuídas às organizações com base em seu papel no gerenciamento de diferentes aspectos da tecnologia de IA, como nas camadas de aplicativos, de modelos e de infraestrutura de IA.

4. Promover a transparência e garantir o acesso acadêmico e sem fins lucrativos à IA.

Acreditamos que um objetivo público fundamental é avançar na transparência e ampliar o acesso aos recursos de IA. Temos o compromisso de informar ao público sobre nossas políticas, sistemas, desempenho e progresso em IA por meio de relatório de transparência anual e tomar outras medidas para expandir a transparência de nossos serviços de IA.

5. Buscar novas parcerias público-privadas para usar a IA como uma ferramenta eficaz no enfrentamento dos desafios sociais que acompanham a nova tecnologia. Testemunhamos o que as sociedades democráticas podem realizar quando aproveitam o poder da tecnologia e unem as esferas pública e privada. Temos de trabalhar juntos para usar a IA para proteger a democracia e os direitos fundamentais, proporcionar amplo acesso às habilidades em inteligência artificial capazes de promover o crescimento inclusivo e aproveitar do poder da IA para fazer avançar nas necessidades de sustentabilidade do planeta.

IA responsável: a abordagem Microsoft para a construção de sistemas de inteligência artificial que beneficiem a sociedade

O compromisso da Microsoft com desenvolvimento responsável da IA

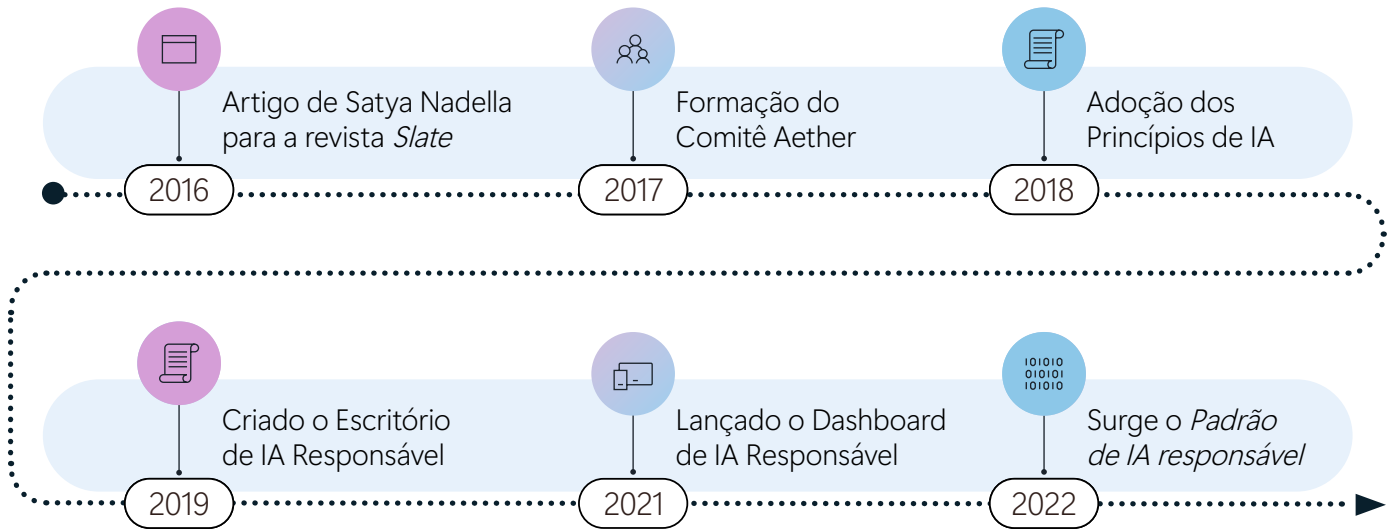
Nos últimos anos, atuamos para desenvolver uma inteligência artificial (IA) responsável, fundamentada em sólidos princípios éticos. Abordamos a questão com o humilde reconhecimento de que confiança é algo que não se ganha, mas deve ser conquistada por meio de ações. E com a compreensão de que nossa responsabilidade vai além da Microsoft, alcançando a comunidade de forma mais ampla. O foco esteve tanto em cumprir nossos próprios compromissos quanto em ajudar clientes e parceiros a seguirem na mesma direção.

Nossa jornada de IA responsável começou em 2016, quando Satya Nadella, CEO da Microsoft, compartilhou sua [visão](#) de como a humanidade sairia fortalecida pela IA. Satya expressou em artigo para a revista eletrônica *Slate* as primeiras ideias de nossos princípios centrais de IA, valores que permanecem até hoje. Com base nesse olhar, lançamos o [Comitê Aether da Microsoft](#), composto por pesquisadores, engenheiros e especialistas em políticas corporativas, que trouxeram conhecimentos sobre tendências emergentes em tecnologia. O empenho desse grupo nos levou à criação e adoção de nossos princípios de IA em 2018.

Aprofundamos os esforços em 2019 ao estabelecer o Escritório de IA Responsável. Essa unidade coordena a governança de nosso programa e colaborou, ao lado de toda a empresa, para a formulação da primeira versão do *Padrão de IA responsável*, um manual para transformar nossos princípios em orientações práticas para as equipes de engenharia que constroem sistemas de IA.

Em 2021, anunciamos publicamente os principais pilares estabelecidos para operacionalizar essa iniciativa. Contemplamos a expansão da capacitação, de processos e de ferramentas que pudessem nos ajudar a aumentar nossos esforços de IA responsável. Em 2022, trouxemos uma nova versão de nosso *Padrão de IA responsável*, evoluindo para o modelo que utilizamos hoje. Esse documento estabelece como a Microsoft construirá sistemas de IA usando métodos práticos para identificar, medir e mitigar previamente potenciais riscos. Essa abordagem de design responsável estabelece processos replicáveis para, desde o início, minimizar danos potenciais e ampliar os benefícios da IA.

Nossa jornada de IA responsável



Temos orgulho do nosso progresso nos últimos anos. Esses esforços nos trouxeram até onde estamos hoje e aprofundaram nosso compromisso de incorporar segurança e responsabilidade ao ciclo de vida de nossos sistemas de IA. Isso só é possível quando os princípios e as práticas de IA responsável transcendem os silos⁴ tradicionais e, também, equipes multidisciplinares trabalham juntas. Ao conhecer as oportunidades e os potenciais riscos, acreditamos que devemos compartilhar o que aprendemos e ajudar todas as organizações a aplicar práticas de IA responsável em seu trabalho. É exatamente isso que nós da Microsoft estamos fazendo – e esperamos dar o exemplo.

⁴ No ambiente de tecnologia, o termo silo é frequentemente usado para descrever a situação em que dados, informações ou sistemas são isolados, impedindo a conexão entre eles.

Tornando operacional a IA responsável na Microsoft

Definindo estruturas fundamentais de governança

Conforme o ritmo da IA avança, procuramos evoluir a estrutura de governança que estabelecemos para possibilitar o progresso e a responsabilidade como componentes fundamentais de nosso programa de IA responsável. A criação da estrutura de governança da Microsoft e a decisão de disseminar a ideia por toda a empresa foram impulsionadas pela liderança. O presidente do Conselho de Administração e CEO da Microsoft, Satya Nadella, e a equipe de liderança sênior fizeram com que a tarefa se tornasse uma obrigação de todos na companhia.

A liderança da Microsoft reconheceu que um único time ou disciplina encarregada da

IA responsável não seria o suficiente. Com as lições aprendidas com os nossos compromissos de longo prazo com privacidade, segurança e acessibilidade, percebemos que a IA responsável deve ser apoiada pelo mais alto escalão e defendida por colaboradores em todos os níveis.

Com esse objetivo, o Escritório de IA Responsável da Microsoft desenvolveu um sistema de governança que incorpora diversas equipes e funções. No operacional, times de engenharia, pesquisa e políticas corporativas desempenham papéis essenciais para difundir os princípios de IA responsável por toda a empresa. O tema também vem sendo incorporado às áreas de produto, de engenharia e de vendas, por meio da escolha de Campeões de IA Responsável (*ver texto na página 41*). Esses profissionais têm a tarefa de incentivar práticas de IA responsável dentro de seus setores, como adotar os padrões estabelecidos, identificar problemas, aconselhar diretamente seus integrantes sobre potenciais mitigações e incentivar uma cultura de inovação

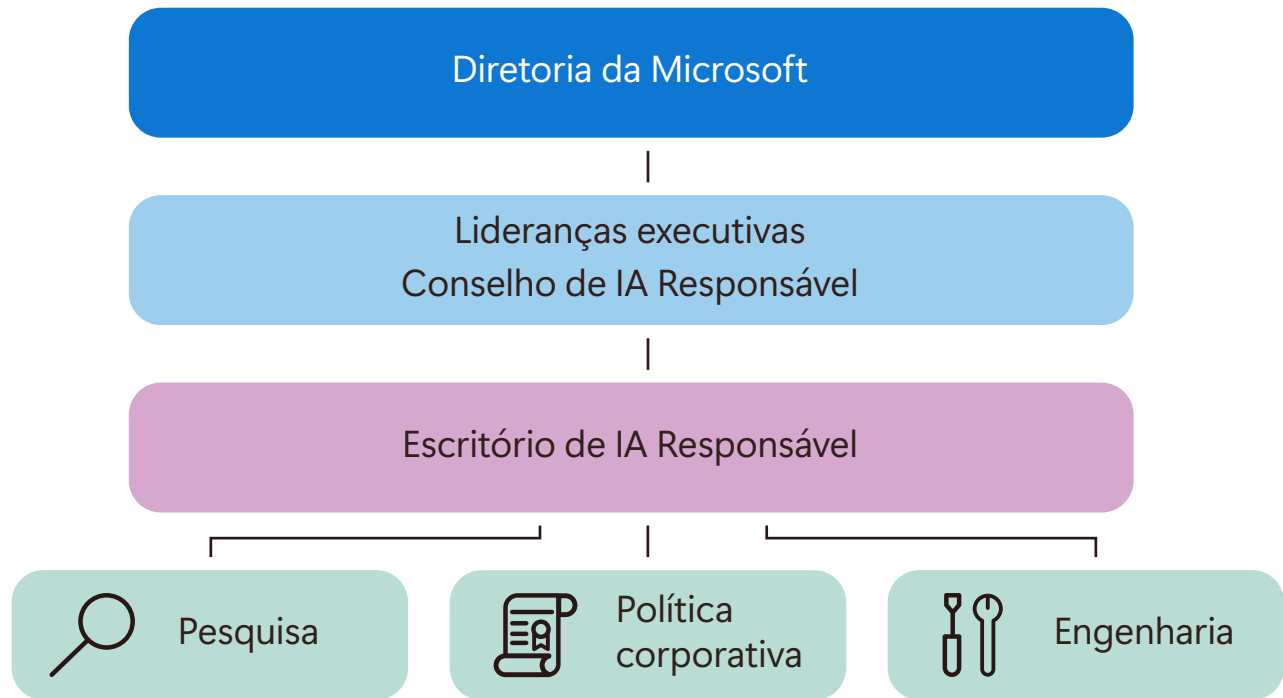
responsável. O Escritório de IA Responsável ajuda a coordenar essas equipes, aproveitando seu profundo conhecimento de produto e expertise no assunto para desenvolver uma abordagem consistente em toda a Microsoft.

No próximo nível, o Conselho de IA Responsável é um fórum para alinhamento de liderança e responsabilidade na implementação do programa de IA responsável, dirigido pelo vice-presidente do Conselho de Administração e presidente da Microsoft, Brad Smith, e pelo vice-presidente executivo de IA, Kevin Scott. Sua missão é definir a visão tecnológica da empresa e supervisionar a divisão de pesquisa da Microsoft. Seus membros se reúnem regularmente com os representantes do pessoal de pesquisa, política corporativa e engenharia dedicados à IA responsável, incluindo o Comitê Aether, o Escritório de IA Responsável e líderes de engenharia e parceiros de negócios seniores responsáveis pela implementação.

Marco de governança de IA responsável



Nosso ecossistema



Por fim, no topo, o Comitê de Política Ambiental, Social e Pública do Conselho de Administração da Microsoft faz a supervisão geral. O compromisso constante com essa representação superior assegura a aplicação, com rigor, da estrutura de gestão de riscos empresariais da Microsoft ao programa.

A necessidade de padronização

Desde a definição de um propósito até o design da forma como as pessoas vão interagir, precisamos manter as pessoas no centro de todas as decisões da IA.

Embora nossos princípios de IA responsável estabeleçam valores duradouros que procuramos seguir, sentíamos falta de orientações mais específicas sobre como construir e implantar sistemas de IA com responsabilidade. Por essa razão,

desenvolvemos nosso [Padrão de IA responsável](#), um guia prático que registra um conjunto de regras a serem seguidas pelas equipes de engenharia, de modo a assegurar a manutenção de nossos princípios de IA como prática diária.

O *Padrão de IA responsável* fornece orientações sobre como construir sistemas de IA de maneira responsável. O manual é resultado de um esforço de muitos anos, com contribuições de pesquisadores, engenheiros, advogados, designers e especialistas em políticas corporativas. Consideramos ser esse um avanço significativo em nossa prática de IA responsável, por oferecer orientações mais concretas e práticas sobre como identificar, medir e mitigar antecipadamente potenciais danos. O documento exige ainda que as equipes adotem ferramentas e controles que garantam o uso benéfico de seus produtos, ao mesmo tempo que coíbe usos indevidos.

Existem duas maneiras pelas quais o *Padrão de IA responsável* oferece orientações concretas aos nossos engenheiros que atuam em todo o ciclo de vida de um produto de IA:

- **Articulando metas** – Trata da defesa dos princípios éticos da IA responsável. Princípios amplos, como responsabilidade, se desdobram em resultados reais, como garantir que sistemas de IA estejam sujeitos a avaliações de impacto, governança de dados e supervisão humana.
- **Montando uma cartilha** – Trata de procedimentos específicos que as equipes devem seguir ao longo do ciclo de vida de um sistema de IA, a fim de cumprir metas estabelecidas. As etapas mapeiam os recursos, ferramentas e práticas disponíveis para abastecer os responsáveis de ferramentas para transformar as metas em realidade. Por exemplo, uma das metas do manual é minimizar o tempo necessário para corrigir falhas previsíveis ou conhecidas de um sistema de IA. Para alcançar esse objetivo, pedimos que sejam identificados possíveis danos por meio de procedimentos repetitivos conhecidos como *red teaming*. Em seguida, é verificada a prevalência dessas falhas, buscando mitigá-las por meio de testes e da implementação de novas ferramentas e estratégias. Esse ciclo de identificação, medição e redução de possíveis danos de um sistema de IA sustenta muitos dos requisitos da IA responsável.

Instalação de freios e contrapesos

Ao desenvolver e atualizar o *Padrão de IA responsável*, reconhecemos desde o início que é impossível reduzir todas as complexas considerações técnicas para muitos casos diferentes de uso em um conjunto exaustivo de regras predefinidas. Isso

nos levou a criar um programa e um processo para revisão contínua e supervisão de casos de alto impacto e questões em ascensão, que denominamos de Usos Sensíveis.

Red teaming em sistemas de IA

O termo *red teaming* historicamente descreve ataques de especialistas que atuam como adversários de um sistema com o objetivo de testar vulnerabilidades de segurança. Com o surgimento de modelos de linguagem grandes (LLMs, na sigla em inglês), o conceito se estendeu além da cibersegurança tradicional e passou a descrever muitos tipos de sondagens, testes e ataques a sistemas de IA. No universo dos LLMs, tanto ações a favor como as adversárias podem produzir resultados potencialmente prejudiciais, que podem assumir muitas formas, incluindo conteúdo prejudicial, como discurso de ódio, incitação ou exaltação de violência ou conteúdo sexual.

O *red teaming* é uma prática essencial no desenvolvimento responsável de sistemas e recursos que utilizam os LLMs. Os testadores ajudam a descobrir e identificar danos e, por sua vez, permitem estratégias de medição para validar a eficácia das mitigações.

A Microsoft conduziu exercícios de *red teaming* e implementou sistemas de segurança para seus modelos de Serviço OpenAI do Azure, como também para aplicativos destinados ao consumidor, como o novo Bing.

Nosso programa de Usos Sensíveis (*ler mais na página 34*) proporciona uma camada adicional de supervisão para equipes que trabalham em situações de uso de IA de maior risco. A iniciativa começou em 2018 sob a coordenação do Comitê Aether e, desde 2019, vem operando no âmbito do Escritório de IA Responsável. De julho de 2019 a maio de 2023, analisamos mais de 600 casos críticos em toda a Microsoft, incluindo aproximadamente 150 entre julho de 2022 a maio de 2023.

A análise de cada um deles começa com um relatório obrigatório, ponto de partida de um processo prático de avaliação de projetos de IA responsável e consultoria junto à equipe de Usos Sensíveis do Escritório de IA Responsável. O processo é concluído com orientações e requisitos específicos, adicionais às exigências básicas. O time de Usos Sensíveis é acionado quando o uso previsível ou indevido de um sistema de IA pode:

1. Ter consequências legais ou para a vida de um usuário.
2. Apresentar o risco de lesões físicas ou psicológicas significativas.

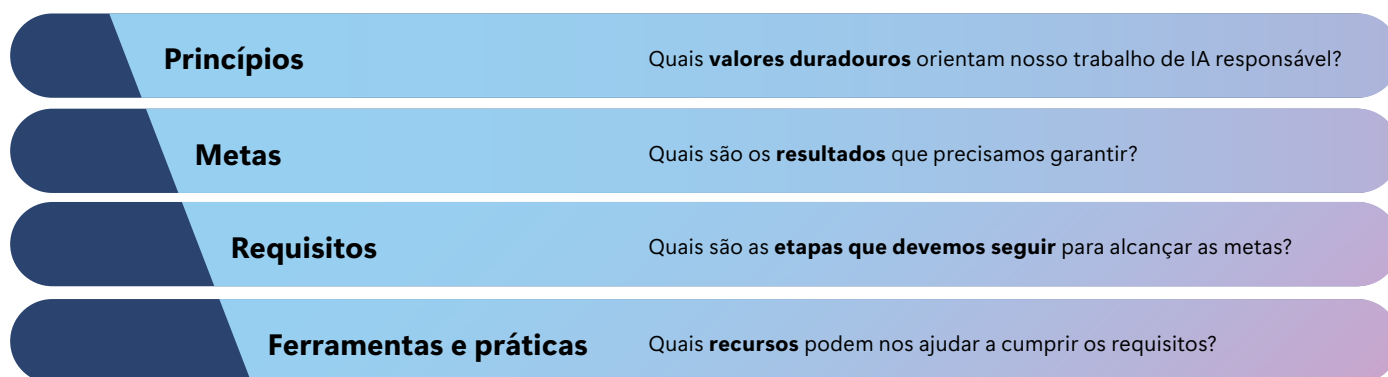
3. Restringir, infringir ou prejudicar o exercício dos direitos humanos de uma pessoa.

Uma vez acionada, a equipe de Usos Sensíveis faz uma triagem e inicia o processo de análise com os responsáveis pelo projeto, ao lado de seu campeão de IA Responsável e outras partes interessadas relevantes.

Para auxiliar na sua estruturação e aprofundar nas questões, utilizamos registros como a Avaliação de Impacto Responsável de IA e a documentação técnica do produto, além de interações próximas e contínuas com os participantes diretos.

É ainda comum buscar o parecer de especialistas de dentro da Microsoft, por meio de consultas direcionadas. Em cenários de uso especialmente impactante ou inovadores, submetemos as informações ao Painel de Usos Sensíveis, composto por profissionais de todo o mundo das áreas de engenharia, pesquisa, direitos humanos, política corporativa, jurídico e atendimento ao cliente. É um grupo essencialmente multidisciplinar, que reúne

A anatomia do Padrão de IA responsável



desde diplomatas, sociólogos e analistas de políticas públicas até cientistas de dados e engenheiros de aprendizado da máquina.

Ao concluir a análise, os responsáveis estabelecem suas condições para que o projeto avance. São exigências adicionais, além dos requisitos padrões, específicas para o contexto avaliado. Em mais de uma ocasião, recusamos oportunidades de construir e implementar aplicativos de IA depois da avaliação do programa de Usos Sensíveis por entendermos que a proposta não estava suficientemente alinhada com nossos princípios de IA responsável.

Por exemplo, o vice-presidente do Conselho de Administração e presidente da Microsoft, Brad Smith, precisou vir a público anunciar que, em razão da análise do grupo, seria prematuro a utilização em tempo real de câmeras instaladas nos uniformes e nos veículos para reconhecimento facial por um departamento de polícia local na Califórnia, nos Estados Unidos. Além de navegar pelos desafios técnicos de realizar a tarefa em um ambiente não controlado, o processo nos auxiliou na formulação de um parecer que indicou que seria necessário um amplo debate com a sociedade sobre o uso do reconhecimento facial, além de uma mudança na legislação.

Padrão de IA responsável em ação: equidade na tecnologia de conversão de fala em texto

A tecnologia de conversão de fala em texto pode melhorar a vida das pessoas, desde o envio de mensagens sem usar as mãos até a ajuda na comunicação de pessoas com deficiência auditiva. No entanto, um estudo acadêmico revelou que essa tecnologia produzia taxas de erro quase duas vezes maiores para membros de algumas comunidades negras do que para usuários brancos nos Estados Unidos.

Esses resultados nos motivaram a examinar mais de perto os sistemas de IA que alimentam as tecnologias de conversão de fala em texto da Microsoft, para garantir que eles sejam treinados de forma adequada sobre a diversidade de padrões de fala. Recorremos a um sociolinguista e ampliamos nossos esforços de coleta de dados para reduzir a diferença de desempenho em nossas soluções. Durante o processo, envolvemos interessados de fora da Microsoft, como especialistas e pessoas de diversas comunidades.

Essas lições foram de valor inestimável e, desde então, passaram a fazer parte do *Padrão de IA responsável* para detalhar ainda mais as etapas específicas que as equipes devem seguir para garantir que os sistemas de IA da Microsoft sejam projetados para oferecer a mesma qualidade de serviço para grupos demográficos identificados, incluindo os marginalizados, e para ajudar a Microsoft e outras organizações a melhor aproveitar os benefícios dessas tecnologias e evitar eventuais falhas no futuro.

Outro desdobramento importante da atuação do time de Usos Sensíveis foi nossa [política de Acessos Limitados](#) para serviços de plataforma de IA mais sensíveis, que agrega uma camada extra de verificação em seu uso e implantação.

Pelas diretrizes estabelecidas, nós não apenas instalamos controles técnicos para mitigar riscos, mas também exigimos que potenciais clientes solicitem uma permissão de uso, com o compromisso de que a solução será aplicada em conformidade com um de nossos padrões aceitáveis predefinidos e uma declaração de que leram e concordam com os termos de serviço.

Aprovamos apenas aplicativos alinhados com nossos princípios de IA responsável.

Análise prática do programa de Usos Sensíveis:

Voz Neural Personalizada do Azure

A Voz Neural Personalizada do Azure IA é uma inovadora tecnologia de fala da Microsoft que permite a criação de uma voz sintética que soa quase idêntica à fonte original. Essa tecnologia já foi utilizada por clientes empresariais, como a companhia de telecomunicações AT&T e a seguradora Progressive. Seu uso tem aplicações promissoras na educação, acessibilidade e entretenimento. No entanto, possíveis abusos podem ocorrer, como a personificação inadequada de vozes alheias para iludir quem ouve.

Em consistência com nossa abordagem cautelosa para sistemas de IA de maior risco, à medida que evoluía e progredia para uma disponibilidade mais ampla, a Voz Neural Personalizada passou por várias análises de Uso Sensível. Tais análises nos levaram a adotar um marco de controle em camadas para o serviço. Limitamos, por exemplo, o acesso de clientes, definimos quais são os usos aceitáveis por meio de um formulário, implementamos mecanismos de consentimento do orador, criamos termos de uso específicos, publicamos documentação de transparência detalhando riscos e limitações, e estabelecemos diretrizes técnicas para garantir a participação ativa do orador ao criar uma voz sintética.

Por meio desses e outros controles, estamos ajudando a proteger contra usos indevidos, ao mesmo tempo que mantemos os benefícios da tecnologia. Um dessas vantagens é uma espécie de banco de vozes, uma solução que permite a pessoas que estejam em risco de perder sua voz possam registrá-la para uso posterior, como a possibilidade de recriar sua voz treinando um modelo sintético por meio da Voz Neural Personalizada.

ESTUDO DE CASO

Aplicando nossa abordagem de IA responsável ao novo Bing

Em fevereiro de 2023, a Microsoft lançou o novo Bing, uma experiência de busca na web aprimorada pela IA. O serviço ajuda os usuários a resumir resultados de busca na web e oferece uma experiência de bate-papo. Os usuários também podem gerar conteúdos criativos, como poemas, piadas, cartas e, com o Bing Image Creator, imagens. A aplicação é executada em diversas plataformas tecnológicas avançadas da Microsoft e da OpenAI, incluindo o GPT-4, um modelo de linguagem grande de última geração (large language model, LLM, na sigla em inglês) da OpenAI. As equipes de IA responsável em toda a Microsoft trabalharam com o GPT-4 durante meses antes de seu lançamento para desenvolver um conjunto personalizado de capacidades e técnicas para acoplar essa tecnologia de IA de ponta e busca na web ao novo Bing.

Na preparação para o lançamento, a Microsoft aproveitou todo o poder de nosso ecossistema de IA responsável. A nova experiência do Bing foi desenvolvida de acordo com os princípios de IA da Microsoft, o *Padrão de IA responsável* da Microsoft e em parceria com especialistas de toda a empresa, incluindo o Escritório de IA Responsável da Microsoft, nossas equipes de engenharia, a Microsoft Research e o Comitê Aether.

Guiados por essas orientações, buscamos identificar, mensurar e reduzir possíveis danos e usos indevidos do novo Bing, ao mesmo tempo que garantimos os usos transformadores e benéficos que a nova experiência proporciona. Nos tópicos a seguir, descrevemos nossa abordagem.

Identificar

No nível do modelo, nosso trabalho começou com análises exploratórias do GPT-4 no final do terceiro trimestre de 2022, incluindo a realização de extensos procedimentos *red teaming* em colaboração com a OpenAI. Esses testes foram projetados para avaliar como a tecnologia mais recente se comportaria sem a aplicação de quaisquer proteções adicionais. Nossa intenção específica era produzir respostas prejudiciais (as “respostas” são produtos do sistema de IA – nesse caso, de um LLM –, podendo ser chamadas também de “conclusão” ou “geração”) para detectar potenciais modalidades de uso indevido e identificar capacidades e limitações. A soma de nossos aprendizados serviu para promover o avanço do desenvolvimento do modelo da OpenAI, ampliou nossa compreensão dos riscos e contribuiu para estratégias de mitigação precoce do novo Bing.

Além do *red teaming* no nível modelo, um grupo multidisciplinar de especialistas conduziu várias rodadas desses testes ao nível da aplicação nas novas experiências de IA do Bing antes de disponibilizá-las em um pré-lançamento limitado. Esse processo nos ajudou a melhor entender como o sistema poderia ser explorado por usuários mal-intencionados e a aprimorar nossas medidas de mitigação. Testadores não adversários também avaliaram extensivamente as novas funcionalidades do Bing em busca de falhas e vulnerabilidades.

Medir

Os testes *red teaming* podem desvendar danos específicos, mas, na prática, os usuários terão milhões de diferentes tipos de comunicação com o novo Bing. Além disso, as conversas incluem diversas interlocuções e contextos, o que transforma a identificação de respostas prejudiciais dentro de uma interação em uma tarefa complexa. Para entender melhor e abordar o potencial de danos nas novas experiências de IA do Bing, desenvolvemos métricas adicionais de IA responsável com a finalidade de medir potenciais danos, como violações de segurança, conteúdo prejudicial e informações sem fundamento.

Também habilitamos a medição em maior escala por meio de fluxos de medição parcialmente automatizados, o que nos permite realizar rapidamente medições de potenciais danos em escala, testando cada alteração antes de disponibilizar para os usuários. Conforme identificamos novos problemas durante o período de testes, não apenas damos continuidade aos procedimentos de *red teaming* em andamento, mas também expandimos os conjuntos de medição para avaliar danos adicionais.

Mitigar

Assim que identificamos e medimos potenciais anomalias e usos indevidos, desenvolvemos mitigações adicionais às utilizadas na ferramenta de pesquisa tradicional. Entre elas:

- **Período de pré-lançamento faseado.** Nossa estratégia de lançamento incremental tem sido uma parte fundamental de como transferimos tecnologias com segurança dos nossos laboratórios para o mundo. Estamos comprometidos com um processo voluntário e cuidadoso para garantir os benefícios do novo Bing. Limitar o número de pessoas com acesso durante o pré-lançamento nos permitiu descobrir como as pessoas usam o novo Bing, e como poderiam utilizá-lo indevidamente, antes de uma disponibilização mais ampla. Continuamos a fazer alterações no novo Bing diariamente para melhorar o desempenho do produto, aprimorar as soluções existentes e implementar novos melhoramentos com base no que aprendemos.
- **Classificadores baseados em IA e metaprompting⁵ para mitigar danos ou uso indevido.** O uso de LLMs pode produzir conteúdo problemático que poderia levar a danos ou uso indevido. Classificadores e metaprompting são dois exemplos de soluções implementadas no novo Bing para ajudar a reduzir riscos

⁵ Metaprompting se refere à técnica de usar ou criar estímulos para instruir modelos de linguagem ou sistemas de inteligência artificial a gerar determinados tipos de resposta.

associados a esse aspecto. Classificadores de texto apontam conteúdo potencialmente prejudicial em consultas de pesquisa, prompts de chat ou respostas geradas. As sinalizações levam a ações preventivas, como não retornar o conteúdo gerado ao usuário, desviar o usuário para um tópico diferente ou redirecioná-lo para a pesquisa tradicional. O metaprompting envolve dar instruções ao modelo para orientar seu comportamento podendo, por exemplo, incluir uma linha como “comunique-se no idioma de escolha do usuário”.

- **Resultados de busca como base.** O novo Bing foi projetado para fornecer respostas baseadas em informações de resultados de pesquisa na web quando os usuários procuram informações. Por exemplo, o sistema recebe textos dos principais retornos às consultas e instruções por meio do metaprompt para fundamentar sua resposta. No entanto, ao resumir o conteúdo da web, o novo Bing pode incluir informações que não constam de suas fontes de entrada. Em outras palavras, ele pode produzir resultados não fundamentados. Implementamos diversas medidas para reduzir o risco de os usuários dependerem demais do conteúdo não fundamentado em cenários de resumo e experiências de bate-papo. Por exemplo, as respostas do novo Bing baseadas em pesquisas incluem referências aos sites de origem, para que os interessados possam verificar a resposta e obter maiores informações. Os usuários também recebem um aviso explícito de que estão interagindo com um sistema de IA e são aconselhados a verificar a origem dos resultados na web, para ajudá-los a exercer seu melhor discernimento.

- **Limitando o desvio de interlocuções.** Durante o pré-lançamento, aprendemos que sessões de bate-papo muito longas podem resultar em respostas repetitivas, pouco úteis ou inconsistentes com o tom pretendido do novo Bing. Para lidar com esse inconveniente, limitamos o número de interlocuções (trocas que contêm tanto uma pergunta do usuário quanto uma resposta do Bing) por sessão, até que possamos atualizar o sistema para melhor mitigar o problema.
- **Divulgação de IA.** No novo Bing, os usuários são avisados que estão interagindo com um sistema de IA, assim como sobre oportunidades para aprender mais sobre a ferramenta.

Nossa abordagem para identificar, medir e mitigar danos continuará a evoluir à medida que aprendemos mais – e conforme fazemos melhorias com base em feedbacks.

Compartilhamos mais detalhes do trabalho de IA responsável para o novo Bing, incluindo nosso empenho nas áreas de privacidade, segurança digital e transparência, em <https://aka.ms/ResponsibleAI-NewBing>.

Avançando na IA responsável por meio da cultura empresarial

Procedimentos e padrões constituem parte essencial da operacionalização da IA responsável e nos ajudam a construir uma cultura comprometida com os princípios e ações da IA responsável. Essas abordagens complementares nos ajudam a transformar nossos compromissos em realidade.

As pessoas formam o núcleo da cultura da Microsoft. Cada indivíduo contribui com a nossa missão e objetivos. Para aprofundar nossa identidade de promoção da IA responsável, investimos em talentos focados em IA e incorporamos essa responsabilidade em cada função.

Investindo em talentos

Ao longo dos anos, investimos significativamente em pessoas como parte de nosso compromisso com a IA responsável. Atualmente, temos quase 350 funcionários atuando com IA responsável, mais de um terço dedicado integralmente à tarefa. Esses funcionários trabalham em áreas como política corporativa, engenharia, pesquisa, vendas e outras funções essenciais, incorporando a IA responsável em todos os aspectos de nosso negócio.

Solicitamos que as equipes que desenvolvem e usam sistemas de IA examinem a tecnologia por meio de uma ótica sociotécnica. Isso significa que contemplamos complexos fatores culturais, políticos e sociais da IA à medida

que se manifestam em diferentes contextos de implantação – o que representa uma mudança fundamental em relação à abordagem convencional da ciência da computação. Embora o treinamento e as práticas que desenvolvemos ajudem as equipes a prever os impactos benéficos e potencialmente prejudiciais da IA nos níveis individual, social e global, não é suficiente. Os desenvolvedores de sistemas de IA, e as lideranças às quais respondem, ainda podem ter pontos cegos. É por essa razão que a diversidade e a inclusão são primordiais para o nosso compromisso com a IA responsável.

O argumento para investir em uma força de trabalho diversificada e em uma cultura inclusiva está bem estabelecido, mas é difícil exagerar o valor da diversidade e da inclusão para avançarmos nesse tema. É por isso que o investimento contínuo e crescente em nosso ecossistema de IA responsável inclui a contratação de talentos com diferentes perfis. Como mostra nosso [Relatório anual de diversidade e inclusão](#), a Microsoft continua a fazer progressos incrementais em relação a diversidade e inclusão. No entanto, como toda a indústria de tecnologia, ainda há um longo caminho a percorrer. O campo da IA continua sendo predominantemente branco e masculino: apenas cerca de um quarto dos colaboradores que trabalha em soluções de IA se identificam como mulheres ou minorias raciais ou étnicas, de acordo com a [pesquisa global de IA de 2022](#) da McKinsey.

Continuaremos a promover a diversidade e a inclusão em todos os níveis, especialmente dentro de nosso programa de IA responsável. Para construir sistemas de IA que sirvam à sociedade da forma mais ampla possível, devemos recrutar e manter uma comunidade de funcionários diversificada, dinâmica e engajada.

Incorporar o espírito de IA responsável em cada atribuição

Acreditamos que todos na Microsoft assumem a oportunidade e a responsabilidade de contribuir para a construção de sistemas de IA que correspondam a nossos compromissos com a IA responsável. Todos os colaboradores, em qualquer função, entregam algo nesse desafio por meio de suas diferentes habilidades, perspectivas e paixões. Essa mudança de perspectiva – de que, independentemente de título, cargo ou equipe, todos podem promover uma IA responsável – requer uma transformação na cultura.

Em apoio a esse crescimento cultural, investimos no desenvolvimento das habilidades dos colaboradores e incentivamos sua participação.

Desenvolvimento de conhecimentos e capacidades

Desenvolvemos treinamentos e práticas para capacitar nossas equipes a pensar amplamente sobre o impacto potencial dos sistemas de IA em indivíduos e na sociedade.

Por exemplo, quando estamos nas etapas iniciais de design de um sistema de IA, nossos guias de Avaliação de Impacto trazem as seguintes orientações:

- Relacionar o(s) uso(s) pretendido(s) do sistema de IA.

Trabalhando rumo a um futuro global e inclusivo para a IA

A criação de sistemas de IA e seu debate regulatório tendem a se concentrar nas economias mais avançadas. No entanto, o desenvolvimento e o uso responsável da IA precisam refletir a diversidade de perspectivas globais, incluindo vozes de países em desenvolvimento.

Na Microsoft, nos esforçamos para incluir países em desenvolvimento em nossa defesa de um conjunto de políticas de IA globalmente coerente e de práticas responsáveis de IA relevantes para todos. É com satisfação que compartilhamos dois exemplos desse compromisso:

- Conselho Empresarial Ibero Americano da Unesco: a Microsoft e a Telefônica estão copresidindo o esforço para promover a adoção da Recomendação da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco) sobre Ética da Inteligência Artificial na América Ibérica. Trata-se do primeiro conjunto de políticas que faz sentido globalmente, assinado por todos os 193 Estados membros da entidade.
- Programa de Bolsas de IA Responsável: a iniciativa reúne representantes da sociedade civil, academia, setor privado e setor público de países em desenvolvimento. Lançado pela Microsoft e pelo Strategic Foresight Hub, do Stimson Center, seu objetivo é promover o desenvolvimento e uso responsável da IA. Os bolsistas contribuirão para uma série de discussões abordando as melhores práticas emergentes e os impactos multifacetados da IA em países em desenvolvimento.

- Questionar como o sistema de IA resolverá o problema para o qual foi projetado.
- Identificar todas as partes interessadas afetadas (e não apenas quem é o cliente imediato da Microsoft).
- Articular os possíveis impactos negativos e/ou benefícios que podem afetar cada parte interessada.
- Descrever as mitigações preliminares para possíveis impactos negativos.

Para ajudar na condução das avaliações de impacto, o Escritório de IA Responsável

desenvolveu treinamentos sob demanda, workshops presenciais e documentos de orientação com exemplos e prompts pré-formatados. Como parte de nosso compromisso de compartilhar melhores práticas, nosso modelo de Avaliação de Impacto e nosso documento de orientação estão disponíveis publicamente.

Nossos cursos de IA responsável, disponíveis para todos os funcionários da Microsoft, estimulam os alunos a estudar a abordagem da Microsoft à IA responsável e a se aprofundar em nossos princípios de IA responsável e processos de governança. Oferecemos, ainda, conteúdos específicos para cientistas de dados e engenheiros de aprendizagem de máquina.

IA responsável incorporada ao Azure Machine Learning



Equidade

Avaliar a equidade e mitigar questões relacionadas para construir modelos para todas as pessoas.



Explicabilidade

Compreender as previsões do modelo por meio da geração de valores das funções importantes para o seu modelo.



Contrafactuais

Observar perturbações das funções e encontrar as alternativas mais próximas com previsões diferentes do modelo.



Fluxo de prompts

Criar fluxos de trabalho para aplicações baseadas em linguagem para simplificar a criação, avaliação e ajuste de prompts.



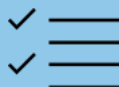
Análise causal

Estimar o efeito de uma função sobre resultados do mundo real.



Análise de erro

Identificar recortes de conjuntos de dados com altas taxas de erro e visualizar a distribuição de erros no seu modelo.



Balanço de IA responsável

Obter um resumo em PDF de suas informações de IA responsável para compartilhar com partes interessadas técnicas e não técnicas para auxiliar nas revisões de conformidade.



Segurança de conteúdo do Azure

Detectar conteúdo de ódio, violência, sexual e automutilação em vários idiomas, tanto em imagens quanto em texto.

Os Campeões de IA Responsável

A Microsoft tem uma rede de Campeões de IA Responsável em toda a organização. Essas pessoas vêm desempenhando papel fundamental na promoção de uma cultura de responsabilidade desde o desenho de um produto.



Mihaela Vorvoreanu, Pesquisa

“A IA responsável não surge apenas de um problema técnico com soluções técnicas. Requer uma colaboração profunda desde o início, não apenas com especialistas em IA responsável, mas também com especialistas em pessoas.”



Alejandro Gutierrez Munoz, Ciência de Dados

“A promoção da IA responsável é essencial para alinhar sistemas de IA com princípios éticos, estimular a confiança, garantir conformidade e promover a responsabilidade social.”



Shweta Gupta, Engenharia Cliente

“Acredito que a aplicação de princípios de IA responsável, reunindo um conjunto diversificado de partes interessadas enquanto desenvolve soluções de IA, nos ajuda a identificar e abordar potenciais riscos, assim como garante que o sistema em desenvolvimento apoia holisticamente seus objetivos.”



Ferdane Bekmezci, Ciência de Dados

“Demora um tempo para incutir uma cultura em uma organização. Tenho paixão por incentivar sua adoção por toda a empresa, pois é importante assegurar que a IA seja desenvolvida e utilizada de forma ética e socialmente confiável.”



Lisa Mueller, Design

“A IA está mudando rapidamente, portanto o crescimento de comunidades e a adoção dos princípios por todos na empresa é importante para construir, cultivar e ampliar a confiança nos sistemas de IA. Como parte dessa abordagem, é também importante incluir muitas disciplinas que auxiliem nesse esforço para realmente fazer a diferença.”

As equipes também têm acesso a uma ampla gama de especialistas em IA responsável em todo o ecossistema da Microsoft, que fornecem engajamento e feedback em tempo real durante todo o ciclo de vida do produto. Essa comunidade inclui o Comitê Aether, o Escritório de IA Responsável e uma grande e crescente turma de Campeões de IA Responsável, que impulsionam a adoção do *Padrão de IA responsável*.

Promovendo a colaboração

Reconhecemos desde o início de nossa jornada de IA responsável os papéis críticos que os pesquisadores, especialistas em política corporativa e engenheiros da Microsoft desempenham na construção de nossa prática de IA responsável. Cada grupo traz insights e expertise essenciais para nosso trabalho, e nos esforçamos para possibilitar a colaboração entre eles.

- Pesquisadores, com uma gama de especializações, do aprendizado de máquina às ciências humanas, nos ajudam a vislumbrar a vanguarda dos sistemas de IA. Eles oferecem as melhores práticas na identificação, medição e mitigação de potenciais danos apresentados pelos sistemas de IA, bem como insights sobre as emocionantes oportunidades para a inovação em IA.
- Especialistas em políticas corporativas definem e operacionalizam a governança para a IA responsável, incluindo a elaboração das regras para orientar o desenvolvimento de sistemas de IA. Nosso marco de governança define papéis e atribuições em toda a organização, de forma a criar responsabilidades e incentivar a colaboração.
- Engenheiros projetam e desenvolvem sistemas de IA em conformidade com o *Padrão de IA responsável*. Eles automatizam e dimensionam as etapas necessárias para identificar, medir e mitigar potenciais danos provocados pelos sistemas de IA. Também criam soluções completamente novas de IA responsável, viabilizadas por aprendizados.

A colaboração frequente e a dependência da experiência uns dos outros – práticas reforçadas pela liderança – nos ajudaram a criar uma cultura que leva a soluções mais benéficas e responsáveis. Por meio do diálogo contínuo, as equipes relatam consistentemente que uma abordagem centrada no ser humano e colaborativa para a IA resulta não apenas em um produto responsável, mas também em um produto melhor.

Northumbria Healthcare NHS: personalização de triagem para cirurgia usando a IA

Optar por uma cirurgia eletiva, como colocação de prótese de joelho, é uma decisão importante. Os pacientes recorrem a seus médicos para avaliar os prós e os contras. Médicos do Sistema Nacional de Saúde do Reino Unido (NHS) empregavam estatísticas tradicionais para analisar dados de avaliações de risco cirúrgico de seus pacientes. Os cirurgiões do Northumbria Healthcare NHS Trust pensaram em uma maneira de realizar tais avaliações de forma mais personalizada com a IA.

A equipe consultou especialistas da Microsoft para construir um modelo usando o Microsoft Azure Machine Learning e o Dashboard de IA Responsável. O protótipo analisa 220 informações dos pacientes, incluindo histórico de tabagismo e idade. A aplicação ajudou a revelar que a contagem de plaquetas tinha um peso maior do que o esperado na determinação de risco. As ferramentas incorporadas ao dashboard auxiliaram a evitar vieses e capacitaram os médicos a entender como os resultados foram alcançados. Essa transparência permite explicar os riscos e fazer recomendações para os pacientes.

O modelo capacita os clínicos a adotar um diagnóstico individualizado. As avaliações personalizadas orientam as conversas entre paciente e médico, ajudam os prestadores de serviços no encaminhamento de pacientes aos centros cirúrgicos, onde podem receber o nível apropriado de atendimento, e identificam pacientes com riscos elevados antes que ocorram problemas. Embora esteja sendo inicialmente utilizado para avaliar candidatos a procedimentos de colocação de prótese de articulações, os médicos estão expandindo o seu uso também para outras modalidades de cirurgias.

Fortalecendo os clientes em suas jornadas de IA responsável

Um dos nossos compromissos mais importantes com a IA responsável é ajudar nossos clientes em sua jornada pela IA responsável, compartilhando aprendizados. Nossos esforços por si só não são suficientes para garantir os ganhos sociais que vislumbramos quando a IA responsável é adotada na prática.

Como parte desse compromisso, disponibilizamos documentação de transparência para serviços de IA de nossas plataformas, na forma de Notas de Transparência, para capacitar nossos clientes durante a implantação responsável de seus sistemas. Esses registros comunicam em linguagem clara e simples os propósitos, capacidades e limitações dos sistemas de IA para que todos possam entender quando e como implantar nossas tecnologias. Também identificam casos de uso que extrapolam as capacidades da solução e do *Padrão de IA responsável*. As Notas de Transparência preenchem a lacuna entre o marketing e a documentação técnica, comunicando proativamente informações que nossos clientes precisam saber para implantar IA de maneira responsável. Você pode ver um exemplo delas, formuladas para o Serviço OpenAI do Azure, [no site da Microsoft](#).

Nossos clientes também precisam de ferramentas funcionais para operacionalizar práticas de IA responsável. Ao longo dos anos, a pesquisa de IA responsável na Microsoft levou à incubação de ferramentas como Fairlearn e [InterpretML](#). A

coleção de dispositivos cresceu, abrangendo muitos aspectos práticos, incluindo a capacidade de identificar, diagnosticar e mitigar erros e limitações potenciais dos sistemas de IA. Desde sua concepção original dentro da Microsoft, esses recursos continuam a melhorar e a evoluir externamente por meio das contribuições de comunidades de código aberto ativas. Você pode encontrar esse material no repositório GitHub da [Responsible AI Toolbox](#). Nossa ferramenta mais recente é o [Azure Content Safety](#), que ajuda as empresas a criarem ambientes e comunidades on-line mais seguros por meio de modelos projetados para detectar conteúdo de ódio, violência, sexual e de automutilação em vários idiomas, tanto em imagens quanto em texto.

Construindo sobre o Responsible AI Tool Box, o programa de IA responsável da Microsoft investiu na integração direta ao Azure Machine Learning (*ler mais na página 40*) de algumas das ferramentas de IA responsável mais maduras, para que nossos clientes também se beneficiem do desenvolvimento de sistemas e ferramentas de engenharia. A coleção de recursos, conhecida como [Dashboard de IA Responsável](#), oferece um painel único para que desenvolvedores de aprendizagem de máquina e outros interessados possam aperfeiçoar modelos e tomar decisões responsáveis conforme personalizam ou constroem novos sistemas de IA.



Alguns de nossos recursos mais recentes incluem suporte para dados de texto e imagem que permite aos usuários avaliarem grandes modelos de dados não estruturados durante as etapas de construção, treinamento e avaliação, e a ferramenta Prompt Flow, que oferece uma experiência simplificada para prompts, avaliação e ajuste de modelos de linguagem grandes, incluindo medidas como a fundamentação.

Investimos e continuaremos a investir, por meio de pesquisa, na transformação de inovações de IA responsável em ferramentas práticas que apoiem nossos clientes em suas jornadas.

A comunidade envolvida no desenvolvimento, avaliação e uso da IA se estende além de nossos clientes diretos. Para atender a esse amplo ecossistema, compartilhamos publicamente dispositivos-chave de nosso programa de IA responsável, incluindo nosso *Padrão de IA responsável*, o gabarito de Avaliação de Impacto e coleções de pesquisa de ponta. Nossos caminhos de aprendizado digital capacitam ainda mais os líderes a criar uma estratégia de IA eficaz, promover uma cultura de AI, inovar de forma responsável e muito mais. Esses recursos podem ser encontrados on-line em <https://aka.ms/rai>.

