

Fortalecendo a governança e a cooperação internacional para a promoção do desenvolvimento benigno da inteligência artificial

Strengthening international governance and cooperation to promote the benign development of artificial intelligence

Tan Tieniu*

►DOI: <https://doi.org/10.14295/principios.2675-6609.2025.173.009>

Depositphotos



A inteligência artificial, como tecnologia estratégica, está mudando a produção, o conhecimento e os estilos de vida humanos a uma velocidade sem precedentes

RESUMO

O desenvolvimento de um sistema de governança da inteligência artificial ainda se encontra atrasado em relação ao avanço revolucionário da tecnologia, tornando cada vez mais proeminentes as questões relativas à governança internacional da IA. Este ensaio procura compreender as tendências globais de desenvolvimento da tecnologia, avaliar de forma científica os riscos e desafios desse processo e elencar medidas que contribuam para fortalecer a governança e a cooperação internacionais, de forma a promover o desenvolvimento da IA em benefício da civilização e do bem-comum de toda a humanidade.

Palavras-chave: Inteligência artificial. Cooperação internacional. Governança internacional. China.

ABSTRACT

The development of an artificial intelligence governance system still lags behind the revolutionary advancement of the technology, making issues related to international AI governance increasingly prominent. This essay seeks to understand global trends in the development of the technology, to scientifically assess the risks and challenges of this process, and to outline measures that can help strengthen international governance and cooperation, in order to promote AI development for the benefit of civilization and the common good of all humankind.

Keywords: Artificial intelligence. International cooperation. International governance. China.

INTRODUÇÃO

Em 25 de abril de 2025, o secretário-geral Xi Jinping destacou, durante a 20ª Sessão de Estudos Coletivos do Birô Político do Comitê Central do Partido Comunista Chinês: “A inteligência artificial, como tecnologia estratégica que conduz uma nova rodada da revolução tecnológica e da transformação industrial, está mudando profundamente a produção e os estilos de vida humanos.”

A tecnologia de IA liberou em enorme medida a capacidade mental humana, transformando fundamentalmente o paradigma da produção de conhecimento e modificando os horizontes da civilização a uma velocidade e escala sem precedentes. Ela está influenciando profundamente o cenário econômico, político e militar global, bem como o desenvolvimento das sociedades, e atraindo cada vez mais a atenção de governos, indústrias e academias em todo o mundo. Em contrapartida, o sistema de governança da IA encontra-se em um estágio significativamente aquém do progresso da tecnologia, tornando as questões sobre a governança internacional do setor cada vez mais proeminentes. Portanto, precisamos entender com exatidão as tendências globais de desenvolvimento da IA, avaliar cientificamente os riscos e desafios que seu desenvolvimento suscita, fortalecer a governança e a cooperação internacionais a respeito do setor e promover seu desenvolvimento em uma direção benigna para a humanidade.

COMPREENDENDO BEM A TENDÊNCIA GLOBAL DE DESENVOLVIMENTO DA INTELIGÊNCIA ARTIFICIAL

Desde sua concepção, em 1956, a inteligência artificial passa por quase 70 anos de altos e baixos. Nos últimos anos, as tecnologias de IA generativa, representadas por ChatGPT, Sora e DeepSeek, tiveram avanços contínuos, criando uma nova onda de entusiasmo em todo o mundo. O desenvolvimento dessa tecnologia apresenta as seguintes tendências básicas:

- 1) *As IAs especializadas estão amadurecendo*: as tecnologias de IA especializadas são aquelas destinadas a campos específicos e tarefas singulares, com demandas claras e limites de aplicação bem definidos, alcançando, assim, avanços significativos em seus respectivos campos. Testes de item único com níveis parciais de inteligência mostram que a IA pode superar a inteligência humana. Por exemplo: características biométricas como rostos, íris e modos de andar podem ser reconhecidas com alta precisão e eficiência, o que tem sido aplicado com sucesso em diversas áreas, incluindo segurança pública, assuntos civis e serviços de socorro; sistemas de diagnóstico assistidos por IA cobrem um número superior a 3 mil doenças e são utilizados em mais de 27 mil instituições de atenção primária à saúde; a Tesla lançou o primeiro automóvel com direção totalmente guiada por inteligência artificial.
- 2) *Modelos de grande porte alcançaram avanços significativos*: tais modelos são modelos os de *machine learning* (aprendizado de máquina), com parâmetros de escala extragrande e estruturas computacionais complexas. Na linha que mais chama a atenção no campo da IA atualmente, os modelos de grande porte estão se tornando

Tendências equivocadas, como o niilismo histórico e o populismo, podem se espalhar no ciberespaço por meio de tecnologias de recomendação algorítmica, impactando a segurança ideológica

cada vez mais uma infraestrutura estratégica. Potencializando o paradigma de desenvolvimento de modelos de grande porte pré-treinados com ajuste fino para tarefas subsequentes, os custos de treinamento e os ciclos de desenvolvimento podem ser reduzidos para permitir uma produção em larga escala. Com o lançamento do ChatGPT, no final de 2022, os modelos de grande porte entraram em um período de crescimento explosivo, adquirindo capacidade de compreensão multimodal e de geração de conteúdos diversos. No início de 2025, surgiu o DeepSeek-R1, oferecendo vantagens como código aberto e baixo custo computacional e tornando os modelos de grande porte acessíveis ao público em geral.

3) *A IA generativa está em ascensão*: trata-se de uma tecnologia capaz de gerar automaticamente textos, imagens, áudio, vídeos, códigos e outros conteúdos com base em algoritmos, modelos e regras. Comparada aos sistemas de IA tradicionais, que só podem processar mediante a inserção de dados, a IA generativa pode criar conteúdo novo por conta própria. Impulsionada por grandes modelos, ela se desenvolveu rapidamente e é amplamente aplicada em questionários profissionais, programação, criação de poesia, pintura e outras áreas, demonstrando uma poderosa capacidade generativa, que está revolucionando o setor. Por exemplo, a revista *Nature* considerou o ChatGPT um dos dez maiores avanços científicos de 2023 — a primeira vez que um programa de computador foi incluído na lista.

4) *A IA incorporada e os robôs humanoides estão chamando atenção*: a IA incorpora a busca desenvolvimento contínuo graças à interação constante da unidade física com o ambiente externo. Ela pode superar as limitações da IA tradicional, que se baseia na combinação entre poder computacional e dados, e dispensa a adaptabilidade aos diferentes cenários, proporcionando uma nova perspectiva para explorar os mecanismos de formação da inteligência humana e um poderoso motor para transformar conquistas científicas e tecnológicas em produtividade no mundo real. Os robôs humanoides, importantes vetores da IA incorporada, oferecem uma plataforma ideal para o estudo do surgimento da inteligência. A pesquisa em inteligência

artificial incorporada cresce em ritmo acelerado e impulsionará ainda mais rápido a mudança de paradigma da IA, da cognição virtual para a inteligência física, por meio da interação em circuito fechado de dados, algoritmos e vetores.

5) *A pesquisa científica impulsionada por IA se desenvolve rapidamente*: ao longo dos últimos séculos, a pesquisa científica passou por transformações, evoluindo da observação experimental, a análise teórica e a simulação computacional para o emprego de paradigmas de uso intensivo de dados. Esse ramo de pesquisas conduz uma nova onda de transformações que mudará fundamentalmente a forma como as descobertas científicas são feitas. Por exemplo, o AlphaFold2 e o AlphaFold3 podem prever com precisão, respectivamente, a estrutura 3D de proteínas e a estrutura e interações das moléculas vivas, acarretando mudanças revolucionárias na pesquisa biomédica. Além disso, a IA tem sido amplamente aplicada no desenvolvimento de vacinas, *design* de *chips*, previsão meteorológica e em muitos outros campos.

6) *A IA está aumentando sua capacitação em diversos setores*: como tecnologia estratégica que encabeça uma nova rodada da revolução tecnológica e da transformação industrial, a IA está mudando profundamente a produção e os estilos de vida humanos, acelerando sua profunda integração com a economia real, aumentando a produtividade social, promovendo a transformação e modernização das indústrias tradicionais, fomentando o desenvolvimento vigoroso de indústrias emergentes, elevando a novos patamares as indústrias do futuro e impulsionando o rápido crescimento da “economia não tripulada”. Ao mesmo tempo, está tendo um impacto significativo em domínios da vida social como transporte, moradia, educação, saúde e cuidados com idosos, fornecendo melhores soluções para promover a equidade social e aumentar o bem-estar e a qualidade de vida humana.

7) *O impacto social da IA é cada vez mais proeminente*: além do rápido desenvolvimento da IA, seu impacto social é abrangente. Por um lado, gera muitos efeitos positivos, como o impulsionamento da economia, provendo os meios de subsistência das pessoas e beneficiando as sociedades; por outro lado, também traz muitos problemas sociais, como riscos de segurança, incertezas legais, anomalias éticas e violação de privacidade. Por exemplo, o uso da tecnologia de IA generativa para criar informações falsas destinadas à difamação, extorsão e ataques maliciosos é um risco significativo; a capacidade de dispositivos e algoritmos inteligentes de explorar e analisar a privacidade das pessoas supera em muito a das tecnologias anteriores, tornando problemas de privacidade, incluindo os que dizem respeito a rastros pessoais, dados de consumo e características biométricas, cada vez mais complexos e difíceis de evitar.

8) *A IA geral continua sendo um objetivo de longo prazo*: a inteligência artificial geral visa desenvolver máquinas com inteligência igual ou até mesmo superior à dos humanos, permitindo inclusive que as máquinas tenham consciência autônoma. Nos últimos anos, à medida que a capacidade dos modelos de redes neurais melhora, o público cultiva expectativas excessivamente altas em relação à tecnologia, enquanto alguns especialistas e acadêmicos fazem previsões irreais sobre o desenvolvimento da IA geral. A inteligência artificial ainda apresenta uma distância significativa em

Foto: Xinhua/Shen Hong



Presidida por Xi Jinping, a 20ª Sessão de Estudos Coletivos do Birô Político do Comitê Central do PCCh discutiu medidas para regulamentar e desenvolver a IA. Pequim, 25 de abril de 2025

relação ao nível de inteligência humana, com deficiências óbvias em compreensão semântica, raciocínio, planejamento, generalização e aprendizado contínuo. Ela também carece de compreensão essencial das leis operacionais espaço-temporais do mundo físico objetivo, e falta-lhe empatia profunda. Ainda há um longo caminho a percorrer para alcançarmos uma verdadeira IA geral.

AVALIANDO CIENTIFICAMENTE OS RISCOS E DESAFIOS DO DESENVOLVIMENTO DA IA

Para Xi Jinping, “a IA gera oportunidades de desenvolvimento sem precedentes, mas também riscos e desafios sem precedentes”. Embora a IA aprofunde a integração entre os mundos digital e físico, em certa medida também torna tênues as fronteiras entre o virtual e o real. As incertezas que desperta, sua opacidade e falta de explicabilidade, características que lhe são inerentes, estão se tornando cada vez mais claras, representando inúmeros riscos e desafios para a sociedade humana.

1) *Riscos e desafios à segurança nacional*: o primeiro risco diz respeito à *segurança ideológica*. Tendências equivocadas, como o niilismo histórico e o populismo, podem se espalhar no ciberespaço por meio de tecnologias de recomendação algorítmica, impactando a segurança ideológica. O segundo risco concerne aos ambientes de *governança nacional*. A IA pode criar informações falsas e confundir a percepção pública por meio de tecnologias *deepfake*, além de usar algoritmos precisos para exacerbar os efeitos das “bolhas de informação” e permitir a manipulação da opinião pública, levando a um enfraquecimento da capacidade explicativa interna. O terceiro risco se relaciona à *segurança militar*. Uma vez formada uma “lacuna ge-

racional” tecnológica na inteligência militar, as potências tecnológicas têm maior probabilidade de se envolver em chantagem política e ataques-surpresa tecnológicos contra países em desenvolvimento. As aplicações militares transversais da IA nos domínios nuclear e espacial também representarão riscos significativos para a segurança militar.

2) *Riscos e desafios para o desenvolvimento industrial*: a maturação da capacidade técnica e das formas de produto da IA, a determinação dos principais cenários de aplicação e o grau de integração industrial estão todos intimamente relacionados ao desenvolvimento das indústrias. Práticas comerciais maduras não podem ser feitas por meio de investimentos maciços de curto prazo em fundos e recursos; o caminho ainda é longo. Tomando o ChatGPT como exemplo, o investimento inicial foi de cerca de US\$ 800 milhões, sendo as despesas diárias com eletricidade de aproximadamente US\$ 50 mil e o custo de resposta de alguns centavos. Diante de custos tão elevados, por um lado, a discussão sobre a monopolização de recursos é cada vez mais importante, visto que gigantes da tecnologia formam uma “hegemonia digital” que continuamente reduz o espaço de sobrevivência das pequenas e médias empresas. Por outro lado, muitas equipes de empresas tradicionais e *startups* estão entrando no mercado às pressas, e alguns governos também anseiam seguir o exemplo. Essa entrada desenfreada torna o desenvolvimento da IA desordenado e o conduz a um grande desperdício de recursos. Além disso, algumas empresas exploram a tendência da IA para as negociações especulativas, aumentando o risco de formação de bolhas de mercado.

3) *Riscos e desafios à governança social*: o primeiro risco é o de *vazamento de dados pessoais*. Por suas características técnicas e cenários de aplicação, a IA pode acessar facilmente grandes quantidades de informações privadas enquanto é utilizada, coletando rastros pessoais, registros de navegação, dados de consumo, características biométricas etc. sem o conhecimento dos usuários, o que representa sérios desafios à proteção de informações pessoais e privacidade. O segundo risco é o de *ataques maliciosos*. Criminosos utilizam tecnologias *deepfake* para criar vídeos e imagens falsos a fim de promover extorsão, ataques maliciosos e outras atividades, ameaçando a segurança e a estabilidade sociais, reputações pessoais e a propriedade. O terceiro risco é o de *“poluição” da informação*. O uso indevido de tecnologias de IA levou à proliferação de conteúdos falsos e de baixa qualidade na internet, o que distorce a percepção de realidade dos modelos, incutindo-lhes vícios irreversíveis.

4) *Riscos e desafios à ética e ao direito*: em primeiro lugar, o desenvolvimento de tecnologias de IA pode desencadear mudanças significativas no *mercado de trabalho*. Tecnologias de inteligência artificial realizam tarefas altamente complexas de forma autônoma, reduzindo consideravelmente as barreiras e os custos de produção e interação de conteúdo. Todos os setores enfrentam o risco de substituição da mão de obra pela IA, o que pode causar pânico pela ameaça de desemprego e desordem social. As fronteiras entre os setores se tornarão ainda mais tênues, rompendo com a atual divisão social do trabalho. Em segundo lugar, *o viés algorítmico e a discriminação* feita por máquinas podem aumentar. Os algoritmos de IA têm limitações

A governança internacional da IA é uma questão crucial do nosso tempo, e diz respeito à construção de um futuro compartilhado para a humanidade. Diante das oportunidades e dos desafios suscitados pela IA, precisamos tanto do reconhecimento de padrões e da prevenção conjunta de riscos por parte dos atores quanto de orientação ética e de salvaguardas legais

inerentes em termos de transparência e robustez, o que exacerbará os vieses e a discriminação no conteúdo gerado, ampliando a injustiça social. Além disso, podem surgir problemas relativos à *ética tecnológica*, tais como gêmeos digitais, interfaces cérebro-computador e simbiose humano-máquina, envolvendo riscos à vida, privacidade, responsabilidade, equidade social e justiça. A aplicação generalizada de robôs humanoides também levantará muitas novas questões no âmbito de um sistema social em que os humanos são os únicos sujeitos, tais como o *status* legal dos “acompanhantes artificiais”. Por último, o sistema jurídico corre o risco de falhar. O estatuto jurídico dos sistemas de IA é questionável e a atribuição de direitos e responsabilidades a eles está indefinida, causando problemas que as leis tradicionais dificilmente conseguem regular. Por exemplo, os direitos de propriedade intelectual do conteúdo gerado por IA são de difícil definição, as leis de trânsito para sistemas de condução autônoma ainda precisam ser aprimoradas e a atribuição de responsabilidade por perdas causadas por IA é complexa.

ACELERANDO A PROMOÇÃO DA GOVERNANÇA E COOPERAÇÃO INTERNACIONAIS EM IA

Nos últimos anos, os temas e as estruturas da governança global da IA vêm se tornando cada vez mais diversificados, revelando uma tendência geral de domínios abrangentes, em múltiplos níveis, e ampla cooperação. Diversas iniciativas, legislações e normas surgiram, abarcando políticas públicas, economia, direito, saúde, vida social e outras áreas. No âmbito das Nações Unidas, países e diversas organizações internacionais têm implementado ativamente a governança e a regulamentação da IA, enquanto entidades e empresas do setor industrial também vêm promovendo a autorregulamentação. A cooperação regional, multilateral e bilateral continuou a se fortalecer. Em 2023, a China lançou a Iniciativa de Governança Global da IA, que propõe soluções construtivas para questões amplamente relevantes

sobre o desenvolvimento e a governança da IA, fornecendo um roteiro para discussões e regulamentações internacionais pertinentes e demonstrando o senso de responsabilidade da China.

No entanto, a governança internacional da IA ainda enfrenta múltiplos desafios. O primeiro é a falta de uma estrutura de governança global unificada. Ainda não há consenso entre os países sobre questões importantes, e muitas diferenças regulatórias são difíceis de coordenar. Iniciativas internacionais não vinculativas têm eficácia limitada, e a fixação de regras de governança com poder vinculante substantivo ainda leva tempo.

O segundo desafio são os interesses divergentes entre os países. As economias desenvolvidas estão mais preocupadas com a segurança tecnológica e as questões éticas, enquanto as economias emergentes e outras enfatizam as oportunidades de desenvolvimento econômico e social. O desequilíbrio no poder de governança exacerba essas divergências.

O terceiro desafio é o da fragmentação da governança. A relação entre o rápido desenvolvimento da tecnologia de IA e a defasada construção de um sistema de governança cria uma contradição. Regulamentações estáticas dificilmente conseguem lidar com riscos dinâmicos, o que significa uma estrutura de governança fraturada e pouco eficiente.

Xi Jinping enfatizou a necessidade de priorizar os seres humanos e garantir que a IA sirva a boas causas, fortalecendo a governança de IA no âmbito das Nações Unidas. Especificamente, o avanço da governança e da cooperação internacional em IA requer progressos nas seguintes áreas:

1) *Defesa do princípio de orientação em prol das pessoas e da tomada de uma direção correta para o desenvolvimento da IA:* independentemente de como a tecnologia evolua, devemos permanecer fiéis às nossas aspirações originais e buscar um desenvolvimento benigno na era da inteligência artificial. Por um lado, as entidades de governança multilateral devem chegar a valores consensuais sobre o desenvolvimento da IA, estabelecer princípios fundamentais de governança, visar ao aprimoramento do bem-estar comum da humanidade e promover o desenvolvimento integral das pessoas, com a premissa de garantir a segurança social e respeitar os direitos e interesses humanos, assegurando que a IA se desenvolva sempre em benefício da civilização. Por outro lado, a comunidade internacional deve formar um consenso sobre os riscos do desenvolvimento da IA e estipular o mais rápido possível orientações jurídicas e éticas, estabelecendo um sistema de avaliação de riscos e de comunicação imediata de alertas sobre a IA para garantir que a pesquisa e o desenvolvimento dessa tecnologia e suas aplicações sejam seguros, controláveis e alinhados com os valores humanos compartilhados, prevenindo danos à segurança nacional, ao interesse público e aos direitos e interesses legítimos de organizações e indivíduos.

2) *Adesão ao princípio de abertura e compartilhamento e promoção de uma IA inclusiva e acessível:* devido às diferenças de capacidade científica e educacional, infraestrutura, estrutura industrial e fundamentos socioeconômicos entre países e regiões, a questão do desenvolvimento desigual da IA tornou-se evidente. Por um lado, devemos aprofundar as iniciativas de código aberto, promover o compartilhamento de tecnologias, talentos e infraestrutura, ajudar os países do Sul Global a fortalecer sua capacidade tecnológica, estabelecer uma plataforma internacional de compartilhamento de tecnologia, romper monopólios tecnológicos e barreiras de

acesso a dados e permitir o justo compartilhamento de recursos de dados e demais benefícios da tecnologia, a fim de reduzir as disparidades no setor. Por outro lado, devemos expandir continuamente o intercâmbio e a cooperação entre organizações acadêmicas, empresas e associações industriais relevantes, construindo um ecossistema de cooperação científica e tecnológica internacional aberto e vantajoso para todos, a fim de eliminar as restrições impostas por certos países à participação de pesquisadores e acadêmicos estrangeiros em eventos de intercâmbio e cooperação internacional e evitar o bloqueio tecnológico e o protecionismo de talentos, que dificultam o intercâmbio e o progresso tecnológico.

3) *Adesão ao princípio da governança por meio da tecnologia e aos métodos de inovação da governança de IA*: os processos de tomada de decisão e a lógica da inteligência artificial são opacos, com características de “caixa preta”. As tecnologias atuais não estão submetidas a mecanismos de autorregulação e controle, criando-se um claro vácuo de governança técnica. Entidades de pesquisa e desenvolvimento em diversos países e regiões devem aprimorar continuamente a explicabilidade e a previsibilidade da IA, criando tecnologias de inteligência artificial auditáveis, supervisionáveis, rastreáveis e confiáveis. Tais entidades devem tratar o poder computacional, a métrica mais intuitiva e quantificável, como objeto de governança, implementando indiretamente a governança da IA. Ao mesmo tempo, deve-se usar a inteligência artificial para eliminar os riscos que ela própria cria, projetando e desenvolvendo ativamente tecnologias e aplicações que possam geri-la de forma eficaz, monitorando e avaliando a implementação de estruturas de governança de IA e aprimorando a capacidade técnica da governança do setor.

4) *Adesão ao princípio da integração de sistemas e à proposta de construção de uma estrutura multilateral de governança de IA*: diante dos diversos desafios à paz e ao desenvolvimento mundial, a comunidade internacional deve construir um amplo consenso global, desenvolver ativamente um mecanismo de governança multilateral aberto, justo e eficaz e promover conjuntamente as tecnologias de IA para que realmente beneficiem a humanidade. Por um lado, países e regiões, bem como organizações internacionais relevantes, devem fortalecer o alinhamento e a coordenação de regras de governança e padrões técnicos, chegar a um consenso sobre questões-chave como segurança da IA, ética, privacidade de dados, regulamentação transfronteiriça e aplicações militares, estabelecer e aprimorar leis e regulamentos e aprimorar as diretrizes éticas e os mecanismos de responsabilização da IA. Devem-se respeitar plenamente as diferenças nacionais, com o aumento da participação dos países em desenvolvimento e a garantia de direitos e oportunidades iguais para que todos os países desenvolvam e utilizem tecnologias de IA. Por outro lado, entidades de governança diversas, como empresas, instituições de pesquisa, organizações sociais e mesmo indivíduos, devem exercer papéis que correspondam à sua própria identidade, participando conjuntamente da construção e do aprimoramento do sistema de governança de IA e viabilizando o desenvolvimento da IA em todo o mundo com ampla consulta, contribuições conjuntas e benefícios compartilhados.

A governança internacional da IA é uma questão crucial do nosso tempo, e diz respeito à construção de um futuro compartilhado para a humanidade. Diante das oportunidades e dos desafios suscitados pela IA, precisamos tanto do reconhecimento de padrões e da prevenção conjunta de riscos por parte dos atores quanto de orientação ética e de salvaguardas legais. Devemos consolidar os consensos e aprofundar a colaboração, sempre pautados por uma orientação focada nas pessoas, garantindo que o progresso tecnológico acompanhe o desenvolvimento da civilização e fazendo da IA uma força positiva para o bem-estar humano.

* Membro da Academia Chinesa de Ciências e secretário do Comitê do Partido Comunista Chinês da Universidade de Nanjing.

Texto cedido pela revista chinesa *Mundo Contemporâneo*. Tradução gentilmente realizada pelo prof. Cristhiano Duarte (UFBA/Brasil e Chapman University/EUA), membro do Conselho Editorial de *Princípios*. Revisão de Ramiro Torres.

► Texto recebido em 21 de agosto de 2025; aprovado em 23 de agosto de 2025.