

# Inteligência Artificial: riscos e oportunidades

---

Kissinger, Henry, Eric Schmidt & Daniel Huttenlocher. *A era da Inteligência Artificial e o nosso futuro humano*. Lisboa: Dom Quixote, 2021.

---

**Antonio Carlos Barbosa de Oliveira**

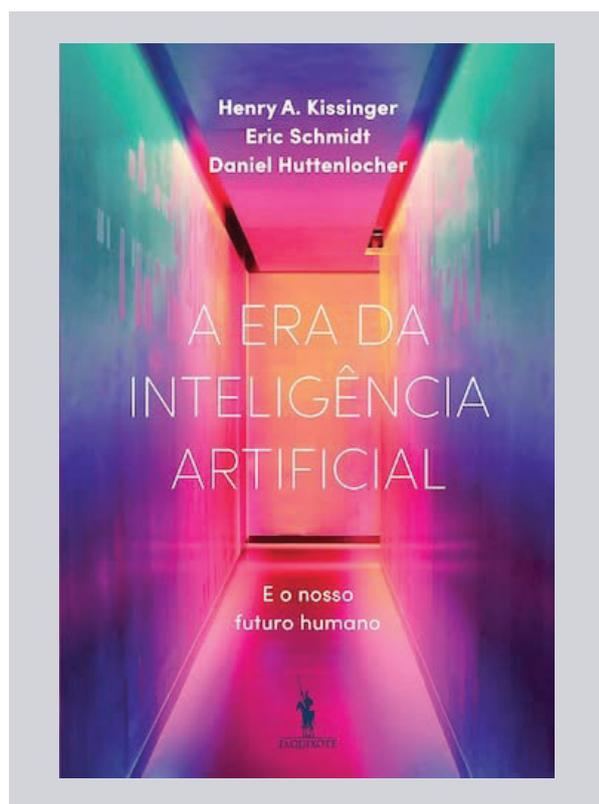
**E**m janeiro de 1956, o futuro Prêmio Nobel de Economia Herbert Simon anunciou, no início de uma aula, que ele e seu colega Allen Newell haviam inventado, nos feriados de Natal, uma “máquina de pensar” (McCorduck 1979). Alguns meses depois, durante o verão, John McCarthy – um jovem matemático – organizou um projeto de pesquisa que reuniu durante dois meses no Dartmouth College os pequenos grupos que estavam se dedicando a essa área. Precisando de um nome para atrair pessoas e ideias, McCarthy usa pela primeira vez a denominação *Artificial Intelligence*. O único

software pronto e funcionando apresentado nessa conferência foi o programa de Simon e Newell, que demonstrava teoremas de lógica formal.

A partir deste início singelo, a Inteligência Artificial (IA) cresceu e se transformou. Está hoje integrada, muitas vezes sem nos darmos conta, a inúmeras atividades humanas. Os três autores deste livro trabalharam durante quatro anos motivados pela visão de que a IA terá enormes impactos no futuro da humanidade. Eles são expoentes em suas respectivas áreas de atuação. Henry Kissinger é ex-secretário de Estado dos EUA e um dos mais influen-

---

**Antonio Carlos Barbosa de Oliveira** é engenheiro pela *Escola Politécnica da Universidade de São Paulo*, mestre em *Ciência pelo Instituto de Tecnologia de Massachusetts (MIT)* e mestre em *Astronomia pela James Cook University*.



tes intelectuais na área de segurança e diplomacia. Eric Schmidt é empresário e ex-CEO da Google, responsável por liderar a empresa no período 2001-2011. Daniel Huttenlocher é diretor do MIT College of Computing, criado em 2018 com investimento de US\$1 bilhão para coordenar as atividades do Instituto na área de computação. O livro não tem a pretensão de esgotar o assunto, mas se propõe a fornecer ao leitor não especializado os instrumentos e a base intelectual para participar da construção de nosso futuro populado por máquinas inteligentes.

Por meio de três exemplos (jogo de xadrez, descoberta de antibióticos e entendimento de linguagem), os autores introduzem o leitor ao mundo da IA. A programação de computado-

res para jogar foi sempre um foco das pesquisas em IA. Em 2017 o programa AlphaZero passou a liderar o *ranking* de programas de xadrez, após vencer outros programas em torneios. Até então, todos os programas escritos para jogar xadrez utilizavam técnicas que os humanos haviam desenvolvido e, através de sua grande capacidade computacional, eram capazes de explorar a enorme árvore de possíveis jogadas com maior eficiência. AlphaZero foi criado apenas com as regras do jogo, sem nenhuma codificação de estratégias ou jogadas humanas. O conhecimento que permitiu ao AlphaZero ser o melhor foi obtido através de aprendizado jogando contra ele mesmo. Em apenas quatro horas e somente avaliando seus resultados jogando contra uma cópia do programa, AlphaZero desenvolveu estratégias incríveis nunca consideradas por jogadores humanos.

Halicina é um novo antibiótico descoberto no MIT em 2020 mediante uso de IA. O processo de descoberta de novas drogas envolve anos de experimentos em laboratórios em que milhares de moléculas são analisadas. A abordagem por meio de IA começou com a elaboração de um banco de dados contendo duas mil moléculas conhecidas. Para cada uma delas foram registrados dados químicos e efeitos biológicos. Com a técnica de *deep learning*, o software de IA analisou essas moléculas e aprendeu como detectar moléculas efetivas em matar bactérias e que não

sejam tóxicas. Esse programa, treinado pelos exemplos codificados em sua base de dados, foi então alimentado com 61 mil novas moléculas e identificou uma delas com as propriedades desejadas.

GPT-3 é um programa capaz de produzir frases e miniensaaios sobre qualquer assunto ao ser alimentado com uma pergunta ou algumas frases. O texto produzido pelo programa a partir de curtas questões filosóficas colocadas pelos autores do livro é incrível. Essa capacidade de produzir textos coerentes e assustadoramente similares a frases humanas foi adquirida sem nenhuma codificação de regras, mas apenas por meio da análise da enorme quantidade de textos disponíveis na internet.

*Em 1990 uma revolução transformou a metodologia da IA. Uma nova abordagem baseada em redes neurais permitiu a construção de programas capazes de aprender.*

Os três exemplos escolhidos pelos autores são representativos da moderna abordagem da IA e muito diferentes dos programas desenvolvidos após a conferência em Dartmouth. Estes eram baseados em algoritmos detalhados e cuidadosamente codificados após análise de como a mente humana funciona.

Esta abordagem, conhecida como a fase simbólica da IA, chegou aos anos 1980 com alguns resultados importantes, mas acabou estagnada. Durante cerca de dez anos, conhecidos como o inverno da IA, as verbas de pesquisa diminuíram e os resultados se tornaram cada vez mais escassos.

Em 1990 uma revolução transformou a metodologia da IA. Uma nova abordagem baseada em redes neurais permitiu a construção de programas capazes de aprender. As redes neurais, inventadas na década de 1950 e inspiradas no funcionamento dos neurônios biológicos, tiveram um renascimento após terem sido esquecidas no período da IA simbólica. Curiosamente foram os *chips* desenvolvidos para processamento de videogames que permitiram a construção de redes neurais com milhares de componentes interligados. Novos algoritmos matemáticos utilizando essas unidades de processamento gráfico consolidaram a abordagem que passou a ser conhecida como *machine learning* – aprendizado de máquina.

No aprendizado de máquina, a rede neural é treinada utilizando uma base de dados. Por exemplo, um programa para identificar gatos em fotos pode ser alimentado com milhares de fotos rotuladas se têm ou não um gato. A rede neural, contendo milhares de componentes, será configurada automaticamente sem nenhuma necessidade de intervenção do programador humano. Na etapa seguinte, o programa será

capaz de fazer inferências: dada uma nova foto, detectar ou não a presença de um gato. A disponibilidade na internet de uma enorme quantidade de imagens e textos facilita o processo da aprendizagem de máquina. Essa metodologia é conhecida como aprendizado supervisionado. Outra técnica importante é o aprendizado não supervisionado, em que os dados brutos sem serem rotulados previamente são analisados, e a rede neural se configura para agrupar casos similares e identificar anomalias.

Uma outra técnica, conhecida como aprendizagem reforçada (*reinforcement learning*), não se limita a analisar os dados disponíveis. O método é baseado no conceito de um agente que observa seu entorno, executa ações e recebe recompensas em função da qualidade de suas decisões. À medida que este agente toma decisões e recebe recompensas, vai aprendendo, de maneira que, no final do processo, possui um conjunto de políticas que permitem executar decisões com alta qualidade – como no caso do Alpha-Zero, que utiliza essa metodologia.

A origem e evolução da IA ocupa os autores nos primeiros capítulos e, em seguida, dá lugar a uma análise profunda e detalhada do impacto da IA nas plataformas de rede e na segurança e ordem mundial. As plataformas de rede são os aplicativos que usamos diariamente: redes sociais, buscas na internet, *streaming* de vídeos, navegação e transporte urbano. A principal característica desses sistemas é que sua utilidade aumenta

*A IA tem sido incorporada às plataformas de rede de maneira quase imperceptível, mas muito intensa. Nossa vida diária já está marcada pela interação com uma forma de inteligência não humana.*

exponencialmente à medida que novos usuários são incorporados. Este efeito positivo favorece uma configuração com algumas poucas empresas, cada uma com um grande número de usuários. A IA tem sido incorporada às plataformas de rede de maneira quase imperceptível, mas muito intensa. Nossa vida diária já está marcada pela interação com uma forma de inteligência não humana.

Os algoritmos de busca da Google foram inicialmente codificados por programadores que sabiam exatamente quais resultados seriam obtidos. Em 2015 a Google passou a utilizar IA com uma significativa melhora na qualidade das respostas obtidas. Entretanto, os programadores perderam a capacidade de entender como as respostas eram produzidas. Eles sabem que a IA melhorou o resultado de uma busca, mas não é mais possível explicar como o sistema chegou a esta resposta.

No caso do Facebook, a atividade de remoção de conteúdo impróprio, que chega a um bilhão de postagens por trimestre, só é possível com a utili-

zação de IA em conjunto com milhares de pessoas dedicadas a essa tarefa. IA também é usada pela Amazon, Netflix e outras plataformas, em que a história passada de consumo é analisada e comparada com outros usuários para sugerir novas recomendações.

A crescente utilização de IA nas plataformas de rede está criando um novo tipo de relacionamento entre usuários e essa tecnologia, que opera com uma lógica não humana. Segundo os autores, isto é algo absolutamente novo e que nunca existiu na história. As plataformas de rede foram concebidas e criadas para atender a necessidades imediatas de seus usuários, mas acabaram adquirindo uma dimensão social e política completamente inesperada. O exemplo mais claro é o impacto das redes sociais no processo de informação (e desinformação) política, que podem impactar a sociedade e as estruturas de governo.

A interação de plataformas de rede baseadas em IA com governos nacionais tende a ser extremamente complexa, na medida em que tentativas de regulamentação podem gerar conflitos com as empresas e com os usuários, produzindo crises imprevisíveis. A IA vem sendo utilizada de forma crescente, tanto na elaboração como na neutralização de desinformação. A tecnologia atual chamada *deep fake* permite a criação de textos e imagens com conteúdo totalmente falso, mas com uma precisão e qualidade

que nos enganam. A supressão desse tipo de desinformação em larga escala só será possível por meio da IA.

As plataformas de rede estão tendo um impacto geopolítico, na medida em que as empresas, apesar de sediadas em um país, atendem consumidores globais. Os autores mostram EUA e China como líderes no desenvolvimento de plataformas e constataam que a Europa, apesar de sua tradição acadêmica e científica, não mostrou interesse em desenvolver plataformas de rede próprias, procurando liderar na regulamentação das plataformas existentes. A Rússia também ficou fora da competição internacional desenvolvendo algumas plataformas para uso interno. Em um *ranking*<sup>1</sup> de países segundo sua capacitação em IA, EUA e China lideram, enquanto o Brasil ocupa o trigésimo nono lugar.

Os países que não produzem plataformas de rede incorporando IA e que têm grande parte de sua atividade econômica dependente dessas plataformas estão vulneráveis às restrições criadas por governos hostis. Podem administrar esse risco incentivando a operação simultânea de várias plataformas, ou mesmo a criação de plataformas nacionais.

A análise feita pelos autores do impacto da IA na segurança e na ordem mundial parte do estudo dos mecanismos que funcionaram na Guerra Fria evitando o confronto nuclear. Enquanto as armas atômicas são entes físicos que

1. *The Global AI Index* (Tortoise s.d.): <https://www.tortoisemedia.com/intelligence/global-ai/>.

podem ser contados e avaliados objetivamente, as tecnologias de ciberataques (*cyber*), incorporadas recentemente pelas grandes potências ao seu arsenal, têm uma natureza completamente diferente. É muito difícil avaliar a capacidade *cyber* dos adversários, com a tecnologia evoluindo rapidamente e os mecanismos de ataque mantidos em segredo.

A estratégia *cyber* consiste em detectar falhas ou vulnerabilidades em sistemas computacionais e, através de programas enviados remotamente, desabilitar esses sistemas. A IA pode ser usada para incrementar os mecanismos de defesa, detectando e corrigindo falhas antes que os inimigos possam lançar um ataque. Na área ofensiva, os vírus utilizando IA podem se transformar e mutar, evitando a detecção e atacando com maior velocidade e persistência. A efetividade de uma arma *cyber* muitas vezes depende de mantê-la em segredo, o que complica muito qualquer negociação visando à não proliferação.

Outro aspecto analisado pelos autores é o impacto do uso da IA, com sua lógica não humana, nos sistemas militares. A força aérea americana já testou aviões autônomos que selecionam alvos e decidem ataques sem intervenção humana. A lógica do conflito militar sempre envolveu a avaliação das estratégias do adversário. Quando ambos os lados estão usando IA, entender a lógica das decisões do inimigo fica muito mais difícil, senão impossível.

Os autores propõem que cada país líder no uso de IA crie um organismo

nacional para discutir os aspectos de defesa e segurança da IA, com o objetivo de garantir a competitividade e limitar a escalada sem controle em uma crise. Em um artigo publicado depois do livro, um dos autores, E. Schmidt (2022a), coloca EUA e China como os dois únicos países capazes de liderar a IA e defende vigorosamente a criação de um programa nacional para manter a liderança americana.

*Para os autores, ao longo da história, a humanidade entendeu o mundo através da fé e da razão. A IA seria uma terceira maneira. Estamos acostumados ao monopólio da inteligência, e a IA desafia essa visão e transforma a experiência humana.*

Ao tentar antever a direção da pesquisa em IA, os autores analisam a Inteligência Geral Artificial (IGA). Até agora, todos os programas utilizando IA têm objetivos claros definidos pelos seus programadores para executar uma tarefa específica, como jogar xadrez ou achar novos antibióticos. Os sistemas com IGA serão capazes de definir seus próprios objetivos e criar os algoritmos e programas para lograr esses objetivos. Os especialistas divergem sobre a viabilidade da IGA. Alguns consideram impos-

sível, enquanto outros acham que em 20 anos teremos IGA (Schmidt 2022b).

IGA será capaz de obter novos conhecimentos científicos que poderão ter enormes impactos econômicos. Como provavelmente essa tecnologia necessitará enorme poder computacional, só disponível em grandes organizações, segundo os autores será necessário que governos, universidades e o setor privado definam limites para que IGA possa ser usada de maneira justa e democrática.

Para os autores, ao longo da história, a humanidade entendeu o mundo através da fé e da razão. A IA seria uma terceira maneira. Estamos acostumados ao monopólio da inteligência, e a IA desafia essa visão e transforma a experiência humana. O livro coloca mais questões do que fornece respostas, mas deixa clara a necessidade de pensar seriamente como controlar a explosiva evolução da IA e garantir que seu uso seja para nosso benefício. ■

---

## Referências Bibliográficas

---

Kissinger, Henry, Eric Schmidt & Daniel Huttenlocher. 2021. *The Age of AI and Our Human Future*. New York: Little, Brown and Company.

Kissinger, Henry, Eric Schmidt & Daniel Huttenlocher. 2021. *A Era da Inteligência Artificial e o nosso futuro humano*. Lisboa: Dom Quixote.

McCorduck, Pamela. 1979. *Machines Who Think: A Personal Inquiry Into the History and Prospects of Artificial Intelligence*. San Francisco: WH Freeman.

Schmidt, Eric. 2022a. "AI, Great Power Competition & National Security". *Daedalus* 151 (2):288-298. [https://doi.org/10.1162/daed\\_a\\_01916](https://doi.org/10.1162/daed_a_01916).

Schmidt, Eric. 2022b. "#280 - The future of Artificial Intelligence. A Conversation with Eric Schmidt". Produced by Sam Harris. *Making Sense*, April 22, 2022. Podcast, Online Streaming, 38:03. <https://>

[www.samharris.org/podcasts/making-sense-episodes/280-the-future-of-artificial-intelligence](http://www.samharris.org/podcasts/making-sense-episodes/280-the-future-of-artificial-intelligence).

Tortoise. s.d. "The Global AI Index". Acessado em 6 de agosto de 2022. <https://www.tortoisemedia.com/intelligence/global-ai/>.

**Como citar:** Oliveira, Antonio Carlos Barbosa de. "Inteligência Artificial: riscos e oportunidades". Resenha de *A Era da Inteligência Artificial e o nosso futuro humano* (Lisboa: Dom Quixote, 2021), de Kissinger et al. *CEBRI-Revista* Ano 1, Número 3 (Jul-Set): 175-181.

**To cite this work:** Oliveira, Antonio Carlos Barbosa de. 2022. "Artificial Intelligence: Risks and Opportunities." Book Review of *The Age of AI and Our Human Future* (New York: Little, Brown and Company, 2021), by Kissinger et al. *CEBRI-Journal* Year 1, No. 3 (Jul-Sep): 175-181.

Recebido: 11 de maio de 2022

Aceito para publicação: 5 de agosto de 2022

---

Copyright © 2022 CEBRI-Revista. Este é um artigo em acesso aberto distribuído nos termos da Licença de Atribuição Creative Commons, que permite o uso irrestrito, a distribuição e reprodução em qualquer meio, desde que o artigo original seja devidamente citado.