

DOMINANDO A

IA

Dominando a IA

Copyright © 2025 STARLIN ALTA EDITORA E CONSULTORIA LTDA.

Copyright ©2024 Jeremy Kahn.

ISBN: 978-85-508-2573-1

Alta Books é uma Editora do Grupo Editorial Alta Books.

Translated from original Mastering AI © 2024 by Jeremy Kahn. ISBN 978-1-6680-5332-4. This translation is published and sold by Simon & Schuster, the owner of all rights to publish and sell the same. PORTUGUESE language edition published by Starlin Alta Editora e Consultoria Eireli, Copyright © 2025 by STARLIN ALTA EDITORA E CONSULTORIA LTDA.

Impresso no Brasil – 1ª Edição, 2025 – Edição revisada conforme o Acordo Ortográfico da Língua Portuguesa de 2009.

Dados Internacionais de Catalogação na Publicação (CIP)
(BENITEZ Catalogação Ass. Editorial, MS, Brasil)

K12d
1.ed. Kahn, Jeremy
Dominando a IA : guia de sobrevivência para o futuro / Jeremy Kahn ; tradução Susane Schmieg. - 1.ed. - Rio de Janeiro : Alta Books, 2025.
336 p; 15,7 x 23 cm.
Título original: Mastering AI.
ISBN 978-85-508-2573-1
1. Cultura digital. 2. Futuro (Perspectivas).
3. Inteligência artificial - Inovações tecnológicas - Aspectos sociais. 4. Tecnologia.
I. Título.
04-2025/218 CDD 006.1

Índice para catálogo sistemático:

1. Cultura digital : Inteligência artificial :
Tecnologia 006.1

Aline Grazielle Benitez - Bibliotecária - CRB-1/3129

Todos os direitos estão reservados e protegidos por Lei. Nenhuma parte deste livro, sem autorização prévia por escrito da editora, poderá ser reproduzida ou transmitida. A violação dos Direitos Autorais é crime estabelecido na Lei nº 9.610/98 e com punição de acordo com o artigo 184 do Código Penal.

O conteúdo desta obra fora formulado exclusivamente pelo(s) autor(es).

Marcas Registradas: Todos os termos mencionados e reconhecidos como Marca Registrada e/ou Comercial são de responsabilidade de seus proprietários. A editora informa não estar associada a nenhum produto e/ou fornecedor apresentado no livro.

Material de apoio e erratas: Se parte integrante da obra e/ou por real necessidade, no site da editora o leitor encontrará os materiais de apoio (download), errata e/ou quaisquer outros conteúdos aplicáveis à obra. Acesse o site www.altabooks.com.br e procure pelo título do livro desejado para ter acesso ao conteúdo.

Suporte Técnico: A obra é comercializada na forma em que está, sem direito a suporte técnico ou orientação pessoal/exclusiva ao leitor.

A editora não se responsabiliza pela manutenção, atualização e idioma dos sites, programas, materiais complementares ou similares referidos pelos autores nesta obra.

Grupo Editorial Alta Books

Produção Editorial: Grupo Editorial Alta Books

Diretor Editorial: Anderson Vieira

Editor da Obra: Gorki Starlin

Vendas Governamentais: Cristiane Mutús

Gerência Comercial: Claudio Lima

Produtor Editorial: Thiê Alves

Tradução: Susane Schmieg

Copidesque: Bernardo Kallina

Revisão: Denise Himpel

Diagramação: Fernanda Buccelli



Rua Viúva Cláudio, 291 – Bairro Industrial do Jacaré
CEP: 20.970-031 – Rio de Janeiro (RJ)
Tels: (21) 3278-8069 / 3278-8419
www.altabooks.com.br – altabooks@altabooks.com.br
Ouvidoria: ouvidoria@altabooks.com.br



**GUIA DE SOBREVIVÊNCIA
PARA O FUTURO**

DOMINANDO A

IA

JEREMY KAHN



ALTA BOOKS

GRUPO EDITORIAL

Rio de Janeiro, 2025

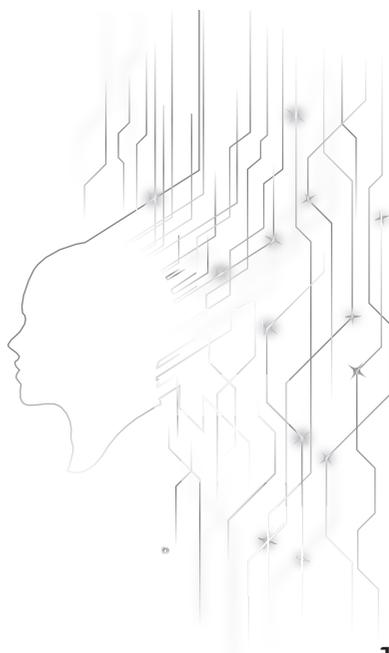
Amostra

Amostra

Para Victoria,
Cordelia e Gabriel

Amostra

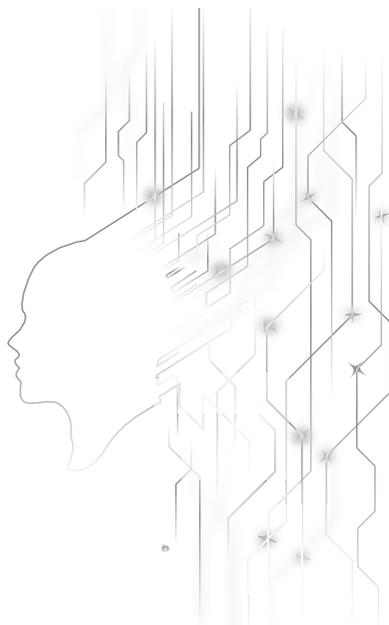
Sumário



INTRODUÇÃO: O EURECA DA IA	1
CAPÍTULO 1 OS ILUSIONISTAS	11
CAPÍTULO 2 A VOZ DENTRO DA SUA CABEÇA	41
CAPÍTULO 3 FALE COMIGO	59
CAPÍTULO 4 TODO MUNDO NO PILOTO AUTOMÁTICO	73
CAPÍTULO 5 OS PILARES DA INDÚSTRIA	95
CAPÍTULO 6 RICO, APENAS PARA TE DESGRAÇARES?	113

CAPÍTULO 7 ARISTÓTELES DE BOLSO	133
CAPÍTULO 8 ARTE E ARTIFÍCIO	151
CAPÍTULO 9 UM MICROSCÓPIO PARA DADOS	175
CAPÍTULO 10 MAIS CALOR DO QUE LUZ	195
CAPÍTULO 11 A CRISE DA CONFIANÇA	207
CAPÍTULO 12 GUERRA NA VELOCIDADE DA MÁQUINA	229
CAPÍTULO 13 O FIM DA LINHA PARA TODOS NÓS	247
CONCLUSÃO: RUMO AO NOSSO FUTURO SUPERPODEROSO	265
AGRADECIMENTOS	269
NOTAS	273
ÍNDICE	321
SOBRE O AUTOR	327

INTRODUÇÃO: O EURECA DA IA



Estes são chamados de *momentos Eureka*: o instante em que uma ideia se materializa nos recônditos do inconsciente e assume uma forma tangível, quase física. As pessoas vivem momentos Eureka, assim como sociedades inteiras passam por momentos de súbita compreensão coletiva que marcam o início de novas eras. Esses momentos ocorrem quando uma tecnologia é combinada a um dispositivo ou software acessível ao público, oferecendo a todos instantaneamente uma nova visão do futuro.

A própria lâmpada, o caricato símbolo do Eureka, foi um desses momentos. Em 31 de dezembro de 1879, Thomas Edison acionou um interruptor em seu laboratório, e a lâmpada incandescente iluminou um admirável mundo novo: alimentado por eletricidade, não por combustível ou vapor. Inventores e cientistas de ambos os lados do Atlântico já competiam há quase uma década para comercializar energia elétrica. Mas muitas de suas invenções eram industriais, técnicas, complexas. A lâmpada, por outro lado, era um dispositivo que qualquer um podia segurar, observar — e entender.

Nas últimas três décadas, passamos por uma série de momentos Eureka no campo da tecnologia digital: a internet existia antes dos navegadores, mas foi só com o lançamento do Netscape Navigator em 1994 que a era da internet realmente começou. Havia

tocadores de MP3 antes do iPod estrear em 2001, mas nenhum desencadeou a revolução da música digital. Havia smartphones antes da Apple lançar o iPhone em 2007, mas não havia aplicativos para ele. Em 30 de novembro de 2022, a inteligência artificial teve seu momento Eureka. Foi o dia em que a OpenAI lançou o ChatGPT.

Como jornalista, acompanho a inteligência artificial desde 2016 e vejo como a tecnologia está avançando sem parar. A IA tem sido um tema que fascina executivos e tecnólogos, porém, antes do ChatGPT, ela ainda não havia alcançado o coração das conversas cotidianas. Para quem acompanha de perto a tecnologia, o progresso parecia constante. Quando o ChatGPT chegou, pensei que seria apenas mais um passo, afinal, não parecia muito diferente de um modelo de IA chamado Instruct GPT, que a OpenAI havia lançado quase um ano antes. Eu já tinha escrito sobre o Instruct GPT, um importante refinamento do GPT-3, o modelo de linguagem de grande escala (LLM) da OpenAI. O Instruct GPT era mais fácil de controlar por meio de instruções baseadas em texto — os chamados prompts — do que o GPT-3, e tinha menos probabilidade de gerar conteúdo racista, sexista, homofóbico ou tóxico. Ele também conseguia executar muitas tarefas de linguagem natural, como traduzir, resumir e codificar. No entanto, o impacto do Instruct GPT foi mínimo.

Então, quando o ChatGPT estreou, fiquei surpreso com a surpresa. E eu não deveria estar. Os *form factors* — as interfaces que usamos para interagir com uma tecnologia, como a barra de pesquisa do Google — são importantes. O Ford T era mais do que um motor de combustão interna e um conjunto de rodas, afinal. E com sua interface simples de chatbot, a OpenAI acertou na mosca. Foi um momento Eureka para mim também.

Durante a maior parte da década anterior, os avanços da IA ocorreram em domínios restritos. Muitas das conquistas, sobre as quais escrevi antes do ChatGPT, envolviam softwares que superavam humanos em jogos. O laboratório de IA de ponta do Google, o DeepMind, criou um software chamado AlphaGo que, em março de 2016, derrotou o campeão mundial Lee Sedol no antigo

jogo de estratégia Go. Foi uma das conquistas mais importantes na ciência da computação, pois o Go possui tantas combinações de movimentos possíveis que um algoritmo não consegue calcular todas, como fazem os softwares de IA com o xadrez. Em vez disso, ele deve agir de acordo com as probabilidades, aprendidas ao longo de suas partidas, para decidir qual movimento é melhor. Mas jogos são apenas jogos. Embora eles representem habilidades que gostaríamos que um sistema de IA dominasse, os jogos, por sua própria natureza, abstraem grande parte da complexidade da vida. Em um tabuleiro de Go, tudo é literalmente preto e branco, e ambos os jogadores podem ver a posição de cada peça.

Mais à frente, a DeepMind apresentou um sistema de IA capaz de vencer os melhores jogadores humanos em um videogame complexo chamado *StarCraft II*, cujo mundo imaginário tem um pouco mais de semelhança com nossa própria realidade confusa. A OpenAI, a mesma empresa de pesquisa de IA que depois criou o ChatGPT, construiu uma equipe de agentes de IA que poderia vencer uma equipe de jogadores humanos em um videogame frenético chamado *Dota 2*. Porém, esses não foram momentos Eureka. Poucas pessoas conheciam *StarCraft II* ou *Dota 2*, e o software que dominava esses jogos não tinha aplicação prática imediata na vida real.

Antes do ChatGPT, a IA realizava seus truques por debaixo dos panos. Ela ajudava produtos conhecidos a fazer coisas como recomendar filmes ou marcar pessoas em fotos nas redes sociais. Esses modelos de IA eram impressionantes, mas não eram tecnologias de uso geral, e a maioria das pessoas não considerava essas funções como sendo IA. Meus editores, às vezes, ignoravam minhas sugestões de matérias sobre esses pequenos avanços, e, pensando bem, eles não têm culpa. Esse tipo de IA parecia muito distante da inteligência artificial que víamos nos filmes, como a senciência malévola de HAL 9000 em *2001: Uma Odisseia no Espaço*, de Stanley Kubrick, o benevolente computador de *Jornada nas Estrelas* ou até mesmo a sedutora assistente digital de *Ela*, de 2013.

Isso começou a mudar em 2018, quando os pesquisadores passaram a progredir rapidamente com os modelos de linguagem de

grande escala (LLMs), o tipo de tecnologia na qual o ChatGPT se baseia. Esses avanços mostraram que um sistema de IA suficientemente grande, treinado apenas para prever palavras em uma frase, poderia ser usado para uma variedade de tarefas, como traduzir, resumir ou responder a perguntas. Antes, cada uma dessas tarefas exigia um sistema de IA diferente. No mundo da IA, isso era um feito significativo, e eu cheguei a publicar artigos na *Bloomberg* e na *Fortune* sobre o potencial impacto de longo alcance que isso poderia ter nos negócios. Mas, mesmo assim, a tecnologia permanecia no backend dos softwares já existentes. Como não estava integrada a um produto de fácil uso voltado para o consumidor, era difícil para a maioria das pessoas — incluindo muitos especialistas — perceberem o quanto o mundo estava prestes a mudar. Essa tendência de desconsiderar o progresso da IA tem um nome: o Efeito Tesler, cunhado pelo cientista da computação Larry Tesler, que certa vez brincou: “IA é tudo aquilo que ainda não foi feito”. O que já foi feito, disse ele, simplesmente tomamos como certo. Além disso, uma vez que um programa de IA consegue executar uma habilidade, deixamos de considerá-la como parte da nossa definição de “inteligência” — uma qualidade amorfa que exige uma combinação de competências semelhantes às humanas, sempre além do alcance do software.

Com o ChatGPT, a IA finalmente superou o Efeito Tesler — e foi um pouco além. De repente, as pessoas entenderam que a IA havia chegado. Você podia falar com ela. Ela respondia de forma convincente e confiante, mesmo que não estivesse sempre correta. Você podia pedir receitas com base nos ingredientes que tinha na geladeira. Você podia fazer com que ela resumisse a sua reunião — em prosa ou, se quisesse, até em versos! Podia pedir ideias para melhorar seus negócios ou conselhos sobre como pedir um aumento de salário. Ela podia até criar um site ou um software personalizado para processar dados e analisá-los. Combinada com outros modelos de IA generativa, ela era capaz de produzir imagens de praticamente qualquer coisa que você descrevesse.

Agora, a IA estava chegando a um novo patamar. Estava mais com a cara da IA de 2001: *Uma Odisseia no Espaço* e de *Jornada nas*

Estrelas. Nossa atitude em relação à tecnologia foi de indiferente a desnorçada da noite para o dia. O que esse momento significava? Que futuro a IA estava descortinando? Será que teríamos um futuro? Surgiram declarações assustadoras, algumas vindas dos próprios executivos e pesquisadores na linha de frente da inovação: que a IA poderia acabar com 300 milhões de empregos ao redor do mundo; que o ChatGPT era como “uma bomba nuclear de bolso” que detonou em meio a um público desavisado; que acabaria com a educação, eliminaria a confiança e destruiria a democracia; que, se não tivéssemos cuidado, a IA poderia até mesmo significar “o fim da linha para todos nós”, como advertiu o cofundador e CEO da OpenAI, Sam Altman.

Este livro é minha tentativa de responder às muitas perguntas suscitadas pelo ChatGPT. Ao pesquisá-lo, fiquei impressionado com o quanto muitas de nossas preocupações, bem como parte do pânico moral que acompanhou a revolução da IA, também fizeram parte da introdução de tecnologias anteriores, como a impressão gráfica, a televisão e a internet. Alguns temores sobre a IA remontam ao início da era dos computadores. Outros surgiram mais recentemente, com a chegada da internet, do GPS, das redes sociais e dos telefones celulares. Há importantes lições a serem tiradas de como nós, coletivamente, enfrentamos os desafios que essas invenções representaram — no modo como fomos transformados mas, talvez mais importante, no modo como não fomos. A IA representa uma aceleração, uma extensão e, em muitos casos, uma exacerbação das tendências que essas tecnologias anteriores já haviam posto em movimento. Quando se trata de isolamento social ou desconfiança na mídia, a IA representa uma diferença de grau, não necessariamente de natureza.

Mas ela difere dessas inovações anteriores de três formas distintas. É uma tecnologia de uso mais amplo do que a maioria das tecnologias anteriores. Isso a torna muito mais próxima da invenção da escrita, da fundição de metais ou da eletricidade do que do telefone, do automóvel ou do avião. Talvez seja “a invenção definitiva” que precisávamos criar, como vários pensadores já disseram. Isso porque ela promete ajudar a criar todas as tecnologias

futuras que possamos vir a precisar, tendo o potencial de impactar em quase todos os aspectos da sociedade.

Em segundo lugar, o ritmo de avanço e adoção da IA é mais rápido do que o das tecnologias anteriores. O ChatGPT atingiu 100 milhões de usuários apenas um mês após o lançamento. Para atingir esse marco, o Facebook, o produto de software de consumo que mais rapidamente se difundiu, levou quase quatro anos e meio. Os gastos com tecnologia estão aumentando, e mais da metade das empresas da Fortune 500 planeja implementar alguma forma de IA similar ao ChatGPT nos próximos anos.

A velocidade com que tudo isso está acontecendo é importante porque nos dá menos tempo em vários aspectos: para nos ajustarmos à nova realidade que a IA está criando; para refletirmos sobre as implicações da IA e sobre como queremos regulá-la; e para agirmos antes que seja tarde demais, antes que danos, sejam pequenos ou grandes, possam ocorrer.

Finalmente, mais do que qualquer tecnologia anterior, a IA toca profundamente na percepção que temos daquilo que nos torna únicos enquanto espécie: nossa inteligência e criatividade. Quando trabalhadores manuais perderam espaço para as máquinas a vapor, isso desvalorizou o trabalho humano. Mas a força física nunca foi nosso grande diferencial: já conhecíamos outras criaturas que poderiam correr mais rápido, escalar mais alto, nadar mais longe e nos subjugar. Mas nada poderia nos superar no campo das ideias. Os microprocessadores e as gerações anteriores de software rivalizavam com algumas de nossas habilidades cognitivas: realizavam cálculos muito mais rápidos e precisamente que nós, e seguiam qualquer conjunto de regras com mais exatidão. Mas fomos nós que *escrevemos* esses processos. Os computadores não desafiavam nosso intelecto, nossa capacidade de raciocinar, inventar novas maneiras de resolver problemas e expressar ideias e emoções. Já a IA desafia a maioria dessas coisas, se não todas. É por isso que ela causa tanto fascínio e pavor. Pela primeira vez, estamos numa posição em que podemos não ser mais a inteligência proeminente do planeta — e talvez por pouco tempo.

Embora isso dê um frio na barriga, temos muito o que esperar de um futuro definido pela IA. Ela pode até nos conceder superpoderes, se for bem projetada e acompanhada das políticas certas. Pode nos livrar de tarefas repetitivas, permitindo que trabalhemos com mais rapidez e inteligência. Isso desencadeará um boom de produtividade sem precedentes, acelerando o crescimento econômico. Em vez de acabar com a educação, como muitos professores e pais temem, a IA pode proporcionar a cada um de nós um tutor pessoal na palma da mão. Ela também nos fará ultrapassar os limites da ciência e da medicina, possibilitando novas curas e tratamentos mais personalizados. Nos ajudará a avançar na compreensão da química, biologia e até da nossa própria história e pré-história. Auxiliará no monitoramento do planeta, na proteção da biodiversidade e no uso mais eficiente das energias renováveis. Para muitos, será uma companheira constante, possivelmente aliviando a solidão e ajudando a desenvolver habilidades interpessoais. Caso usada corretamente, ela pode até mesmo fortalecer a democracia.

Mas também existem perigos graves. Se fizermos as escolhas erradas, a IA pode nos debilitar individual e coletivamente. Nossas habilidades sociais podem atrofiar. No trabalho, podemos acabar ficando fragilizados e desmotivados, servos da máquina, em vez de mestres. Sem políticas governamentais adequadas, a desigualdade se ampliará e o poder se concentrará. A proliferação de conteúdo sintético pode aniquilar a confiança e comprometer o já frágil tecido da democracia. Além disso, o alto consumo de energia da IA pode prejudicar o combate às mudanças climáticas. A tecnologia pode ajudar a projetar armas biológicas, tornando as guerras ainda mais mortais para civis e militares. Nas mãos erradas, pode se tornar uma ferramenta de terror. Se automatizarmos as decisões relacionadas a armas nucleares, haverá o risco de uma catástrofe global. Uma futura superinteligência pode até representar uma ameaça à nossa espécie.

Nas páginas a seguir, você também verá argumentos sobre o impacto da IA que não recebem tanta atenção quanto as teorias sobre o colapso da democracia ou o fim do mundo. O principal