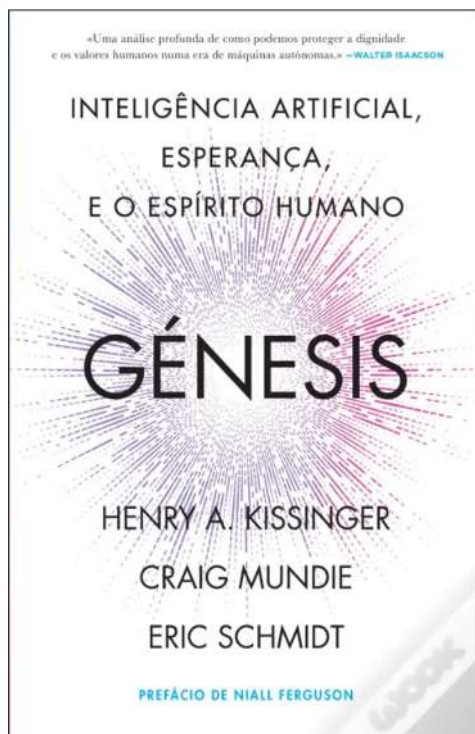


# Génesis - Inteligência Artificial, Esperança e o Espírito Humano



José Braga

Quando o termo «inteligência artificial» (IA) foi criado por John McCarthy (1927-2011), em 1955, o cientista da computação definiu-a como «a ciência e a engenharia de fazer máquinas inteligentes». Estava a nascer algo cujas implicações, passados 70 anos, ainda não são completamente compreendidas, mas que se torna importante estudar sobretudo desde que a «OpenAI» lançou o «ChatGPT» em 2022, criando uma «IA de massas». Este livro vem contribuir para esse entendimento, levantar questões e fomentar a discussão, sendo traduzido para português, de forma competente, no mesmo ano em que foi editado nos E.U.A.

Os autores refletiram ou trabalharam na área. Henry Kissinger (1924- 2024) foi Secretário de Estado dos E.U.A. e laureado com Prémio Nobel da Paz em 1973. Figura polémica, este foi o seu último livro, escrito já ao chegar ao século de vida. Eric Schmidt foi presidente e CEO da Google, e Craig Mundie foi diretor de pesquisa e estratégia da Microsoft. Os autores apresentam o resultado de um diálogo entre gerações diferentes de pensadores. A obra continua o livro de 2021, «A Era da Inteligência Artificial», editado pela mesma editora.

«Génesis» possui um «Prefácio» do historiador Niall Ferguson, biógrafo de Kissinger, professor da universidade de Oxford, lem-

## «Génesis»

Henry Kissinger, Eric Schmidt e Craig Mundie; Edições D. Quixote, 2024, 255 p.; ISBN 978-972-20-8379-9.

brando que «Os avanços tecnológicos podem ter tanto consequências benéficas como malélicas, dependendo da forma como decidimos coletivamente explorá-los» (p. xxii) e que se a primeira bomba nuclear surgiu 5 anos depois da descoberta da fissão nuclear (1938), só em 1951 esta foi usada para a produção de energia. Assim, uma evolução não é linear nem previsível e o maior perigo colocado pela IA é declararmos demasiado cedo, ou definitivamente, que a compreendemos totalmente. Deve-se, pois, ter sempre uma atitude crítica e cautelosa. As tecnologias possuem impactos que os seus inventores não previram, e esse padrão irá manter-se, pelo que os cientistas devem encontrar formas de incluir salvaguardas em todos os sistemas de IA. Isto «...vai exigir a resolução não de um, mas de dois “problemas de alinhamento”: o alinhamento técnico dos valores e intenções humanos com as ações da IA, e o alinhamento diplomático dos humanos uns com os outros.» (p. 4), como se lembra na «Introdução».

Neste volume, abarca-se o impacto da IA em oito áreas diferentes da atividade e do pensamento humanos, procurando-se uma estratégia que permita equilibrar benefícios e riscos. A tese dos autores é de que a IA deve complementar a inteligência humana e deve submeter-se a ela.

Na «Parte I- No Começo», sublinha-se como a descoberta humana se pode livrar do perigo físico: existem restrições biológicas e limitações humanas para explorar o espaço, as profundezas oceânicas ou o interior do planeta e do ser humano. A IA também não se intimida pela extensão da realidade ou profundidade dos problemas. Para se fazer exploração é necessária inteligência, recursos, ambiente e colaboradores certos. A competição e colaboração também possuem um papel, tal como a capacidade de integrar conhecimentos numa ponte entre disciplinas. A IA permite tornar a inteligência coletiva facilmente acessível. Será capaz de aprender e talvez mesmo aperfeiçoar-se a si mesma. Pode fazer avançar o ritmo das descobertas científicas.

Mostra-se como as tecnologias melhoraram as funções corporais humanas: a velocidade distingue a IA da forma de aprendizagem e das capacidades mentais humanas que podem estar a aproximar-se de um limite biológico. A mente humana é confrontada com o facto de já não ser o único, ou o melhor, sistema de inteligência, passando a surgir informação sem explicação ou justificação, fontes ou outros meios de prova, sendo opaca mesmo para os seus produtores, o que é um grave problema. Levantam-se questões para a Filosofia da Ciência, na medida em que surge outra entidade capaz de ler a realidade, como já aconteceu com a demonstração do teorema das quatro cores, em 1976. Além disso, a IA inspira-se no cérebro humano. Consegue ingerir mais informação e responder a mais pedidos simultâneos, resolvendo problemas mais complexos (precisão) e de maior dimensão (escala) do que os solucionáveis pelo intelecto. A IA (ou as IA's) pode começar a associar a experiência à compreensão como fazem os humanos.

A «Parte II- Os Quatro Ramos» traz reflexões sobre a aplicação da IA na política e diplomacia, segurança e defesa, Economia e Ciência. A IA pode ser aplicada à decisão política, abrindo possibilidades de diversificar os modos de governação e inovar politicamente, estas vantagens devem ser incorporadas nas instituições lentamente de modo que a IA seja aceite como complemento à decisão humana. Imune ao medo e preferências pessoais, a IA pode dar objetividade à tomada de decisões estratégicas, mas estas devem preservar a subjetividade humana, importante para exercer policiamento ou praticar a defesa.

Sabotagem, sequestro de sistemas de comunicação ou criação de informação falsa tornam-se mais fáceis e de difusão mais rápida, sendo necessárias medidas prévias de prevenção e de educação, daí a importância da familiarização dos jovens com este fenómeno nas escolas. A guerra poderá evoluir criando apenas baixas mecânicas, sendo a força psíquica do Homem decisiva (lembre-se o uso de «drones» na guerra russo/ucraniana). Mas devemos usar a IA para fomentar a prosperidade, eliminar a pobreza e a desigualdade. Terá impactos significativos no mundo do trabalho, aumentando a produtividade. Pode ser usada para criar substitutos sustentáveis para matérias-primas não renováveis e gerar uma nova base para a riqueza e bem-estar humanos: «A IA poderia ser direcionada para fazer o que já não temos de fazer, precisamente para podermos fazer o que queremos fazer» (p. 164). Devemos concentrar-nos na distribuição dos seus benefícios: entre as páginas 162 e 181, os autores discorrem sobre os efeitos da IA em educação, Medicina e Ciência. São páginas que devem ser lidas e discutidas por professores, investigadores e alunos.

Na «Parte III- A Árvore da Vida» traçam-se cenários de engenharia biológica para a evolução do Ser Humano, sendo a IA ameaça ou parceira. Deve-se alinhar a IA com os valores humanos, legislando o seu uso, prevendo sanções para quem a aplicar de forma ilegal, incorreta, imoral. Chegaremos ao ponto em que será necessário controlar a própria IA. Como a puniremos, se for necessário? Irá a própria IA competir com o ser humano? Os sistemas de controlo devem ser capazes de lidar com utilizações em tempo real e aprender. Será necessária uma IA supervisora? Quem a supervisionará?

O alinhamento entre a IA e o Homem pode ser feito de duas formas: sistemas baseados em instruções pré-programadas que funcionam como regras, ou fomentar a «aprendizagem por reforço», aprendendo a IA a partir da interação com os humanos, adaptando-se a circunstâncias específicas. Não há substituto da experiência, da participação ou da interação, sendo por isso as relações humanas insubstituíveis, mais ainda por permitirem incutir comportamentos éticos.

Os autores apresentam como «Conclusão» a importância de diferentes equilíbrios dinâmicos entre controlo e utilidade, da dignidade humana e o surgimento de uma nova parceira para o Homem. «(...) é necessário duvidarmos de nós mesmos para procurar conhecimento, mas confiarmos em nós mesmos para agir. (...) na era da IA, isto é ainda mais urgente.» (p. 226). «Notas», com referências bibliográficas para quem quiser aprofundar algum tema e um «Índice Remissivo» encerram a obra.

Este trabalho traça cenários astutos, embora, na perspetiva de 2025, excêntricos, caso da Inteligência Artificial Geral (p. 76-78). A IA também perturba a confiança no método científico. São questões para as quais a História e Filosofia da Ciência podem contribuir fornecendo perspetivas, pois existem precedentes e paralelos históricos estudados que podem servir para um melhor entendimento de cenários futuros, e, sobretudo, para traçar cenários de Prospetiva: a questão-chave que anima «Génesis», é a do equilíbrio entre controlar ou ser controlado pela IA.

Outro aspeto central de «Génesis» está na recusa dos autores em adotar uma postura unidimensional. Em vez de crença na salvação tecnológica ou medo da dominação digital, defendem uma abordagem equilibrada. Pedem regulamentação, ética e cooperação internacional, reconhecendo a IA não apenas como uma questão tecnológica, mas algo complexo. «Génesis» leva-nos a considerar a IA não como uma abstração, mas como algo que já está presente na nossa vida. É um livro com um vasto público-alvo: estudantes, professores e qualquer pessoa interessada em algo cada vez mais importante na vida quotidiana.



José Braga, Professor do Ensino Básico e Secundário, membro do Centro Interuniversitário de História das Ciências e da Tecnologia (CIUHCT).