

As habilidades da Inteligência artificial



Por **YUVAL NOAH HARARI***

A Inteligência artificial invadiu o sistema operacional da civilização humana

O medo da Inteligência artificial (IA) tem assombrado a humanidade desde o início da era do computador. Até agora, esses temores se concentravam nas máquinas que usavam meios físicos para matar, escravizar ou substituir pessoas. Porém, nos últimos dois anos, surgiram novas ferramentas de Inteligência artificial que ameaçam a sobrevivência da civilização humana de uma forma inesperada. A Inteligência artificial adquiriu algumas habilidades notáveis para manipular e gerar linguagem, seja com palavras, sons ou imagens. Dessa forma, a Inteligência artificial invadiu o sistema operacional da nossa civilização.

A linguagem é o material de que é feita quase toda a cultura humana. Os direitos humanos, por exemplo, não estão inscritos em nosso DNA. Em vez disso, são artefatos culturais que criamos ao contar histórias e escrever leis. Os deuses não são realidades físicas. Em vez disso, são artefatos culturais que criamos ao inventar mitos e escrever escrituras.

O dinheiro também é um artefato cultural. As cédulas são apenas pedaços de papel colorido e, atualmente, mais de 90% do dinheiro nem sequer são cédulas, são apenas informações digitais em computadores. O que dá valor ao dinheiro são as histórias que banqueiros, ministros da fazenda e gurus da criptomoeda nos contam sobre ele. Sam Bankman-Fried, Elizabeth Holmes e Bernie Madoff não eram particularmente bons em criar valor real, mas todos eles eram contadores de histórias extremamente competentes.

O que aconteceria quando uma inteligência não humana se tornasse melhor do que o ser humano médio para contar histórias, compor melodias, desenhar imagens e escrever leis e escrituras? Quando as pessoas pensam no ChatGPT e em outras novas ferramentas de Inteligência artificial, geralmente são atraídas para exemplos como crianças em idade escolar que usam Inteligência artificial para escrever suas redações. O que acontecerá com o sistema escolar quando as crianças fizerem isso? Mas a esse tipo de pergunta escapa um panorama mais geral. Esqueça as redações escolares. Pense na próxima corrida presidencial americana, em 2024, e tente imaginar o impacto das ferramentas de Inteligência artificial que podem ser criadas para produzir conteúdo político em massa, notícias falsas e escrituras para novos cultos.

Nos últimos anos, o culto QAnon se aglutinou em torno de mensagens anônimas on-line, conhecidas como “Q drops”. Os seguidores coletaram, reverenciaram e interpretaram esses Q drops como um texto sagrado. Embora, até onde sabemos, todos os Q drops anteriores tenham sido compostos por humanos e os bots tenham apenas ajudado a disseminá-los, no futuro poderemos ver os primeiros cultos da história cujos textos reverenciados foram escritos por uma inteligência não humana. As religiões ao longo da história reivindicaram uma fonte não humana para seus livros sagrados. Em breve, isso poderá se tornar realidade.

Em um nível mais prosaico, em breve poderemos nos ver conduzindo longas discussões on-line sobre aborto, mudanças climáticas ou a invasão russa da Ucrânia com entidades que pensamos serem humanos, mas que na verdade são

a terra é redonda

Inteligência artificial. O problema é que é totalmente inútil gastarmos tempo tentando mudar as opiniões declaradas de um *bot* de Inteligência artificial, enquanto a Inteligência artificial poderia aprimorar suas mensagens com tanta precisão que teria uma boa chance de nos influenciar.

Por meio de seu domínio da linguagem, a Inteligência artificial poderia até mesmo formar relacionamentos íntimos com as pessoas e usar o poder da intimidade para mudar nossas opiniões e visões de mundo. Embora não haja nenhuma indicação de que a Inteligência artificial tenha consciência ou sentimentos próprios, para promover uma falsa intimidade com os humanos, basta que a Inteligência artificial consiga fazer com que eles se sintam emocionalmente ligados a ela.

Em junho de 2022, Blake Lemoine, um engenheiro do Google, afirmou publicamente que o *chatbot* LaMDA, no qual ele estava trabalhando, tinha se tornado senciente. Essa afirmação polêmica custou-lhe o emprego. O mais interessante desse episódio não foi a afirmação do Sr. Lemoine, que provavelmente era falsa. Em vez disso, foi sua disposição de arriscar seu emprego lucrativo em prol do *chatbot* de Inteligência artificial. Se a Inteligência artificial pode influenciar as pessoas a arriscarem seus empregos por ela, o que mais ela poderia induzi-las a fazer?

Em uma batalha política por mentes e corações, a intimidade é a arma mais eficiente, e a Inteligência artificial acaba de ganhar a capacidade de produzir em massa relacionamentos íntimos com milhões de pessoas. Todos nós sabemos que, na última década, as mídias sociais se tornaram um campo de batalha para controlar a atenção humana. Com a nova geração de Inteligência artificial, a frente de batalha está mudando da atenção para a intimidade. O que acontecerá com a sociedade humana e a psicologia humana quando a Inteligência artificial lutar contra a Inteligência artificial em uma batalha para fingir relacionamentos íntimos conosco, que podem ser usados para nos convencer a votar em determinados políticos ou comprar determinados produtos?

Mesmo sem criar uma “falsa intimidade”, as novas ferramentas de Inteligência artificial teriam uma imensa influência em nossas opiniões e visões de mundo. As pessoas podem vir a usar um único conselheiro da Inteligência artificial como um oráculo completo e onisciente. Não é de se admirar que o Google esteja apavorado. Por que se preocupar em pesquisar se posso simplesmente perguntar ao oráculo? As indústrias de notícias e publicidade também deveriam estar apavoradas. Por que ler um jornal se posso simplesmente pedir ao oráculo para me informar sobre as últimas notícias? E qual é a finalidade dos anúncios, quando posso simplesmente pedir ao oráculo que me diga o que comprar?

E mesmo esses cenários não capturam o panorama geral. O que estamos falando é potencialmente o fim da história humana. Não o fim da história, apenas o fim de sua parte dominada pelos humanos. A história é a interação entre biologia e cultura; entre nossas necessidades biológicas e desejos por coisas como comida e sexo e nossas criações culturais, como religiões e leis. A história é o processo pelo qual as leis e as religiões moldam a comida e o sexo.

O que acontecerá com o curso da história quando a Inteligência artificial assumir o controle da cultura e começar a produzir histórias, melodias, leis e religiões? Ferramentas anteriores, como a imprensa e o rádio, ajudaram a disseminar as ideias culturais dos seres humanos, mas nunca criaram novas ideias culturais próprias. A Inteligência artificial é fundamentalmente diferente. A Inteligência artificial pode criar ideias completamente novas, uma cultura completamente nova.

No início, a Inteligência artificial provavelmente imitará os protótipos humanos com os quais foi treinada em sua infância. Mas, a cada ano que passa, a cultura da Inteligência artificial irá audaciosamente para onde nenhum ser humano foi antes. Durante milênios, os seres humanos viveram dentro dos sonhos de outros seres humanos. Nas próximas décadas, poderemos nos encontrar vivendo dentro dos sonhos de uma inteligência alienígena.

O medo da Inteligência artificial tem assombrado a humanidade apenas nas últimas décadas. Mas, há milhares de anos, os seres humanos são assombrados por um medo muito mais profundo. Sempre apreciamos o poder das histórias e imagens para manipular nossas mentes e criar ilusões. Consequentemente, desde os tempos antigos, os seres humanos temem estar presos em um mundo de ilusões.

a terra é redonda

No século XVII, René Descartes temia que talvez um demônio malicioso o estivesse aprisionando em um mundo de ilusões, criando tudo o que ele via e ouvia. Na Grécia antiga, Platão contou a famosa Alegoria da Caverna, na qual um grupo de pessoas é acorrentado dentro de uma caverna por toda a vida, diante de uma parede em branco. Uma tela. Nessa tela, eles veem projetadas várias sombras. Os prisioneiros confundem as ilusões que vêem com a realidade.

Na Índia antiga, os sábios budistas e hindus indicaram que todos os humanos viviam presos em *Maya* – o mundo das ilusões. O que normalmente consideramos realidade muitas vezes são apenas ficções em nossa própria mente. As pessoas podem travar guerras inteiras, matando outras e querendo ser mortas, por causa de sua crença nessa ou naquela ilusão.

A revolução da Inteligência artificial está nos colocando frente a frente com o demônio de Descartes, com a caverna de Platão e com *Maya*. Se não formos cuidadosos, poderemos ficar presos atrás de uma cortina de ilusões, que não conseguiremos rasgar, ou mesmo perceber que está ali.

É claro que o novo poder da Inteligência artificial também pode ser usado para bons propósitos. Não vou me estender sobre isso, porque as pessoas que desenvolvem a Inteligência artificial já falam bastante sobre ela. O trabalho de historiadores e filósofos como eu é apontar os perigos. Mas, certamente, a Inteligência artificial pode nos ajudar de inúmeras maneiras, desde encontrar novas curas para o câncer até descobrir soluções para a crise ecológica. A questão que enfrentamos é como garantir que as novas ferramentas de Inteligência artificial sejam usadas para o bem e não para o mal. Para fazer isso, primeiro precisamos avaliar as verdadeiras capacidades dessas ferramentas.

Desde 1945, sabemos que a tecnologia nuclear poderia gerar energia barata para o benefício dos seres humanos, mas também poderia destruir fisicamente a civilização humana. Por isso, reformulamos toda a ordem internacional para proteger a humanidade e garantir que a tecnologia nuclear fosse usada principalmente para o bem. Agora temos que lidar com uma nova arma de destruição em massa que pode aniquilar nosso mundo mental e social.

Ainda podemos regulamentar as novas ferramentas de Inteligência artificial, mas precisamos agir rapidamente. Enquanto as bombas nucleares não podem inventar bombas nucleares mais potentes, a Inteligência artificial pode criar exponencialmente Inteligência artificial ainda mais poderosa. O primeiro passo crucial é exigir verificações de segurança rigorosas antes que ferramentas de Inteligência artificial poderosas sejam liberadas para o domínio público.

Assim como uma empresa farmacêutica não pode liberar novos medicamentos antes de testar seus efeitos colaterais de curto e longo prazos, as empresas de tecnologia não devem liberar novas ferramentas de Inteligência artificial antes que elas sejam consideradas seguras. Precisamos de um equivalente da *Food and Drug Administration* para novas tecnologias, e precisamos disso para ontem.

A desaceleração das implementações públicas de Inteligência artificial não fará com que as democracias fiquem atrás de regimes autoritários mais cruéis? É exatamente o contrário. As implementações não regulamentadas de Inteligência artificial criariam o caos social, o que beneficiaria os autocratas e arruinaria as democracias. A democracia é uma conversação, e as conversas dependem da linguagem. Quando a Inteligência artificial invade a linguagem, ela pode destruir nossa capacidade de ter conversas significativas, destruindo assim a democracia.

Acabamos de encontrar uma inteligência alienígena aqui na Terra. Não sabemos muito sobre ela, exceto que ela pode destruir nossa civilização. Devemos por fim à implementação irresponsável de ferramentas de Inteligência artificial na esfera pública e regular a Inteligência artificial antes que ela nos regule. E a primeira regulamentação que eu sugeriria é tornar obrigatório que a Inteligência artificial revele que é uma Inteligência artificial. Se eu estiver conversando com alguém e não puder saber se é um humano ou uma Inteligência artificial, será o fim da democracia.

Este texto foi gerado por um humano. Ou será que não?

Yuval Noah Harari é professor de história na Universidade Hebraica de Jerusalém. Autor, entre outros livros, de *Sapiens*

a terra é redonda

- uma breve história da humanidade (*Companhia das Letras*).

Tradução: **Fernando Lima das Neves**.

Publicado originalmente no portal da revista [The economist](#).

A Terra é Redonda existe graças aos nossos leitores e apoiadores.
Ajude-nos a manter esta ideia.

CONTRIBUA