

As 3 etapas da inteligência artificial e por que a 3ª pode ser fatal

Author, Redação Role, BBC News Mundo

29 de maio de 2023



Desde o seu lançamento no final de novembro de 2022, o ChatGPT, o chatbot que usa inteligência artificial (também chamada de IA) para responder perguntas ou gerar textos sob demanda dos usuários, tornou-se o aplicativo de internet que mais cresce na história.

Em apenas dois meses atingiu 100 milhões de usuários ativos. O popular aplicativo TikTok levou nove meses para atingir esse marco. E para o Instagram dois anos e meio, segundo dados da empresa de monitoramento de tecnologia Sensor Town.

"Nos 20 anos que acompanhamos a internet, não conseguimos lembrar de um crescimento mais rápido de um aplicativo de internet para o consumidor", disseram analistas do banco suíço UBS, que relataram o recorde em fevereiro.

A enorme popularidade do ChatGPT, desenvolvido pela empresa OpenAI, com o apoio financeiro da Microsoft, gerou todo tipo de discussões e especulações sobre o impacto que a inteligência artificial generativa já está tendo e terá em nosso futuro próximo.

É o ramo da IA que se dedica a gerar conteúdo original a partir de dados existentes (geralmente retirados da internet) em resposta a instruções de um usuário.

Os textos (de ensaios, poesia e piadas a códigos de computador) e imagens (diagramas, fotos, obras de arte de qualquer estilo e muito mais) produzidos por IAs geradoras como ChatGPT, DALL-E, Bard e AlphaCode - para citar apenas alguns dos mais conhecidos - são, em alguns casos, tão indistinguíveis do trabalho humano, que já foram utilizadas por milhares de pessoas para substituir o seu trabalho habitual.

De estudantes que os usam para fazer o dever de casa a políticos que confiam a eles seus discursos - o deputado democrata Jake Auchincloss lançou o recurso no Congresso americano - ou fotógrafos que inventam instantâneos de coisas que não aconteceram (e até

ganham prêmios por isso, como o alemão Boris Eldagsen, que ganhou o primeiro lugar no último Sony World Photography Award por uma imagem criada por IA).

Este texto poderia ter sido digitado por uma máquina e você provavelmente não saberia.

O fenômeno levou a uma revolução nos recursos humanos, com empresas como a gigante da tecnologia IBM anunciando que deixará de contratar pessoas para preencher cerca de 8.000 vagas que podem ser gerenciadas pela IA.

Um relatório do banco de investimentos Goldman Sachs estimou no final de março que a IA poderia substituir um quarto de todos os empregos humanos hoje, embora também crie mais produtividade e novos empregos.



Quanto mais a IA avança, maior sua capacidade de substituir nossos trabalhos

Se todas essas mudanças representam uma sobrecarga para você, prepare-se para um fato que pode ser ainda mais desconcertante.

E é que, com todos os seus impactos, o que estamos vivenciando agora é apenas a primeira etapa do desenvolvimento da IA.

Segundo os especialistas, o que pode vir em breve - a segunda etapa - será muito mais revolucionário.

E a terceira e última, que pode ocorrer logo depois disso, é tão avançada que alterará completamente o mundo, mesmo à custa da existência humana.

Os três estágios

As tecnologias de IA são classificadas por sua capacidade de imitar características humanas.

1. Inteligência Artificial Estreita (ANI)

A categoria mais básica de AI é mais conhecida por sua sigla: ANI, para Artificial Narrow Intelligence

É assim chamado porque se concentra estritamente em uma única tarefa, realizando trabalhos repetitivos dentro de um intervalo predefinido por seus criadores.

Os sistemas ANI geralmente são treinados usando um grande conjunto de dados (por exemplo, da Internet) e podem tomar decisões ou realizar ações com base nesse treinamento.

Um ANI pode igualar ou exceder a inteligência e a eficiência humanas, mas apenas naquela área específica em que opera.

Um exemplo são os programas de xadrez que usam IA. Eles são capazes de vencer o campeão mundial dessa disciplina, mas não podem realizar outras tarefas.



ANI pode superar os humanos, mas apenas em uma área específica

É por isso que também é conhecido como "IA fraca" .

Todos os programas e ferramentas que usam IA hoje, mesmo os mais avançados e complexos, são formas de ANI. E esses sistemas estão por toda parte.

Os smartphones estão repletos de aplicativos que usam essa tecnologia, desde mapas GPS que permitem localizar qualquer lugar do mundo ou saber o tempo, até programas de música e vídeo que conhecem seus gostos e fazem recomendações.

Também assistentes virtuais como Siri e Alexa são formas de ANI. Como o buscador Google e o robô que limpa sua casa.

O mundo dos negócios também usa muito essa tecnologia. É usado nos computadores internos dos carros, na fabricação de milhares de produtos, no mundo financeiro e até nos hospitais, para fazer diagnósticos.

Sistemas ainda mais sofisticados, como carros sem motorista (ou veículos autônomos) e o popular ChatGPT, são formas de ANI, pois não podem operar fora do intervalo predefinido por seus programadores e, portanto, não podem tomar decisões por conta própria.

Eles também não têm autoconsciência, outro traço da inteligência humana.

No entanto, alguns especialistas acreditam que sistemas programados para aprender automaticamente (aprendizado de máquina), como ChatGPT ou AutoGPT (um "agente autônomo" ou "agente inteligente" que usa informações do ChatGPT para executar determinadas

subtarefas de forma autônoma) podem passar para o próximo estágio de desenvolvimento.

2. Inteligência Artificial Geral (AGI)

Esta categoria – Inteligência Artificial Geral – é alcançada quando uma máquina adquire habilidades cognitivas no nível humano.

Ou seja, quando você consegue realizar qualquer tarefa intelectual que uma pessoa faz.



O AGI tem a mesma capacidade intelectual de um ser humano.

O AGI também é conhecido como "IA forte" .

Tal é a crença de que estamos prestes a atingir esse nível de desenvolvimento, que em março passado mais de 1.000 especialistas em tecnologia pediram às empresas de IA que parassem de treinar, por pelo menos seis meses, os programas mais poderosos do que o GPT-4, a versão mais recente do ChatGPT.

"Sistemas de IA com inteligência que competem com os humanos podem representar riscos profundos para a sociedade e a humanidade", alertaram em carta aberta o cofundador da Apple, Steve Wozniak, e o proprietário da Tesla, SpaceX Neuralink, entre outros. os co-fundadores da Open AI antes de renunciar ao conselho devido a desentendimentos com a liderança da empresa).

A carta em que mais de 1.000 especialistas pedem para parar a inteligência artificial por ser uma "ameaça à humanidade".

Na carta, publicada pela organização sem fins lucrativos *Future of Life Institute*, os especialistas afirmam que, se as empresas não concordarem rapidamente em interromper seus projetos, "os governos devem intervir e instituir uma moratória" para que medidas de segurança possam ser projetadas e implementadas.

Embora isso seja algo que - por enquanto - não aconteceu, o governo dos Estados Unidos convocou os proprietários das principais empresas de IA - Alphabet, Anthropic, Microsoft e OpenAI - para acordar "novas ações para promover a inovação responsável da IA".

"A IA é uma das tecnologias mais poderosas do nosso tempo, mas para aproveitar as oportunidades que ela apresenta, devemos primeiro mitigar seus riscos", disse a Casa Branca em comunicado em 4 de maio.

O Congresso dos EUA, por sua vez, convocou o CEO da OpenAI, Sam Altman, na terça-feira para responder a perguntas sobre o ChatGPT.

Durante a audiência no Senado, Altman disse que é "crucial" que sua indústria seja regulamentada pelo governo, pois a IA se torna "cada vez mais poderosa".

Carlos Ignacio Gutiérrez, pesquisador de políticas públicas do Future of Life Institute, explicou à BBC News Mundo, serviço em espanhol da BBC, que um dos grandes desafios que a IA apresenta é que "não existe um colegiado de especialistas que decida como regulá-la, como acontece, por exemplo, com o Painel Intergovernamental sobre Mudanças Climáticas (IPCC)".

Na carta dos especialistas, eles definiram quais eram suas principais preocupações.

"Devemos desenvolver mentes não humanas que possam eventualmente nos superar em número, ser mais espertos que nós, nos tornar obsoletos e nos substituir?", eles perguntaram.

"Devemos arriscar perder o controle de nossa civilização?"

O que nos leva ao terceiro e último estágio da IA.

3. Superinteligência Artificial (ASI)

A preocupação desses cientistas da computação tem a ver com uma teoria bem estabelecida de que, quando chegarmos à AGI, logo depois chegaremos ao último estágio do desenvolvimento dessa tecnologia: a Superinteligência Artificial, que ocorre quando a inteligência sintética supera a humana.

O filósofo da Universidade de Oxford e especialista em IA, Nick Bostrom, define superinteligência como "um intelecto muito mais inteligente do que os melhores cérebros humanos em praticamente todos os campos, incluindo criatividade científica, sabedoria geral e habilidades sociais".

A teoria é que, quando uma máquina atinge inteligência equivalente à dos humanos, sua capacidade de multiplicar essa inteligência exponencialmente por meio de seu próprio aprendizado autônomo fará com que ela nos ultrapasse amplamente em pouco tempo, atingindo o ASI.

"O ser humano para ser engenheiro, enfermeiro ou advogado tem que estudar muito. O problema do AGI é que ele é imediatamente escalável", diz Gutiérrez.

Isso se deve a um processo chamado autoaperfeiçoamento recursivo que permite que um aplicativo de IA "se aprimore continuamente".

Embora haja muito debate sobre se uma máquina pode realmente adquirir o tipo de inteligência ampla que um ser humano possui - especialmente quando se trata de inteligência emocional - é uma das coisas que mais preocupa aqueles que acreditam que estamos perto de alcançar o AGI.

Recentemente, o chamado "padrinho da inteligência artificial" Geoffrey Hinton, pioneiro na investigação de redes neurais e aprendizado profundo que permitem que as máquinas aprendam com a experiência, assim como os humanos, alertou em entrevista à BBC que poderíamos estar próximos desse marco.

"Atualmente (as máquinas) não são mais inteligentes do que nós, até onde vejo. Mas acho que em breve poderão ser", disse o homem de 75 anos, que acabou de se aposentar do Google.

Extinção ou imortalidade

Existem, em geral, dois campos de pensamento em relação à ASI: há os que acreditam que essa superinteligência será benéfica para a humanidade e os que acreditam no contrário.

Entre eles estava o famoso físico britânico Stephen Hawking, que acreditava que as máquinas superinteligentes representavam uma ameaça à nossa existência.



O físico britânico Stephen Hawking acreditava que a IA superinteligente poderia levar ao 'fim dos humanos'

"O desenvolvimento da inteligência artificial completa pode significar o fim dos humanos", disse ele à BBC em 2014, quatro anos antes de morrer.

Uma máquina com esse nível de inteligência "descolaria por conta própria e se redesenharia em um ritmo crescente", disse ele.

"Os humanos, que são limitados pela lenta evolução biológica, não seriam capazes de competir e seriam superados", previu.

No entanto, do lado oposto, há previsões mais positivas.

Um dos maiores entusiastas da ASI é o autor e inventor futurista americano Ray Kurzweil, pesquisador de IA do Google e cofundador da Singularity University do Vale do Silício ("singularidade" é outro nome para a era em que as máquinas se tornam superinteligentes).

Kurzweil acredita que os humanos serão capazes de usar IA superinteligente para superar nossas barreiras biológicas, melhorando nossas vidas e nosso mundo.

Em 2015 chegou a prever que até o ano de 2030 os humanos conseguirão alcançar a imortalidade graças aos nanobots (robôs extremamente pequenos) que atuarão dentro de nossos corpos, reparando e curando qualquer dano ou doença, inclusive as causadas pela passagem do tempo.

Em sua declaração ao Congresso na terça-feira, Sam Altman, da OpenAI, também estava otimista sobre o potencial da IA, observando que ela poderia resolver "os maiores desafios da humanidade, como a mudança climática e a cura do câncer".

No meio estão pessoas, como Hinton, que acreditam que a IA tem um enorme potencial para a humanidade, mas acham o atual ritmo de desenvolvimento, sem regras e limites claros, "preocupante".

Em um comunicado enviado ao The New York Times anunciando sua saída do Google, Hinton disse que agora se arrependia do trabalho que havia feito porque temia que "maus agentes" usassem IA para fazer "coisas ruins".



Geoffrey Hinton disse à BBC que os avanços na IA o 'assustam'

(...) As máquinas poderiam eventualmente "criar subobjetivos como: 'Preciso obter mais energia', o que representaria um risco existencial", observou ele.

Ao mesmo tempo, o especialista britânico-canadense disse que, a curto prazo, a IA trará muito mais benefícios do que riscos, por isso "não devemos parar de desenvolvê-la".

"A questão é: agora que descobrimos que funciona melhor do que esperávamos alguns anos atrás, o que fazemos para mitigar os riscos de longo prazo de coisas mais inteligentes do que assumirmos o controle?"

Guitérrez concorda que a chave é criar um sistema de governança de IA antes de desenvolver uma inteligência que possa tomar suas próprias decisões.

"Se essas entidades são criadas com motivação própria, o que significa quando não estamos mais no controle dessas motivações?", questiona.

O especialista aponta que o perigo não é apenas que uma AGI ou ASI, seja por motivação própria ou controlada por pessoas com "maus objetivos", inicie uma guerra ou manipule o sistema financeiro, produtivo, infraestrutura energética, transporte ou qualquer outro sistema que agora é informatizado.

Uma superinteligência poderia nos dominar de uma forma muito mais sutil, adverte.

"Imagine um futuro onde uma entidade tenha tanta informação sobre cada pessoa no planeta e seus hábitos (graças às nossas buscas na internet) que poderia nos controlar de maneiras que não perceberíamos", diz ele.

"O pior cenário não é que existam guerras entre humanos e robôs. O pior é não percebermos que estamos sendo manipulados porque compartilhamos o planeta com uma entidade muito mais inteligente do que nós."