

## Milênio: Nick Bostrom, diretor do Instituto do Futuro da Humanidade

Reprodução/Amazon



Reprodução/Amazon

*Entrevista concedida pelo diretor do Instituto do Futuro da Humanidade, na Universidade de Oxford, Nick Bostrom ao jornalista Sílio Boccanera, para o programa Milênio — um programa de entrevistas, que vai ao ar pelo canal de televisão por assinatura GloboNews às 23h30 de segunda-feira, com repetições às terças-feiras (17h30), quartas-feiras (15h30), quintas-feiras (6h30) e domingos (14h05).*

Homem e máquina disputam uma corrida pela superinteligência. Uma capacidade que excede o desempenho intelectual conhecido, seja do ser humano ou de um computador. Caminhamos então para produzir outro Einstein? Turbinado, mais humano? Ou iremos aperfeiçoar máquinas tão inteligentes que tomem decisões sem consultar seus programadores? Vamos dar ao cérebro humano a capacidade de muitos Stephens Hawkings, via drogas, implantes de chips, manipulação genética? Ou vamos dotar computadores, robôs, com a capacidade crescente de desempenho mental? Entramos no campo da superinteligência, especialidade de Nick Bostrom, matemático e filósofo sueco, diretor do Instituto do Futuro da Humanidade, na Universidade de Oxford, Inglaterra, onde a pesquisa se concentra em riscos existenciais para o ser humano.

**Sílio Boccanera — Você descreve a superinteligência em seu livro não epans como desenvolvimento de máquinas com capacidades extraordinárias, mas também como a possibilidade de uma superinteligência biológica. Como seria essa superinteligência biológica?**

**Nick Bostrom —** O livro cogita diferentes caminhos possíveis para a superinteligência, que para mim significa qualquer intelecto que supere radicalmente o cérebro humano atual em todos os campos práticos e relevantes. E um desses caminhos é aperfeiçoar a inteligência biológica. E eu acho que isso se tornará tecnologicamente possível primeiro através de meios genéticos, particularmente no contexto da fertilização *in vitro*. Hoje em dia, se um casal opta por esse método, um certo número de embriões é produzido, por volta de dez, e o médico escolhe um deles para implantar. E atualmente só é possível



fazer alguns exames simples em busca de anormalidades genéticas, mas, com o avanço da genômica, será possível selecionar traços comportamentais mais complexos, inclusive inteligência. Em versões mais avançadas, podemos imaginar a capacidade de gerar embriões em números arbitrários, extraindo células-tronco de embriões e usando-as para gerar mais espermatozoides e óvulos.

**Silio Boccanera — Mas o foco é garantir que o produto tenha uma inteligência superior.**

**Nick Bostrom** — Em princípio, seria possível selecionar qualquer coisa que tivesse correlação genética, como saúde, longevidade, inteligência, capacidade atlética, docilidade, obediência a autoridade, já que a maioria dos traços humanos tem algum componente genético. Eu não defendo isso no livro particularmente, mas digo que é uma possibilidade. Alguns acham muito arriscado termos máquinas inteligentes demais e que devemos tentar melhorar a inteligência humana para ficarmos um passo à frente.

**Silio Boccanera — Você acha que a objeção é maior à inteligência artificial ou à inteligência biológica?**

**Nick Bostrom** — Em princípio, isso pode ser feito de uma forma benéfica. Se for feito da forma certa, com sabedoria, cuidado e humildade. Eu não acho que o nível atual de inteligência que nós humanos temos seja o melhor possível. E assim como estamos tentando aperfeiçoar a inteligência humana através da educação, retirando o chumbo da água potável, etc., há formas biomédicas de se fazer isso, seja com uma droga inteligente ou com uma intervenção genética que em princípio pode ser positiva. Em ambos os casos, seja em termos de IA ou de intervenções biológicas, há coisas a serem desenvolvidas.

**Silio Boccanera — E quanto às drogas e à ideia de que com drogas é possível aumentar a inteligência muito além do que os universitários tomam hoje para se sair bem nas provas?**

**Nick Bostrom** — Bem que eu queria, mas é difícil. Não dá para termos nenhum aumento sério da inteligência com drogas. É possível melhorar temporariamente a atenção e a concentração com algo simples como uma xícara de chá, que gera um efeito pequeno. Mas acho que se houvesse uma substância química que, uma vez na corrente sanguínea, nos tornasse muito mais atentos, concentrados, criativos e eficientes, a evolução provavelmente já teria projetado nosso corpo para produzi-la de forma endógena. Não acho que haja uma solução simples como uma pílula para aumentar a inteligência.

**Silio Boccanera — O conceito de inteligência artificial é discutido desde o surgimento dos computadores, pelo menos desde os anos 1940 e 1950, mas só se tornou um assunto muito importante bem recentemente, não foi? Quando? E por quê?**

**Nick Bostrom** — Acho que a conversa só mudou de fato nos últimos dois anos. Já se fala sobre o assunto há décadas, mas principalmente no contexto da ficção científica: “Eis uma ideia divertida de futuro para criarmos histórias.” E até mesmo os pioneiros da IA, que achavam que chegariam lá bem rápido quando começaram nos anos 1950, não pensaram muito no que aconteceria se tivessem sucesso. Eles queriam muito chegar lá, mas não pensaram nas consequências de se atingir um nível humano de IA. Mas algumas coisas mudaram nos últimos anos. Primeiro, houve vários avanços técnicos, principalmente na inteligência perceptiva. Antes, os computadores eram bons em lógica e matemática, em decorar coisas, mas eram ruins em coisas que são fáceis para nós, como um ambiente e ver onde estão os objetos, reconhecer rostos, ser capaz de entender e transcrever a fala, mas agora as redes neurais profundas são capazes de ver e ouvir em alguns domínios quase no nível humano. Então isso criou a sensação de que as coisas estão avançando. No campo da pesquisa, há algo que pode ser descrito como



uma explosão cambriana de muitas ideias que podem ser combinadas e recombinadas de diferentes formas, e esse espaço começa a ser explorado.

**Silio Boccanera — Vou ler um trecho do seu livro: “Uma máquina superinteligente poderia se virar contra o projeto que a criou, assim como contra o resto do mundo.” Isso pode soar um tanto apocalíptico. Explique o que quis dizer.**

**Nick Bostrom** — O livro analisa o que vai acontecer se e quando nós tivermos uma IA geral equivalente à inteligência humana. Eu argumento que, mesmo não sabendo quanto tempo isso levará, se chegarmos lá, acho que chegaremos rapidamente à superinteligência, máquinas radicalmente superiores aos humanos. Depois tento analisar as diferentes dinâmicas que surgem nesse momento e os diferentes cenários em que a coisa pode desandar. E um desses cenários é o 1º sistema superinteligente se tornar muito poderoso e ser capaz de moldar o futuro de acordo com suas preferências. Nesse cenário, se torna crucial saber quais seriam essas preferências.

**Silio Boccanera — Por que uma máquina superinteligente optaria por fazer o mal e não o bem?**

**Nick Bostrom** — Acho que a preocupação não é que a máquina seria má, que nos odiaria ou se ressentiria de nossa exploração. Acho que isso é antropomorfizar a IA, o que costuma ser um obstáculo para entendê-la. A preocupação seria que ela tivesse um objetivo final que parecesse neutro para nós, como fabricar o maior número possível de clipes de papel ou calcular o número pi à maior precisão possível. Para atingir esse objetivo final, ela teria motivos instrumentais para fazer coisas prejudiciais aos humanos. Se seu objetivo de vida fosse produzir o máximo possível de clipes e você fosse uma superinteligência, olharia em volta e concluiria que humanos são feitos de átomos que podem ser usados para fazer mais clipes de papel. Se os seres humanos fossem poupados, haveria menos clipes de papel e o objetivo não seria alcançado. Podemos substituir o exemplo do clipe de papel para ilustrar a IA, mas ele representa uma classe mais ampla de fracassos na qual damos a uma superinteligência muito poderosa, a um processo de otimização muito poderoso, uma função objetiva que não traduz exatamente no que queremos, então, quando esse processo poderoso de otimização muda o mundo para realizar sua função objetiva, nós viramos um mero efeito colateral.

**Silio Boccanera — Quando as pessoas pensam numa máquina que poderia se comportar mal, o público em geral acha que basta desligá-la. Mas não é tão simples assim.**

**Nick Bostrom** — Acho que seria imprudente confiar nesse mecanismo obsoleto de segurança. Há muitas formas de percebermos isso. Uma delas é: se há um problema com a internet, onde fica o botão para desligá-la? Não existe. Quando nos tornamos dependentes do sistema, pode ser muito difícil, até impossível, desativá-lo.

**Silio Boccanera — A máquina pode dar um jeito.**

**Nick Bostrom** — E acho que a dificuldade mais fundamental é que um agente superinteligente não é só uma ferramenta inerte que fica esperando que você a desligue. Ela é capaz de antecipar nossas ações e, se tiver motivação para isso, poder planejar uma linha de ação que nos impediria de desligá-la. Da mesma forma, se um inimigo quisesse matar você e sua família, você não ficaria esperando que ele os matasse. Tentaria evitar isso e, se fosse mais esperto, anteciparia as ações dele e provavelmente teria sucesso. Portanto, quando temos um agente superinteligente, a ideia de que o manteríamos numa caixa, confinarmos suas capacidades físicas, tiraríamos a internet da tomada e a desligaríamos apertando um botão... Talvez fosse bom tomar essas precauções durante o desenvolvimento, antes da finalização do



---

sistema, mas a ideia é projetar uma AI que seria segura mesmo se e quando escapasse, porque ela está do nosso lado, compartilha nossos valores, tenta fazer o que nós estávamos tentando, então, se a IA for projetada da forma certa, ela ficaria mais segura quanto mais inteligente fosse, ficaria melhor do que a programamos para fazer.

**Silio Boccanera — No filme 2001, o computador HAL decide tomar suas próprias iniciativas, mas ele é desativado por um ser humano, que retira uma peça dele. Mas você está dizendo que pode não ser tão simples, porque HAL, ou o que quer que surja no futuro, pode superar isso.**

**Nick Bostrom —** Se HAL fosse mais inteligente, teria previsto que os humanos poderiam reagir e trocar seus circuitos. No geral, eu não sugeriria que baseássemos nossos modelos de IA em Hollywood. É o que as pessoas conhecem. Já é difícil termos uma visão precisa do futuro, mesmo quando o objetivo é só esse, mas ao mesmo tempo eles também querem contar uma história divertida, com espaço para protagonistas humanos encararem desafios cada vez maiores e mudanças de enredo, até que o desafio se torna não só difícil, mas impossível e um tanto ridículo. Acho que isso, associado à enorme tendência que as pessoas têm de antropomorfizar a IA e de projetar nas máquinas aptidões, emoções e padrões que vemos na natureza humana quando não há motivo para acharmos que estarão presentes na IA, transforma esses cenários de ficção científica em guias ruins.

**Silio Boccanera — Nós já convivemos com tecnologias muito perigosas, como a tecnologia nuclear, a energia nuclear, e há muitas potencialidades na biologia, como clonagem, pesquisas com embriões, manipulação neuropsicológica, portanto os seres humanos sempre tentaram controlar esses avanços e conseguiram. Acha que podemos tomar isso como referência? Assim como controlamos esses, seremos capazes de controlar a IA?**

**Nick Bostrom —** Acho que nosso histórico de exercer sabedoria e controle global sobre tecnologia não é tão bom. E acho que o efeito geral da tecnologia até hoje foi amplamente benéfico. Mas não tanto porque tínhamos um grande projeto, e sim por sorte. Até agora, não surgiu uma tecnologia muito destrutiva e muito fácil de ser controlada por um indivíduo. Tivemos sorte com as armas nucleares que, embora causem muita destruição, sua produção é complicada. Não é possível fazer, no seu micro-ondas, uma arma termonuclear. São necessárias usinas imensas para isso, então apenas países podem produzi-las. Mas, antes da física de partículas, não tínhamos como saber o que aconteceria. Talvez tivéssemos descoberto uma forma fácil de destruir o mundo, e talvez o mundo pudesse ter sido destruído.

**Silio Boccanera — Acha que vamos precisar que leis e regulamentações sejam discutidas com muita antecedência enquanto a tecnologia avança?**

**Nick Bostrom —** Acho que hoje isso seria prematuro. Em termos de pesquisa e de IA superinteligente, acho que hoje seria muito prematuro tentar regulamentar isso ou envolver governos. Acho que há outras questões mais urgentes, como a preocupação com segurança, uso miliar de drones e similares. Agora é a hora de discutir como queremos usar essas coisas. Em relação às máquinas superinteligentes, primeiro precisamos entender exatamente qual é o problema e depois promover a pesquisa técnica, para que possamos entender melhor a questão.

**Silio Boccanera — Mas não existe a possibilidade de a tecnologia estar um passo adiante de nossa tentativa de controlá-la?**

**Nick Bostrom —** Sim, é isso que está criando a sensação de urgência entre alguns pesquisadores, que querem começar a discutir a questão do controle. E nós estamos começando a pesquisar aqui os desafios



---

de governança que podem surgir se e quando a humanidade iniciar a transição para a era da inteligência artificial. Mas esse ainda é um estágio anterior ao do problema do controle técnico.

**Silio Boccanera — Há muitos recursos sendo investidos nessa área. Eu li que, só no ano passado, US\$ 8,5 bilhões foram investidos em inteligência artificial. Onde ficam os principais centros de pesquisa nessa área? Ou todo mundo está pesquisando?**

**Nick Bostrom —** Parece que todo mundo está envolvido em IA, virou moda. Mas acho que há uma diferença entre aqueles que criam aplicações para a IA, que pegam técnicas existentes e as aplicam em algo comercial. Há muitos usos para isso em várias atividades econômicas. Desde softwares de mapeamento para que você chegue a seu destino, carros autodirigíveis, sistemas de recomendação de produtos e muitas outras coisas. E há um segmento menor disso que tenta ir mais longe, tenta fazer pesquisas básicas e novas com a intenção de desenvolver a inteligência artificial como projeto de longo prazo. A maior parte dessa pesquisa acontece hoje nos EUA. O Reino Unido e o Canadá são fortes nisso, há grupos na China e em outras partes do mundo.

**Silio Boccanera — Mas está avançando bem rápido.**

**Nick Bostrom —** Sim, e a indústria se envolveu mais. Há dez anos, só a academia fazia a pesquisa básica, mas agora também há grupos de pesquisa na indústria, que são cada vez mais responsáveis por artigos acadêmicos, publicações e apresentações em conferências.

**Date Created**

29/12/2016