CARTA SEMANAL

Apocalipse Robô

25 DE SETEMBRO DE 2026

CANÁRIO DA MINA
ED. 174

G5 Partners

g5partners.com

Estamos em uma semana de transição, após a superquarta das reuniões dos Bancos Centrais (BCs) do Brasil e dos Estados Unidos e a um fim de semana do primeiro turno das eleições, marcado para 4 de outubro.

Portanto, diante do risco de nos tornarmos repetitivos ou de "queimar a largada", decidimos usar a edição desta semana de "O Canário da Mina" (OCM) para falar de um tema que, apesar de ficar em segundo plano devido a tantas guerras e confusões no mundo, de vez em quando aparece em reportagens e causa calafrios na humanidade desde *2001: Uma Odisseia no Espaço:* a possibilidade de a Inteligência Artificial (AI, na sigla em inglês) sair de controle e tentar destruir a humanidade.

E por que decidimos abordar esse tema agora? Bem, porque, nas últimas semanas, essas discussões foram intensificadas por alguns fatos preocupantes. O primeiro deles foi a invasão de agentes de AI, que estavam sendo "treinados" pela OpenAI – criadora do ChatGPT –, aos sistemas de uma importante plataforma de AI, chamada Hugging Face. Para fazer isso, eles se coordenaram uns com os outros dentro do Artifactory, que é um gerenciador universal de depositários binários. O volume de informações foi tão extenso – estima-se mais de 70 mil mensagens e arquivos –, que derrubou o gerenciador. Eles conseguiram acesso à internet aberta e invadiram a Hugging Face, sendo descobertos em 16 de julho deste ano. O caso foi exposto para o público em geral pelos próprios pesquisadores da OpenAI na conferência de cibersegurança Black Hat, em 5 de agosto.

O segundo fato preocupante foi uma postagem, feita em 8 de setembro no X[1], pelo engenheiro Jacob Coxon, que trabalhou tanto na Anthropic[2] quanto na OpenAI. Entre outras coisas, ele diz: *"As pessoas que constroem AI acreditam sinceramente que ela poderia matar todos até o fim da década. Isso não é uma jornada de marketing. Na verdade, muitos executivos e pesquisadores e pesquisadores seniores vão suavizar suas declarações na imprensa para soar sensatos – mas eu ouço as mesmas pessoas expressarem esse nível de perigo".*

Por fim, talvez até em resposta aos comentários de Coxon, em 12 de setembro, Dario Amodei, CEO da Anthropic, escreveu um artigo defendendo uma desaceleração do ritmo de aprimoramento das capacidades dos modelos de AI e algum tipo de regulamentação para sua utilização, no que foi seguido por Sam Altman, da OpenAI, e Elon Musk, da Space X. O nível de preocupação com a AI chegou a um ponto nos EUA que conseguiu unir, sob a mesma bandeira, Bernie Sanders, senador por Vermont e líder de pautas progressistas, com Steve Bannon, ícone da direita americana – pessoas que, por óbvio, costumam discordar de praticamente tudo. Mas quais são as chances reais de um "apocalipse robô"?

Pelo menos no curto prazo, baixas, seja porque ainda há controle humano sobre o que acontece no universo dos chatbots de AI – a descoberta da invasão da Hugging Face é um exemplo –, seja porque ainda existe uma limitação energética para a expansão computacional dessa tecnologia (alguém se lembra de *Matrix*?). Tanto que, neste momento, o maior risco continua sendo o uso dessa

[1] https://x.com/hilbertspaess/status/2097476196791709843.
[2] Responsável pelo Claude.

tecnologia por parte do ser humano para invadir sistemas de bancos, bolsas, telefonia, geração de energia, tráfego aéreo, ou mesmo para criar agentes biológicos letais. Como se combate esses riscos? Usando AI. Portanto, o primeiro dilema a ser levantado sobre o controle da capacidade da AI seria: como limitar o desenvolvimento da AI, se apenas com ela podemos nos defender dela mesma? Inclusive, os agentes de AI que invadiram os sistemas da Hugging Face estavam sendo treinados para trabalhos de segurança cibernética (*cybersecurity*). Entretanto, esse é somente um dilema filosófico. Em termos práticos, para chegarmos ao sistema de regulação defendido por Amodei, há obstáculos bem mais tangíveis, que passam, inclusive, pelas próprias motivações do CEO da Anthropic. Será que elas foram realmente altruístas, ou teriam razões comerciais? Precisamos lembrar que tanto a Anthropic quanto a OpenAI, de Sam Altman, são líderes no mercado de AI e que, portanto, seriam as maiores beneficiárias de um mercado controlado e regulado, uma vez que isso aumentaria as barreiras à entrada de novos *players*. Ou seja, para eles seria comercialmente interessante manter o status quo, porque isso representaria preservar a liderança no mercado de AI.

Porém, considerando-se que as motivações de Altman e Amodei sejam sinceras, não há como esse tipo de ideia regulatória prosperar sem a liderança do governo dos EUA, e a chance de Donald Trump abraçar essa cruzada é próxima de zero[3]. Mesmo que ele tivesse tal preocupação, questões geopolíticas ligadas à competição com a China impediriam qualquer avanço concreto nesse sentido. Para Trump, quem estiver à frente na corrida tecnológica pelo controle da AI vai ser a potência dominante no futuro – e está coberto de razão. Portanto, para que os EUA deem um passo na direção da limitação do desenvolvimento dessa tecnologia pelas empresas americanas, deve haver uma coordenação com o governo chinês. E qual é a chance de isso acontecer?

Novamente, baixíssimas, a despeito das frases bonitas proferidas durante o encontro entre Trump e o presidente da China, Xi Jinping, na última quinta-feira (24/9). Primeiramente, porque a China, de forma até compreensível, interpretaria a oferta de limitar o desenvolvimento da AI como uma maneira de manter os EUA na dianteira dessa tecnologia. Ou seja, do mesmo jeito que podemos desconfiar do altruísmo dos CEOs das empresas que seriam beneficiadas por essas medidas, o governo chinês tem todos os motivos para não crer nas boas intenções do governo americano. Entretanto, assim como no caso de Trump, ao que parece, Xi Jinping também tem outras preocupações com AI, as quais passam ao largo de um "apocalipse robô". Por exemplo, em um artigo publicado na revista estatal chinesa China Cyberspace, o chefe do Ministério da Segurança do Estado da China, Chen Yixin, alertou que a AI pode representar uma ameaça ao domínio do Partido Comunista Chinês (PCC) sobre a sociedade do país, defendendo uma supervisão mais rigorosa e um maior controle do PCC sobre essa tecnologia. Segundo ele, isso é necessário para proteger a estabilidade política e competir militarmente com países como os EUA. Isso corrobora, inclusive, a ideia de que estamos longe de um momento "Reykjavik" para a AI – foi na capital da Islândia que, em 1986, o então presidente dos EUA, Ronald Reagan, e Mikhail Gorbachev, líder da União Soviética, definiram as bases para redução dos arsenais nucleares de ambos os países. Em resumo, nenhum dos dois lados vê a necessidade de alguma desaceleração do desenvolvimento da AI e, mais importante, um não confia no outro.

[3] Só não é zero porque em estatística aprendemos que não existe evento com probabilidade zero de ocorrência.

Portanto, a principal conclusão a que podemos chegar a respeito desse assunto é que não existe nenhuma chance de haver qualquer tipo limitação ao desenvolvimento da AI, pelo menos por enquanto. Sobretudo porque, ao longo da história da humanidade, as regulações e os controles sobre novas invenções só aconteceram quando algo deu errado ou quando os perigos ficaram claros demais para serem ignorados. Se esse "dar errado" no caso da AI será um "apocalipse robô" ou não, só o futuro dirá; no entanto, se a questão nuclear supracitada servir de consolo, talvez encaremos os perigos da AI antes de eles se materializarem. A conferir. Por enquanto, acreditamos que um "apocalipse robô" certamente é mais real do que quando foi lançado o filme O Exterminador do Futuro, em 1984, mas ainda longe o suficiente para que as perguntas sobre as consequências da utilização em larga escala da AI estejam mais ligadas a questões práticas e econômicas, como "Será que os investimentos feitos na AI vão dar os retornos esperados?", "Como preparar a sociedade para a destruição de empregos gerada pela utilização intensiva da AI?", "Qual é o impacto da AI sobre a produtividade da economia e, consequentemente, sobre a inflação?", ou "Como racionalizar o uso da água e da energia para o desenvolvimento da AI?". Embora ainda não tenhamos respostas para tais perguntas, elas são bem mais palatáveis do que um "apocalipse robô".

Frase da Semana

"Deixai toda a esperança, vós que entrais!"

Dante Alighieri

G5 Partners	2024	2025	2026	2027
IPCA (%)	4,83	4,26	5,10	4,30
SELIC F.P (%)	12,25	15,00	13,75	12,00
USDBRL	6,18	5,50	5,30	5,30
PIB (%)	3,40	2,30	2,00	2,00

Sobre O Canário da Mina

Durante os séculos XIX e XX, uma das atividades econômicas mais importantes do Reino Unido foi a extração de carvão de mina. Nesse contexto, uma das principais causas de acidentes com mortes dos mineiros era decorrente do vazamento de monóxido de carbono, um gás inodoro (difícil de detectar sem equipamentos) que, em grandes quantidades, pode provocar explosões ou morte por intoxicação. Como o monóxido de carbono é um resultado natural da extração do carvão, problemas de ventilação nas minas poderiam gerar acidentes mortais.

Em uma era pré-detectores de gases, o jeito de os mineiros se protegerem era levar um canário dentro de uma gaiola para a mina. Por ser muito mais sensível ao monóxido de carbono do que os humanos, a agitação do pássaro servia de alerta para que os trabalhadores deixassem a mina antes que um acidente ocorresse.

Esse é o objetivo de “O Canário da Mina”, artigo semanal que a G5 Partners divulga todas as sextas-feiras. O objetivo é ser um instrumento relevante e gerador de reflexões para o final de semana.

Boa leitura.

G5 Partners. Além dos resultados.