Azeem Azhar: quebras de códigos e outros problemas de IA

As ameaças continuam chegando

Paul Krugman - 31 de outubro de 2026

TRANSCRIÇÃO:
Paul Krugman em Conversa com Azeem Azhar

(gravado em 29/10/26)

Paul Krugman: Olá a todos. Hoje estou conversando com Azeem Azhar, escritor, pesquisador, investidor da Exponential View, que é uma das pessoas a quem recorro quando não entendo o que está acontecendo com a IA – algo que acontece praticamente o tempo todo. Houve um artigo realmente marcante que a Azeem escreveu outro dia sobre o "hack PHP" com o título "O momento Wuhan."

Por que você não me explica o que quer dizer com isso? Vamos tentar entender o que está acontecendo agora com a IA.

Azhar: Sim. Bem, aquele "hack do PHP" foi apenas um exemplo particular de uma série de coisas que estão acontecendo agora porque os modelos de IA se tornaram muito bons. E o motivo pelo qual chamei de "momento Wuhan" é que há doenças infecciosas o tempo todo. Agora surgiu uma doença com características um pouco diferentes e ela teve um impacto muito maior do que todas as outras que circulavam.

Então, digo que modelos de IA ficaram mais capazes — a experiência com o uso do ChatGPT mostrou que ele ficando cada vez mais inteligente. Os modelos de IA ficaram bons o suficiente para serem usados na construção de códigos, mas também para serem direcionados a encontrar vulnerabilidades, fraquezas e bugs em software.

E como sistemas de IA são incansáveis, é simplesmente possível fazê-los trabalhar sem que se esgotem. Pois, eles não são como pesquisadores de segurança humanos ou hackers humanos. Eles conseguem encontrar todas essas vulnerabilidades, muitas vulnerabilidades, rapidamente. Podem pesquisá-las de maneiras novas, não humanas, conectando-as umas às outras. E esse é um novo regime em que nos encontramos. E parece diferente de onde estávamos há um ou dois anos, quando essas fraquezas e vulnerabilidades eram descobertas metodicamente por hackers ocasionais, mas, mais comumente, por equipes de segurança. Então esse é o momento.

Krugman: Então, veja o que entendi: basicamente o mundo funciona hoje é que há software por toda parte. Tudo está rodando em softwares que são complicados, mas talvez que não estejam cuidadosamente montados. Me diga se eu estiver errado nisso, mas acho que a coisas está complicada pois se tem programas de computador mais novos sobrepostos a programas mais antigos... E só Deus sabe onde estão as vulnerabilidades.

Azhar: Bem, programas de computador e sua tecnologia é uma indústria realmente grande. E, de alguma forma, conseguiu escapar por 50 ou 60 anos de testes de segurança mais refinados.

Krugman: Certo.

Azhar: E então, sabe, é por isso que existem versões beta e correções de bugs. A cultura do desenvolvimento de software manteve por muito tempo uma mentalidade de ir mexendo no software conforme a necessidade. E não há um modelo realmente seguro de construção de programas. Não é um problema de IA, não é um problema de redes sociais ou internet.

Era um problema nos anos 70 e 80. Passamos de mainframes centralizados para computadores pessoais de tal modo que surgiram muitas, muitas oportunidades para que softwares fossem desenvolvidos de forma descentralizada. E aqui estamos, décadas depois! É um milagre ver a internet funcionar, dado o número de bugs e problemas.

Mas isso também tem sido um benefício, porque ela, por não ser uma catedral fixa, é bem ágil. Um desenvolvedor de software comparou isso a um bazar lotado e bem movimentado. Ele tem a capacidade de responder, de se adaptar e de se manter bastante resiliente. Foi isso o que aconteceu com a internet em grande parte, exceto por alguns grandes ataques cibernéticos nos últimos 25 anos. Então, aquilo que se mostrou como uma característica é também uma falha reconhecida pelos engenheiros de software. Que falha: não haver um modelo de segurança acordado para criar software. Ora, essa também foi a razão pela qual tivemos um florescimento do software. Conseguimos continuar caminhando mesmo quando bugs são encontrados e problemas são descobertos.

Krugman: Certo. E muito disso é baseado em PHP. Você provavelmente deveria contar para as pessoas o que é. Quer dizer, na verdade, eu nunca tinha ouvido falar até semana passada, mas claro...

Azhar: Bem, você certamente já usou. Na realidade, 70% dos sites rodam em PHP. É uma linguagem de programação muito fácil de usar e, por isso, bem barata também. Se no final dos anos 90 você queria escrever um software, usava PHP. É simplista; não tem a curva de aprendizado do C++ ou algo assim. E esse uso se estendeu cada vez mais.

O Facebook usou muito, e assim o PHP se tornou parte da arquitetura da internet. E como muitos outros sistemas que usamos, como Python, que é outra linguagem, ou BIND, que é um software que transforma endereços de internet amigáveis como Substack.com nos números que os roteadores precisam. Esses são softwares que também são open source, sabe, são escritos por colaborações de voluntários que têm seus próprios sistemas de governança, que compartilham esse código em repositórios. E parece um milagre que funcione, mas temos evidências e provas de que funciona.

Krugman: Ok. Quão relevante é o fato de que é antigo? O PHP é coisa dos anos 90. É tudo feito assim. O que tínhamos era bom o suficiente, mas agora termos esses modelos de IA incansáveis, certo? Temos agora, portanto, um exército de especialistas em descobrir falhas, certo?

Azhar: Certo, sim. Quero dizer, toda a configuração não estava montada exatamente em pernas de pau, porque tudo funcionava. Também é preciso reconhecer que é realmente difícil escrever um software extremamente preciso e

infalível em uma empresa — seja em ambiente corporativo, consumidor ou cliente — porque os padrões mudam.

Se você quer entender o motivo, pense na experiência de voar num avião antigo que precisa atravessar o Atlântico. Você abre o entretenimento a bordo e vê que o software é muito lento quando comparado ao seu smartphone moderno. Eis que ele foi projetado e bloqueado há 18 anos. Há 10 anos antes a companhia aérea o comprou e ele ainda está em vigor.

Então todas essas coisas realmente fizeram parte do trade-off que nós, como clientes e consumidores, tivermos de fazer para ter algo que se adaptasse e respondesse às nossas demandas. Nenhum de nós realmente quer usar o tipo de software que opera uma usina nuclear, certo? Porque ele está preso nos anos 70 ou 80. E isso significa que existem muitas maneiras diferentes pelas quais esses códigos podem travar. E quando eles falham dão oportunidade para que se possa descobrir porque eles falharam. E é isso que grandes hackers fazem.

Modelos de IA ultrapassaram um limiar no início deste ano, quando a Anthropic lançou um modelo chamado Claude Mythos. É o melhor modelo de cibersegurança do mundo. Não foi amplamente disponibilizado. Na verdade, apenas um pequeno número de organizações tem acesso a ele. E eles têm corrido por aí tentando corrigir ou corrigir esses bugs e fechar brechas. A primeira parte é encontrá-los. Mas uma das coisas que acontece com a IA é que as capacidades deslizam para modelos mais baratos muito, muito rápido.

E assim, em algum momento deste verão, os modelos abertos, também conhecidos como modelos open-source, começaram a se aproximar das capacidades que o Mythos tinha. E assim, eles podem agora ser direcionados por hackers ou agentes mal-intencionados para encontrar vulnerabilidades antes que os mocinhos as encontrem com o fim de tentar explorá-las. E vimos um aumento, um aumento rápido no número dessas vulnerabilidades que estão sendo descobertas. E esse do PHP foi um exemplo disso.

E acho que o que eu escrevia no meu ensaio era o seguinte: olha, o que sabemos sobre essas tecnologias é que elas ficam mais baratas; e que, à medida que ficam mais baratas, são usadas muito, muito mais, certo? Todos nós tiramos muito mais fotos digitais do que fotos em filme hoje em dia de 40 anos atrás. Provavelmente a mesma coisa acontece na esfera dos programas de computador.

E isso significa que, se houver vulnerabilidades, por mais obscuras que sejam, é bem possível que eles sejam encontrados e explorados. Vou mencionar apenas um dado que eu vi. Houve um software específico que foi testado para vulnerabilidades trilhões de vezes por humanos rodando suítes de teste. E ainda assim, um modelo recente de IA conseguiu, apesar disso, encontrar vulnerabilidades graves que poderiam ser exploradas. Então, mesmo que tenhamos pesquisado muito, nunca podemos garantir que essas coisas sejam 100% robustas.

Krugman: E isso parece uma corrida armamentista constante. Agora você vai usar IA para tentar corrigir ou proteger, mas parece que a vantagem inerente vai ser com os vilões, certo? Vai ser sempre difícil, ou estou errado?

Azhar: Bem, acho que os vilões se movem mais rápido do que todo mundo em geral. Lembre-se, os bandidos só precisam encontrar uma vulnerabilidade em um só lugar. Os mocinhos têm que defender todas as vulnerabilidades o tempo todo. Então, acho justo dizer que há muitas organizações do setor privado e governamentais que estão à frente, mas também há bilhões de computadores conectados à internet e trilhões e trilhões de linhas de código e sistemas que não são atualizados há anos e que podem conter vulnerabilidades. E é por isso que o Vice-Diretor de Segurança da Informação da Anthropic disse que serão 18 meses bem difíceis.

E acho que o que aludi no meu ensaio foi que há aquele período em que, assim como foi com a disseminação do que chamávamos, acredite ou não, 2019-nCoV antes de chamarmos SARS-CoV-2, as coisas podem se espalhar mais rápido do que conseguimos descobrir o que podemos fazer a respeito. E acho que esse é o risco que surge nesse período mais curto.

Krugman: O que fizemos para lidar com a disseminação do coronavírus, algo que tinha consequências políticas graves, foram os fechamentos (lockdowns). Você sugeriu que veríamos muita desconexão parcial na web aberta. Então, podemos fazer isso hoje, permanecer em jardins murados?

Azhar: Bem, pode-se argumentar, Paul, acho que muitos de nós passamos muito tempo em jardins murados. Se você mora no Instagram ou no Facebook, não está vivendo na web aberta. Não é a web aberta pela qual me apaixonei em 1992, onde você vai de um lugar para outro. E sabe, há muitos sítios que exigem autenticação por senha para acessar.

Mas acho que é de se esperar que o número de agentes de IA vai rapidamente ultrapassar o número de humanos. Acho que a quantidade de navegação feita por bots na internet hoje em dia é maior do que a dos humanos. Sempre foi alta porque os mecanismos de busca precisam rastrear, e as pessoas que compram coisas no eBay usam bots de precisão de preços para monitorar os anúncios. Mas isso vai se inclinar enormemente para as máquinas nos próximos anos. A internet e as empresas na internet vão evoluir para atender a essa demanda.

Pode-se fazer uma comparação. Na Quinta Avenida, em 1895, havia mais pessoas passeando do que automóveis mesmo nas pistas. Mas em 1920, as pessoas não andavam mais no meio da rua. Então acho que a internet precisaria começar a filtrar o que esses sistemas de IA precisam. Eles se movem muito rápido. Eles são gatilhos rápidos. Eles funcionam em milissegundos. Nós trabalhamos em segundos. Eles têm requisitos técnicos completamente diferentes. E isso significa que isso ocuparia um trecho diferente do ciberespaço em que queremos estar.

E, em particular, se quisermos estar realmente, realmente seguros e robustos diante das ameaças, podemos escolher: "Quero ficar em um enclave que seja seguro." E acho que um bom exemplo disso é o que a Apple faz com a Apple Store. Em alguns países, você pode usar lojas de terceiros — mas na maioria das vezes você precisa pegar aplicativos da Apple Store, porque a Apple quer dizer: "Verificamos se isso é seguro."

Enquanto no Android, você pode ir a qualquer lugar. E é por isso que há muito mais vírus e problemas de segurança no Android. Então, vejo essa possível jornada em

que a internet precisa atender às necessidades econômicas das empresas que usam agentes, o que significa quão rápido é, seja transações, pagamento conforme o uso, validação e assim por diante. Mas pode não ser assim que vamos querer operar na internet. Vai parecer que não é para nós, da mesma forma que a rodovia 101 na Califórnia não é para os humanos caminharem.

(A sequência não foi revisada minimamente)

Krugman: Então vai ser um mundo em que, em vez de poder deixar um bot ir e fazer compras, você terá que provar isso, ou você sabe, oferecer alguma evidência de que você é você, ou só certas coisas estarão disponíveis; só que um ambiente muito mais instável, é isso?

Azhar: Acho que você vai acabar tendo que validar seu bot, dar sua aprovação ao bot, e o bot terá um conjunto de regras. Ele pode gastar $50 sem te enviar o link de aprovação ou seja lá o que for o sistema. E a verdade é que, acho que não sabemos nada agora como isso seria. Quero dizer, as pessoas estão experimentando, e minha impressão é que as editoras em particular, e também os varejistas, vão começar a publicar listagens para consumo por agentes de máquinas da forma que precisam ver o material, que é muito mais estruturado, muito menos emotivo, só com mais dados que eles podem navegar. E eles farão isso porque você ou eu provavelmente usaremos um agente no nosso iPhone ou um agente da Meta para interagir com onde sabemos que estamos comprando ou navegando. Quer dizer, realmente acho que ainda não chegamos lá, mas acho que as pessoas da indústria geralmente acham que essa é a direção que seguimos.

E eu também tenho um pouco de experiência porque tenho um bot que acompanha e fica atento aos shows que eu possa querer ir, porque não gosto de ir aos sites de locais porque todos são diferentes uns dos outros. Então o bot conhece meus gostos musicais, tem meu catálogo musical, e verifica toda semana mais ou menos se tem algum ingresso que eu queira comprar e me avisa sobre isso. Então, não vou mais visitar esses sites de locais, mas vou pagar pelo ingresso.

Krugman: Lembro como fiquei assustado quando fui um dos primeiros compradores de livros na Amazon, ou seja, um comprador físico, e quando a Amazon começou a recomendar músicas para mim. Na verdade, nunca comprei música deles, mas as recomendações eram bem boas mesmo naquela época. E foi meio chocante perceber que, em algum nível, pelos livros que comprei, eles descobriram que música eu provavelmente gostava. E é um mundo muito assustador para alguém da minha idade.

Azhar: Acho assustador; é assustador para todos nós quando você recebe essas sugestões. Quer dizer, provavelmente dá para fazer isso no ChatGPT. Se você tirar uma foto da minha estante e colocar no ChatGPT e dissesse, sabe, "Recomende alguns livros que esse cara possa gostar", acho que vai fazer um bom trabalho. Lembro a mim mesmo que é matemática, é isso que é. É como atravessar uma enorme matriz e encontrar algumas probabilidades.

Krugman: Ninguém ouvindo isso vai saber do que estou falando, mas me choca que basicamente a multiplicação de matrizes com algumas não linearidades pode fazer tudo isso. É absolutamente bizarro que funcione.

Azhar: É tão estranho. É tão estranho.

Krugman: Ok. Mas isso é um pouco menos apocalíptico do que o que entendi ao ler seu texto. Mas acho que ainda pode ser muito, muito assustador. Você acha que eventualmente teremos encontrado uma forma razoavelmente fluida de fazer funcionar?

Azhar: Olha, o que é um apocalipse? Se olharmos para o que aconteceu com a COVID, a COVID poderia ter sido muito, muito pior? Sim. A COVID foi realmente ruim? Sim. Foi apocalíptica e horrível para muitas, muitas pessoas, individualmente e também para grupos inteiros? Sim. Quero dizer, imagine os jovens que fizeram a faculdade durante o lockdown, sozinhos, longe de casa pela primeira vez. Quer dizer, isso é simplesmente horrível. Então acho que esses são resultados ruins. Acho que o que estamos enfrentando aqui é que parece que em 18 a 20 meses as pessoas já terão percebido isso, e esses caras não vão dormir muito, mas enquanto isso coisas realmente desconfortáveis podem acontecer. Você pode ter um banco que simplesmente não se antecipe, e seu dinheiro em um banco é só uma entrada em um banco de dados. Não está nas mãos do Jimmy Stewart, certo? E o dinheiro é literalmente só uma linha em um banco de dados financeiro da Oracle. E isso seria bem apocalíptico para as pessoas.

E acho que não podemos descartar o que os estatísticos chamam de riscos de cauda, mas não conseguimos ver esses riscos de cauda, certo? Eles são os "desconhecidos desconhecidos" de Donald Rumsfeld. E o que sabemos é que a internet, com todos os bilhões de máquinas conectadas, agora está muito mais porosa e muito mais vulnerável do que era há dois ou três meses. E isso significa que algo bem ruim pode acontecer? Muitas coisas como redes elétricas às vezes estão conectadas ou acessíveis pela internet. Bem, agora é mais provável do que há um ano por causa disso. E é aí que precisamos ser muito, muito proativos.

Conversei com formuladores de políticas e, há alguns meses, fui lá e disse: olha, vocês precisam se unir e ter uma parceria público-privada onde o Estado incentive ativamente empresas pequenas, médias e grandes que não são empresas de tecnologia a fazerem as perguntas certas, para garantir que essas coisas estejam sendo corrigidas proativamente, porque haverá um momento em que você atravessar esse vale estranho, e as vulnerabilidades vão aumentar enormemente. Então, quero dizer, isso precisa ser levado muito, muito a sério.

Se eu não parecer que estou correndo por aí em pânico, é porque eu modelo coisas importantes em Sully Sullenberger, o piloto daquele avião que pousou no Hudson. Sabe, ele tinha 115 almas naquele avião e não entrou em pânico. E tem uma ótima entrevista para a mídia com Katie Couric onde ela fica perguntando como ele estava se sentindo e se ele estava entrando em pânico? E ele é frio como um pepino porque, em momentos como este, quando as coisas importam, queremos usar nossas evidências, queremos usar dados, queremos usar nosso julgamento, queremos pensar com calma. Então acho que é sério. As pessoas devem agir rápido.

Krugman: Ok, mas quantas pessoas em posições de liderança importantes no mundo te lembram Sully Sullenberger?

Azhar: Bem, quero dizer, a democracia é o melhor sistema que temos. Infelizmente, é isso que é.

Krugman: Eu sei. Ou como Churchill disse, "O pior sistema, exceto todos os outros." Mas enfim, sim.

Ok. Antes de eu ler sobre o apocalipse hack ou a possibilidade disso, a grande pergunta que todos nós temos feito é sobre uma bolha de IA. Para muitos de nós, nossa experiência formativa nisso foi a bolha de TI dos anos 90, que cada vez mais estou começando a achar que não é um modelo muito bom para onde estamos agora. Mas você tem sido, sabe, não um entusiasta de cara louca, mas basicamente tem dito: "Calma. Há muitas evidências de que essa bolha pelo menos não é tão ruim assim." Quer falar sobre por que está se posicionando do lado mais calmo aqui?

Azhar: Sim, ainda não é uma bolha tão ruim, talvez seja a melhor forma de dizer. Uma coisa que notei no debate nos últimos anos é que ele se divide entre discussões descritivas — os fatos do mundo e a discussão normativa — e sentimentos: os *deveres* do mundo. E às vezes as pessoas acham muito difícil separar isso. E acho que o debate fica extremamente acalorado, muito apaixonado, mas provavelmente não muito útil. E eu diria que tudo o que eu digo realmente é descritivo neste ponto. Não estou dizendo que deveria acabar de um jeito porque o mundo seria melhor ou de outro, certo? Então acho que isso é bastante importante.

Bolhas e essa expansão excessiva de investimentos em qualquer tecnologia são realmente, muito comuns, como você sabe pelo seu próprio trabalho. A verdadeira dor surge quando os clientes não compram o suficiente para pagar o financiamento do que for necessário para construir as fábricas, a infraestrutura. E esse descompasso é o que cria bolhas. É o que criou as bolhas ferroviárias nos EUA. É o que, de certa forma, estava acontecendo com as pontocom e com as telecomunicações. Nem toda onda de investimento em uma nova tecnologia termina em bolha desastrosa. A maioria termina, e na maioria das vezes acontece porque o financiamento fica exótico demais, se torna circular demais, grande demais. Mas nem todo mundo acaba. Automóveis nos EUA, por exemplo, não acabaram em bolha. Eles acabaram com muitas empresas falindo e havendo consolidação, mas empresas geracionais também saíram disso.

Então, quando olho para o que está acontecendo com a IA, a pergunta que faço — e olhamos os dados — é: qual é o peso, o lastro? Como a complexidade da dívida e do financiamento que está sendo construído contrasta com a forma como os clientes, principalmente as empresas, estão gastando? E um pesa mais do que o outro? E acho que muito disso pode ser medido. Muito disso pode ser movido por fatos, não por opinião. Parte disso é desconhecido, e é por isso que estou calmo. Tenho uma visão que, para simplificar, quando comecei no trabalho, minha primeira hipoteca era de 85.000 libras, ou seja, uns 110 mil dólares, isso foi há 30 anos, e foi muito difícil para mim pagar. Minha hipoteca hoje é substancialmente maior do que isso, e é bem fácil para mim pagar. E se eu pegasse a hipoteca de hoje e me desse aos 24 anos, eu teria ficado esmagado pelo peso disso.

Então, a escala da dívida, as obrigações financeiras, precisam ser colocadas em proporção a como os clientes estão gastando e como podem gastar. E é por isso que estou um pouco mais calmo, mais otimista agora, embora eu ache que a

métrica da bolha está maior hoje do que era há um ano, sabe, pelos motivos que tenho certeza que vamos explorar.

Krugman: Sim. Primeiro houve um debate sobre o quão produtiva essa tecnologia é. Obviamente é incrível. Ainda é difícil acreditar que basicamente a aritmética matricial possa soar tanto como uma pessoa, mas ainda há algum argumento. Ainda há pessoas dizendo que, ok, isso é incrível, mas não conseguimos ver isso nos números agregados de produtividade. Você tem abordado isso de alguma forma?

Azhar: É muito difícil ver isso em números agregados. Você é um economista muito maior do que eu, então não sei quando esperaria ver isso em números agregados. E ouvi um economista na China — estive na Ásia algumas semanas atrás — dizendo que estão começando a ver alguns números em alguns setores. Ainda acho muito justo dizer que eles não aparecem em números agregados. Eles aparecem em pequenos números de empresas fora da indústria de tecnologia. Dependendo da pesquisa que você acredita, entre 5 e 10% das grandes empresas estão vendo talvez um benefício de 6%, aumento de vendas, custos menores.

Krugman: Sim.

Azhar: E isso pode aumentar. O que não sabemos é se os outros 90% vão se transformar nesses 10% em um ano, dois anos, cinco anos, 10 anos, ou algum dia. E podem não aumentar. E não sabemos qual porcentagem desses 10% vai receber cada vez mais benefícios e, portanto, gastar cada vez mais. Mas acho que seria— com algo assim, é um pouco como um médico tentando entender uma doença complexa. Você não só pega o pulso do paciente. Você faz muitos exames e exames de sangue e tenta pintar um quadro.

Então, certamente, não há muita coisa nas estatísticas agregadas, certamente não nos EUA. Há algumas evidências do setor, há algumas evidências de pesquisas de empresas. Eu converso bastante com executivos. O que os executivos me dizem é que isso é mais difícil do que pensavam. Mais deles estão tendo sucesso. Todos pretendem continuar gastando no próximo ano para tentar mais. Quando isso aparece em um número que um economista como você possa dizer: "Aqui está um número"? Quero dizer, acho que essa é uma pergunta realmente importante. Se demorar muito, provavelmente é o caso de que as receitas não vão crescer rápido o suficiente para pagar a hipoteca.

Krugman: Ok, voltando às receitas em um segundo. Onde eu olharia para o boom de TI dos anos 90, sabe, em 2005 estávamos falando sobre produtividade total dos fatores. No total, estávamos algo acima de 10% do que a tendência anterior esperava, o que é muito, embora tenha meio que paralisado depois disso. Mas estamos, eu diria que, em termos de timing, estamos algo como em 1999.

Azhar: Hum-hum.

Krugman: Os benefícios de produtividade foram muito mais evidentes nos primeiros quatro ou cinco anos dentro da bolha da internet. Então não sei o que pensar disso.

Azhar: Sim, minha lembrança é que foi ainda mais cedo com TI. Achei que o aumento de velocidade real foi naquele período de 94 a 98. Aconteceu de coincidir,

acho, com Bill Clinton como presidente. Então ele conseguiu fazer todo tipo de coisa com o orçamento e tal. Foi bem rápido e sustentado. E a questão é: onde estamos com essa coisa do ChatGPT? Acho que o tempo deveria começar provavelmente um ano depois do ChatGPT, que é quando Microsoft e Google lançaram versões corporativas e corporativas do que estavam fazendo. Então já foram três anos. E a questão é: quão rápido esperaríamos que isso avançasse, e se TI é uma comparação justa? E com TI, as pessoas tiveram que comprar computadores, por exemplo, o que é um custo muito maior do que pagar uma assinatura do ChatGPT.

Então, acho que há um argumento razoável para dizer que isso ainda não é comprovadamente mais rápido que TI e pode muito bem ser mais lento, certo? Acho que essa seria uma posição razoável para se manter hoje. E tenho algumas ideias do porquê disso. Como, qual é a causa da inércia dentro das empresas, e isso significa que [o retorno sobre o investimento] não é tão rápido quanto esperávamos ou pensávamos?

Krugman: Mas a segunda pergunta é: suponha que isso faça grandes coisas para o crescimento agregado da produtividade. Realmente é ótimo para a economia. A questão tem que ser quanto disso é capturado pelas empresas, certo? No fim, a internet era ótima, mas você pode olhar aqueles antigos anúncios da Qwest dos anos 90 sobre esse novo mundo maravilhoso da internet banda larga, e tudo nos anúncios se tornou realidade, mas a Qwest não existe mais, certo? Então, quanto devemos nos preocupar que a Anthropic, a OpenAI e talvez alguns outros estejam ocupados fazendo grandes investimentos e produzindo coisas ótimas, mas que todo mundo acabe comprando produtos chineses que são 90% tão eficazes quanto 90% e 90% mais baratos?

Azhar: Bem, acho que precisamos separar empresas como OpenAI e Anthropic das empresas que realmente pagam pelos investimentos de capital nos data centers, que são as empresas de infraestrutura. E essas empresas vão ganhar dinheiro, seja a Anthropic vendendo para você ou usando um modelo chinês, porque o modelo ainda precisa rodar em um computador.

Krugman: Ok.

Azhar: E esse computador é uma GPU provavelmente pertencente à Microsoft ou Amazon. Então eles ainda vão ganhar dinheiro na camada de infraestrutura. Então tentamos separar quanto dinheiro vai para a camada de infraestrutura de quanto vai para as empresas de laboratório modelo como OpenAI e Anthropic. Mesmo que as pessoas migrem para esses modelos mais baratos, suspeito que haja fluxo de capital e negócios suficientes, pelo menos pelos próximos anos.

No entanto, o que ouvimos constantemente da indústria é que há uma enorme subcapacidade de computação. Pouco antes de gravarmos isso, Microsoft e Google fizeram seus anúncios de resultados e disseram a mesma coisa: simplesmente não temos capacidade computacional suficiente para atender todos os nossos clientes. Então separo OpenAI e Anthropic. E acho que quando olho para a Anthropic, a empresa publicou um documento chamado S-1, que é o documento de marketing antes de abrir capital. E analisamos e desconstruímos isso de forma forense. Fiquei bastante surpreso com coisas que vi. Por exemplo, margens são

uma parte muito importante para saber se um negócio pode sobreviver ou não. E a margem bruta da Anthropic pelos custos de computação que paga é de cerca de 85%. Esse número era cerca de menos 200% há dois anos, quando muita gente dizia que a empresa ia morrer em um ou dois anos. (E essas pessoas não se corrigiram, aliás.)

Então acho que é uma surpresa que as margens sejam tão fortes assim, que eles estejam lucrando dessa forma. Claro, seu ponto permanece. Pode haver deflação ou competição de preços de modelos mais baratos, mas no fim das contas, o projeto de infraestrutura está principalmente nas mãos das empresas de infraestrutura, e elas ainda vão ganhar dinheiro servindo modelos de peso aberto.

Krugman: Então o sistema de peso aberto—sou vergonhosamente fraco exatamente sobre o que estamos falando aqui, mas tudo bem, eu acho. Embora os modelos de peso aberto, pelo menos pelo que entendo, basicamente fazem muito menos computação do que o molho completo, mas mesmo assim a demanda por computação ainda está disparando.

Azhar: Então, o motivo pelo qual modelos open-weight são mais baratos para nós é que eles são efetivamente disponibilizados para qualquer um baixar, como softwares open-source. Então você não está pagando margem de lucro para a OpenAI ou para a Anthropic. Você está pagando apenas a margem que a empresa de infraestrutura vai colocar nesse modelo para atendê-lo. Você desintermediou uma camada de margem. Quanto a se eles realmente são mais baratos em tarefas, essa é uma questão muito disputada, onde podem ser mais baratos pelo token, mas você precisa de mais tokens para realizar uma tarefa, então eles acabam sendo mais caros.

E minha visão, novamente, é um pouco otimista sobre isso. Tenho quase certeza de que muitas cargas de trabalho serão transferidas para modelos de peso aberto. Mas também acho que Anthropic e OpenAI se assemelham às empresas de software do passado: havia alternativas open-source mais baratas ao Microsoft Office, como o LibreOffice, mas ninguém usava. A maioria das empresas simplesmente foi para o negócio de marca, e assim também. E acho que existe um modelo mental simples que diz que, mesmo que percam muita participação de mercado, ainda podem persistir como negócios porque haverá as razões típicas pelas quais as pessoas comprariam software para comprar seus produtos. Acho que a grande questão é: quanto de mercado elas perderiam e o que acontece com seu poder de precificação? E a verdade é que não acho que alguém possa esperar por mais de alguns meses para ter alguma certeza. Por isso volto a todos os dados que vejo e vejo todos os dados que podem contradizer os argumentos que estou fazendo, só para garantir que estou me calibrando.

Krugman: Mas essa é a ideia, que o ChatGPT e o Claude vão acabar sendo o Microsoft Words do futuro. Basicamente, produtos inferiores que todo mundo usa porque ninguém se complica por comprar.

Azhar: Bem, sim, pode ser isso, mas acho que ainda são produtos melhores. Na minha equipe, usamos muitos modelos de peso aberto para o trabalho que é feito porque muitas vezes é bom o suficiente, rápido o bastante, barato o bastante. Curiosamente, eu uso especialmente um modelo de peso aberto para um dos meus

agentes porque o modelo não é tão inteligente em máquinas quanto os modelos de última geração. Então ele faz o que peço com mais frequência. Se eu trocasse por um modelo caro e de última geração do ChatGPT, ele muitas vezes faria mais do que eu pedia, certo? E eu diria: "Não, mas não era isso que eu queria. Eu só queria que fosse feito o simples aquilo. Eu queria que você me trouxesse ovos. Eu não queria que você me fizesse um bolo inteiro." E então há todas essas coisas, Paul, que simplesmente não entendemos agora. Temos que sentir nosso caminho, e por cima disso tem uns 5 trilhões de dólares em cima.

Krugman: Sim, isso mesmo. Quer dizer, há alguns lugares aparentemente tentando usar IA para controle de tráfego aéreo, e é como se você pudesse receber a resposta: "Desculpe, você estava certo, já havia um avião parado na pista. Gostaria de algum conselho sobre como lidar com eventos com vítimas em massa?"

Azhar: Certo. Bem, felizmente, ainda não chegamos a isso em nível de massa, mas claro que houve histórias horríveis de pessoas que foram derrubadas por psicose de IA, até o fim da vida, o que, claro, é um desfecho horrível. Mas acho que há um risco nesse debate sobre uma bolha ou um boom, e isso tende a apresentar uma dicotonomia. E acho que há outros desfechos. Há um resultado em que não há uma bolha que estoura, mas um período de retornos muito, muito baixos para as empresas que fizeram esses investimentos. Então não é bom para os investidores deles. Mas você não vê esse tipo de paralisação catastrófica como vimos no mercado imobiliário dos EUA, quando o início do mercado imobiliário parou completamente após a crise financeira global.

Uma das razões pelas quais acho importante pensar em outros cenários além da catástrofe financeira é que, se empresas americanas e europeias começarem a obter resultados e pagarem o suficiente como clientes desses negócios para que possam pagar essa hipoteca metafórica que contraíram, onde está o peso disso? O peso disso provavelmente recai sobre a força de trabalho. E poderíamos nos cegar — ficando tão empolgados com os terminais da Bloomberg ficando vermelhos à medida que os preços das ações caem — para começar a tomar medidas proativas sobre a possibilidade de que a tecnologia realmente funcione, proporcione maior produtividade, que as empresas, como resultado, reduzam custos, e parte disso recai sobre a força de trabalho. E acho que isso é importante demais para deixar para os gritos sensacionalistas sobre calamidades financeiras.

Krugman: Sim, estou principalmente preocupado com o quão intensivo em capital isso é. Definitivamente há uma competição por capital e esses data centers basicamente não empregam pessoas. Você pega oito bilhões de dólares que poderiam estar criando muitos empregos e acaba com cem trabalhadores de manutenção. Mas deixe-me perguntar, já que uma das coisas que tenho ouvido, e não tenho como avaliar, é que presumivelmente haverá muitos investimentos que podem causar prejuízo. Sabe, haverá investimentos que as empresas vão lamentar de ter feito ou a empresa deixa de existir, o que pode ser aceitável. Tivemos isso no boom de TI dos anos 90 porque a infraestrutura não existia. Mas depois estava. Então tivemos muita fibra que ficou escura por anos, mas que depois toda voltou a ser útil. Mas a história que algumas pessoas estão me contando, e não faço ideia se é verdade, é que esses investimentos vão se desvalorizar muito rápido. Que eles vão se desvalorizar financeiramente porque os chips ficam ultrapassados, mas

talvez ainda mais porque literalmente, fisicamente, se deterioram. Isso é uma preocupação?

Azhar: Este é outro exemplo em que temos dados para nos ajudar.

Krugman: Ok.

Azhar: Então, esses chips, claro, ficam obsoletos como qualquer outro chip, e tendem a ser usados por empresas de infraestrutura em depreciação por entre cinco e seis anos. As evidências vêm de declarações públicas e regulatórias feitas por pessoas, então podemos assumir que são verdadeiras. E o que eles dizem é que até chips de oito ou nove anos ainda estão sendo alugados a preços premium porque a demanda é muito alta. Uma GPU de 2026, um desses chips de computador para IA, estará em uso em 2046 como uma linha ferroviária de 1888 está em uso hoje? Não, não será. Tenho certeza disso. Mas, no momento, eles se pagam em dois ou três anos.

Krugman: Ok.

Azhar: E o que acontece se uma dessas empresas menores de infraestrutura falir? Falamos das menores, que não são grandes nomes conhecidos, e claro que estão em condições financeiras mais instáveis. Então, se uma delas falir, alguns investidores, seus acionistas, obviamente perderão dinheiro. As pessoas que emprestaram dinheiro podem perder dinheiro, mas os chips não serão perdidos. Elas serão imediatamente assumidas. Na verdade, muitos contratos têm o que é conhecido como cláusula de interrupção no serviço. Então, mesmo que a entidade legal que os possui falhe, o cliente ainda recebe atendimento e atenção. E enquanto estivermos com falta de computação, como parece estar, essas coisas serão retomadas. E acho que podemos esperar que, dessas centenas de pequenas empresas de infraestrutura de menor qualidade, haja muita consolidação nos próximos anos porque os mercados não apoiam tantos players individuais.

Krugman: Certo. Então, a analogia lamentável que as pessoas fizeram é com o fracking, onde um dos grandes problemas foi que os poços fraturados declinaram muito mais rápido do que na contabilidade. Mas você está dizendo que isso não parece realmente um problema dos chips. Quero dizer, isso é um pouco reconfortante.

Azhar: Bem, isso não é um problema agora, e é o que os dados dizem. Mas a questão sobre semicondutores e infraestrutura como essa é que temos falta de oferta agora, mas no momento em que há excesso de oferta, os preços podem desabar. E o que acontece com as empresas de chips em geral, e isso aconteceu ao longo de 50 anos, é que elas entram em períodos em que a demanda cresce, então há falta de oferta, preços sobem, e então eles vão construir fábricas, e quando a fábrica é construída, a demanda já esfriou, e agora eles têm supercapacidade, precisam reduzir seus preços e perdem todo o dinheiro. Então os investidores em semicondutores estão realmente conscientes desse ciclo, e no momento a questão que as pessoas modelo precisam construir em suas cabeças — e há muitas variáveis — é por quanto tempo a demanda por computação de IA e esses chips vai superar a oferta que é fornecida e a oferta disponível? E se isso mudar porque existe um novo algoritmo muito mais eficiente que todo mundo usa,

e a demanda não aumentar, então você pode entrar em uma situação de superoferta que teria muitos dos efeitos que vimos no xisto.

Mas acho que há uma diferença com o xisto, porque o que aconteceria com o xisto seria que você teria o preço externo do petróleo ou gás que está definindo um preço de recuperação, mas mesmo que o preço seja alto, você também tem a curva de esgotamento do poço, onde cada barril subsequente — ou o que quer que você esteja extraindo — é mais caro. Então você tem esse tipo de redução econômica. E não tenho conhecimento suficiente para saber quão rápido um poço franqueado valeria a pena. Mas certamente agora, as GPUs que estão sendo compradas estão sendo quitadas em alguns anos ou dois anos e meio.

Agora, há uma questão maior, certo? IA não é uma coisa de fada aérea que simplesmente vive na nuvem. A nuvem é física. É aço, suor, energia e prédios enormes. E como você disse, prédios que não empregam ninguém. Podemos falar sobre o que poderíamos fazer a respeito. Então há uma corrida para conseguir tudo, exceto os chips, que você precisa para data centers, como terra, resfriamento e, mais importante, energia. Minha equipe e eu achamos que há uma lacuna emergente de energia. Ou seja, a quantidade de energia que os chips precisarão comparada à energia que a rede dos EUA e da Europa podem fornecer. Isso não é tão importante este ano, provavelmente é administrável no ano que vem, mas começa a se tornar bastante significativo no ano seguinte. E isso levanta todo tipo de questão, porque ninguém vai comprar 100 bilhões de dólares em chips em 2028 só para não conectá-los por seis meses. Então, acho que há esses riscos adicionais que tornam essa questão bastante difícil de encontrar uma resposta definitiva.

Krugman: Sim, é aí que eu tenho entrado, dizendo que, além dos data centers e das conexões com a rede ou outras fontes de energia, há muitos pontos de veto político. E frequentemente há governos locais e cidadãos irritados onde isso se torna um problema real.

Azhar: Bem, vamos falar sobre os cidadãos irritados. Então acho que isso é algo realmente importante que está acontecendo nos EUA. É um ato de engajamento democrático e político. Muitas vezes, as pessoas que observam os EUA pensam: "Ninguém vota, ninguém se importa, todo mundo está sintonizado na TV." E isso não é o que estamos vendo agora. Acho que os data centers são menos populares que o ICE, que é um dos departamentos governamentais menos populares. Essa é uma parte importante, eu acho, de como durante uma mudança tecnológica radical, [precisa haver um] contrato social, um acordo que seja negociado.

Como você disse corretamente, para cada 100 milhões de dólares investidos em um data center, eles empregam 25 vezes menos pessoas do que uma fábrica avançada. Então, não são empregadores locais. E um acordo que foi negociado no Reino Unido para dar permissão era essencialmente para cada megawatt de energia — e um grande data center poderia consumir 100, 200 ou 500 megawatts — o operador do data center tem que pagar $100.000 para a comunidade local, e o valor total foi, só olhando minhas notas, cerca de 500 milhões de libras, cerca de 758 milhões de dólares. Agora, acho que essa ideia de construir uma licença local para operar é extremamente acessível tanto para as empresas de infraestrutura quanto para as empresas de IA que elas atendem porque as margens são muito

boas. Eles podem absolutamente pagar isso. Intermediar em nível local, se necessário, esse tipo de acordo acho que ajuda de certa forma. E não é só sobre dinheiro. Precisamos voltar ao ponto que você mencionou: data centers empregam muitas pessoas durante a construção, mas depois não empregam muitas pessoas. Então, o que as comunidades locais podem fazer em termos de fundos de desenvolvimento e acesso? Conheço melhor o Reino Unido do que os EUA. Mas fundos assim podem estimular o empreendedorismo local e a construção de negócios que são geradores e sustentáveis para uma comunidade.

Krugman: Sim, uma das coisas que me intriga um pouco sobre a reação pública é por que os construtores dos data centers não estão dizendo: "Vamos ao Condado de Pleasantville, vamos realizar assembleias públicas abertas e vamos contratar o pagamento de uma grande quantia de dinheiro ao tesouro do condado." Não conheço nenhum exemplo disso, mas parece algo tão óbvio a se fazer neste momento.

Azhar: Há alguns exemplos. Acho que certamente há um conjunto de precedentes no Condado de Loudoun, que é o ponto quente da Virgínia. E sei que o campus da Microsoft no estado de Washington fez compromissos sobre o que aconteceria com as contas de água e os preços da eletricidade e investiu dinheiro para a reurbanização local de diferentes formas. Então acho que o dinheiro está lá. E isso significa que metade da batalha já foi resolvida, porque não é como se você estivesse pedindo para alguém que vai ganhar muito dinheiro para contribuir com uma comunidade local. E acho que a questão é: o que você realmente faz com isso? Você faz como fizemos no Reino Unido com a North Sea Oil, devolve como cortes de impostos e Deus sabe o que acontece 50 anos depois? Ou você consegue usar isso para injetar dinamismo econômico comercial que possa sobreviver quando os trabalhadores da construção saírem?

Krugman: Ok, última pergunta: o que te assusta? Qual é o verdadeiro medo que você tem? Obviamente, sabemos que o hacking é algo que parece bastante alarmante, mas mais adiante, deixando de lado talvez a extinção da raça humana, com o que você realmente se preocupa aqui?

Azhar: A primeira é a extensão do hacking para uma guerra cinzenta organizada pelo Estado, em múltiplas camadas, desde a polusão do ambiente informacional até ataques generalizados e cada vez mais significativos ao que chamamos de infraestrutura nacional. Estou pensando em redes elétricas, sistemas financeiros, sistemas de água, onde não são grupos criminosos individuais, mas são mediados e gerenciados por um Estado. Acho isso realmente, realmente sério, porque não vemos bombas explodindo. Não achamos que estamos em um período de conflito. Especialmente na Europa, onde há muito mais sabotagem acontecendo também, acho que estamos chegando lá, e as pessoas precisam entender que esse risco é real e existe. Então acho que essa é uma das grandes coisas que me preocupam.

Mas a segunda coisa, como falamos, são os empregos. E acho que o que acontece com a força de trabalho, será uma elevação muito suave da temperatura. Não serão cem mil empregos sendo lançados de uma vez. Será uma redução nos salários, menos novas vagas. E antes que você perceba, a força de trabalho parece diferente do que era. E precisamos estar atentos a isso.

E, por fim, estou super, super preocupado com o que vai acontecer com nossos cérebros com a podridão cerebral e dependendo de, sabe, "não é isso, é aquilo", ChatGPT-ismos, e perdendo faculdades críticas. Comprei todas as canetas-tinteiro da minha equipe. Todos nós escrevemos com caneta-tinteiro. Minha mesa está coberta de páginas de anotações manuscritas. Eu incentivo eles a trabalharem sem computador algumas horas por dia. Leio livros bem antigos, um livro de 1794, coisas que são anteriores ao ChatGPT. Eu uso as ferramentas, as ferramentas de IA, o tempo todo, mas ao mesmo tempo, sou muito deliberado em como tento proteger minha mente e passar minhas ideias de pensamento e processamento para meus filhos que, na idade que eles têm—bem, na verdade dois deles mal usam IA, mas eles não vão ser tão receptivos. Mas acho que precisamos realmente pensar sobre o que há no pensamento humano que precisamos continuar cultivando e permitir que ele floresça junto com essas tecnologias que, seja uma bolha que estoure e atrasar por alguns anos, seja ela funcionando ou suave, estará conosco. E precisamos falar sobre isso e realmente agir a respeito quanto antes.

Krugman: Ok, acho que esse é um bom ponto de encerramento. Nossa, coisa assustadora. Fascinante, mas eu não esperava essa aqui.

Azhar: Paul, essa é a natureza da IA. Nenhum de nós prevê nada disso chegando. Quer dizer, essa é uma parte triste.

Krugman: Talvez aconteça que todas as pessoas que dizíamos "aprenda a programar" deveriam ter dito "aprenda a escrever cursiva".

Azhar: Talvez não tenha sido uma má ideia. Valeu.