

OpenAI: nova versão do ChatGPT processa voz e imagem em tempo real

GPT-4o estará aberto a quem não é assinante, mas com limite de uso. Chatbot está agora mais próximo de um assistente virtual

JULIANA CAUSIN
jcausin@globo.com.br

A OpenAI vai abrir para todos os usuários sua versão mais poderosa do ChatGPT, alimentada pelo GPT-4o, novo modelo de inteligência artificial (IA) da empresa — o “o” é de omni, prefixo de origem grega que significa “tudo”. A atualização permite que o robô processe imagem, texto e voz em tempo real e de forma mais eficiente. O objetivo é permitir uma interação humano-computador “muito mais natural”, afirmou ontem a OpenAI.

— Uma parte importante da nossa missão é poder disponibilizar nossas ferramentas avançadas de IA para todos gratuitamente. — afirmou Mira Murati, diretora de Tecnologia da OpenAI, no evento que apresentou o GPT-4o.

Elas acrescentou que o novo modelo é “muito mais rápido” que a versão anterior, o GPT-4, além de ter maior capacidade de processar texto, imagem e áudio.

Segundo Mira, quem assina o serviço poderá processar “cinco vezes” mais conteúdo do que os usuários da versão gratuita, que enfrentaram um limite. Ela explicou que, quando o limite for atingido,

o ChatGPT mudará automaticamente para GPT-3.5 (a atual versão gratuita).

Depois do evento em que apresentou o GPT-4o, a OpenAI divulgou uma série de demonstrações das novas funções do sistema. Em vídeos curtos, funcionários da companhia contam piadas ao ChatGPT, que responde com risada, pedem para a ferramenta cantar “parabéns” a um aniversariante e escutam conselhos do robô sobre como melhorar o visual para participar de uma entrevista.

O GPT-4o é o novo melhor modelo de todos os tempos. É inteligente, é rápido e é nativamente multimodal”, escreveu o CEO da OpenAI, Sam Altman, nas redes sociais.

CONSELHOS E TRADUÇÕES É a primeira vez que um modelo de inteligência artificial da OpenAI já nasce multimodal — com treinamento de texto, áudio e vídeo. A voz do GPT-4o também foi aprimorada e está menos robótica. Com esse salto, a empresa leva para o ChatGPT características de assistentes pessoais, como a Alexa, da Amazon, e a Siri, da Apple, mas com mais funcionalidades.

Outra novidade apresentada ontem foi o aplicativo para computador. Os usuários,

além de acessar o robô pela web ou pelo celular, será possível ter o ChatGPT instalado em seu PC.

O app será ativado com um comando simples, que abrirá uma janela do GPT-4o no computador. O usuário poderá compartilhar a tela com a ferramenta, fazer perguntas, enviar informações e conversar com o sistema. A versão para macOS está disponível desde ontem, e a para Windows será lançada até o fim do ano, segundo a OpenAI.

Os novos recursos de texto e imagem do GPT-4o já estão disponíveis no ChatGPT. Já as interações com voz serão lançadas “em breve”, segundo a empresa.

Como GPT-4o, a ferramenta irá funcionar com capacidades de fala que envolvem interações em tempo real. Na demonstração dos novos recursos, um executivo da OpenAI pediu ao ChatGPT conselhos para se acalmar em uma apresentação importante. O assistente de voz foi capaz de ouvir e analisar a respiração dele, depois oferecer sugestões sobre como seria possível melhorar.

— Diminua um pouco o ritmo, Mark, você não é um aspirador de pó — brincou a voz feminina da ferramenta. Antes do GPT-4o, já era

possível enviar mensagens de voz para o ChatGPT, mas este demorava até 5,4 segundos para responder. Com a atualização, a ideia é que os usuários tenham uma conversa fluida com a ferramenta. As respostas vão demorar, em média, 320 milissegundos para serem dadas, tempo “sensivelmente ao tempo de resposta humano em uma conversa”, de acordo com a OpenAI.

O ChatGPT turbinado é capaz de fazer traduções e intermediar conversas em diferentes línguas. No evento, Mira Murati falou em italiano com Mark Chen, um dos chefes de pesquisa da OpenAI. Ele pediu que a ferramenta traduzisse o que fosse falado em italiano para o inglês, e vice-versa.

Segundo a OpenAI, o GPT-4o tem bom desempenho em cerca de 50 idiomas. Em outra demonstração, a

equipe da OpenAI mostrou que o ChatGPT será capaz de “enviar” a partir da câmera do celular. Uma equação simples foi exibida, e a IA deu conselhos sobre como resolvê-la.

CHATGPT VAI ‘LER’ EMOÇÕES O sistema também poderá interpretar emoções a partir de imagens. No evento, Barret Zoph, diretor de Pós-Treinamento da OpenAI, abriu a câmera, sorriu para o ChatGPT e perguntou “que tipo de emoção” ele estava sentindo. Em um primeiro momento, a ferramenta o confundiu com uma superfície de madeira, apresentada minutos antes. Depois se corrigiu.

— Parece que você está se sentindo muito feliz e alegre, com um grande sorriso e talvez até um toque de entusiasmo — respondeu a IA. E qualquer usuário poderá enviar imagens — de documentos a fotografias e captu-

ras de tela — ao ChatGPT e pedir que ele analise o conteúdo. Antes, essa funcionalidade era restrita a assinantes. Segundo a OpenAI, será possível tirar a foto de um cardápio de restaurante em outra língua e pedir ao GPT-4o para traduzi-lo e até dar recomendações.

Nas redes, Altman disse que “parece a IA dos filmes”, citando “Ela”, em que um homem solitário se apaixona por um chatbot de voz feminina: “ainda é um pouco surpreendente para mim que seja real”.

As inovações do ChatGPT, no entanto, mostram que a OpenAI vem sendo pressionada a ampliar sua base de usuários. Quando a ferramenta foi lançada, em novembro de 2022, foi considerada aquela que mais rapidamente atingiu a marca de 100 milhões de usuários ativos. Nos últimos tempos, no entanto, tem oscilado bastante, segundo a consultoria Similarweb.



Ampliando fronteiras. O aplicativo da OpenAI não estará mais restrito a celulares: terá uma versão para computadores

FT: mídia britânica alerta Apple por bloqueio de anúncios

Em carta, associação de jornais do Reino Unido afirma que plano da empresa para novo iOS afetaria sustentabilidade do setor

Um grupo de jornais do Reino Unido alertou a Apple que qualquer movimento para adotar uma ferramenta conhecida como web eraser (“apagador da web”), capaz de bloquear anúncios, colocaria em risco a sustentabilidade do jornalismo, afirmou no domingo o diário de negócios

britânico Financial Times. A mídia especializada já noticiou que a Apple está se preparando para inserir uma ferramenta de privacidade em seu browser, o Safari, na próxima atualização de seu sistema operacional, o iOS. Essa funcionalidade removeria anúncios e outros conteúdos indesejados. Na semana passada, a agência Bloom-

berg informou que Apple negocia com a OpenAI, dona do ChatGPT, para incluir características de inteligência artificial (IA) no novo iPhone. Já o New York Times relatou que a empresa quer apresentar uma atualização do assistente virtual Siri na Conferência Mundial de Desenvolvedores, em junho. A Associação de Mídia Jour-

nalística (NMA, pela sigla em inglês), que representa 900 publicações no Reino Unido, enviou uma carta à Apple, afirmando que o “apagador da web” poderia prejudicar as receitas digitais do setor. Entre os veículos membros da associação estão The Times, The Guardian e The Daily Telegraph. De acordo com o FT, a carta

afirma que o jornalismo profissional demanda recursos, “e a publicidade é uma fonte crucial de receita”. A Apple não comentou o assunto. A NMA ressalta ainda que “bloquear anúncios é uma tática brusca, que frustra a capacidade de criadores de conteúdo de bancarem seu trabalho, e ainda pode levar consumidores a perderem

informações importantes.”

A associação alerta também que há o risco de as ferramentas de IA serem usadas para remover ou alterar determinados Na carta, a NMA reivindica ainda uma reunião com a Apple para discutir as implicações do “apagador da web”.

Em 2021, a Apple introduziu o iOS uma ferramenta de transparência de rastreamento. Os usuários passaram a ter de dar permissão para que aplicativos rastreassem suas atividades na web. Como a maioria recusou, as redes sociais registraram queda na receita de publicidade direcionada.

Setor de radiodifusão discute lei em defesa da liberdade de expressão

Objetivo é garantir que ‘big techs’ sejam responsabilizadas por desinformação

A Associação Internacional de Radiodifusão (AIR) apresentou ontem um texto-base para criar uma legislação em defesa da liberdade de expressão e da democracia nas novas tecnologias digitais.

Milhões de pessoas no mundo se informam em plataformas digitais e redes sociais. Muitas vezes consomem, sem saber, material produzido por jornalistas profissionais e difundido, originalmente, por empresas de rádio e televisão que vai parar nas plataformas digitais sem que o veículo original que produziu a notícia seja remun-

nerado por isso. Também há muita desinformação, fake news, sem que a plataforma seja responsabilizada.

Rádios e TVs têm atividades regulamentadas por lei e podem ser responsabilizadas na Justiça pelo que publicam, diferentemente das grandes empresas de tecnologia, que administram as plataformas e redes sociais. As big techs não têm as atividades regulamentadas; podem publicar conteúdo sem revelar quem produziu e mesmo divulgar informações falsas, as fake news.

A AIR, que representa 17 mil emissoras das Américas

do Norte, Central e do Sul, reuniu-se em Washington a fim de sugerir novas regras para proteger a produção de notícias e a sociedade. A ideia é que o texto sirva de base para criar leis que regulamentem o setor na região.

‘REGRAS TRANSPARENTES’

O presidente da AIR e vice-presidente de Relações Internacionais do Grupo Globo, Paulo Tonet Camargo, disse que o documento é um divisor de águas e apresenta os pontos indispensáveis que as leis dos países devem conter. Os pontos são muito claros. Começando pela



Tonet Camargo. “É fundamental que se tenha um mínimo de regulação”

responsabilidade sobre a distribuição dos conteúdos, passando por regras transparentes de publicidade, por regras de concorrência e também a remuneração do conteúdo jornalístico, que é monetiza-

do pelas plataformas, sem nenhuma remuneração a quem faz efetivamente o conteúdo jornalístico — afirma Tonet Camargo.

O documento destaca que “as redes sociais se tornaram fóruns de debate para

defensores da igualdade e da liberdade, mas também, e com cada vez mais frequência, têm se tornado bolhas de isolamento e fontes de desinformação que representam ameaças verdadeiras à democracia e à estabilidade das nações.”

O texto reforça a necessidade de construir um marco regulatório para as plataformas digitais que corrija as assimetrias existentes entre as big techs e os meios de comunicação social.

— Então, é fundamental que se tenha um mínimo de regulação. Não de conteúdo. Jornais, jamais admitiremos regular conteúdo ou impedir qualquer conteúdo que fosse publicado. Agora, quem publica, quem tem isso como atividade empresarial, tem que ser responsável sobre o produto que entrega. Qualquer negócio no mundo assim — afirma Tonet Camargo. (Com JN)