OPINIÃO

Deepfakes, phishing automatizado e a nova era do cibercrime com IA

Fernando Galdino (*)

Deepfakes, phishing automatizado e a nova era do cibercrime com IA

A inteligência artificial deixou de ser apenas uma aliada da cibersegurança. Em 2025, ela também se tornou uma ferramenta poderosa nas mãos de criminosos, dando origem a uma nova geração de golpes digitais. Deepfakes quase perfeitos, ataques de phishing automatizados e sistemas que exploram falhas em tempo recorde já são realidade, e estão forçando empresas e pessoas a repensarem completamente sua segurança online.

Em um caso que ganhou repercussão internacional, um funcionário de uma empresa em Hong Kong foi enganado por criminosos que usaram deepfakes para simular uma videoconferência com o diretor financeiro da companhia. Convencido de que falava com seus superiores, ele autorizou a transferência de US$25 milhões. A fraude foi tão bem executada que incluiu a reprodução de vozes, rostos e até os gestos de outros membros da equipe, todos criados por inteligência artificial. É o tipo de golpe que parecia coisa de cinema — e que agora acontece de verdade.

Hoje, 78% das pessoas não conseguem distinguir um deepfake bem feito de uma gravação real. Ao mesmo tempo, o chamado "crime cibernético como serviço" está transformando o submundo digital: ferramentas de invasão, tutoriais e suporte técnico circulam em plataformas clandestinas, permitindo que qualquer pessoa execute ataques complexos com poucos cliques. O tempo médio para explorar uma falha caiu de 32 para apenas 5 dias — uma janela mínima para que empresas consigam reagir.

Manipulação digital, eleições e fraudes invisíveis

A inteligência artificial também tem sido usada para distorcer a percepção pública em grande escala. As campanhas de desinformação evoluíram de boatos nas redes para estratégias automatizadas, com vídeos, áudios e conteúdos criados para enganar até os mais atentos. Estima-se que, até 2028, metade das empresas adotará tecnologias específicas para lidar com esse tipo de manipulação — um salto frente à baixa adoção atual.

O impacto vai muito além do ambiente digital. Em contextos eleitorais, deepfakes têm sido usados para simular falas de políticos e sabotar campanhas. No mercado financeiro, há relatos de investidores que tomaram decisões com base em conteúdos falsos. E no universo corporativo, marcas se tornaram alvo de campanhas orquestradas com o objetivo de derrubar reputações em questão de horas.

Outro sinal claro de transformação é o avanço do chamado Ransomware 3.0. Em vez de sequestrar arquivos, os criminosos simplesmente roubam os dados e ameaçam divulgá-los publicamente. Como não há criptografia envolvida, os backups — antes considerados linha de defesa fundamental — se tornam ineficazes. Esse novo modelo já domina os ataques de ransomware e mira, principalmente, setores como saúde, finanças e infraestrutura crítica.

Frente a esse novo cenário, cresce a urgência de estratégias mais inteligentes de defesa. Autenticação multifator resistente a deepfakes, monitoramento contínuo com apoio de IA e treinamentos práticos sobre engenharia social são passos fundamentais. Mas o desafio não é só técnico. Ele exige uma nova cultura digital, em que até uma videoconferência precisa ser questionada — porque o que parece real pode não ser.

A inteligência artificial que ajuda a diagnosticar doenças também pode imitar pessoas com precisão assustadora. A mesma tecnologia que automatiza processos administrativos está sendo usada para enganar, fraudar e invadir. E, nesse jogo, quem ainda aposta apenas em antivírus e senhas fortes está vários passos atrás.

O cibercrime movido por IA não é mais um risco distante. Ele já está em operação, entre nós, muitas vezes de maneira silenciosa, moldando a forma como nos relacionamos com a tecnologia. A pergunta já não é se seremos alvos, mas quando. E, mais do que nunca, será preciso aprender a duvidar daquilo que parece verdadeiro demais.

(*) Diretor de portfólio da SEK.

Cuidado com os chatbots

O Google publicou uma avaliação sobre a confiabilidade dos chatbots de inteligência artificial (IA); os resultados estão longe de serem animadores.

Vivaldo José Breternitz (*)

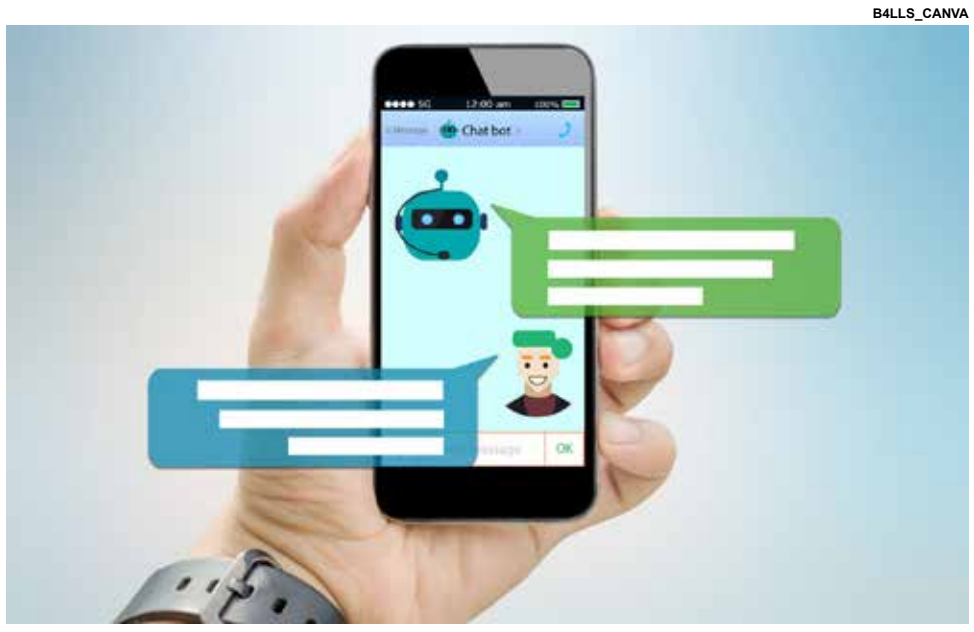
Utilizando sua recém-lançada FACTS Benchmark Suite, uma ferramenta que mede a precisão das respostas dadas por chatbots, o Google descobriu que mesmo os melhores modelos de IA não conseguem chegar a 70% de acertos. O sistema com melhor desempenho, o Gemini 3 Pro, atingiu 69% de precisão, enquanto sistemas de outras empresas, como OpenAI, Anthropic e xAI obtiveram pontuações ainda mais baixas.

A conclusão é simples e desconfortável: esses chatbots ainda erram aproximadamente uma em cada três respostas, mesmo quando essas respostas parecem convincentes.

A nova ferramenta é importante porque a maioria dos testes de IA existentes se concentra em saber se um modelo pode concluir uma tarefa, e não se a informação que ele produz é realmente verdadeira.

Para setores como finanças, saúde e direito, essas falhas podem ser extremamente custosas; uma resposta convincente, mas que contém erros, pode causar danos reais, especialmente quando os usuários presumem que o chatbot sabe do que está falando.

B4LLS_CANVA



Os chatbots tiveram o pior desempenho em tarefas que envolvem a leitura de gráficos, diagramas ou imagens, com acertos frequentemente abaixo de 50%. Isso é preocupante, pois se um chatbot ler erroneamente um raio-X, um gráfico de vendas ou extrair o número errado de um documento, pode gerar erros que são fáceis de ignorar, mas difíceis de reverter.

A principal conclusão não é que os chatbots sejam inúteis, mas que a confiança cega é um risco. Os dados do Google sugerem que a IA está se tornando mais precisa, mas ainda exige verificação, salvaguardas e supervisão humana antes de ser tratada como uma ferramenta totalmente confiável.

(*) Doutor em Ciências pela Universidade de São Paulo, é professor, consultor e diretor do Fórum Brasileiro de Internet das Coisas – vjnitz@gmail.com.

"Prompt", "LLM", "embedding": ranking revela os termos de IA que mais geram dúvidas nos brasileiros

Com o boom das ferramentas de inteligência artificial no dia a dia dos brasileiros, nos últimos anos, um novo glossário tem feito parte da rotina das pessoas — com expressões como "LLM" e "tokens" circulando com naturalidade em reuniões, planejamentos e trocas rápidas dentro e fora da internet.

Aparentemente, no entanto, muitas dessas palavras ainda geram questionamentos: quando indagados sobre as expressões relacionadas à IA que menos entendem em 2025, pergunta feita pela Adapta, o maior ecossistema de IA do Brasil, 500 usuários de tecnologias como o ChatGPT e Gemini apresentaram uma lista de mais de 90 termos e siglas que, em 2025, não conseguem explicar.

Entre os mais mencionados, por exemplo, está a palavra "prompt", que apareceu em 12,6% das respostas e descreve os comandos e instruções fornecidos às ferramentas. Além dela, "fine-tuning", "embeddings" e "overfitting" também estiveram presentes mais de uma vez, revelando que, para muitos, as causas de confusão geralmente são os termos vindos do inglês.

Mas, afinal, diante de tantas dúvidas, quais outras expressões se repetiram nas respostas dos entrevistados? O que significam e como entendê-las de modo simples e sem grandes mistérios? As explicações podem ser encontradas no ranking abaixo, elaborado pela especialista em ferramentas e cursos de IA. Confira!

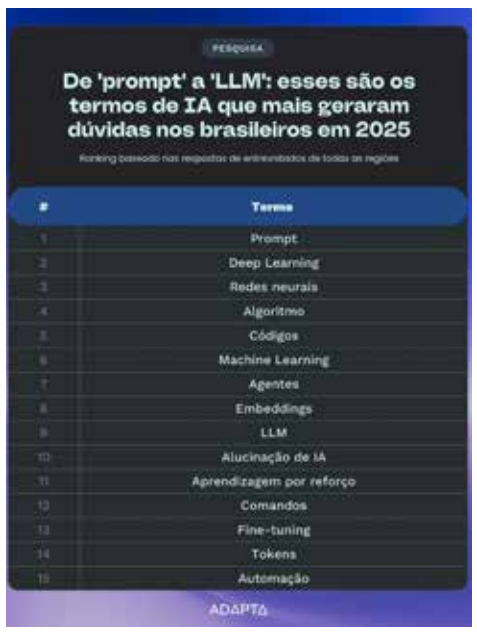
Presente nas respostas de 12,6% dos brasileiros entrevistados, "prompt" foi o termo mais citado no levantamento — sendo, portanto, aquele que mais gera dúvidas entre os usuários. A Adapta explica que não há segredo: de forma geral, a palavra se refere ao comando, instrução ou pergunta enviada ao modelo para que ele produza uma resposta específica.

É a forma como o usuário "orienta" a inteligência artificial, seja para criar textos, resolver problemas, analisar informações ou gerar ideias.

Exemplo prático: Se um analista de marketing quiser usar a IA para criar um conteúdo sobre tendências de consumo, um bom prompt seria: "Crie um artigo sobre as principais tendências de consumo no Brasil para 2025, com foco no impacto das redes sociais e das compras por influência digital." Já um prompt vago como "Crie um artigo sobre tendências" pode resultar em uma resposta mais genérica, sem o foco desejado.

"Como a qualidade do resultado depende diretamente da clareza e do contexto oferecidos no prompt, compreender o termo e saber utilizá-lo bem se tornou indispensável para quem trabalha com IA" explica Eduardo Coelho, Head de Marketing da Adapta.

Reprodução



Os fundamentos gerais da IA, de "machine learning" a "redes neurais"

Além dos famosos "prompts", a ampla menção a expressões como "machine learning", "deep learning" e "redes neurais" deixa evidente um ponto: mesmo quem utiliza IA diariamente ainda tem dúvidas sobre os conceitos que dão origem a essas tecnologias. Somados, os três termos estiveram presentes em mais de 9% das respostas — sinal de que a própria base da inteligência artificial segue sendo um terreno pouco claro para muitos brasileiros.

De maneira simples, "machine learning" (ou aprendizado de máquina) é o campo que permite aos sistemas aprenderem padrões a partir de grandes conjuntos de dados. Dentro dele está o "deep learning", uma técnica mais avançada que utiliza estruturas chamadas "redes neurais" (citadas por 3,27%) para reconhecer imagens, identificar padrões complexos ou prever comportamentos.

Já o termo "algoritmo", também recorrente no levantamento, se refere ao conjunto de regras e cálculos que orienta a IA em cada etapa, seja para analisar um texto, sintetizar informações ou gerar uma imagem. Em resumo, são alguns dos fundamentos que definem a inteligência artificial como a conhecemos hoje e explicam por que ela é capaz de resolver tarefas cada vez mais sofisticadas.

O que é um "token" no contexto da IA?

Se há dúvidas sobre a teoria, há ainda mais incertezas quando o assunto é o funcionamento interno das ferramentas. Termos como "tokens", "embeddings", "LLM" e "códigos" apareceram repetidamente entre os entrevistados (juntos, representando mais de 5% das respostas), revelando que muitos usuários ainda não entendem bem como a IA processa textos e o conhecimento.

"Tokens", por exemplo, podem ser entendidos como pequenos pedaços de texto, como sílabas ou fragmentos de palavras, que o modelo utiliza para compreender e gerar frases. Já os "embeddings" seriam formas matemáticas de representar informações: uma espécie de "tradução" de palavras e conceitos para números, permitindo que o sistema identifique relações de sentido.

As "LLMs" (Large Language Models), que também apareceram nas respostas, são os grandes modelos de linguagem responsáveis por interpretar comandos, reconhecer contexto e produzir textos coerentes em segundos. Por trás de tudo isso estão códigos e arquiteturas complexas que operam sem que o usuário perceba… mas que explicam por que os resultados são tão naturais e, em muitos casos, surpreendentes.

Outros termos difíceis de se explicar

Além da teoria e da estrutura técnica, o levantamento também mostrou que muitos brasileiros têm dúvidas sobre outras expressões que surgem no uso cotidiano das ferramentas. Entre elas estão "fine-tuning", "agentes", "aprendizagem por reforço", "automação" e até mesmo "alucinação de IA".

No universo aplicado, "fine-tuning" é o processo de ajustar um modelo já treinado para um uso mais específico, como adaptar uma IA para responder dúvidas de clientes de um único setor. Já a "aprendizagem por reforço" diz respeito a treinar modelos com base em tentativas e erros, permitindo que eles aprendam comportamentos ideais ao longo do tempo.

Os chamados "agentes de IA", por sua vez, são sistemas capazes de executar tarefas de forma autônoma: pesquisar dados, organizar informações, acionar ferramentas externas ou realizar fluxos completos sem intervenção humana. Já a automação sintetiza a aplicação direta dessas capacidades no dia a dia, desde acelerar tarefas administrativas até otimizar processos de atendimento.

Mas e quanto à "alucinação de IA", dúvida comum entre as centenas de respondentes? Nesse caso, estamos falando de um processo que ocorre quando o modelo apresenta informações incorretas ou inventadas. O fenômeno, nesse sentido, reforça a importância de orientar bem as ferramentas, verificar conteúdos sensíveis e manter a supervisão humana nos processos críticos.

News@TI

Solução transforma celular de empreendedores em maquininha

@O PicPay anuncia o lançamento do PicPay Tap, solução que transforma o celular do empreendedor em uma maquininha de pagamento. A ferramenta permite que qualquer empreendedor aceite pagamentos com cartão ou carteira digital por aproximação (NFC) diretamente no celular sem precisar de um terminal físico. Assim, quem vende ganha mobilidade, reduz custos, e oferece mais formas de pagamento para o cliente final — tudo isso sem mensalidade ou taxa de adesão.

Empresas & Negócios **José Hamilton Mancuso (1936/2017)** **Laurinda Machado Lobato (1941-2021)** Responsável: **Lilian Mancuso**

Editorias
Economia/Política: J. L. Lobato (lobato@netjen.com.br); *Ciência/Tecnologia:* Ricardo Souza (ricardosouza@netjen.com.br); *Livros:* Ralph Peter (ralphpeter@agenteliterarioralph.com.br);
Comercial: comercial@netjen.com.br
Publicidade Legal: lilian@netjen.com.br

Webmaster/TI: Fabio Nader; *Editoração Eletrônica:* Ricardo Souza.
Revisão: Maria Cecília Camargo; *Serviço informativo:* Agências Brasil, Senado, Câmara, EBC, ANSA.

Artigos e colunas são de inteira responsabilidade de seus autores, que não recebem remuneração direta do jornal.

Colaboradores: Claudia Lazzarotto, Eduardo Moisés, Geraldo Nunes e Heródoto Barbeiro.

ISSN 2595-8410

Jornal Empresas & Negócios Ltda
Administração, Publicidade e Redação: Rua Joel Jorge de Melo, 468, cj. 71 – Vila Mariana – São Paulo – SP – CEP.: 04128-080
Telefone: (11) 3106-4171 – E-mail: (netjen@netjen.com.br)
Site: (www.netjen.com.br). CNPJ: 05.687.343/0001-90
JUCESP, Nire 35218211731 (6/6/2003)
Matriculado no 3º Registro Civil de Pessoa Jurídica sob nº 103.