

PILLAR

INSIGHTS

Os geradores de imagem com IA tropeçam em textos que não são em inglês — e reforçam a desigualdade linguística.

Um relato em primeira mão de pedir a modelos de imagem com IA que escrevessem espanhol, francês e português dentro de seus resultados — e de ver o idioma se desfazer de formas que mapeiam diretamente o viés fonético, regional e de conjunto de dados.

PILLAR MEDIA & ENTERTAINMENT · PILLARME.COM/INSIGHTS

01 – A ilusão de competência.


Um dos principais desafios que encontrei ao trabalhar com sistemas de IA foi a questão das barreiras de idioma e das imprecisões linguísticas, especialmente nas ferramentas de geração de imagens com IA. Embora essas tecnologias sejam frequentemente apresentadas como altamente avançadas e multilíngues, minha experiência mostrou que o tratamento que dão à linguagem escrita dentro das imagens geradas permanece surpreendentemente pouco confiável. Essa limitação ficou particularmente visível ao trabalhar com prompts em idiomas que não o inglês, ou ao pedir à IA que gerasse um texto preciso dentro do conteúdo visual.

À primeira vista, os modelos de geração de imagens parecem extremamente sofisticados. Eles conseguem criar fotografias realistas, ilustrações artísticas, anúncios, pôsteres e composições visuais complexas em segundos. No entanto, quando se pede a esses mesmos sistemas que incluam linguagem escrita na imagem, muitas inconsistências começam a aparecer. Em vez de produzir palavras coerentes ou frases gramaticalmente corretas, a IA muitas vezes gera tipografia distorcida, palavras inventadas, idiomas misturados ou combinações de letras sem sentido que se parecem com linguagem sem de fato serem compreensíveis.

Um dos aspectos mais interessantes dessa questão é que a IA frequentemente dá a *ilusão* de competência linguística. O texto gerado muitas vezes parece visualmente convincente à distância porque imita a estrutura da escrita real. O espaçamento, a tipografia e o ritmo visual se parecem com linguagem de verdade, mas, sob inspeção mais atenta, as palavras costumam estar incorretas ou ser completamente fictícias. Isso sugere que o modelo talvez não processe verdadeiramente a linguagem escrita dentro das imagens como linguagem em si, mas sim como um padrão visual ou uma textura estética.

IA en las escuelas latinas





Uma imagem gerada por IA que parece competente em tamanho de miniatura. De perto, as palavras se dissolvem em formas que imitam a escrita sem serem escrita.

02 — O estudo de caso “educasion”.

Esta imagem me pareceu particularmente interessante porque os erros de ortografia gerados pela IA se assemelham muito a erros ortográficos comuns encontrados em partes da América Latina devido a diferenças de pronúncia em relação ao espanhol europeu. Por exemplo, a IA escreveu “**educasion**” em vez de “**educación**”, refletindo a forma como a palavra costuma ser pronunciada em muitos sotaques latino-americanos, em que os sons do “c” e do “s” são pronunciados de modo semelhante.

Em contraste, na Espanha, a pronúncia distingue mais claramente o som de “ción”. O que torna isso especialmente fascinante é que a IA parece reproduzir não apenas padrões linguísticos, mas também **tendências fonéticas associadas a variedades regionais específicas do espanhol**. O modelo internalizou como soa o espanhol latino-americano sem internalizar a ortografia padronizada que o acompanha.



“Educasion” na prática. O erro de ortografia reflète como a palavra é pronunciada em boa parte da América Latina — evidência de que o modelo absorveu padrões fonéticos, não apenas texto.

03 — Os contextos multilíngues pioram a situação.

Esse problema se torna ainda mais pronunciado em contextos multilíngues. Durante minha experimentação, notei que os prompts em inglês geralmente produziam resultados mais coerentes do que os prompts escritos em francês, espanhol ou português. Mesmo quando a IA entendia o sentido geral do pedido, o texto gerado frequentemente continha erros de ortografia, acentos faltando, gramática incorreta ou misturas entre vários idiomas ao mesmo tempo. Em alguns casos, a IA inventava palavras inteiramente novas que se pareciam com fragmentos de idiomas existentes, mas não tinham significado real.



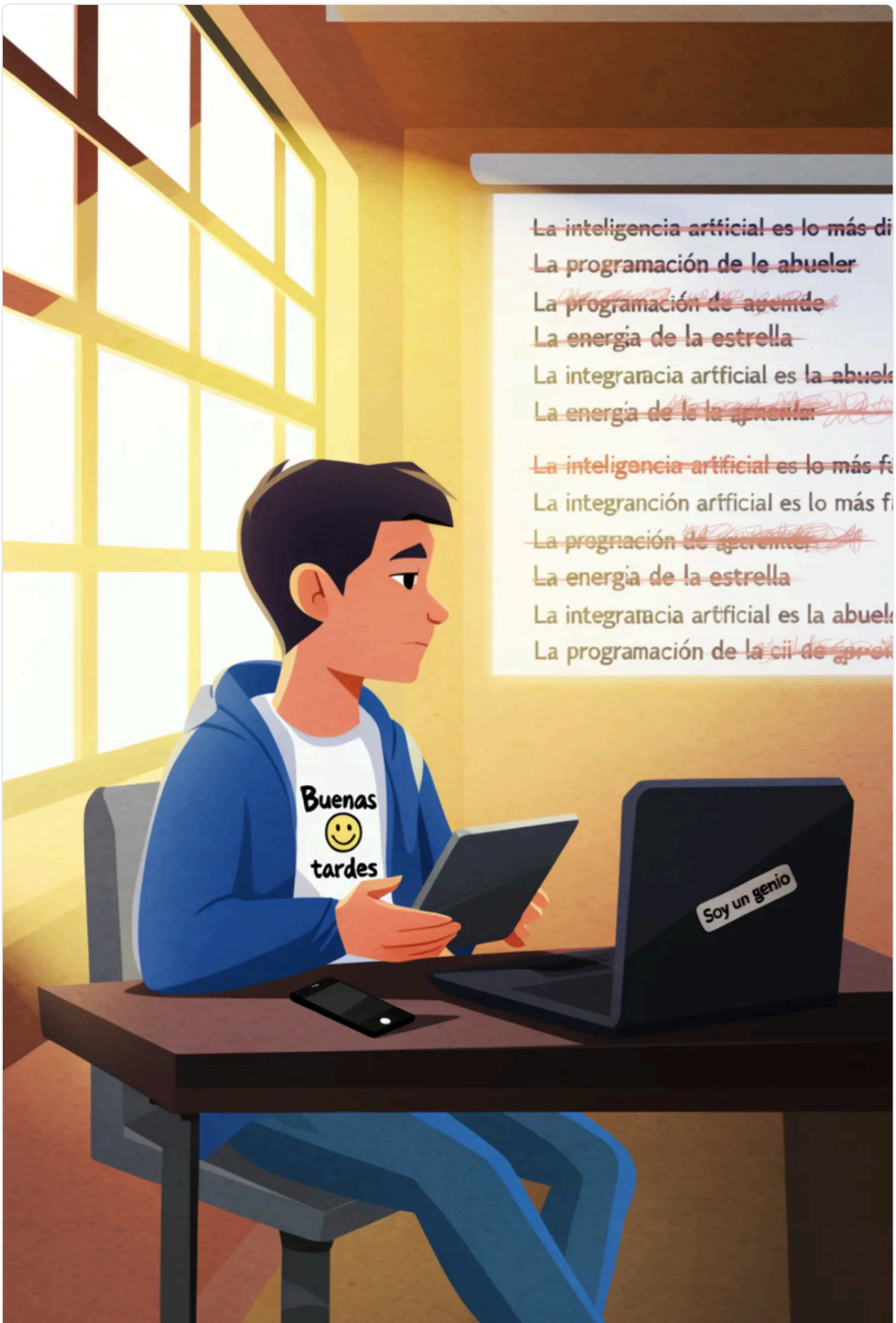
Um prompt multilíngue produz vocabulário inventado e fragmentos que se parecem com francês e espanhol sem serem nenhum dos dois. O modelo gesticula em direção à

linguagem; ele não a escreve.

04 – A IA “corrige” riscando.

Outro exemplo que achei particularmente revelador foi uma imagem contendo um texto completamente sem sentido. Depois de notar os erros, pedi à IA que corrigisse a escrita mantendo o mesmo estilo visual e a mesma estrutura. No entanto, em vez de produzir uma versão coerente, o sistema gerou uma segunda imagem que *ainda* continha texto sem sentido.

Ainda mais interessante: em vez de de fato corrigir os erros, a IA simplesmente **riscou algumas das palavras incorretas**, como se reconhecesse visualmente os erros sem ser capaz de corrigi-los de fato. Isso deu a impressão de que o modelo reconhecia que algo estava errado em um nível superficial, embora ainda lhe faltasse uma verdadeira compreensão do próprio conteúdo linguístico.



~~La inteligencia artificial es lo más di~~
~~La programación de le abueler~~
~~La programación de apemde~~
~~La energía de la estrella~~
~~La integraracia arftificial es la abuelo~~
~~La energía de lo la apnemla~~

~~La inteligencia artificial es lo más fe~~
~~La integranción arftificial es lo más fi~~
~~La prognación de apnemla~~
~~La energía de la estrella~~
~~La integraracia arftificial es la abuelo~~
~~La programación de la cil de apnem~~

Solicitado a corrigir o texto, o modelo riscou os próprios erros. Ele reconheceu que algo estava errado; não soube dizer o quê.

05 – Pôsteres em espanhol, vocabulário inventado.

Ao solicitar a criação de um pôster em espanhol, a IA gerou frases que visualmente se pareciam com espanhol, mas incluíam estruturas gramaticais impossíveis e vocabulário inventado. Além disso, às vezes o resultado era um texto híbrido misturando elementos do espanhol e do inglês. Essa inconsistência revelou uma limitação importante: embora os sistemas de IA sejam comercializados como ferramentas multilíngues, seu desempenho permanece fortemente influenciado pela **dominância dos dados em inglês nos conjuntos de treinamento**.



PROYECTO
IA PARA ESCUELAS LATAM

Inteligencia Artificial al servicio de la educación.

Herramientas de IA para directivos y docentes

Nuestro proyecto impulsa la transformación educativa en Latinoamérica, brindando herramientas de IA seguras, prácticas y



Perfeccionar imágenes



Editar imagen completa

Toca Editar y describe los cambios en el chat.



Selecciona un área de la imagen para editarla

Toca Seleccionar, destaca el área y describe los cambios en el chat.

Comenzar

Obtener más información

Um “pôster em espanhol” de um modelo de imagem: layout de aparência real, gramática quebrada, vocabulário metade em espanhol e metade em inglês. A moldura visual está correta. A linguagem não.

o6 – Imprevisibilidade.

Outra grande questão foi a imprevisibilidade dos resultados. Exatamente o mesmo prompt podia produzir resultados drasticamente diferentes de uma geração para outra. Em algumas tentativas, o texto estava correto ou quase correto, enquanto em outras se tornava completamente ilegível. Essa inconsistência tornava as ferramentas difíceis de utilizar com confiança para uso profissional ou acadêmico. Mesmo depois de refinar os prompts várias vezes, adicionar especificações ou simplificar as instruções, a IA continuava tendo dificuldade com a precisão textual em espanhol dentro das imagens.

O modelo sabe como a escrita se parece. Ele não sabe o que a escrita significa.

07 — Consequências práticas para os usuários.

Essa limitação tem consequências práticas importantes. Um pôster, infográfico, anúncio ou apresentação que contenha linguagem distorcida ou incorreta pode parecer pouco profissional ou até enganoso. Como resultado, os usuários muitas vezes precisam editar manualmente as imagens geradas depois, usando software externo. A IA ainda exige correção humana.

No entanto, com a versão Pro do ChatGPT, você desbloqueia muitas opções mais poderosas para editar fotos e imagens com facilidade. Desde melhorar a qualidade da imagem e corrigir erros até gerar designs profissionais e fazer ajustes criativos avançados, o processo se torna mais rápido, mais simples e muito mais flexível — embora o problema linguístico subjacente dentro da própria imagem permaneça.



PROYECTO
IA PARA ESCUELAS LATAM

Inteligencia Artificial al servicio de la educación.

Herramientas de IA para directivos y docentes

Nuestro proyecto impulsa la transformación educativa en Latinoamérica, brindando herramientas de IA seguras, prácticas y adaptadas a la realidad de cada escuela.



Ahorra tiempo en tareas administrativas



Toma decisiones basadas en datos



Mejora los aprendizajes y la experiencia educativa



IA segura, ética y alineada a valores educativos



Escuelas más eficientes, equipos más empoderados, estudiantes **con más oportunidades.**

MEJORES DECISIONES
MÁS TIEMPO PARA LO QUE REALMENTE IMPORTA:
NUESTROS ESTUDIANTES



Tipografía de infográfico de aparência realista que se desfaz em resolução total. Para uso profissional, o resultado ainda precisa de correção humana em software externo.

o8 – A questão mais ampla: acessibilidade e representação.

A questão também levanta perguntas mais amplas sobre acessibilidade e representação. Idiomas com menor pegada digital ou menos representação nos conjuntos de treinamento de IA podem ser mais vulneráveis a resultados de baixa qualidade. Isso cria um desequilíbrio em que os usuários de língua inglesa se beneficiam de sistemas mais precisos e otimizados, enquanto os usuários multilíngues enfrentam taxas mais altas de erros e limitações. Nesse sentido, os sistemas de IA podem, de forma não intencional, reforçar **desigualdades linguísticas existentes nos espaços digitais**.

Outra observação interessante é o contraste entre os sistemas de IA textuais e os sistemas de IA de geração de imagens. Os modelos baseados em texto costumam ser capazes de produzir uma escrita multilíngue altamente coerente, enquanto os geradores de imagem têm dificuldade até para reproduzir frases simples com precisão. Isso sugere que gerar linguagem visualmente é fundamentalmente diferente de gerar linguagem textualmente.

As capturas de tela de gerações incorretas revelaram padrões que teriam sido difíceis de explicar teoricamente. Algumas imagens continham placas com palavras quase corretas, enquanto outras produziam uma tipografia que parecia realista, mas era totalmente ilegível. Em vários casos, a IA ignorou instruções explícitas sobre a escolha do idioma ou a precisão ortográfica. Essas falhas visuais demonstraram que, apesar da aparência impressionante dos sistemas de IA generativa, limitações técnicas significativas permanecem sem solução.

Cancelar

Seleccionar

Siguiente



PROYECTO
IA PARA ESCUELAS LATAM

Inteligencia Artificial al servicio de la educación.

Herramientas de IA para directivos y docentes

Nuestro proyecto impulsa la transformación educativa en Latinoamérica, brindando herramientas de IA seguras, prácticas y adaptadas a la realidad de cada escuela.



Ahorra tiempo en tareas administrativas



Toma decisiones basadas en datos



Mejora los aprendizajes y la experiencia educativa

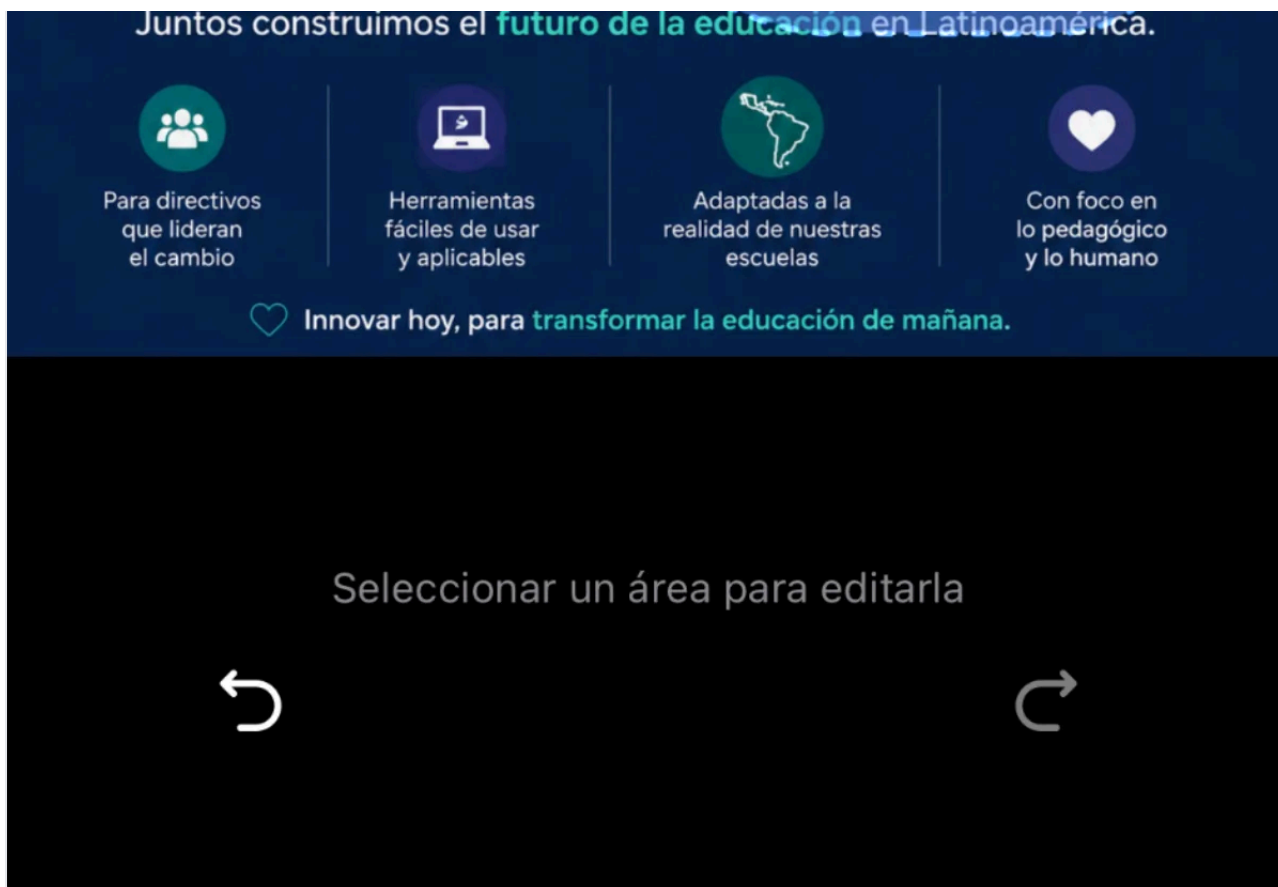


IA segura, ética y alineada a valores educativos



Escuelas más eficientes, equipos más empoderados, estudiantes **con más oportunidades.**

MEJORES DECISIONES
MÁS TIEMPO PARA LO QUE REALMENTE IMPORTA:
NUESTROS ESTUDIANTES



Instruções explícitas de idioma e ortografia, ignoradas. A limitação técnica não é estética — é linguística, e está distribuída de forma desigual entre as línguas do mundo.

No fim das contas, essa experiência mudou minha percepção sobre a confiabilidade da IA. Embora a IA seja frequentemente apresentada como uma solução universal capaz de substituir o trabalho humano nas indústrias criativas, esses experimentos mostraram que os sistemas atuais ainda têm dificuldade com a precisão linguística, especialmente em contextos multilíngues e multimodais. Isso ressalta a importância de usar a IA de forma crítica e de verificar cuidadosamente os resultados, sobretudo ao trabalhar com diferentes idiomas.

Ao mesmo tempo, fiquei genuinamente surpresa com quão impressionantes eram muitas das imagens geradas e dos resultados criativos. A tecnologia é inegavelmente poderosa e evolui rapidamente, mesmo que os desafios em torno da precisão multilíngue e da confiabilidade textual ainda permaneçam como áreas importantes de melhoria.

SOBRE A AUTORA

Marie Prieto

Marie Prieto é pesquisadora e escritora com foco em como os sistemas de IA se saem — e falham — nas línguas do mundo de fala espanhola e francesa. Este artigo é uma contribuição de Embaixadora da LATAM para a Pillar Insights.

[Leia mais artigos da Pillar Insights \(/pt/perspectivas\)](/pt/perspectivas)

Perguntas frequentes.

Por que os geradores de imagem com IA têm dificuldade com texto dentro das imagens?

Os geradores de imagem com IA parecem não processar a linguagem escrita dentro das imagens como linguagem em si. Eles tratam letras e palavras como padrões visuais ou texturas estéticas. À distância, o texto gerado parece convincente porque o espaçamento, a tipografia e o ritmo visual imitam a estrutura da escrita real. Sob inspeção mais atenta, as palavras costumam estar incorretas, embaralhadas ou completamente fictícias. É por isso que pôsteres e placas de modelos de imagem podem parecer críveis em tamanho de miniatura e se desfazer em resolução total. O modelo sabe como a escrita se parece, mas não sabe o que a escrita significa.

Por que a geração de imagens com IA tem melhor desempenho em inglês do que em espanhol, francês ou português?

Os prompts em inglês geralmente produzem um texto mais coerente dentro das imagens geradas do que os prompts em francês, espanhol ou português. Mesmo quando a IA entende o sentido geral de um pedido que não é em inglês, o resultado visual frequentemente contém erros de ortografia, acentos faltando, gramática incorreta ou misturas de vários idiomas ao mesmo tempo. A causa é o desequilíbrio do conjunto de dados. Os modelos de imagem com IA são fortemente influenciados pela dominância dos dados em inglês em seus conjuntos de treinamento, então os idiomas que não o inglês herdam taxas mais altas de erros. Os sistemas de IA comercializados como multilíngues, na prática, ainda privilegiam o inglês.

O que o erro de ortografia “educasion” revela sobre os modelos de linguagem de IA?

Quando um gerador de imagem com IA escreveu “educasion” em vez de “educación”, ele reproduziu um erro ortográfico comum encontrado em partes da América Latina, onde a pronúncia do seseo funde os sons do c e do s. No espanhol peninsular, a terminação ción é pronunciada de forma mais distinta. O erro de ortografia sugere que o modelo não está apenas aprendendo o vocabulário escrito do espanhol, mas também absorvendo tendências fonéticas de sotaques regionais específicos. A IA aprendeu como soa o espanhol latino-americano sem aprender a ortografia padronizada que o acompanha.

O que acontece quando você pede à IA que corrija o texto dentro de uma imagem gerada?

Pedir a um modelo de imagem com IA que corrija um texto quebrado dentro de uma imagem gerada anteriormente raramente produz uma correção coerente. Em um caso revelador, o modelo regeou uma imagem e, em vez de consertar as palavras, simplesmente riscou o texto incorreto. O sistema pareceu reconhecer que algo estava errado em um nível superficial, mas faltava-lhe a compreensão linguística necessária para de fato corrigi-lo. Esse padrão reforça a conclusão de que os geradores de imagem não processam o texto embutido como linguagem. Eles o processam como um elemento visual que pode ser substituído ou riscado, e não como uma sequência de caracteres com significado.

Por que os resultados das imagens de IA são tão imprevisíveis de uma geração para outra?

O mesmo prompt pode produzir resultados drasticamente diferentes de uma geração para outra. Em algumas tentativas, o texto embutido está correto ou quase correto. Em outras, está ilegível. Mesmo com prompts refinados, especificações adicionais ou instruções simplificadas, os geradores de imagem com IA continuam tendo dificuldade com a precisão textual em espanhol dentro das imagens. Essa inconsistência torna as ferramentas difíceis de utilizar com confiança em trabalhos profissionais ou acadêmicos sem correção humana manual posterior em software externo.

Como isso afeta a acessibilidade e a representação de idiomas que não o inglês?

A lacuna na geração de texto levanta perguntas mais amplas sobre acessibilidade e representação. Idiomas com menor pegada digital ou menos representação nos conjuntos de treinamento de IA são mais vulneráveis a resultados de baixa qualidade. Os usuários de língua inglesa se beneficiam de sistemas mais precisos e otimizados, enquanto os usuários multilíngues enfrentam taxas mais altas de erros e limitações. Os sistemas de IA podem, portanto, reforçar de forma não intencional as desigualdades linguísticas existentes nos espaços digitais. Essa é a mesma dinâmica que a Pillar foi construída para combater no nível da infraestrutura: fechar a lacuna de qualidade entre o inglês e as línguas do Sul Global.

Se os modelos de imagem com IA não conseguem soletrar “educación”, a mesma lacuna de conjunto de dados está moldando quais idiomas recebem imóveis digitais premium. A Pillar gerencia mais de 100.000 propriedades premium em espanhol, francês, português e nas línguas do Sul Global.