

PILLAR

INSIGHTS

# Les générateurs d'images par IA peinent avec les textes non anglophones — et renforcent l'inégalité linguistique.

*Un témoignage de première main : demander à des modèles d'images par IA d'écrire en espagnol, en français et en portugais dans leurs productions — et regarder la langue se déliter d'une manière qui correspond directement aux biais phonétiques, régionaux et de jeux de données.*

PILLAR MEDIA & ENTERTAINMENT · PILLARME.COM/INSIGHTS

## 01 – L'illusion de compétence.

---


L'un des principaux défis que j'ai rencontrés en travaillant avec les systèmes d'IA était la question des barrières linguistiques et des imprécisions de langue, en particulier dans les outils de génération d'images par IA. Bien que ces technologies soient souvent présentées comme très avancées et multilingues, mon expérience a montré que leur traitement de la langue écrite à l'intérieur des images générées reste étonnamment peu fiable. Cette limite est devenue particulièrement visible lorsque je travaillais avec des prompts dans des langues autres que l'anglais, ou lorsque je demandais à l'IA de générer un texte précis au sein d'un contenu visuel.

À première vue, les modèles de génération d'images paraissent extrêmement sophistiqués. Ils peuvent créer des photographies réalistes, des illustrations artistiques, des publicités, des affiches et des compositions visuelles complexes en quelques secondes. Cependant, lorsqu'on demande à ces mêmes systèmes d'inclure de la langue écrite dans l'image, de nombreuses incohérences commencent à apparaître. Au lieu de produire des mots cohérents ou des phrases grammaticalement correctes, l'IA génère souvent une typographie déformée, des mots inventés, des langues mélangées, ou des combinaisons de lettres dénuées de sens qui ressemblent à de la langue sans être réellement compréhensibles.

L'un des aspects les plus intéressants de ce problème est que l'IA donne fréquemment l'*illusion* d'une compétence linguistique. Le texte généré paraît souvent visuellement convaincant à distance parce qu'il imite la structure d'une véritable écriture. L'espacement, la typographie et le rythme visuel ressemblent à une langue réelle, mais à y regarder de plus près, les mots sont souvent incorrects ou totalement fictifs. Cela laisse penser que le modèle ne traite peut-être pas véritablement la langue écrite à l'intérieur des images comme une langue en soi, mais plutôt comme un motif visuel ou une texture esthétique.

# IA en las escuelas latinas





*Une image générée par IA qui paraît compétente en taille vignette. De près, les mots se dissolvent en formes qui imitent l'écriture sans en être.*

## 02 – L'étude de cas « educasion ».

---

Cette image m'a semblé particulièrement intéressante parce que les fautes d'orthographe générées par l'IA ressemblent fortement aux erreurs orthographiques courantes que l'on trouve dans certaines parties de l'Amérique latine, en raison des différences de prononciation avec l'espagnol européen. Par exemple, l'IA a écrit « **educasion** » au lieu d'« **educación** », reflétant la façon dont le mot est souvent prononcé dans de nombreux accents latino-américains, où les sons « c » et « s » se prononcent de manière similaire.

En revanche, en Espagne, la prononciation distingue plus nettement le son « ción ». Ce qui rend cela particulièrement fascinant, c'est que l'IA semble reproduire non seulement des schémas linguistiques, mais aussi des **tendances phonétiques associées à des variétés régionales spécifiques de l'espagnol**. Le modèle a intériorisé la sonorité de l'espagnol latino-américain sans intérioriser l'orthographe normalisée qui l'accompagne.



« Educasion » en conditions réelles. La faute d'orthographe reflète la façon dont le mot est prononcé dans une grande partie de l'Amérique latine — la preuve que le modèle a absorbé des schémas phonétiques, et pas seulement du texte.

### 03 — Les contextes multilingues aggravent la situation.

Ce problème devient encore plus prononcé dans les contextes multilingues. Au cours de mon expérimentation, j'ai remarqué que les prompts en anglais produisaient généralement des résultats plus cohérents que les prompts rédigés en français, en espagnol ou en portugais. Même lorsque l'IA comprenait le sens général de la demande, le texte généré contenait fréquemment des fautes d'orthographe, des accents manquants, une grammaire incorrecte, ou des mélanges entre plusieurs langues à la fois. Dans certains cas, l'IA inventait des mots entièrement nouveaux qui ressemblaient à des fragments de langues existantes mais n'avaient aucun sens réel.



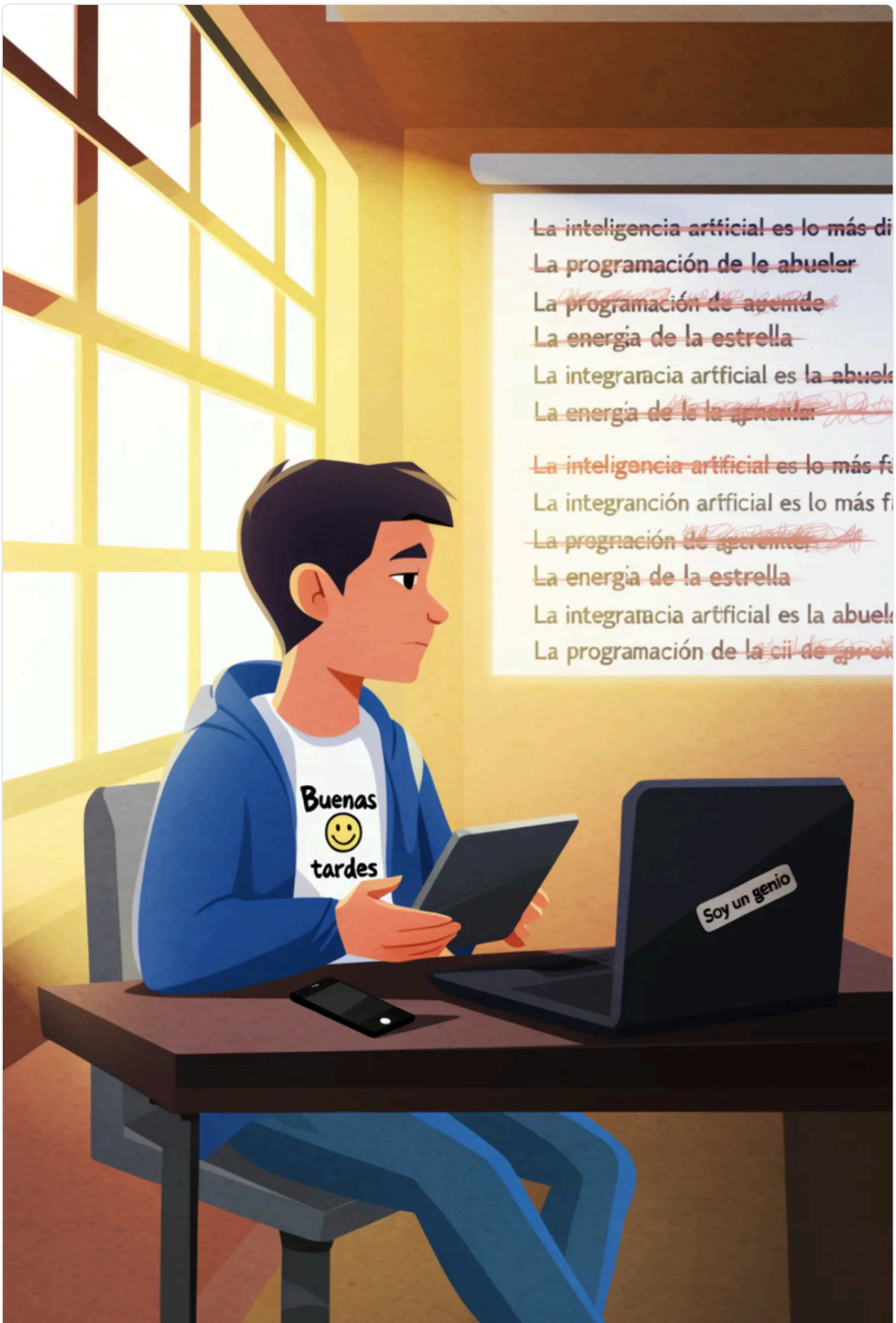
*Un prompt multilingue produit un vocabulaire inventé et des fragments qui ressemblent au français et à l'espagnol sans être ni l'un ni l'autre. Le modèle fait signe vers la langue ; il ne l'écrit pas.*

## 04 – L'IA « corrige » en barrant.

---

Un autre exemple que j'ai trouvé particulièrement révélateur était une image contenant un texte totalement absurde. Après avoir remarqué les erreurs, j'ai demandé à l'IA de corriger le texte tout en conservant le même style visuel et la même structure. Cependant, au lieu de produire une version cohérente, le système a généré une seconde image qui contenait *encore* un texte dénué de sens.

Plus intéressant encore, plutôt que de corriger réellement les erreurs, l'IA s'est contentée de **barrer certains des mots incorrects**, comme pour reconnaître visuellement les erreurs sans être capable de les corriger correctement. Cela donnait l'impression que le modèle reconnaissait qu'il y avait un problème à un niveau superficiel, tout en manquant encore d'une véritable compréhension du contenu linguistique lui-même.



*Prié de corriger le texte, le modèle a barré ses propres erreurs. Il a reconnu qu'il y avait un problème ; il n'a pas su dire lequel.*

## 05 – Affiches en espagnol, vocabulaire inventé.

---

Lorsque je demandais la création d'une affiche en espagnol, l'IA générait des phrases qui ressemblaient visuellement à de l'espagnol mais comportaient des structures grammaticales impossibles et un vocabulaire inventé. De plus, cela aboutissait parfois à un texte hybride mêlant des éléments espagnols et anglais. Cette incohérence révélait une limite importante : bien que les systèmes d'IA soient commercialisés comme des outils multilingues, leur performance reste fortement influencée par **la domination des données en langue anglaise dans les jeux d'entraînement.**



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**

*Une « affiche en espagnol » issue d'un modèle d'images : une mise en page d'apparence réelle, une grammaire brisée, un vocabulaire moitié espagnol moitié anglais. Le cadre visuel est correct. La langue ne l'est pas.*

## o6 – Imprévisibilité.

---

Un autre problème majeur était l'imprévisibilité des résultats. Exactement le même prompt pouvait produire des résultats radicalement différents d'une génération à l'autre. Dans certaines tentatives, le texte était correct ou presque correct, tandis que dans d'autres il devenait totalement illisible. Cette incohérence rendait les outils difficiles à utiliser de façon fiable pour un usage professionnel ou universitaire. Même après avoir affiné les prompts à plusieurs reprises, ajouté des précisions ou simplifié les instructions, l'IA continuait à peiner avec la précision textuelle de l'espagnol à l'intérieur des images.

*Le modèle sait à quoi ressemble l'écriture. Il ne sait pas ce que l'écriture signifie.*

## 07 — Conséquences pratiques pour les utilisateurs.

---

Cette limite a d'importantes conséquences pratiques. Une affiche, une infographie, une publicité ou une présentation contenant une langue déformée ou incorrecte peut paraître peu professionnelle, voire trompeuse. Par conséquent, les utilisateurs doivent souvent retoucher manuellement les images générées par la suite à l'aide d'un logiciel externe. L'IA exige encore une correction humaine.

Cependant, avec la version Pro de ChatGPT, vous débloquez bien d'autres options puissantes pour retoucher des photos et des images avec facilité. De l'amélioration de la qualité d'image et de la correction des erreurs à la génération de designs professionnels et aux ajustements créatifs avancés, le processus devient plus rapide, plus simple et bien plus flexible — même si le problème de langue sous-jacent à l'intérieur de l'image elle-même demeure.



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

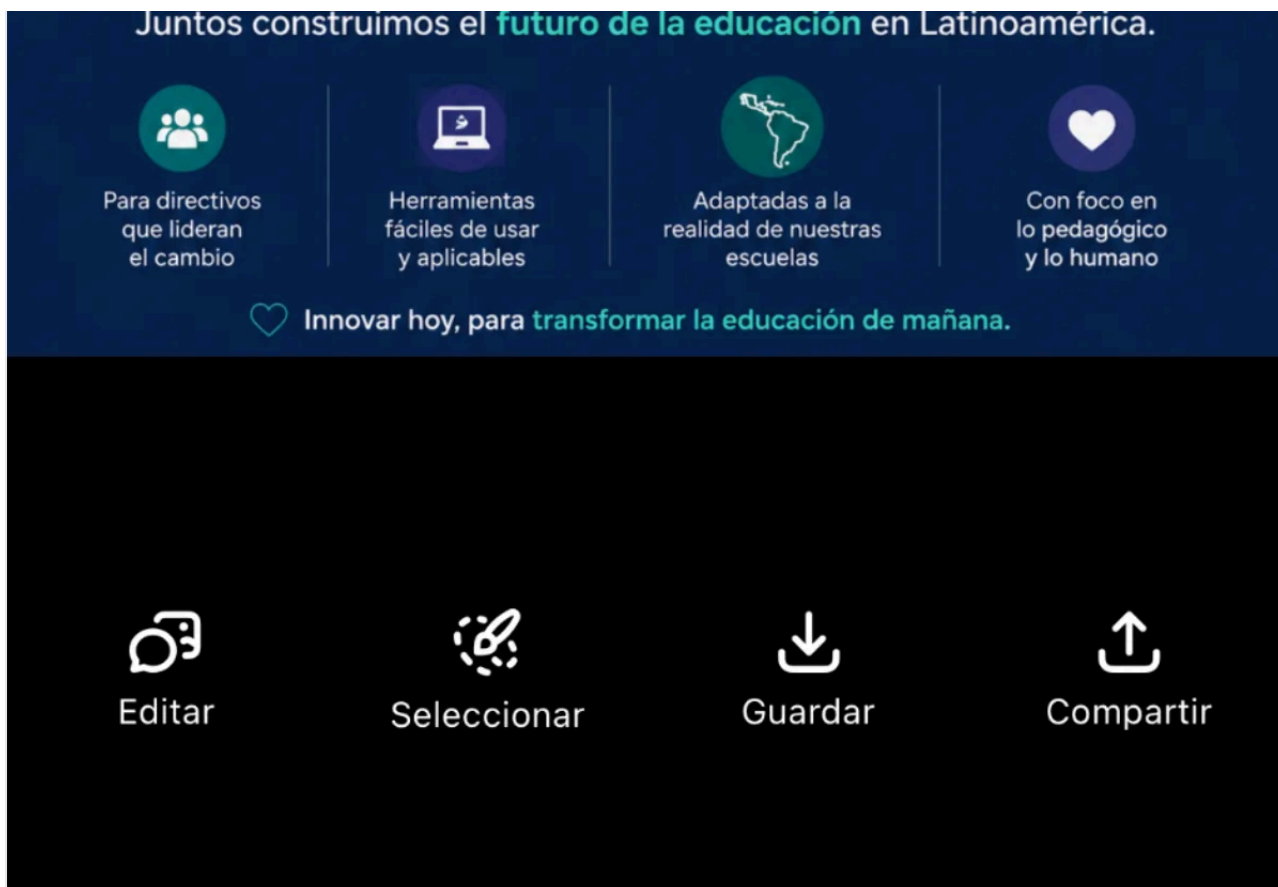


IA segura, ética y alineada a valores educativos



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

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



*Une typographie d'infographie d'apparence réaliste qui s'effondre en pleine résolution. Pour un usage professionnel, le résultat nécessite encore une correction humaine dans un logiciel externe.*

## o8 – La question plus large : accessibilité et représentation.

Le problème soulève aussi des questions plus larges sur l'accessibilité et la représentation. Les langues à plus faible empreinte numérique ou moins représentées dans les jeux d'entraînement de l'IA peuvent être plus vulnérables à des résultats de mauvaise qualité. Cela crée un déséquilibre où les utilisateurs anglophones bénéficient de systèmes plus précis et optimisés, tandis que les utilisateurs multilingues subissent des taux plus élevés d'erreurs et de limitations. En ce sens, les systèmes d'IA peuvent involontairement renforcer **les inégalités linguistiques existantes dans les espaces numériques.**

Une autre observation intéressante est le contraste entre les systèmes d'IA textuels et les systèmes d'IA de génération d'images. Les modèles fondés sur le texte sont souvent capables de produire une écriture multilingue très cohérente, tandis que les générateurs d'images peinent à reproduire avec exactitude même de simples phrases. Cela laisse penser que générer la langue visuellement est fondamentalement différent de générer la langue textuellement.

Les captures d'écran de générations incorrectes ont révélé des schémas qu'il aurait été difficile d'expliquer théoriquement. Certaines images contenaient des panneaux avec des mots presque corrects, tandis que d'autres produisaient une typographie qui paraissait réaliste mais était totalement illisible. Dans plusieurs cas, l'IA a ignoré des instructions explicites concernant le choix de la langue ou l'exactitude orthographique. Ces échecs visuels ont démontré que, malgré l'apparence impressionnante des systèmes d'IA générative, d'importantes limites techniques demeurent non résolues.

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

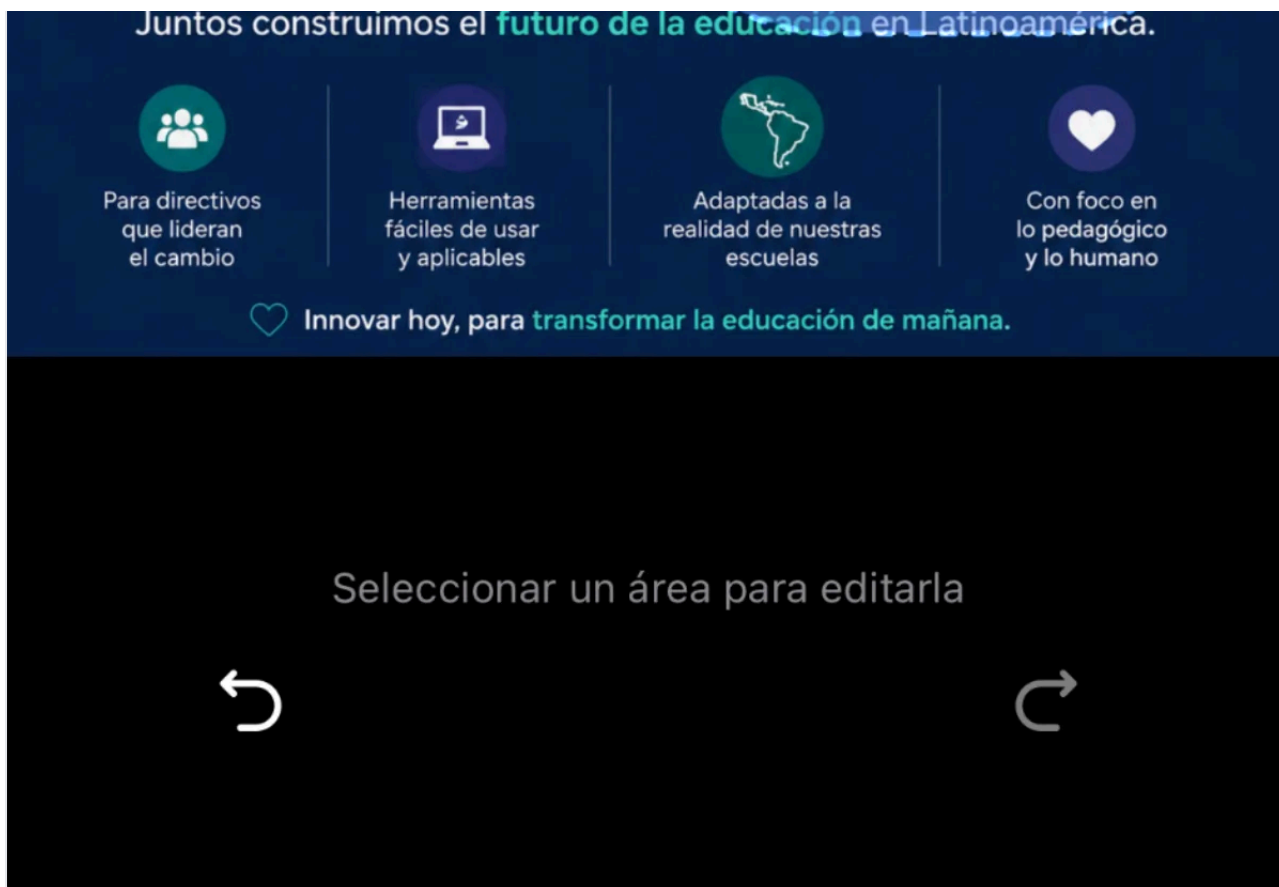


IA segura, ética y alineada a valores educativos



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

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



*Des instructions explicites de langue et d'orthographe, ignorées. La limite technique n'est pas esthétique — elle est linguistique, et elle est répartie de manière inégale entre les langues du monde.*

Au final, cette expérience a changé ma perception de la fiabilité de l'IA. Bien que l'IA soit souvent présentée comme une solution universelle capable de remplacer le travail humain dans les industries créatives, ces expérimentations ont montré que les systèmes actuels peinent encore avec la précision linguistique, en particulier dans les contextes multilingues et multimodaux. Cela souligne l'importance d'utiliser l'IA de manière critique et de vérifier soigneusement les résultats, surtout lorsqu'on travaille à travers différentes langues.

En même temps, j'ai été sincèrement surprise par la qualité impressionnante de nombreuses images générées et résultats créatifs. La technologie est indéniablement puissante et évolue rapidement, même si les défis liés à l'exactitude multilingue et à la fiabilité textuelle demeurent d'importants axes d'amélioration.

## À PROPOS DE L'AUTEURE

## Marie Prieto

Marie Prieto est chercheuse et rédactrice qui s'intéresse à la façon dont les systèmes d'IA fonctionnent — et échouent — à travers les langues du monde hispanophone et francophone. Cet article est une contribution d'Ambassadrice LATAM à Pillar Insights.

[Lire d'autres articles Pillar Insights \(/fr/perspectives\)](#)

## Foire aux questions.

---

### **Pourquoi les générateurs d'images par IA peinent-ils avec le texte à l'intérieur des images ?**

Les générateurs d'images par IA ne semblent pas traiter la langue écrite à l'intérieur des images comme une langue en soi. Ils traitent les lettres et les mots comme des motifs visuels ou des textures esthétiques. À distance, le texte généré paraît convaincant parce que l'espacement, la typographie et le rythme visuel imitent la structure d'une véritable écriture. À y regarder de plus près, les mots sont fréquemment incorrects, brouillés ou totalement fictifs. C'est pourquoi les affiches et la signalétique issues de modèles d'images peuvent paraître crédibles en taille vignette et s'effondrer en pleine résolution. Le modèle sait à quoi ressemble l'écriture, mais il ne sait pas ce qu'elle signifie.

### **Pourquoi la génération d'images par IA est-elle plus performante en anglais qu'en espagnol, en français ou en portugais ?**

Les prompts en anglais produisent généralement un texte plus cohérent à l'intérieur des images générées que les prompts en français, en espagnol ou en portugais. Même lorsque l'IA comprend le sens général d'une demande non anglophone, le résultat visuel contient fréquemment des fautes d'orthographe, des accents manquants, une grammaire incorrecte, ou des mélanges de plusieurs langues à la fois. La cause est le déséquilibre des jeux de données. Les modèles d'images par IA sont fortement influencés par la domination des données en langue anglaise dans leurs jeux d'entraînement, de sorte que les langues non anglophones héritent de taux d'erreurs plus élevés. Les systèmes d'IA commercialisés comme multilingues privilégient encore l'anglais en pratique.

## Que révèle la faute « educasion » sur les modèles de langue d'IA ?

Lorsqu'un générateur d'images par IA a écrit « educasion » au lieu d'« educación », il a reproduit une erreur orthographique courante que l'on trouve dans certaines parties de l'Amérique latine, où la prononciation seseo confond les sons c et s. En espagnol péninsulaire, la terminaison ción se prononce plus distinctement. La faute laisse penser que le modèle n'apprend pas seulement le vocabulaire écrit de l'espagnol, mais absorbe aussi les tendances phonétiques d'accents régionaux spécifiques. L'IA a appris la sonorité de l'espagnol latino-américain sans apprendre l'orthographe normalisée qui l'accompagne.

---

## Que se passe-t-il lorsqu'on demande à l'IA de corriger le texte à l'intérieur d'une image générée ?

Demander à un modèle d'images par IA de corriger un texte brisé à l'intérieur d'une image précédemment générée produit rarement une correction cohérente. Dans un cas révélateur, le modèle a régénéré une image et, au lieu de réparer les mots, s'est contenté de barrer le texte incorrect. Le système semblait reconnaître qu'il y avait un problème à un niveau superficiel, mais lui manquait la compréhension linguistique nécessaire pour réellement le corriger. Ce schéma renforce la conclusion selon laquelle les générateurs d'images ne traitent pas le texte intégré comme une langue. Ils le traitent comme un élément visuel qui peut être remplacé ou barré, et non comme une suite de caractères porteuse de sens.

---

## Pourquoi les résultats des images par IA sont-ils si imprévisibles d'une génération à l'autre ?

Le même prompt peut produire des résultats radicalement différents d'une génération à l'autre. Dans certaines tentatives, le texte intégré est correct ou presque correct. Dans d'autres, il est illisible. Même avec des prompts affinés, des précisions supplémentaires ou des instructions simplifiées, les générateurs d'images par IA continuent de peiner avec la précision textuelle de l'espagnol à l'intérieur des images. Cette incohérence rend les outils difficiles à utiliser de façon fiable pour un travail professionnel ou universitaire sans correction humaine manuelle ultérieure dans un logiciel externe.

---

## Comment cela affecte-t-il l'accessibilité et la représentation des langues non anglophones ?

L'écart de génération de texte soulève des questions plus larges sur l'accessibilité et la représentation. Les langues à plus faible empreinte numérique ou moins représentées dans les jeux d'entraînement de l'IA sont plus vulnérables à des résultats de mauvaise qualité. Les utilisateurs anglophones bénéficient de systèmes plus précis et optimisés, tandis que les utilisateurs multilingues subissent des taux plus élevés d'erreurs et de limitations. Les systèmes d'IA peuvent donc involontairement renforcer les inégalités linguistiques existantes dans les espaces numériques. C'est la dynamique même que Pillar a été conçu pour contrer au niveau de l'infrastructure : résorber l'écart de qualité entre l'anglais et les langues du Sud global.

---

Si les modèles d'images par IA ne savent pas orthographier « educación », le même écart de jeux de données façonne quelles langues obtiennent un immobilier numérique premium. Pillar gère plus de 100 000 propriétés premium à travers l'espagnol, le français, le portugais et les langues du Sud global.