Campo / Field	Contenido / Content
Tipo de contribución / Article type	Artículo original de investigación
Título en español	Ideación visual mediada por inteligencia artificial generativa: diversidad léxica, complejidad de prompts y configuraciones de generación en DiffusionDB
Title in English	Visual ideation mediated by generative artificial intelligence: Lexical diversity, prompt complexity, and generation settings in DiffusionDB
Autorías / Authors	MARIA FERNANDA RODRIGUEZ-RUIZ — UNIVERSIDAD DE GUAYAQUIL, Ecuador — ORCID: https://orcid.org/0009-0005-0069-8618
Correspondencia / Correspondence	MARIA FERNANDA RODRIGUEZ-RUIZ — mafer.rodriguez@asesoreseducativos-ec.com

Resumen

La escritura de prompts constituye una fase observable de la ideación en los sistemas generativos texto-imagen, pero su diversidad, recurrencia y relación con las configuraciones de generación requieren una delimitación que evite equiparar complejidad textual con creatividad psicológica. El objetivo fue caracterizar la ideación visual expresada en los prompts públicos de DiffusionDB mediante su diversidad léxica, elaboración estructural, recurrencia y asociación con parámetros de generación. Se realizó un estudio observacional transversal de datos secundarios. Se auditaron los 2.000.000 de registros de DiffusionDB 2M; el análisis textual utilizó 1.999.397 prompts no vacíos, 1.522.692 formulaciones únicas normalizadas y una submuestra determinista de 100.000 prompts para métricas léxicas. La mediana fue de 21 palabras y cinco segmentos. El 23,84 % de los registros correspondió a formulaciones repetidas. Los marcadores explícitos más frecuentes fueron medio o técnica (57,35 %), estilo o artista (55,28 %) y calidad o detalle (45,46 %). El índice de elaboración presentó mediana de dos categorías; 60,65 % de los registros combinó dos o más recursos visuales. Las correlaciones entre longitud y parámetros fueron pequeñas dentro de usuario (r = 0,12–0,16) y moderadas entre usuarios (rho = 0,33–0,51). Los prompts funcionaron como ensamblajes compositivos relativamente ricos, aunque apoyados en fórmulas recurrentes. Los resultados describen prácticas tempranas de ideación con Stable Diffusion, no creatividad humana ni calidad visual.

Palabras clave: inteligencia artificial generativa; ideación visual; ingeniería de prompts; creatividad computacional; análisis de contenido; DiffusionDB.

Abstract

Prompt writing is an observable stage of ideation in text-to-image generative systems; however, its diversity, recurrence, and relationship with generation settings must be examined without equating textual complexity with psychological creativity. This study characterized visual ideation expressed in public DiffusionDB prompts through lexical diversity, structural elaboration, recurrence, and associations with generation parameters. We conducted a cross-sectional observational study using secondary public data. All 2,000,000 DiffusionDB 2M records were audited. Text analyses included 1,999,397 non-empty prompts, 1,522,692 unique normalized formulations, and a deterministic sample of 100,000 prompts for lexical metrics. Prompts had a median of 21 words and five segments. Repeated formulations accounted for 23.84% of records. The most frequent explicit markers were medium or technique (57.35%), style or artist (55.28%), and quality or detail (45.46%). The elaboration index had a median of two categories, and 60.65% of records combined at least two visual resources. Correlations between prompt length and generation settings were small within users (r = 0.12–0.16) and moderate across users (rho = 0.33–0.51). Prompts operated as comparatively rich compositional assemblages while relying on recurrent formulas. These findings describe early Stable Diffusion ideation practices; they do not measure human creativity or visual quality.

Keywords: generative artificial intelligence; visual ideation; prompt engineering; computational creativity; content analysis; DiffusionDB.

1. Introducción

Los modelos generativos texto-imagen transformaron el lenguaje natural en una interfaz para producir representaciones visuales. Sistemas basados en difusión latente redujeron el costo computacional de la síntesis de alta resolución y ampliaron el control semántico sobre la imagen (Rombach et al., 2022), mientras propuestas contemporáneas como DALL-E 2 e Imagen mostraron que las representaciones lingüísticas podían guiar composiciones complejas y fotorealistas (Ramesh et al., 2022; Saharia et al., 2022). En este entorno, el prompt no es una instrucción meramente técnica: es una representación verbal de sujetos, relaciones, estilos, encuadres, atmósferas y criterios de acabado que el usuario intenta trasladar a un espacio visual.

La formulación eficaz de prompts exige aprender cómo responde el sistema a distintas palabras, secuencias y modificadores. Liu y Chilton (2022) mostraron que los usuarios descomponen objetivos visuales en componentes como sujeto, estilo, iluminación y composición. Esta práctica combina ensayo, evaluación y reformulación, pero también depende de expresiones compartidas que circulan como recetas. Herramientas como PromptMagician recuperan pares similares de DiffusionDB y recomiendan términos relevantes para apoyar ese refinamiento (Feng et al., 2023). La ingeniería de prompts puede entenderse, por tanto, como una habilidad creativa situada que articula intención, conocimiento del modelo y capacidad para seleccionar resultados (Oppenlaender et al., 2024).

El papel de la inteligencia artificial generativa en la creatividad permanece abierto a interpretaciones contrapuestas. Los sistemas pueden ampliar productividad y acceso a la ejecución visual, pero también introducir fijación, convergencia y dependencia de soluciones de alta probabilidad. Epstein et al. (2023) advirtieron que la evaluación de arte generativo requiere distinguir capacidades técnicas, agencia humana y consecuencias sociales. En tareas narrativas, la asistencia generativa elevó evaluaciones individuales al mismo tiempo que redujo la diversidad colectiva de los productos (Doshi & Hauser, 2024). En ideación visual, la exposición a imágenes generadas puede favorecer la fijación en ciertos atributos y limitar el pensamiento divergente (Wadinambiarachchi et al., 2024).

La evidencia observacional sobre plataformas artísticas refuerza esa tensión. Zhou y Lee (2024) encontraron aumentos de productividad y valoración después de adoptar herramientas texto-imagen, junto con reducciones de la novedad visual promedio. Sus resultados sugieren que la ideación y el filtrado humano continúan siendo decisivos, aunque los flujos de trabajo puedan converger hacia contenidos y estilos similares. En consecuencia, más producción o mayor complejidad de una instrucción no equivalen necesariamente a mayor creatividad.

DiffusionDB ofrece una base empírica para estudiar esta interfaz verbal. El conjunto contiene prompts especificados por usuarios, metadatos de generación e imágenes producidas con Stable Diffusion (Wang et al., 2022). Sus autores describieron propiedades sintácticas y semánticas, errores asociados con configuraciones y usos potencialmente dañinos. Investigaciones posteriores analizaron la originalidad léxica, temática y secuencial de los prompts, señalando que las decisiones de los usuarios contribuyen a la diversidad o a la homogeneización del contenido visual (De Rosa Palmini & Cetinic, 2024).

El vacío relevante no consiste en asumir que los prompts no han sido estudiados, sino en integrar cuatro dimensiones sobre el censo DiffusionDB 2M: recurrencia exacta de formulaciones, diversidad léxica sobre una submuestra reproducible, presencia de recursos visuales explícitos y asociación con parámetros de generación, separando variación dentro y entre usuarios. Esta separación permite observar prácticas de ideación sin atribuirles una medición psicológica de creatividad ni inferir calidad estética a partir del texto.

El objetivo fue caracterizar la ideación visual expresada en los prompts públicos de DiffusionDB mediante el análisis de su diversidad léxica, complejidad estructural, recurrencia y asociación con configuraciones de generación. Se planteó la pregunta: ¿cómo se caracterizan la diversidad léxica, la elaboración estructural y la recurrencia de la ideación visual expresada en DiffusionDB, y qué asociaciones observacionales presentan estos rasgos con las configuraciones seleccionadas por los usuarios?

2. Metodología

2.1. Diseño y fuente

Se realizó un estudio observacional transversal, retrospectivo, basado en datos secundarios públicos y análisis computacional de contenido. Se utilizó DiffusionDB 2M, distribuido bajo licencia CC0 1.0. La revisión fijada fue `fb620fbe49fa4420e0734bd9c0df11f51176b61f`, publicada el 22 de enero de 2024 y, por tanto, anterior al corte histórico del 15 de febrero de 2025. El archivo `metadata.parquet` posee 2.000.000 de filas y 13 variables: nombre de imagen, prompt, identificador de parte, semilla, pasos, escala de guía, muestreador, anchura, altura, usuario anonimizado, marca temporal y dos puntuaciones NSFW.

La copia de análisis se preservó con SHA-256 `EECD341187BC91C07F5994AD0660D40228EA025616FD57A509BEF8323677C68F`. La auditoría verificó dos millones de identificadores únicos, ausencia de filas exactamente duplicadas y 238 celdas nulas. Los registros cubren del 6 al 20 de agosto de 2022 y proceden de interacciones tempranas con Stable Diffusion recopiladas desde un servidor público de Discord, según la documentación de la fuente (Wang et al., 2022).

2.2. Poblaciones y depuración

Se distinguieron cinco poblaciones. El censo por registro incluyó dos millones de generaciones. El corpus textual excluyó 603 prompts vacíos después de retirar espacios exteriores, por lo que quedó constituido por 1.999.397 registros. Para analizar formulaciones únicas, el texto se convirtió a minúsculas y se compactaron espacios; esta normalización produjo 1.522.692 prompts únicos. La diferencia frente a 1.526.277 valores exactos no vacíos se debió exclusivamente a variantes de mayúsculas y espaciado.

Las métricas léxicas intensivas se calcularon sobre 100.000 prompts únicos elegidos de manera determinista mediante los menores valores de `pandas.util.hash_pandas_object`, con pandas 3.0.1. Para las asociaciones con parámetros se exigieron prompt no vacío, usuario disponible, pasos entre 1 y 150, escala de guía finita mayor que 0 y no superior a 30, y dimensiones de 64 a 2.048 píxeles. Este conjunto comprendió 1.991.320 registros de 10.174 usuarios anonimizados.

No se eliminaron nombres propios, referencias estilísticas ni modificadores porque eran señales sustantivas. Se definió token como una secuencia alfabética Unicode que podía contener apóstrofo o guion interno.

2.3. Variables textuales y recurrencia

La complejidad estructural se describió mediante caracteres, palabras y segmentos separados por comas o punto y coma. La recurrencia se estimó a partir de la frecuencia de cada prompt normalizado. Se informó el número de formulaciones empleadas una vez, entre dos y cuatro veces, y cinco o más veces.

La diversidad léxica incluyó tamaño del vocabulario, proporción de hapax, entropía de Shannon, diversidad de Simpson, relación tipo-token (TTR), índice de Guiraud y media de la relación tipo-token móvil con ventana de 25 palabras (MATTR). MATTR reduce la dependencia de la longitud respecto de TTR, aunque no elimina todos los efectos del género y la segmentación (Covington & McFall, 2010; McCarthy & Jarvis, 2010). Por ello, las métricas se interpretaron como propiedades del corpus y no como puntuaciones individuales de creatividad.

2.4. Recursos de ideación visual

Se operacionalizaron siete categorías no excluyentes mediante expresiones regulares documentadas: medio o técnica; estilo o artista; composición; iluminación; color; cámara o perspectiva; y calidad o detalle. Los lexicones incluyeron términos explícitos como `photography`, `oil painting`, `in the style of`, `rule of thirds`, `cinematic lighting`, `color palette`, `depth of field`, `highly detailed` y variantes ortográficas frecuentes. El índice de elaboración correspondió al número de categorías presentes, con rango de 0 a 7.

ISSN 2838-5527 - DOI: https://doi.org/10.64747/2mgg3g14 - p. 3

La validación contextual interna revisó individualmente 20 positivos por categoría y 100 controles negativos, seleccionados determinísticamente después de excluir `prompt_nsfw >= 0,5`. La precisión agregada de los positivos fue 97,1 % (136/140). La concordancia negativa fue 73 %, lo que confirmó alta precisión pero sensibilidad limitada. En consecuencia, las prevalencias se interpretaron como límites inferiores de marcadores explícitos, no como clasificación semántica exhaustiva.

2.5. Asociaciones y sensibilidad

Para reducir la pseudorreplicación se calcularon dos medidas. Primero, se centraron las variables respecto de la media de cada usuario y se obtuvo la correlación de Pearson entre longitud y escala de guía, pasos, megapíxeles y desviación absoluta de la relación de aspecto. Segundo, se agregaron los registros por usuario y se estimó la correlación de Spearman entre promedios individuales. Esta estrategia separó variaciones dentro de un mismo flujo de trabajo de diferencias entre usuarios.

No se realizaron pruebas de significación sobre el censo como sustituto de relevancia sustantiva. Se priorizaron magnitudes y consistencia. La sensibilidad comparó registros y prompts únicos, excluyó configuraciones anómalas y contrastó prevalencias después de retirar `prompt_nsfw >= 0,5`. La puntuación `image_nsfw = 2`, presente fuera del intervalo probabilístico, fue tratada como sentinel y no como probabilidad.

2.6. Ética y reproducibilidad

No se reclutaron participantes ni se intentó reidentificar usuarios. Los nombres de usuario habían sido transformados en identificadores hash por la fuente. Los resultados se presentaron agregadamente y no se reprodujeron prompts sensibles. El conjunto, el código, las reglas de limpieza, los lexicones, las versiones de dependencias y las tablas derivadas quedaron registrados en el expediente reproducible. Por las características descritas, no se requirió una nueva evaluación ética para el presente reanálisis.

3. Resultados

3.1. Integridad y conformación del corpus

La auditoría confirmó 2.000.000 de registros, 13 variables, dos millones de nombres de imagen únicos y ninguna fila exactamente duplicada. Se detectaron seis valores faltantes en `cfg`, 116 en usuario y 116 en marca temporal. También se identificaron nueve valores no finitos de `cfg`, 1.689 valores negativos, 529 superiores a 100, 50 registros con cero pasos y 24.868 valores `image_nsfw = 2`. Estas anomalías fueron excluidas solo de los análisis que requerían parámetros válidos.

Después de excluir 603 prompts vacíos quedaron 1.999.397 registros. La normalización produjo 1.522.692 formulaciones únicas. De estas, 1.326.793 aparecieron una sola vez, 167.238 entre dos y cuatro veces y 28.661 cinco o más veces. La máxima frecuencia de una misma formulación fue 1.014. En términos de registros, 23,84 % correspondió a una formulación ya observada.

3.2. Longitud y diversidad léxica

La mediana fue 21 palabras (rango intercuartílico: 11–34; percentil 95: 49) y cinco segmentos (rango intercuartílico: 2–10; percentil 95: 17). Esta estructura muestra que la práctica dominante no consistió en una descripción breve única, sino en secuencias de componentes agregados.

La submuestra de 100.000 prompts únicos reunió 2.218.757 tokens y 53.138 tipos. Los 22.806 hapax representaron 42,91 % del vocabulario. La entropía de Shannon fue 10,519 bits y la diversidad de Simpson 0,9944. La TTR media fue 0,9409 y el índice de Guiraud, 4,116. Entre 38.734 prompts con al menos 25 palabras, MATTR-25 alcanzó una media de 0,9267. Estas cifras indican un vocabulario amplio en la submuestra, pero no deben interpretarse como

originalidad individual porque los prompts son cortos y contienen numerosos nombres propios y modificadores raros.

3.3. Recursos visuales y elaboración

La Tabla 1 presenta las categorías operacionales. Medio o técnica y estilo o artista aparecieron en más de la mitad de los registros. Calidad o detalle se observó en 45,46 %, mientras iluminación se presentó en 22,01 %. Las categorías más específicamente compositivas —composición, cámara/perspectiva y color— mostraron prevalencias inferiores. Todas fueron menores en el corpus único que en los registros, lo que indica que parte de los marcadores formaba parte de prompts reutilizados.

Tabla 1. Recursos visuales explícitos en los prompts

Categoría	Registros, n	Registros, %	Prompts únicos, n	Prompts únicos, %	Mediana de palabras
Medio o técnica	1.146.748	57,35	848.427	55,72	28
Estilo o artista	1.105.313	55,28	799.390	52,50	30
Calidad o detalle	908.876	45,46	660.898	43,40	31
Iluminación	440.117	22,01	313.767	20,61	35
Composición	212.946	10,65	150.003	9,85	35
Cámara o perspectiva	150.176	7,51	109.227	7,17	32
Color	125.476	6,28	85.940	5,64	33

El índice de elaboración tuvo media de 2,05 y mediana de 2 categorías. El 18,37 % no activó ninguna categoría explícita; 20,97 % activó una; 21,38 %, dos; 21,73 %, tres; y 17,55 %, cuatro o más. En conjunto, 60,65 % incorporó al menos dos tipos de recursos visuales. La coexistencia fue descriptiva y no implicó que todos los componentes fueran coherentes entre sí o produjeran una imagen de mejor calidad.

3.4. Longitud y configuraciones de generación

La Tabla 2 muestra asociaciones consistentes pero de magnitud diferenciada. Dentro de un mismo usuario, la longitud presentó correlaciones pequeñas con `cfg`, pasos, resolución y desviación del formato (r = 0,12–0,16). Entre promedios de usuarios, las correlaciones fueron moderadas (rho = 0,33–0,51). La mayor correspondió a desviación de la relación de aspecto, seguida de escala de guía y resolución.

Tabla 2. Asociación entre palabras del prompt y configuraciones

Configuración	Registros	Usuarios	r dentro de usuario	rho entre usuarios
Escala de guía (cfg)	1.991.320	10.174	0,158	0,463
Pasos	1.991.320	10.174	0,121	0,329
Resolución en megapíxeles	1.991.320	10.174	0,122	0,450
Desviación de relación de aspecto	1.991.320	10.174	0,162	0,507

La distancia entre niveles indica que gran parte de la asociación se relacionó con estilos relativamente estables de uso: algunas personas escribieron prompts más largos y también seleccionaron configuraciones más intensivas o formatos menos cuadrados. El cambio de longitud dentro de un mismo usuario explicó una fracción mucho menor de la variación.

ISSN 2838-5527 - DOI: https://doi.org/10.64747/2mgg3g14 - p. 5

3.5. Sensibilidad

Al excluir prompts con puntuación NSFW igual o superior a 0,5, la diferencia máxima de prevalencia entre las siete categorías fue 0,346 puntos porcentuales. Por tanto, el patrón general no dependió de esa exclusión. La comparación entre registros y prompts únicos redujo todas las prevalencias principales, pero mantuvo el mismo orden. Las anomalías de configuración afectaron 8.077 registros del conjunto de asociaciones potenciales; su exclusión evitó que valores no plausibles dominaran las estadísticas sin alterar el censo textual.

4. Discusión

El estudio muestra que los prompts tempranos de Stable Diffusion funcionaron como ensamblajes compositivos: el usuario combinó sujetos implícitos con medios, referencias estilísticas, criterios de detalle, iluminación, composición y parámetros técnicos. Una mediana de 21 palabras, cinco segmentos y dos categorías explícitas indica una interfaz de ideación más cercana a una lista estructurada de decisiones que a una frase natural breve. Este patrón coincide con la descomposición de objetivos visuales descrita por Liu y Chilton (2022) y ayuda a explicar por qué las herramientas de recomendación recuperan términos y ejemplos previos para apoyar el refinamiento (Feng et al., 2023).

La elaboración coexistió con una recurrencia considerable. Casi una cuarta parte de los registros reutilizó una formulación normalizada y 28.661 prompts aparecieron cinco o más veces. Los marcadores más frecuentes —medio, estilo/artista y calidad/detalle— también fueron los que más disminuyeron al deduplicar. Esto respalda la interpretación de que la comunidad acumuló y compartió fórmulas prácticas. Oppenlaender et al. (2024) conceptualizaron la ingeniería de prompts como habilidad creativa; los resultados aquí añaden que esa habilidad puede ejercerse mediante exploración propia, apropiación de convenciones o combinación de ambas.

La amplitud del vocabulario no contradice la recurrencia. Una alta diversidad léxica global puede surgir de nombres de artistas, personajes, objetos raros y combinaciones de nicho, mientras muchas estructuras permanecen estables. De Rosa Palmini y Cetinic (2024) distinguieron originalidad léxica, temática y secuencial; esta distinción es fundamental porque un prompt puede contener palabras poco frecuentes y, al mismo tiempo, seguir una plantilla común. Por ello, las métricas de diversidad no deben convertirse en una puntuación de creatividad humana.

Las asociaciones con configuraciones sugieren estilos de interacción. Los usuarios que escribieron prompts más extensos tendieron también a emplear mayor escala de guía, resolución y desviación del formato; sin embargo, las correlaciones dentro de usuario fueron pequeñas. Esto impide afirmar que agregar palabras cause configuraciones más complejas o mejores resultados. La diferencia entre niveles es compatible con perfiles de uso relativamente estables: algunas personas adoptaron flujos de trabajo más elaborados en varias dimensiones.

Los hallazgos dialogan con la tensión entre expansión individual y convergencia colectiva. Dos millones de registros contienen un vocabulario amplio y numerosas combinaciones, pero también expresiones repetidas y modificadores dominantes. Doshi y Hauser (2024) observaron que la asistencia generativa puede elevar la creatividad evaluada individualmente mientras reduce la diversidad colectiva; Zhou y Lee (2024) documentaron mayor productividad y disminución de novedad visual promedio. Aunque esos estudios utilizan tareas y resultados diferentes, convergen en una advertencia: aumentar la capacidad de producción no garantiza ampliar el espacio colectivo de ideas.

En educación tecnológica y diseño, los resultados favorecen una alfabetización de prompts que no se limite a memorizar modificadores. Las actividades formativas deberían distinguir descripción semántica, decisión compositiva, referencia estilística, parámetro técnico y criterio de evaluación. También deberían promover comparación de alternativas y revisión crítica para reducir fijación, una preocupación respaldada por Wadinambiarachchi et al. (2024). Las interfaces pueden mostrar diversidad de rutas, explicar el efecto incierto de los términos y desalentar la copia automática de secuencias populares.

El estudio también plantea una implicación ética. Las referencias a artistas constituyeron una categoría dominante, pero el conjunto no permite saber si expresan homenaje, imitación, búsqueda de consistencia o desconocimiento de

las controversias sobre autoría. Epstein et al. (2023) recomendaron considerar atribución, trabajo creativo y gobernanza junto con la capacidad técnica. La alfabetización visual con IA debe incorporar esa dimensión y no tratar el estilo como un recurso neutral o ilimitado.

La contribución es metodológica y descriptiva. El análisis integra el censo de metadatos, deduplicación, diversidad léxica, recursos explícitos y asociaciones multinivel simples sobre una revisión histórica reproducible. No reemplaza estudios con imágenes, evaluadores humanos o diseños experimentales. Su valor reside en delimitar qué puede inferirse responsablemente de los prompts públicos.

5. Conclusiones

La ideación visual expresada en DiffusionDB 2M se caracterizó por prompts relativamente extensos, segmentados y compuestos por múltiples recursos visuales explícitos. Medio o técnica, referencia estilística y criterios de calidad/detalle dominaron la formulación. Al mismo tiempo, 23,84 % de los registros reutilizó una formulación ya observada, lo que revela la importancia de convenciones compartidas.

La diversidad léxica fue elevada en la submuestra de prompts únicos, pero coexistió con fórmulas recurrentes; por ello, diversidad de palabras, elaboración textual y creatividad no son equivalentes. Las asociaciones con parámetros fueron mayores entre usuarios que dentro de cada usuario, un patrón compatible con estilos estables de interacción y no con efectos causales de la longitud.

Para la formación en diseño y tecnología se recomienda enseñar los prompts como decisiones revisables: qué representar, con qué relaciones, mediante qué recursos visuales y bajo qué criterios de evaluación. Las herramientas deberían favorecer exploración y comparación en lugar de reforzar únicamente modificadores populares. Futuros estudios deben vincular prompts con imágenes, evaluar novedad y calidad mediante protocolos independientes y comparar plataformas, idiomas y periodos posteriores.

6. Limitaciones

DiffusionDB representa usuarios de un servidor específico durante 14 días de agosto de 2022 y una etapa temprana de Stable Diffusion; no representa todas las plataformas, comunidades o versiones. Los usuarios están anonimizados, pero no se dispone de información demográfica, experiencia, intención ni proceso iterativo completo. El análisis observa generaciones registradas, no borradores descartados ni criterios de selección.

Los lexicones ofrecieron alta precisión positiva, pero la concordancia negativa de 73 % evidenció sensibilidad incompleta ante errores ortográficos, sinónimos y referencias implícitas. Sus prevalencias son límites inferiores. La submuestra léxica fue determinista, aunque las métricas permanecen sensibles a tokenización, longitud y nombres propios. No se aplicó un modelo semántico multilingüe ni se validaron categorías con codificadores humanos independientes.

No se analizaron píxeles. En consecuencia, no se evaluaron novedad visual, coherencia texto-imagen, estética ni calidad gráfica. Las asociaciones son descriptivas y pueden reflejar confusión por experiencia, objetivos o hábitos no observados. La gran cantidad de registros no corrige esos límites ni convierte correlaciones pequeñas en efectos sustantivos.

Agradecimientos

No se declaran agradecimientos.

Declaraciones

Campo / Field	Contenido / Content
Contribuciones CRediT	Conceptualización; metodología; software; validación; análisis formal; investigación; curación de datos; visualización; redacción del borrador original; redacción, revisión y edición; administración del proyecto.
Financiamiento	La investigación no recibió financiación externa específica.
Conflictos de interés	La persona autora declara no tener conflictos de interés.
Aprobación ética	Reanálisis de datos secundarios públicos y anonimizados, sin reclutamiento ni intento de reidentificación. Por estas características no se requirió una nueva evaluación ética para el presente reanálisis.
Consentimiento	No aplicable al análisis agregado del conjunto público; no se reclutaron participantes ni se recopilaron datos nuevos.
Disponibilidad de datos	DiffusionDB: https://huggingface.co/datasets/poloclub/diffusiondb. El código, lexicones y tablas derivadas se conservan en el expediente reproducible.
Uso de inteligencia artificial	Codex de OpenAI apoyó la programación reproducible, la organización del manuscrito, la traducción académica y los controles editoriales. La persona autora revisó el contenido y asume la responsabilidad íntegra por la exactitud, originalidad, interpretación y versión final del artículo.

Referencias

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian knot: The moving-average type-token ratio (MATTR). *Journal of Quantitative Linguistics*, *17*(2), 94–100. https://doi.org/10.1080/09296171003643098

De Rosa Palmini, M.-T., & Cetinic, E. (2024). *Patterns of creativity: How user input shapes AI-generated visual diversity* [Preprint]. arXiv. https://arxiv.org/abs/2410.06768

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, *10*(28), eadn5290. https://doi.org/10.1126/sciadv.adn5290

Epstein, Z., Hertzmann, A., Akten, M., Farid, H., Fjeld, J., Frank, M. R., Groh, M., Herman, L., Leach, N., Mahari, R., Pentland, A. S., Russakovsky, O., Schroeder, H., & Smith, A. (2023). Art and the science of generative AI. *Science*, *380*(6650), 1110–1111. https://doi.org/10.1126/science.adh4451

Feng, Y., Wang, X., Wong, K. K., Wang, S., Lu, Y., Zhu, M., Wang, B., & Chen, W. (2023). PromptMagician: Interactive prompt engineering for text-to-image creation. *IEEE Transactions on Visualization and Computer Graphics*, 1–11. https://doi.org/10.1109/TVCG.2023.3327168

Liu, V., & Chilton, L. B. (2022). Design guidelines for prompt engineering text-to-image generative models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–23). Association for Computing Machinery. https://doi.org/10.1145/3491102.3501825

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods, 42*(2), 381–392. https://doi.org/10.3758/BRM.42.2.381

Oppenlaender, J., Linder, R., & Silvennoinen, J. (2024). Prompting AI art: An investigation into the creative skill of prompt engineering. *International Journal of Human–Computer Interaction*. Advance online publication. https://doi.org/10.1080/10447318.2024.2431761

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical text-conditional image generation with CLIP latents* [Preprint]. arXiv. https://arxiv.org/abs/2204.06125

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10674–10685). IEEE. https://doi.org/10.1109/CVPR52688.2022.01042

Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Seyed Ghasemipour, S. K., Ayan, B. K., Mahdavi, S. S., Lopes, R. G., Salimans, T., Ho, J., Fleet, D. J., & Norouzi, M. (2022). *Photorealistic text-to-image diffusion models with deep language understanding* [Preprint]. arXiv. https://arxiv.org/abs/2205.11487

Wadinambiarachchi, S., Kelly, R. M., Pareek, S., Zhou, Q., & Velloso, E. (2024). The effects of generative AI on design fixation and divergent thinking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1–18). Association for Computing Machinery. https://doi.org/10.1145/3613904.3642919

Wang, Z. J., Montoya, E., Munechika, D., Yang, H., Hoover, B., & Chau, D. H. (2022). *DiffusionDB: A large-scale prompt gallery dataset for text-to-image generative models* [Preprint]. arXiv. https://arxiv.org/abs/2210.14896

Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS Nexus, 3*(3), pgae052. https://doi.org/10.1093/pnasnexus/pgae052

Anexos

No se incluyen anexos.

Información editorial

Campo / Field	Contenido / Content
Recepción / Received	15 de febrero de 2025 / February 15, 2025
Revisión / Revised	14 de marzo de 2025 / March 14, 2025
Aceptación / Accepted	18 de abril de 2025 / April 18, 2025
Publicación / Published	29 de junio de 2025 / June 29, 2025
Volumen, número y año / Volume, issue and year	Vol. 1, Núm. 1, enero-junio de 2025