

	SERIE DE RESÚMENES DE CURSOS ABRA	Octubre 2024
	“SESIÓN 2: Principales herramientas de generación de imágenes con IA”	

1. Introducción a la sesión: Clase práctica de generación de imágenes y objetivos

Rodrigo Rojo abre la sesión dando una calurosa bienvenida a los asistentes y explicando que esta será una **sesión puramente práctica** en la que los participantes aprenderán a generar imágenes utilizando varias herramientas de **Inteligencia Artificial (IA)**. Tras una breve recapitulación de las clases anteriores, donde se abordaron temas teóricos sobre el funcionamiento de la IA en la generación de imágenes, Rodrigo explica que hoy será una oportunidad para **aplicar esos conocimientos** directamente en la práctica.

Rodrigo enfatiza que uno de los objetivos principales de la clase es que los participantes **se familiaricen** con varias plataformas de IA para la creación de imágenes, como **MidJourney, Leonardo y DALL-E**. Destaca que, a pesar de que estas herramientas tienen características y funcionalidades diferentes, es importante que los usuarios **experimenten** con todas ellas para descubrir cuál se adapta mejor a sus necesidades.

El instructor hace hincapié en que estas herramientas permiten a los usuarios generar **imágenes creativas** sin necesidad de tener habilidades avanzadas de diseño o ilustración. Según él, la IA abre nuevas oportunidades tanto para creadores amateurs como para profesionales que buscan **optimizar su flujo de trabajo**.

Puntos clave:

- Esta sesión será **totalmente práctica** y se centrará en la generación de imágenes con IA.
- Rodrigo busca que los participantes **exploren y experimenten** con diferentes plataformas para descubrir cuál les resulta más útil.
- Se recalca que la IA facilita la creación visual sin necesidad de **habilidades técnicas** avanzadas, democratizando el acceso a la creatividad.

2. Novedades recientes en el uso de herramientas de IA: Cambios significativos y nuevas funciones

Rodrigo se toma unos minutos para actualizar a los asistentes sobre los **cambios recientes** en el ámbito de la IA, destacando los **anuncios** más importantes de empresas como **Google y OpenAI** que impactan directamente en las herramientas de generación de imágenes. Menciona que las últimas actualizaciones han permitido que muchas funciones que antes estaban disponibles solo para usuarios de **cuentas pagas** se abran ahora para las **cuentas gratuitas** de ChatGPT y otras plataformas.

Por ejemplo, OpenAI lanzó el modelo **GPT-4 Turbo**, que incluye la capacidad de **generar imágenes directamente desde ChatGPT**. Este nuevo modelo está disponible para usuarios de pago, pero Rodrigo anticipa que **pronto se abrirá** para todos, permitiendo que más personas accedan a esta tecnología. Esto significa que las herramientas para crear imágenes estarán al alcance de un mayor número de personas, democratizando aún más el acceso a la IA.

Además, habla sobre el lanzamiento del modelo **Gemini** de Google, que promete competir con las herramientas actuales como **DALL-E y MidJourney**, ofreciendo funciones mejoradas para la creación de imágenes y otras aplicaciones de IA. Rodrigo menciona que aunque el modelo

	SERIE DE RESÚMENES DE CURSOS ABRA	Octubre 2024
	“SESIÓN 2: Principales herramientas de generación de imágenes con IA”	

Gemini todavía está en una fase de prueba limitada, su desarrollo muestra hacia dónde va la **evolución de la IA** en los próximos meses.

Puntos clave:

- **GPT-4 Turbo** y **Gemini** son nuevos modelos que permiten **generar imágenes** directamente en ChatGPT y otras plataformas, democratizando el acceso.
- Las funciones que antes estaban limitadas a **cuentas pagas** se están abriendo gradualmente a **usuarios gratuitos**.
- Google y OpenAI siguen siendo los líderes en innovación, y sus nuevos lanzamientos marcarán la pauta para el futuro de la **creación visual con IA**.

3. Herramientas de IA para la creación de imágenes: Características, diferencias y ejemplos

Rodrigo presenta en profundidad las **herramientas más populares** para la creación de imágenes con IA y explica las diferencias clave entre ellas. El objetivo es que los participantes comprendan cómo cada herramienta tiene sus **ventajas y limitaciones**, y cuándo conviene usar una sobre otra, dependiendo del tipo de proyecto que tengan en mente. Rodrigo divide su análisis en tres herramientas principales:

1. **MidJourney**: Descrita por Rodrigo como una de las plataformas más potentes para generar **imágenes fotorrealistas**. Explica que **MidJourney** ha ganado una enorme popularidad entre los creadores visuales gracias a su capacidad para crear imágenes extremadamente detalladas y con un **alto nivel de realismo**. Sin embargo, también menciona que esta herramienta tiene un costo, y que, aunque se puede usar de manera limitada en su versión gratuita, las características más avanzadas requieren una **suscripción**. Rodrigo sugiere utilizar MidJourney cuando se necesita una imagen rápida con **impacto visual**.
2. **Leonardo**: Esta herramienta está basada en **Stable Diffusion**, lo que significa que ofrece un **mayor control** sobre los elementos individuales de la imagen que se está creando. Rodrigo la recomienda para proyectos en los que se necesita una mayor **precisión en los detalles**. Además, menciona que **Leonardo** permite a los usuarios **ajustar parámetros finos**, como la iluminación, los colores, y la posición exacta de los objetos dentro de la imagen. Es una opción ideal para quienes buscan más **flexibilidad y personalización** en sus creaciones.
3. **DALL-E**: Desarrollada por **OpenAI**, **DALL-E** se ha integrado directamente en ChatGPT y permite generar imágenes de manera sencilla a partir de **prompts textuales**. Rodrigo destaca que esta herramienta es excelente para quienes buscan crear imágenes **rápidamente** sin necesidad de una configuración compleja. Sin embargo, también señala que DALL-E, aunque útil, no ofrece el mismo nivel de **control detallado** que Leonardo o MidJourney, por lo que es más adecuada para proyectos que no requieran un alto nivel de detalle.

Rodrigo hace énfasis en que todas estas herramientas son útiles en distintos contextos y que es importante conocer sus **fortalezas** para elegir la mejor opción según las necesidades del proyecto.

	SERIE DE RESÚMENES DE CURSOS ABRA	Octubre 2024
	“SESIÓN 2: Principales herramientas de generación de imágenes con IA”	

Puntos clave:

- **MidJourney** es ideal para generar imágenes **fotorrealistas** y tiene una **gran capacidad visual**, aunque requiere una suscripción para acceder a funciones avanzadas.
- **Leonardo** ofrece **mayor control** sobre los detalles y permite **ajustar parámetros finos**, ideal para quienes buscan personalizar sus creaciones.
- **DALL-E**, integrada en **ChatGPT**, es rápida y eficiente para generar imágenes **creativas** sin demasiada complejidad, aunque ofrece menos control sobre los detalles.

4. Generando imágenes con ChatGPT y Copilot: Demostración en vivo y primeras impresiones

Rodrigo dedica esta sección a una **demostración en vivo** de cómo generar imágenes utilizando **ChatGPT** y **Copilot**. Explica que **Copilot** es una funcionalidad de **Microsoft Office** que permite a los usuarios generar contenido visual y otros tipos de documentos directamente desde un prompt, integrando capacidades de IA en aplicaciones tradicionales como Word, Excel y PowerPoint.

Para demostrar la capacidad de estas herramientas, Rodrigo pide a los asistentes que le sugieran un tema para una imagen. Un participante propone crear una escena que incluya a personajes como **Pinocho, El Principito, Caperucita Roja y un perrito** caminando juntos por un escenario similar al camino amarillo del Mago de Oz. Rodrigo ingresa esta solicitud en **Copilot**, y rápidamente la herramienta genera una imagen basada en la descripción proporcionada.

El primer resultado incluye algunos personajes de manera aproximada, pero presenta **errores de interpretación** que son comunes en herramientas de IA que manejan múltiples elementos en una escena compleja. Rodrigo señala que es habitual que la IA **malinterprete algunos detalles** o incluya personajes que no estaban en el prompt inicial. Explica que este es un aspecto en el que las IAs siguen mejorando, pero que es importante **iterar y ajustar** los prompts para obtener mejores resultados.

Rodrigo también señala que, además de **generar la imagen**, los usuarios pueden continuar refinando su solicitud hasta llegar a una versión más precisa. La flexibilidad de herramientas como **Copilot Designer** permite a los usuarios realizar ajustes manuales a las imágenes generadas, mejorando aún más el resultado final.

Puntos clave:

- Rodrigo realiza una **demostración en vivo** con **ChatGPT** y **Copilot**, generando una imagen basada en un prompt con varios personajes.
- El primer resultado tiene algunos **errores de interpretación**, lo que es común en escenas complejas, pero se puede mejorar iterando en el prompt.
- **Copilot Designer** permite ajustar manualmente las imágenes generadas para hacerlas más precisas y adaptarlas a las necesidades del usuario.

	SERIE DE RESÚMENES DE CURSOS ABRA	Octubre 2024
	“SESIÓN 2: Principales herramientas de generación de imágenes con IA”	

5. Profundizando en el prompting: Cómo lograr mejores resultados con instrucciones detalladas

Uno de los conceptos más importantes de la clase es el **prompting**, y Rodrigo dedica una parte significativa del tiempo a enseñar a los participantes cómo formular **instrucciones claras y precisas** para obtener los mejores resultados posibles de las IAs generativas de imágenes. Explica que el **éxito** en la creación de imágenes depende en gran medida de la **calidad** del prompt que el usuario ingresa en la herramienta.

Rodrigo ilustra cómo un **prompt mal estructurado** puede generar resultados inesperados o poco precisos. Utiliza ejemplos para mostrar que, cuando no se especifican detalles clave, como la posición de los personajes, la IA puede asumir ciertas decisiones que no corresponden con lo que el usuario tenía en mente. Por el contrario, un prompt bien diseñado, que incluya **detalles visuales y emocionales**, puede producir imágenes mucho más cercanas a la visión del usuario.

Para lograr esto, Rodrigo recomienda incluir en el prompt descripciones claras sobre elementos como:

- **Posición de los personajes** en la escena.
- **Ropa** o vestimenta específica que deben llevar.
- **Interacción** entre los personajes, si se espera que se comuniquen o interactúen de alguna manera.
- **Colores predominantes** o esquemas de color deseados para la escena.
- **Estilo visual**, como ilustración, fotografía, pintura o dibujo.

Rodrigo también menciona que algunos **detalles emocionales** pueden mejorar la comprensión de la IA. Por ejemplo, términos como "caminando alegremente" o "mirando al horizonte con asombro" pueden ayudar a que la IA interprete la actitud o el estado de ánimo de los personajes.

Puntos clave:

- El **prompting** es esencial para obtener imágenes de alta calidad, y un prompt mal diseñado puede generar resultados inesperados.
- Es importante incluir **detalles específicos** sobre la posición, vestimenta, interacción y colores para que la IA interprete mejor la solicitud.
- Las **emociones** también pueden añadirse al prompt para ayudar a la IA a captar mejor la atmósfera de la escena.

6. Ejercicio práctico: Generar una imagen compleja con múltiples personajes y ajustes en los prompts

Rodrigo propone un **ejercicio práctico** en el que los participantes colaboran para generar una imagen compleja, utilizando **DALL-E** como herramienta principal. La idea es crear una imagen que incluya varios personajes en una escena detallada, para que los asistentes puedan ver cómo los prompts influyen en el resultado.

El ejercicio comienza con un **prompt simple** que describe a **El Principito, Pinocho, Caperucita Roja y un perrito** caminando por el famoso camino amarillo del **Mago de Oz**. Sin embargo, el

	SERIE DE RESÚMENES DE CURSOS ABRA	Octubre 2024
	“SESIÓN 2: Principales herramientas de generación de imágenes con IA”	

primer resultado no es satisfactorio. Los personajes aparecen desproporcionados y algunos detalles visuales no coinciden con lo esperado. Rodrigo utiliza este ejemplo para explicar que las IAs todavía tienen dificultades cuando se trata de **manejar múltiples personajes** en una misma imagen.

Para mejorar el resultado, Rodrigo ajusta el prompt añadiendo detalles específicos. Por ejemplo:

- El **tipo de ropa** que debe llevar cada personaje.
- La posición exacta de los personajes en la escena (por ejemplo, "Pinocho en el centro, El Principito a su izquierda").
- La **interacción** entre los personajes, describiendo cómo deben caminar juntos y mirar hacia el horizonte.
- Descripción detallada del **perrito** y su comportamiento, indicando que debe caminar junto a los personajes.

Cada vez que ajusta el prompt, Rodrigo genera una nueva versión de la imagen, mostrando cómo las **iteraciones** ayudan a mejorar gradualmente el resultado. También menciona que, aunque las herramientas de IA son poderosas, es normal tener que realizar varias iteraciones antes de obtener una imagen que cumpla con las expectativas.

Puntos clave:

- Rodrigo guía a los participantes a través de un **ejercicio práctico**, generando una imagen compleja con múltiples personajes.
- Los resultados mejoran con cada **iteración**, a medida que se ajustan los detalles del prompt.
- Las herramientas de IA aún tienen limitaciones al manejar escenas complejas, pero la **iteración constante** es clave para mejorar los resultados.

7. Desafíos de la IA en la generación de imágenes detalladas: Limitaciones comunes

Rodrigo dedica una sección importante de la clase a discutir las **limitaciones actuales** de la IA en la creación de imágenes detalladas y complejas. A pesar de los avances recientes, las IAs aún presentan ciertos **desafíos** que los usuarios deben tener en cuenta al generar imágenes. Estos desafíos pueden frustrar a los usuarios nuevos, pero Rodrigo explica que, con la práctica, se pueden mitigar.

1. **Proporciones y anatomía:** Aunque las IAs han mejorado significativamente en la generación de imágenes humanas, todavía cometen errores cuando se trata de **proporciones anatómicas**. Esto es especialmente evidente en las manos, donde la IA a menudo produce dedos adicionales o deformidades. Rodrigo muestra ejemplos de imágenes donde las manos y los pies aparecen desproporcionados o con formas incorrectas.
2. **Inconsistencias en los detalles:** Las herramientas de IA a veces generan **detalles ilógicos o inconsistentes**, especialmente cuando se maneja una escena compleja con muchos elementos visuales. Por ejemplo, la IA puede generar objetos que parecen estar flotando en el aire o sombras que no tienen sentido en la escena. Esto ocurre porque la IA todavía no comprende completamente el contexto tridimensional de la imagen.

	SERIE DE RESÚMENES DE CURSOS ABRA	Octubre 2024
	“SESIÓN 2: Principales herramientas de generación de imágenes con IA”	

3. **Carga semántica mal interpretada:** En algunos casos, la IA da más importancia a ciertas palabras del prompt y genera imágenes que no reflejan con precisión la intención del usuario. Rodrigo explica que esto ocurre cuando la IA no puede identificar qué elementos son los más importantes en el prompt. Por ejemplo, en el ejercicio de El Principito y Pinocho, la IA puede concentrarse demasiado en elementos del paisaje y descuidar los personajes.

Rodrigo señala que estos problemas son comunes, pero que es importante **ajustar el prompt** y tener paciencia. Con la evolución de las herramientas de IA, estos desafíos serán menos frecuentes, pero mientras tanto, el ajuste manual y la **revisión crítica** de las imágenes generadas son fundamentales para obtener buenos resultados.

Puntos clave:

- La IA aún tiene dificultades para generar proporciones anatómicas **correctas**, especialmente con partes del cuerpo como las manos.
- Las **inconsistencias visuales** en detalles como sombras, objetos flotantes o proporciones incorrectas siguen siendo un desafío.
- La **carga semántica** puede causar que la IA dé prioridad a ciertos elementos sobre otros, generando resultados inesperados.

8. Reflexión final y próximos pasos: Lo que hemos aprendido y qué esperar en las siguientes sesiones

Para concluir la clase, Rodrigo reflexiona sobre todo lo que los asistentes han aprendido durante la sesión práctica. Desde los fundamentos del **prompting** hasta la creación de imágenes complejas con múltiples personajes, los participantes han podido experimentar de primera mano con las herramientas de IA y comprender tanto sus **fortalezas** como sus **limitaciones**.

Rodrigo recalca que el uso de IA para la generación de imágenes es una habilidad que requiere **práctica y experimentación** constante. Aunque las herramientas actuales son muy poderosas, el verdadero dominio de estas plataformas viene con la experiencia y el aprendizaje de cómo ajustar y mejorar los resultados con cada iteración. Aconseja a los participantes que sigan practicando fuera de la clase y que exploren otras plataformas de IA para seguir perfeccionando su técnica.

Finalmente, menciona que en la **próxima sesión** se profundizará en **estrategias avanzadas de prompting**, abordando cómo formular prompts aún más detallados y creativos para obtener resultados de calidad profesional. Rodrigo invita a los participantes a revisar los materiales compartidos y estar preparados para **aplicar técnicas más complejas** en las siguientes clases.

Puntos clave:

- Los participantes han aprendido los **fundamentos del prompting** y han experimentado con herramientas como **DALL-E** y **Copilot**.
- Rodrigo anima a seguir practicando fuera de clase y a perfeccionar las habilidades de **creación de imágenes** con IA.

	SERIE DE RESÚMENES DE CURSOS ABRA	Octubre 2024
	“SESIÓN 2: Principales herramientas de generación de imágenes con IA”	

- En la próxima sesión, se abordarán **estrategias avanzadas de prompting** para mejorar aún más los resultados creativos.