ACADEMIA HUMANITARIA · HUMANITAIRES JOBS

Mini MBA IA y acción humanitaria: trabajar con Claude y Gemini

Material del curso · Semana 1 : Comprender los asistentes de IA generativa: Claude, Gemini y sus límites

Un programa de exactamente 3 meses (12 semanas, 144 horas) para utilizar los asistentes de IA Claude (Anthropic) y Gemini (Google) en el trabajo humanitario, de forma útil, segura y responsable: funcionamiento y límites, redacción de peticiones, protección de datos, ética y principios humanitarios, informes, propuestas, análisis de datos y MEAL, traducción y comunicación con las comunidades, investigación y verificación de fuentes, desinformación y seguridad, productividad y automatización, política de uso y despliegue en una organización. Ejercicios corregidos, tres exámenes mensuales con estudios de caso y un diploma verificable de la Academia Humanitaria. Programa independiente, sin asociación con Anthropic ni con Google.

https://humanitairejobs.org/es/academia/mini-mba-ia-claude-gemini/semaine-1/

Documento gratuito de la Academia Humanitaria. Puede imprimirse y compartirse libremente con fines de formación, sin modificaciones.

Semana 1 : Comprender los asistentes de IA generativa: Claude, Gemini y sus límites

Objetivos

- Explicar de forma sencilla cómo un gran modelo de lenguaje produce una respuesta y por qué puede equivocarse con total seguridad.
- Describir las principales funciones de Claude (Anthropic) y de Gemini (Google), y las diferencias entre versiones gratuitas, de pago y profesionales.
- Distinguir las tareas en las que la IA realmente ayuda a un trabajador humanitario de aquellas en las que es arriesgada o inútil.

Qué es un asistente de IA generativa

Claude, desarrollado por la empresa estadounidense **Anthropic**, y Gemini, desarrollado por **Google**, son **asistentes conversacionales** basados en grandes modelos de lenguaje (en inglés *large language models*, LLM). Se les escribe una petición en lenguaje corriente; responden con texto y, según la versión, con tablas, código, imágenes o documentos.

Un modelo de lenguaje ha sido entrenado con enormes cantidades de texto para **predecir la continuación más probable** de un texto, fragmento a fragmento (los «tokens», unas tres cuartas partes de una palabra en inglés, a menudo menos en francés o en árabe). Después se ajustó para seguir instrucciones, rechazar ciertas peticiones y adoptar un tono útil.

Consecuencia práctica: cuando responde de memoria, el modelo no «consulta» una base de hechos verificados. Produce un texto **plausible**. Muy a menudo, plausible y exacto coinciden; a veces no.

> Este programa es independiente. La Academia Humanitaria no tiene ninguna asociación con Anthropic ni con Google. Las funciones descritas son las conocidas en noviembre de 2026; estas herramientas cambian rápido: compruebe siempre las condiciones y los ajustes vigentes.

Las «alucinaciones»: errores dichos con seguridad

Se llama **alucinación** a una respuesta falsa presentada como verdadera: una cifra inventada, una referencia a un artículo que no existe, una resolución de la ONU mal citada, una organización confundida con otra. El tono es tan seguro como el de una respuesta exacta.

Los riesgos aumentan cuando se piden:

- **cifras precisas** (número de personas desplazadas, montos, fechas) sin proporcionar la fuente;
- **referencias** (títulos de informes, artículos de ley, números de resoluciones);

- información **reciente**, posterior a la fecha de fin de los datos de entrenamiento, si la búsqueda web no está activada;
- temas **locales o poco comunes**: un municipio de un país poco documentado, una lengua con poca presencia en Internet.

Los riesgos disminuyen cuando uno mismo **proporciona el documento fuente** (informe, tabla, acta) y pide al modelo que trabaje únicamente a partir de él, citando los pasajes utilizados.

Ejemplo. Ante la pregunta «¿Cuántas personas desplazadas hay en la prefectura X en 2025?», un asistente sin búsqueda web puede dar una cifra creíble y falsa. La buena práctica: proporcionar la última tabla publicada por la fuente oficial (por ejemplo la Matriz de Seguimiento de los Desplazamientos de la OIM) y pedir una síntesis de esa tabla.

Claude y Gemini: panorama de funciones

En noviembre de 2026, ambos asistentes ofrecen funciones comparables, con nombres diferentes:

- **Conversación** en la web, en el ordenador y en el teléfono, en numerosas lenguas, entre ellas el francés, el inglés, el español, el portugués y el árabe.
- **Análisis de archivos**: PDF, documentos, imágenes, hojas de cálculo. Ambos pueden leer un informe largo y extraer lo esencial.
- **Búsqueda web** con enlaces a las fuentes, y un modo de **investigación en profundidad** que consulta numerosas páginas para producir un informe.
- **Espacios de trabajo reutilizables**: los *Proyectos* de Claude (documentos de referencia e instrucciones comunes a varias conversaciones); los *Gems* de Gemini (asistentes personalizados con instrucciones).
- **Creación de documentos y de código**: los *artefactos* de Claude, el *Canvas* de Gemini.
- **Integración con herramientas**: Gemini está integrado en Google Workspace (Gmail, Docs, Sheets, Meet) según la suscripción; Claude puede conectarse a servicios como Google Drive o Gmail mediante conectores, según la suscripción.

Google ofrece también **NotebookLM**, una herramienta que responde únicamente a partir de los documentos que usted sube, con citas: muy útil para trabajar sobre un conjunto de informes.

Ambos existen en **versión gratuita** (límites de uso), como **suscripción individual** y como **planes para organizaciones** (equipos, empresas, educación), con compromisos distintos sobre los datos. La semana 3 detalla este punto esencial.

Lo que la IA hace bien y lo que hace mal

Para un trabajador humanitario, la IA generativa es sobre todo un **asistente de redacción, síntesis y estructuración**. Hace bien:

- resumir un documento largo que usted proporciona;

- proponer un primer borrador (nota, correo, esquema de informe, términos de referencia) a partir de sus elementos;
- reformular para un público diferente (donante, comunidad, autoridades);
- traducir un texto corriente entre lenguas bien dotadas, con revisión;
- explicar una fórmula de hoja de cálculo, corregir una macro, estructurar datos;
- actuar como revisor crítico: «¿qué debilidades vería un donante en esta sección?».

Hace mal, o de forma arriesgada:

- proporcionar hechos, cifras y referencias sin fuente;
- decidir en lugar de un responsable (selección de beneficiarios, contratación, nivel de seguridad);
- tratar situaciones en las que un error puede dañar gravemente a una persona;
- comprender un contexto local sutil sin que se lo expliquen.

> Regla de oro del programa: la IA prepara, el ser humano verifica y decide. Usted sigue siendo responsable de todo lo que firma, envía o publica.

Ventana de contexto, memoria y fecha de conocimiento

Tres nociones evitan muchos malentendidos:

- **Ventana de contexto**: la cantidad de texto que el modelo puede tener en cuenta en una conversación (sus mensajes, sus respuestas y los archivos). Es muy grande en Claude y Gemini (varios cientos de páginas), pero no ilimitada; en una conversación muy larga, los detalles del principio pueden tenerse menos en cuenta. Para un tema nuevo, abra una conversación nueva.
- **Memoria**: por defecto, una conversación no conoce las demás. Algunas versiones ofrecen memoria entre conversaciones, que se puede activar, consultar o borrar en los ajustes. No suponga nunca que el asistente «recuerda» a su organización.
- **Fecha de conocimiento**: el modelo conoce el mundo hasta una fecha determinada. Más allá, no sabe nada, salvo si hace una búsqueda web. Para una crisis en curso, active la búsqueda o proporcione sus documentos.

Ejemplo. Un informe de situación de 40 páginas cabe sin dificultad en la ventana de contexto. El expediente completo de un programa de tres años (cientos de documentos) no cabe: hay que seleccionar los documentos útiles, o utilizar una herramienta pensada para un corpus (Proyectos de Claude, NotebookLM).

Elegir entre Claude y Gemini

No existe un asistente «mejor» en términos absolutos. Ambos son de muy buen nivel y evolucionan cada pocos meses. Los criterios útiles para una organización humanitaria:

- **Su entorno de trabajo**: una organización que ya usa Google Workspace dispone de Gemini directamente en Gmail, Docs y Sheets; una organización con Microsoft 365 utilizará

más bien las integraciones disponibles para ese entorno, o las aplicaciones web.

- **Las condiciones de protección de datos** del plan accesible a su organización (semana 3).
- **La calidad en sus tareas reales**: pruebe ambos en 5 a 10 tareas representativas (una nota, una síntesis, una traducción al árabe, un análisis de tabla) y compare con una rejilla sencilla.
- **El coste** y los posibles descuentos para organizaciones sin ánimo de lucro (a verificar con cada proveedor).
- **La disponibilidad en su país**: algunos servicios no están disponibles en todas partes; compruébelo antes de construir un proceso sobre ellos.

Muchos profesionales utilizan ambos: uno como herramienta principal, el otro para una segunda lectura de un texto importante.

Ejercicio: detectar los riesgos en una respuesta de IA

Situación. Una oficial de programas pide a un asistente, sin búsqueda web ni documentos: «Redacta un párrafo de contexto sobre la inseguridad alimentaria en la región de Kaga para una propuesta a un donante». La respuesta cita «1,2 millones de personas en fase 3 o superior según la CIF de marzo de 2026», menciona «el informe de 2025 del PMA titulado *Kaga Food Outlook*» y concluye que «los hogares encabezados por mujeres están un 40 % más afectados».

Trabajo solicitado. 1) ¿Qué elementos tienen alto riesgo de alucinación? 2) ¿Cómo rehacer correctamente la petición? 3) ¿Qué verificar antes de utilizar el párrafo?

Corrección. 1) Los tres elementos factuales: la cifra de la CIF y su fecha, el título del informe del PMA (quizá inexistente) y el porcentaje sobre los hogares encabezados por mujeres. Ninguno tiene fuente proporcionada. 2) Descargar el análisis oficial más reciente de la CIF y los informes pertinentes, dárselos al asistente (o activar la búsqueda web con enlaces) y pedir: «A partir únicamente de estos documentos, redacta un párrafo de contexto de 150 palabras; cita la fuente y la página de cada cifra; escribe "no disponible" si falta una información». 3) Abrir cada fuente citada, verificar cada cifra, su fecha y su zona geográfica; eliminar toda afirmación sin fuente.

Para recordar

- Un modelo de lenguaje produce el texto más plausible, no necesariamente el más exacto.
- Las alucinaciones afectan sobre todo a las cifras, las referencias, los hechos recientes y los contextos locales.
- Proporcionar sus propios documentos y exigir citas reduce mucho los errores.
- Claude y Gemini tienen funciones comparables; la elección depende del entorno, los datos, la calidad probada y el coste.
- La IA prepara, el ser humano verifica y decide.

Cuestionario

1. Un gran modelo de lenguaje produce su respuesta:
   A. Consultando una base de hechos verificados
   B. Prediciendo la continuación más probable del texto
   C. Copiando una página web al azar
   D. Consultando a un experto humano

2. Una «alucinación» es:
   A. Una respuesta falsa presentada con seguridad
   B. Un error que impide responder
   C. Una negativa a responder
   D. Una respuesta demasiado larga

3. ¿Qué petición presenta el mayor riesgo de error?
   A. Resumir un informe proporcionado
   B. Reformular un correo
   C. Explicar una fórmula de hoja de cálculo
   D. Dar el número exacto de desplazados de un municipio sin fuente

4. La herramienta de Google que responde únicamente a partir de los documentos subidos, con citas, se llama:
   A. Gems
   B. Canvas
   C. NotebookLM
   D. Artefactos

5. ¿Quién es responsable de un informe redactado con ayuda de la IA?
   A. El proveedor de la IA
   B. Nadie
   C. El donante
   D. La persona y la organización que lo firman

Respuestas del cuestionario : 1. B · 2. A · 3. D · 4. C · 5. D