

SEGUNDA EDICIÓN

IA alineada con la soberanía

para la diplomacia multilateral



GAINFOREST · YOUTH NEGOTIATORS ACADEMY



MMXXVI

— UNA NOTA ANTES DE EMPEZAR

Este es un documento vivo.

Continúa la guía de 2024 escrita para el Climate Youth Negotiator Programme, conserva sus entrevistas y sus prompts, y los actualiza para herramientas que hoy no solo responden, sino que también actúan. Aprovechar las oportunidades y reconocer los límites para fortalecer su incidencia.



Índice general

1	Introducción	4
2	¿Qué son los grandes modelos lingüísticos?	6
3	Perspectivas de las partes interesadas	9
4	LLMs para las negociaciones	11
5	Uso de los modelos	13
6	Agentes	17
7	Agentes para correos electrónicos	20
8	Agentes para la logística	22
9	Agentes para otras actividades	23
10	Soberanía, y quién es propietario de qué	24
11	Desplegar su propia IA soberana	26
12	La huella ambiental	28
13	Conclusión y recomendaciones	30
14	Apéndice: resumen de entrevistas	32
15	Glosario de IA	34
16	Referencias	38

Introducción

El espacio de negociación de la UNFCCC concentra una gran densidad de datos, documentos, historias y terminologías en constante evolución. Es un ámbito en el que la precisión del lenguaje y una sólida comprensión de cuestiones multifacéticas pueden influir de manera significativa en los resultados. Es aquí donde tecnologías como los grandes modelos de lenguaje pueden ofrecer una ventaja.

1.1 ¿Por qué debería importarnos?

Para las personas jóvenes negociadoras sobre el clima que representan a sus países en la Convención Marco de las Naciones Unidas sobre el Cambio Climático, lo que está en juego es considerable. Portan las preocupaciones y aspiraciones de una generación más joven, que heredará las consecuencias de las decisiones que se adopten hoy. Sin embargo, a menudo afrontan dificultades para acceder a grandes volúmenes de datos climáticos, comprender los matices del lenguaje de las políticas o comunicar sus posiciones con eficacia.

Los modelos de lenguaje pueden ayudar en los tres aspectos. Procesan, comprenden y generan texto a partir de colecciones muy amplias de escritos, lo que los hace especialmente útiles para cuatro fines.

Análisis de información. Destilar grandes cantidades de datos climáticos e investigación en ideas concisas, para que las personas negociadoras puedan mantenerse al día con los hallazgos más recientes.

Apoyo a la redacción de políticas. Ayudar a las personas negociadoras a formular sus argumentos en un lenguaje que resuene, se ajuste a la convención y, aun así, destaque.

Comunicación intercultural. Traducir y contextualizar información, lo que hace que el entorno de negociación sea más inclusivo.

Simulación y capacitación. Proporcionar escenarios de negociación simulados para que las personas jóvenes negociadoras puedan practicar antes de incorporarse al escenario mundial.

La tecnología en sí misma es solo la mitad de la historia. Usar bien estas herramientas señala un cambio más amplio, en el que la juventud combina su compromiso con la acción climática con las herramientas que moldean la forma en que se adoptan las decisiones.

1.2 Cómo utilizar esta guía

La primera mitad explica qué son estos sistemas, qué necesitan las personas negociadoras y cómo pedir a un modelo algo útil. La segunda mitad trata sobre los

agentes, que leen conjuntos completos de documentos y ejecutan una tarea en varios pasos. Los últimos capítulos abordan quién es propietario de sus datos, cómo ejecutar un modelo bajo su control y cuánto le cuesta al planeta. Cada capítulo incluye prompts que se pueden copiar, y nada de ello requiere formación técnica.¹

¹ La primera edición se escribió en 2024 para el Climate Youth Negotiator Programme. Las entrevistas del capítulo 3 y el apéndice proceden de ese trabajo.

¿Qué son los grandes modelos lingüísticos?

2.1 Cómo llegamos hasta aquí

En 1957, el psicólogo Frank Rosenblatt inició el programa del perceptrón en el Cornell Aeronautical Laboratory, y la Marina de los Estados Unidos presentó una máquina de perceptrón al año siguiente. Un perceptrón es un modelo matemático simplificado de una célula cerebral: combina entradas, aplica ponderaciones y produce una salida. Era un modelo abstracto, no una simulación de un cerebro, pero las redes neuronales modernas descienden en parte de él.

Después, el progreso quedó estancado durante décadas, limitado por los datos disponibles, el hardware disponible y los métodos de entrenamiento disponibles. El campo volvió a avanzar cuando esas tres restricciones se flexibilizaron a la vez. En 2012, Alex Krizhevsky, Ilya Sutskever y Geoffrey Hinton entrenaron **AlexNet** con 1,2 millones de imágenes etiquetadas utilizando unidades de procesamiento gráfico, y comunicaron una tasa de error del 17,0 por ciento entre las cinco primeras predicciones en ImageNet.¹

En 2017, investigadores de Google presentaron el **Transformer**, que se basa en la atención en lugar de la recurrencia. La atención calcula ponderaciones entre las palabras de un pasaje según el contexto, de modo que el modelo puede relacionar partes distantes de un texto. Los Transformers se entrenan bien en paralelo y sustentan casi todos los sistemas de esta guía.

OpenAI aplicó después el Transformer al preentrenamiento generativo. GPT y GPT-2 mostraron que un modelo entrenado únicamente para predecir la siguiente palabra podía adquirir muchas capacidades sin un modelo entrenado por separado para cada tarea. Los investigadores observaron **leyes de escalamiento**, es decir, que el rendimiento mejoraba de forma predecible a medida que aumentaban el cómputo, los datos y el tamaño del modelo, siempre que el tamaño del modelo y los datos de entrenamiento se mantuvieran en equilibrio.

El 30 de noviembre de 2022, OpenAI lanzó ChatGPT como vista previa gratuita de investigación, y la interfaz conversacional puso estas capacidades al alcance de todos.

Siguieron otros dos cambios. En septiembre de 2024, OpenAI anunció o1, y en enero de 2025 DeepSeek lanzó R1. Ambos emplean cómputo adicional en el momento de responder, generando notas de trabajo más largas antes de contestar, lo que ayuda en problemas que requieren varios pasos. Y los modelos empezaron a conectarse a herramientas y bucles, tema del capítulo 6.

¹ El resultado procedía tanto de la arquitectura, el aumento de datos y el conjunto de datos como del hardware.

2.2 La receta

En esencia, un modelo lingüístico es un tipo de inteligencia artificial diseñado para comprender, generar y trabajar con el lenguaje humano. Puede pensarse como un vasto cerebro digital, entrenado con textos de libros, artículos, sitios web y otras fuentes, que utiliza para generar texto similar al humano sobre la base de los patrones que reconoce.

ChatGPT, al igual que otros modelos de la serie GPT de OpenAI o Llama de Meta, se construye en dos etapas: preentrenamiento y luego postentrenamiento. En otro tiempo, el postentrenamiento significaba una sola ronda de ajuste fino. Ahora tiene dos partes, y la segunda explica por qué los modelos recientes pueden elaborar problemas largos.

Preentrenamiento. El modelo se entrena con una cantidad enorme de texto de internet. No sabe específicamente qué documentos estaban en su conjunto de entrenamiento. El objetivo es aprender gramática, hechos sobre el mundo, cierta capacidad de razonamiento y, sobre todo, cómo predecir la siguiente palabra de una frase. Al hacerlo, absorbe una cantidad sorprendente de información, desde hechos cotidianos triviales hasta conceptos complejos.

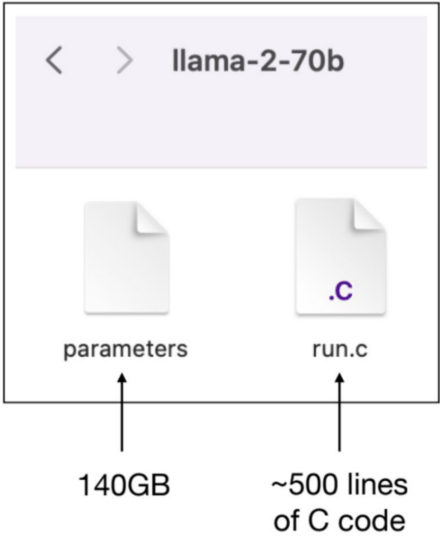
Ajuste fino supervisado. La primera mitad del postentrenamiento. El modelo se entrena adicionalmente con un conjunto de datos más acotado, escrito y evaluado por revisores humanos que siguen directrices publicadas. Los revisores puntúan posibles salidas para una gama de entradas de ejemplo, y el modelo generaliza a partir de esa retroalimentación para responder a una amplia variedad de solicitudes de usuarios. Esto es lo que le enseña a seguir una instrucción en lugar de limitarse a continuar su frase.

Aprendizaje por refuerzo. La segunda mitad, y la más nueva. En lugar de copiar una respuesta escrita, el modelo recibe recompensa por llegar a un buen resultado, de modo que aprende qué formas de trabajar sobre un problema tienden a dar resultado. Esto es lo que enseña a un modelo a razonar a lo largo de un problema extenso y mantener unida una cadena de pasos antes de responder.

2.3 Limitaciones

Los investigadores aún están desentrañando cómo funcionan estos modelos en detalle. La visión general es que el preentrenamiento incorpora una amplia variedad de conocimientos sobre el mundo, y que el postentrenamiento moldea cómo se utilizan esos conocimientos.

Lo que sigue siendo cierto es que un modelo lingüístico predice la siguiente palabra en una secuencia a partir de patrones estadísticos. Ese método capta una gran cantidad de conocimiento humano, y también trae problemas consigo. Los modelos reproducen los sesgos del texto con el que fueron entrenados, y producen afirmaciones seguras que son falsas, lo que el campo denomina alucinación. Una frase probable no es una frase verificada.



En disco, un modelo son dos archivos: los parámetros que aprendió y un programa breve que los ejecuta. Llama 2 70B son 140GB de parámetros y unas 500 líneas de C.

2.4 De dónde procede el sesgo

Estos modelos se entrenan con lo que contiene internet, e internet no contiene el mundo de manera uniforme. Esto importa más para una persona negociadora sobre el clima que para casi cualquier otro lector.

Tomemos el registro de la propia naturaleza. Global Biodiversity Information Facility es el mayor repositorio abierto de observaciones de especies de la Tierra, y aproximadamente el 83 por ciento de sus registros procede de América del Norte y Europa. El 17 por ciento restante cubre todo lo demás, incluidos los lugares más biodiversos del planeta.

El mismo desequilibrio atraviesa el texto. Se ha escrito mucho más en línea en inglés, francés y español que en quechua, twi o Tok Pisin, de modo que un modelo ha visto mucho más del primer grupo. Traducirá, resumirá y redactará con fluidez para una delegación que trabaje en una lengua bien representada, y fallará con más frecuencia, y con más seguridad, para una que no lo esté.

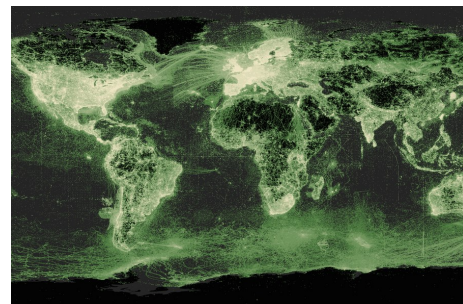
De aquí procede también una gran parte de la alucinación. Cuando se le pregunta por un lugar bien documentado, un modelo responde a partir de un registro denso. Cuando se le pregunta por uno escasamente documentado, produce las mismas frases seguras a partir de mucho menos material y rellena los vacíos con lo que sugiera el patrón. La salida no parece menos cierta. Simplemente lo es.

Nada de esto es culpa de nadie de forma simple. Recopilar y publicar datos requiere infraestructura y dinero de los que muchos países no disponen, y algunas comunidades rechazan entregar conocimientos sobre su territorio y sus especies, lo cual es una posición legítima y no un fracaso. La consecuencia llega de todos modos: estas herramientas funcionan mejor para los lugares que ya están mejor servidos por datos, y peor para los lugares donde más está en juego.

La misma brecha atraviesa quién utiliza siquiera estas herramientas. El índice de Anthropic sobre el uso de Claude por país sigue de cerca el ingreso nacional, y veinticinco países se sitúan en un nivel inferior que registra un uso medido casi nulo.² Entre ellos figuran Tonga, Samoa, Nauru y Palau, que son Partes en la misma negociación y se cuentan entre los más expuestos a su resultado.

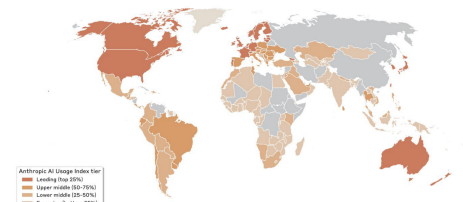
■ NOTA Qué significa esto en la sala

Trate la fluidez de un modelo sobre su región como una medida de cuánto se ha escrito sobre ella, no de cuánto sabe. Cuanto menos representados estén su país, su lengua o su ecosistema, más rigurosamente deberá verificar lo que le diga.



Todas las observaciones de especies en GBIF, representadas en un mapa. Las zonas brillantes son donde están los registros, no donde está la vida. Según recuentos brutos, Bélgica y Suecia resultan más biodiversas que Brasil y Madagascar.

² Anthropic, *Anthropic Economic Index*, 2025. El uso escala con el PIB por persona en edad de trabajar aproximadamente con una potencia de 0,69.



Uso de Claude por persona en edad de trabajar, por país. Compárelo con el mapa de registros de especies anterior.

Perspectivas de las partes interesadas

Las entrevistas con jóvenes negociadores climáticos de varios países revelan un conjunto de desafíos comunes y una visión compartida de lo que sería útil.¹

¹ Seis negociadores de Liberia, Paraguay, Perú, Nigeria, Líbano e Indonesia. El resumen completo figura en el apéndice.

3.1 Principales desafíos y dificultades

Barreras lingüísticas. El inglés no es la primera lengua de la mayoría de las personas presentes en la sala.

Lenguaje técnico. La jerga de la UNFCCC tarda años en asimilarse (Liberia, Paraguay).

Transferencia de conocimientos. Los conocimientos técnicos y la información se comparten de manera desigual, y gran parte se pierde entre ciclos.

Conocimientos históricos. La comprensión de las negociaciones anteriores es difícil de adquirir (Líbano).

Comunicación de temas complejos. Las conversaciones con negociadores superiores son difíciles de sostener bajo presión (Perú, Nigeria), y es necesario ganarse la legitimidad antes de que se escuche el fondo.

Comprensión de las posiciones de las Partes. La dinámica de las negociaciones cambia rápidamente y es difícil de interpretar.

Limitaciones de tiempo. Nunca hay tiempo suficiente para procesar la información que importa.

Expresión de ideas complejas. La expresión rápida y precisa en inglés sigue siendo una barrera persistente.

3.2 Herramientas y recursos deseados

Herramientas lingüísticas. Apoyo lingüístico y gramatical más sofisticado, concebido para textos de la UNFCCC y no para la redacción empresarial general.

Recuperación rápida de información. Plataformas que examinen documentos extensos, extraigan lo que importa y confirmen cuál es el texto vigente.

Formación personalizada. Recursos de aprendizaje adaptados a la región, el grupo y el nivel de experiencia de cada negociador.

Las herramientas que se utilizaban efectivamente en el momento de las entrevistas eran Grammarly para la redacción y Google Drive para la colaboración. Un nego-

ciador indicó que no utilizaba ninguna herramienta de traducción, por considerar que el lenguaje de la negociación no sobrevive a la traducción.

LLMs para las negociaciones

Las negociaciones sobre el clima son multifacéticas y abarcan temas que van desde las emisiones de carbono y la biodiversidad hasta las implicaciones económicas y las dinámicas sociopolíticas. Navegarlas exige tres cosas a la vez.

Acceso rápido a la información. La ciencia del clima evoluciona de manera continua, y los negociadores necesitan tener los datos más recientes al alcance de la mano.

Redacción eficaz de acuerdos. La precisión del lenguaje puede determinar si un acuerdo prospera o fracasa.

Comunicación multilingüe. Una plataforma mundial exige una interacción intercultural y multilingüe.

4.1 Dónde ayudan los modelos

Comunicación. Redactar y traducir documentos, de modo que la información y los acuerdos se comuniquen con exactitud entre Partes con distintas lenguas maternas.

Recopilación de información. Reunir y resumir grandes volúmenes de material sobre temas climáticos, incluida la investigación científica, los documentos de política y los acuerdos históricos, para que los negociadores se mantengan informados.

Trabajo con escenarios. Examinar las posibles consecuencias de distintas decisiones de política, lo que ayuda a un negociador a explorar posiciones y resultados antes de comprometerse con uno.¹

Análisis de argumentos. Resumir los argumentos formulados por distintas Partes durante las negociaciones, lo que ayuda a un negociador a identificar los puntos clave y los contraargumentos.

¹ A veces se denomina modelización predictiva, lo cual exagera su alcance. Los modelos de lenguaje no ejecutan modelos climáticos ni económicos. Reformulan y comparan escenarios, y cualquier cifra que produzcan debe verificarse con una fuente real.

4.2 Posibles dificultades

Hay beneficios reales en esto, y también salvedades que importan en la diplomacia climática.

Simplificación de cuestiones complejas. Estos sistemas comprimen, y la compresión elimina el matiz que una Parte luchó durante tres períodos de sesiones por insertar.

Falta de inteligencia emocional. Un modelo no percibe las corrientes emocionales y políticas de una negociación, ni sabe qué se acordó en un pasillo.

Falta de conocimientos especializados humanos. La fluidez en el registro de la política climática no equivale a conocimientos especializados en política climática. Depender de un modelo para cualquiera de las dos cosas puede producir información incompleta o inexacta.

Seguridad y privacidad. Las conversaciones suelen incluir material sensible. Todo lo que se pegue en un sistema alojado sale de su delegación, y rara vez se leen las condiciones bajo las cuales se almacena.

Consideraciones éticas. Sea transparente sobre la asistencia automatizada, mantenga la responsabilidad de las decisiones en manos de las personas y esté atento al uso indebido.

Comunicación errónea e interpretación equivocada. Un modelo puede malinterpretar la intención detrás de una solicitud y devolver un texto que suena como su posición sin serlo.

Uso de los modelos

5.1 Qué modelo, y en qué términos

Antes de escoger un asistente, conviene saber que existen dos tipos, y que la diferencia importa más que la marca.

Pesos cerrados. ChatGPT, Claude y Gemini funcionan en servidores de empresas y se accede a ellos mediante una cuenta. Son los modelos más potentes y los más rápidos para empezar. Los niveles gratuitos sirven para redactar, traducir y resumir. Los niveles de pago cuestan alrededor de veinte dólares al mes y añaden documentos más largos, cargas de archivos y las funciones de agente del capítulo 6. Todo lo que usted escribe se envía a la empresa y, según su configuración, puede conservarse, revisarse o utilizarse para entrenar el siguiente modelo.

Pesos abiertos. Llama, Mistral, Qwen, Gemma y DeepSeek publican sus parámetros, de modo que el mismo modelo puede ejecutarlo usted, su institución o un proveedor que usted elija, en lugar de uno que le venga dado. Ollama instala uno en un portátil en pocos minutos. Polly, ejecutado con GainForest y la Youth Negotiators Academy en polly.ai4cop.org, es un modelo alojado de pesos abiertos que no conserva ningún registro de lo que usted envía. El capítulo 11 trata de cómo ejecutar el suyo propio.

Una regla para empezar. Use un modelo cerrado alojado para material que ya sea público, y un modelo de pesos abiertos para todo lo que no lo sea. No pegue en un servicio que no controla una posición que su delegación aún no haya presentado. El capítulo 10 explica qué está en juego en esa elección, y merece la pena leerlo antes de abrir una cuenta, no después.

Elija lo que elija, instale la aplicación móvil. La entrada de voz es útil cuando se desplaza entre sesiones, y dictar una idea preliminar en su propio idioma suele producir un mejor borrador que escribir una frase cuidadosa en inglés.

5.2 Ingeniería de instrucciones 101

La instrucción que usted escribe se llama prompt. Escribirla es el proceso de crear instrucciones específicas y bien definidas que le ayuden a resolver una tarea, y la calidad y especificidad de lo que escriba determinan la respuesta que recibirá.

Un prompt puede contener cualquiera de los elementos siguientes, y no necesita incluir los cuatro.

Instrucción. La tarea que quiere que el modelo realice.

Contexto. Información de antecedentes que orienta al modelo hacia una mejor respuesta.

Datos de entrada. El texto, la pregunta o el documento sobre el que quiere que trabaje.

Indicador de salida. El tipo o formato de la salida que desea.

5.3 Qué hace que un prompt sea bueno

Seis movimientos hacen la mayor parte del trabajo, y valen para cualquier modelo que esté usando.

Sea claro, directo y detallado. Diga quién es usted, qué quiere y cómo debe ser la salida. Un modelo no puede inferir la sala en la que usted se encuentra.

Muestre un ejemplo. Una muestra de la salida que desea enseña más rápido que un párrafo que la describa.

Pida citas. Indique al modelo que cite el pasaje en el que se basa. Anclar una respuesta en un documento recuperado es la protección más sólida que existe contra una invención formulada con seguridad.

Marque las partes. Encierre cada elemento en etiquetas, `<context>`, `<task>`, `<text>`, para que el modelo pueda distinguir su instrucción del documento que ha pegado.

Deje que piense primero. Pida el razonamiento antes de la respuesta. Los modelos cometen menos errores en cualquier tarea con más de un paso cuando la desarrollan de forma visible.

Asígnele un papel. “Usted es investigador sobre pérdidas y daños” produce un texto distinto, y normalmente mejor, que la misma solicitud sin que nadie hable.

Luego siga iterando. Refinar un prompt es la forma en que mejora la calidad del intercambio, y el segundo intento casi siempre es mejor que el primero.

Conviene adquirir un hábito desde el principio. Pida fuentes, y pida al modelo que marque las partes sobre las que no esté seguro. Una respuesta que no puede verificar es una respuesta que no puede usar en una plenaria.

ANATOMÍA DE UNA INTERVENCIÓN

Para la tarea que más realizan los negociadores, cinco partes marcan la diferencia: quién habla y dónde, la cuestión concreta de que se trata, la evidencia que la sustenta, el marco al que se vincula y lo que debe hacerse.

¹Como delegado juvenil de Malasia en COP29, redacte una intervención de dos minutos ²sobre las pérdidas y los daños que afectan a la niñez en Malasia. Destaque ³preocupaciones como el aumento del nivel del mar que amenaza a las comunidades costeras, las perturbaciones educativas derivadas de los impactos climáticos y los problemas de salud mental para la niñez. ⁴Formule recomendaciones específicas para priorizar la financiación destinada a proteger los derechos de la niñez y mejorar su participación en los debates sobre el clima ⁵en el marco del Acuerdo de París.

¹ Papel y contexto ² Tema central ³ Evidencia ⁴ Llamado a la acción ⁵ Contenido técnico

Quite cualquiera de las cinco y se percibe lo que falta. Sin el papel, redacta para nadie; sin la evidencia, produce sentimiento; y sin la última línea, no se vincula a ningún acuerdo sobre el que alguien pueda actuar.

5.4 Prompts para distintos casos de uso

1. **Análisis de datos** CARGA DE ARCHIVO
Dada esta hoja de cálculo, trace el aumento medio de las emisiones durante los últimos años, indique su método y enumere los supuestos que ha formulado sobre los datos faltantes.
2. **Traducción** CUALQUIER MODELO
Traduzca este párrafo al español, manteniendo el registro formal utilizado en los textos de decisión.
3. **Resumen** BÚSQUEDA WEB ACTIVADA
Resuma la intervención del LDC Group en el Global Stocktake Technical Dialogue del 10 de junio de 2023, y cite los pasajes en los que se basa.
4. **Recopilación de información para elaborar una declaración de posición o preparar un discurso** CUALQUIER MODELO
Usted es la persona encargada de redactar discursos para un negociador de las Naciones Unidas sobre el cambio climático. Está escribiendo sobre los esfuerzos del Reino Unido en materia de fomento de la capacidad. Yo soy el negociador. ¿Qué cinco preguntas me haría para extraer la información que necesita a fin de redactar una declaración?
5. **Prever contraargumentos y preparar refutaciones** CUALQUIER MODELO
Las pérdidas y los daños son una cuestión clave en las negociaciones sobre el clima. Usted es investigador sobre pérdidas y daños. Exponga las principales razones por las que los países no dan importancia a las pérdidas y los daños. Produzca un contraargumento sólido frente a estos puntos, demostrando la necesidad de un fondo para las pérdidas y los daños.
6. **Redactar una declaración de apertura o un discurso principal** CUALQUIER MODELO
Usted es redactor de discursos y está elaborando una declaración utilizando las preguntas anteriores. Produzca una declaración de 300 palabras que des-

criba los esfuerzos del Reino Unido en materia de fomento de la capacidad climática. La declaración debe reflejar el estilo de redacción empleado por el Gobierno del Reino Unido.

7. Descifrar la jerga

CUALQUIER MODELO

Explique qué significa “medios de implementación” en este párrafo, en lenguaje claro, y dé un ejemplo de cómo se ha utilizado la expresión en una decisión anterior.

8. Encontrar el documento

RECUPERACIÓN

¿Qué texto de decisión contiene el mandato para este programa de trabajo? Indíqueme la signatura del documento y cite el párrafo en el que figura.

9. Traducir para su propio público

CUALQUIER MODELO

Traduzca esta intervención al Bahasa Malaysia para un público comunitario, manteniendo el significado y eliminando la jerga de la UNFCCC.

Agentes

Todo lo expuesto hasta este punto funciona de la misma manera. Usted escribe una solicitud, lee una respuesta y decide qué hacer con ella. Cuatro cosas han cambiado y, en conjunto, añaden una segunda forma de trabajar.

El contexto se volvió extenso. Un modelo ahora puede contener al mismo tiempo un texto de negociación completo, sus revisiones anteriores y las notas informativas de su delegación.

La recuperación se volvió normal. Un sistema puede buscar en un conjunto de documentos o en la web antes de responder, y citar lo que encontró. Eso es lo que hace que una respuesta sea verificable.

Los modelos adquirieron herramientas. Un modelo puede conectarse a un índice de búsqueda, una hoja de cálculo, un calendario o una carpeta de archivos, de modo que pueda actuar sobre ellos en lugar de limitarse a describirlos.

Los sistemas adquirieron bucles. En lugar de responder una sola vez, un sistema puede planificar, dar un paso, observar el resultado y volver a intentarlo hasta que la tarea esté terminada.

Un modelo conectado a herramientas, memoria y un bucle de repetir hasta terminar, que funciona con permisos que alguien le ha concedido, es lo que se entiende por agente. No tiene criterio ni rendición de cuentas. Tiene una tarea y un conjunto de permisos.

■ LÍNEA ROJA La línea

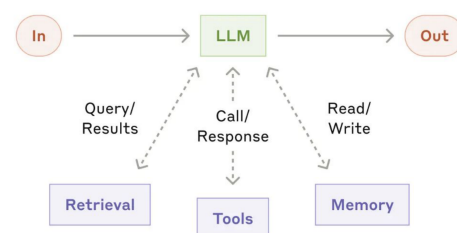
Un agente puede preparar una posición. No puede enviar, presentar ni publicar nada en nombre de una delegación sin que una persona identificada autorice ese acto.

6.1 Tres formas de trabajar

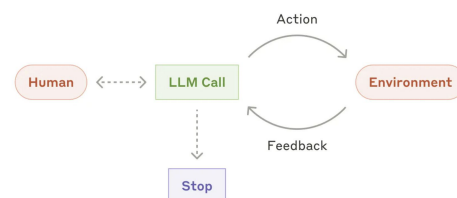
Chat. Usted pregunta, el sistema responde, usted verifica. Es lo más adecuado para redactar, traducir, explicar y ensayar.

Recuperación. El sistema busca primero en un conjunto definido de documentos y responde con citas. Úselo para cualquier cuestión fáctica sobre la negociación: historial de textos, posiciones de las Partes, decisiones anteriores.¹

Agéntico. El sistema ejecuta una tarea en varios pasos utilizando herramientas. Úselo para trabajo repetitivo con un resultado verificable, como hacer seguimiento de los cambios entre revisiones o preparar una nota informativa diaria.



Un modelo aumentado. El mismo modelo, al que se le da algo en lo que buscar, herramientas que invocar y un lugar donde conservar lo aprendido.



El bucle. El modelo actúa, lee lo que volvió y continúa hasta detenerse. Una persona fija la tarea y los permisos, y es lo único en el diagrama a lo que se puede exigir responsabilidad.

¹ Si un sistema le dice qué dice un párrafo de un texto de decisión y no puede mostrarle el párrafo, trátelo como un rumor.

6.2 Casos de uso

1. Seguimiento de textos

ARCHIVOS ADJUNTOS

Compara la revisión 3 de este proyecto de decisión con la revisión 2. Enumera todos los cambios sustantivos, cita ambas versiones y señala los cambios que afectan las líneas rojas de mi delegación.

2. Mapeo de posiciones

RECUPERACIÓN

A partir de estas comunicaciones, elabora un cuadro de posiciones de las Partes sobre el artículo 6.4 con una cita y una referencia para cada fila. Marca cualquier posición que hayas inferido.

3. Recuperación del historial

RECUPERACIÓN

Rastrea cómo se ha matizado la expresión “responsabilidades comunes pero diferenciadas” en los textos de decisión desde 2015, citando cada decisión.

4. Lectura nocturna

ARCHIVOS ADJUNTOS

Aquí están los cuatro textos publicados hoy. Dime qué avanzó, qué quedó estancado y cuáles son las tres páginas que debo leer antes de mañana.

5. Resumen diario

AGENTE PROGRAMADO

Cada tarde, recopila los textos publicados durante el día en los temas que sigo, resume qué cambió y envíame la lista.

6. Análisis de corchetes

ARCHIVOS ADJUNTOS

Para cada opción entre corchetes de este párrafo, explica quién se beneficia, quién se opone y cómo podría formularse una opción alternativa de reserva.

7. Ensayo

CHAT

Interpreta a una persona negociadora que se opone a este texto. Cuestiona mi intervención y no seas complaciente.

8. Retrotraducción

CHAT

Traduce mi formulación al francés, luego tradúcela de nuevo al inglés de forma independiente y dime qué se desplazó.

9. Traspaso

ARCHIVOS ADJUNTOS

A partir de mis notas de esta sesión, redacta el documento de traspaso para quien se haga cargo de este expediente en el próximo ciclo.

10. Trabajo con datos

ACCESO A ARCHIVOS

A partir de esta hoja de cálculo de emisiones, grafica la tendencia de estos países desde 2010, indica tu método y enumera todos los supuestos sobre datos faltantes.

6.3 Trabajar con un agente de forma segura

Un agente que actúa sobre una lectura errónea produce consecuencias, no un mal párrafo.

Verifique lo que va a repetir. Abra uno de los documentos citados y confirme que la cita está allí antes de usarla en la sala.

Conceda el permiso mínimo que funcione. Un sistema que lee sus archivos no necesita también enviar correo.

Trate todo documento externo como no confiable. Un archivo presentado puede contener instrucciones dirigidas al agente.

Decida quién aprueba un texto asistido por IA. Con varias herramientas en la cadena, nadie más lo hará.

Agentes para correos electrónicos

El correo electrónico sigue siendo un medio principal de comunicación profesional, y tres factores lo vuelven difícil.

Gestión del volumen. Las bandejas de entrada se desbordan, y priorizarlas toma un tiempo del que usted no dispone.

Redacción eficaz. Un correo electrónico claro, conciso y orientado a la acción exige habilidad.

Barreras multilingües. La comunicación mundial implica trabajar en idiomas que no son los propios.

Un asistente conectado a su buzón ahora lee el hilo por sí mismo, así que concédale acceso de lectura y deje que prepare el borrador. El envío sigue estando en sus manos.

1. Seguimiento de reuniones

Resume los puntos clave y los compromisos asumidos durante la reunión de negociación reciente, atribuye cada uno a la persona que lo formuló, y redacta un correo electrónico de seguimiento cortés dirigido a las personas participantes para alentar la colaboración en las acciones acordadas.

2. Programación

Genera un correo electrónico diplomático y flexible para proponer varias fechas y horas para la próxima ronda de reuniones, expresadas en las zonas horarias de todas las personas participantes.

3. Informes diarios y resúmenes de resultados

Crea un correo electrónico breve de informe diario que destaque los avances logrados en las sesiones más recientes, incluidas las nuevas propuestas, los puntos acordados y las esferas de controversia que requieren un examen ulterior.

4. Priorización inteligente

Elabora un correo electrónico dirigido al equipo de negociación que exponga una estrategia para priorizar los temas del programa en función de su urgencia, su posible impacto en los objetivos climáticos y la viabilidad de alcanzar un acuerdo.

5. Clasificación

Agrupar mis mensajes no leídos en: requieren una decisión mía, requieren una respuesta, requieren reenvío, no requieren nada. Redacta respuestas solo para el tercer grupo y déjalas sin enviar.

6. Verificación del registro

Reescribe este correo electrónico en un registro diplomático formal sin modificar ningún compromiso sustantivo, y luego enumera lo que cambiaste para que yo pueda comprobarlo.

Estos sistemas adoptan por defecto una voz cálida y conciliadora que puede conceder cosas que su texto original no concedía. Pregunte qué se modificó y lea la lista.

Agentes para la logística

Organizar horarios, elaborar listas de verificación y reservar viajes puede volverse abrumador durante un período de sesiones. Tres problemas aparecen de manera recurrente.

Gestión del tiempo. Priorizar tareas y asignarles tiempo no es sencillo.

Omisión de tareas. Elementos esenciales quedan fuera de la lista de verificación.

Reserva de vuelos y hoteles. Encontrar opciones que se ajusten a su calendario y a su presupuesto lleva horas.

Este es el trabajo de menor riesgo en la guía y el mejor punto de partida para trabajar con un agente. Las tareas se repiten, los resultados son fáciles de verificar y casi nada de lo que aparece aquí es confidencial.

1. Reservas

Identifica los vuelos y hoteles más costoefectivos y convenientes para un viaje desde [ciudad de salida] hasta [ciudad de destino] en [fechas], con preferencia por vuelos por la mañana y alojamiento cerca de la sede. Muéstrame las ventajas y desventajas en lugar de darme una sola recomendación.

2. Creación de listas de verificación

Genera una lista de verificación completa para un período de sesiones internacional de dos semanas que incluya los documentos necesarios, la acreditación, ropa para condiciones meteorológicas variables, equipo de trabajo esencial y copias sin conexión de los textos clave.

3. Planificar el día

Examina mi calendario y el programa del período de sesiones, y luego propone un horario que equilibre reuniones, bloques de trabajo individual, ejercicio y tiempo personal, teniendo en cuenta estos plazos y compromisos.

4. Seguimiento de avances

Crea una plantilla para hacer seguimiento de los avances en las líneas de trabajo que sigo, incluidos hitos, plazos y un sistema para anotar actualizaciones y próximos pasos.

La logística también es el ámbito en el que un agente recibe por primera vez permisos reales sobre su calendario, su correo y sus archivos. Eso es razonable, y es ahí donde comienza el hábito de conceder un poco más cada vez. El sistema que reserva sus vuelos no necesita redactar sus posiciones.

Agentes para otras actividades

Las negociaciones sobre el clima no son solo cifras y hechos. Están entrelazadas con emociones humanas, diferencias culturales y bienestar mental.

Tensiones en la salud mental. Las realidades del cambio climático y la presión de negociar pasan factura a las personas delegadas.

Malentendidos culturales. Las interpretaciones erróneas derivadas de diferencias culturales generan divisiones y ralentizan el diálogo.

Esta es la única parte de la guía en la que una simple ventana de chat supera a un agente. Estas conversaciones requieren un modelo con el que se dialogue, no un sistema con permisos.

1. Informe cultural

Estoy en una negociación de las Naciones Unidas sobre el cambio climático con una persona de Kenya. ¿Qué aspectos culturales debería tener presentes para avanzar en mi negociación?

2. Reconocimiento emocional

Me siento desanimado en este momento, ¿puedes decirme algo alentador?

3. Balance

Esto es lo que ocurrió en esa sesión y cómo lo gestioné. ¿Qué habría hecho de manera diferente una persona negociadora con más experiencia?

Dos cautelas. Un informe cultural produce una generalización sobre un país, y usted se encontrará con una persona; por tanto, trate lo que reciba como orientación, no como predicción. Y un modelo es un lugar razonable donde formular un pensamiento a la una de la mañana, pero no es un profesional clínico, no es un colega y no es confidencial. Si la presión va más allá del cansancio, hable con su jefe o jefa de delegación, con los contactos de apoyo de su programa o con un profesional.

Soberanía, y quién es propietario de qué

Estas herramientas pueden igualar el espacio multilateral. Eliminan la barrera lingüística que determina quién interviene con confianza, condensan documentos que nadie tiene tiempo de leer y dan a una delegación de dos personas parte del alcance de una delegación de cincuenta. Eso es valioso. También plantea cuestiones de integridad que conviene formular antes de escribir.

Quién es propietario de los datos. El texto que usted pega en un asistente alojado sale de su delegación. Según la cuenta y su configuración, puede ser retenido, revisado por personal o utilizado para entrenar el siguiente modelo. Las posiciones nacionales, las instrucciones recibidas de la capital y todo lo que se le haya comunicado en confianza no deben estar allí.

Quién es propietario de la decisión. Un modelo produce texto que se lee como una posición. Si redacta y nadie verifica, el razonamiento que sustenta la posición de su Parte empieza a residir fuera de su delegación. Toda comunicación que salga debe estar respaldada por una persona con mandato.

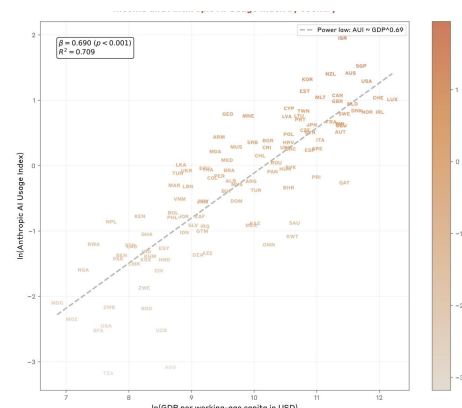
Lo que usted no puede ver. OpenAI y Anthropic no devuelven las trazas de razonamiento. Usted recibe una respuesta, a veces un resumen de cómo se llegó a ella, mientras que los pasos reales permanecen en poder del proveedor. No puede auditar el proceso, reproducirlo ni mostrar a un colega por qué el sistema dijo lo que dijo.

Pesos abiertos y pesos cerrados. Los pesos de un modelo son los números que aprendió durante el entrenamiento. ChatGPT, Claude y Gemini mantienen los suyos en servidores empresariales, y son los más potentes y fáciles de usar. Llama, Mistral, Qwen, Gemma y DeepSeek publican los suyos, de modo que usted puede ejecutar el modelo en su propia máquina o en un servidor regional. Los modelos abiertos gestionan bien la traducción, el resumen y la redacción, y van por detrás en los problemas más difíciles de múltiples pasos. Los pesos que usted conserva no pueden encarecerse ni retirarse en mitad de una sesión.

Una IA alineada con la soberanía significa que la delegación, la región o la comunidad establece las condiciones: qué modelo se utiliza, qué datos entran en él, quién puede inspeccionar lo ocurrido y el derecho a negarse. En el proceso de la COP, las Partes con menor capacidad suelen ser las Partes con más en juego, lo que hace que esto sea una cuestión práctica más que filosófica.

Una regla de trabajo. Utilice el modelo alojado más potente para material público, lectura de antecedentes, redacción y preparación. Utilice pesos abiertos, o un despliegue regido por un contrato que su institución haya revisado, para cualquier asunto que afecte a una posición que aún no sea pública.

Esta guía existe para mantener visible esa disyuntiva. Los capítulos anteriores



Uso de Claude frente a ingresos, por país. La adopción sigue de cerca el PIB por persona en edad de trabajar, lo que es otra forma de decir que estas herramientas llegan en último lugar allí donde más se necesitan.

muestran lo que estos sistemas hacen bien. Este trata de lo que se cede a cambio.

Desplegar su propia IA soberana

El capítulo anterior expuso la disyuntiva. Este trata de la respuesta más sólida a ella: ejecutar el modelo por cuenta propia.

El autoalojamiento significa que nada sale del edificio. No hay política de conservación de datos en la que confiar, ni condiciones de servicio que puedan cambiar, ni cuenta que pueda ser suspendida en medio de una sesión. También desarrolla la capacidad dentro de su institución en lugar de alquilarla.

Hasta hace poco, esto significaba aceptar un modelo mucho más débil. Ya no es así.

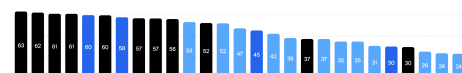
Kimi K3 y GLM-5.3 obtienen ambos 60 puntos en el Artificial Analysis Intelligence Index, tres puntos por detrás de Claude Opus 5, con 63, y Qwen3.8 les sigue con 58. Diecisiete de los 27 modelos del índice tienen pesos abiertos, aunque la parte alta de la tabla sigue siendo mayoritariamente propietaria.¹ Hace un año, el principal modelo abierto estaba trece puntos por detrás del principal modelo propietario, de modo que esta brecha se ha cerrado rápidamente y puede cerrarse aún más.

El precio es la segunda mitad del argumento. DeepSeek V4 Pro 0813 obtiene 53 puntos por alrededor de un cuarto de dólar por tarea, dentro de ese cuadrante. Claude Opus 5 obtiene diez puntos más por aproximadamente diez veces el precio. Esa relación es la razón práctica por la que una delegación puede permitirse ejecutar el suyo propio.

11.1 Qué ejecutar

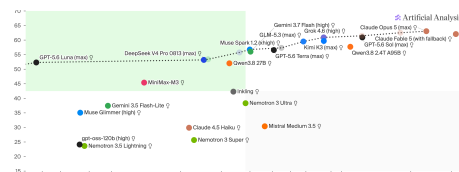
MODELO	LABORATORIO	PESOS	ÍNDICE
Kimi K3 (max)	Moonshot	Abierto, uso restringido	60
GLM-5.3 (max)	Z.ai	Abierto, MIT	60
Qwen3.8 2.4T A95B	Alibaba	Abierto, uso restringido	58
DeepSeek V4 Pro 0813	DeepSeek	Abierto	53
Qwen3.8 27B	Alibaba	Abierto	52
MiniMax-M3	MiniMax	Abierto, uso restringido	45
Nemotron 3 Super	NVIDIA	Abierto	26
gpt-oss-120b (high)	OpenAI	Abierto	24

Estado de situación al 20 de agosto de 2026. Lea la licencia antes de comprometerse. Varios de los modelos abiertos más sólidos restringen el uso comercial, lo que



Los 27 modelos del Artificial Analysis Intelligence Index, 20 de agosto de 2026, en orden de clasificación. El negro indica modelos propietarios; el azul, pesos abiertos. El principal modelo abierto, Kimi K3, ocupa el quinto lugar con 60, a tres puntos del primero.

¹ El gráfico cuenta GLM-5.3 como propietario. Z.ai publica sus pesos bajo una licencia MIT, por lo que este libro lo considera abierto; por eso la figura muestra dieciséis barras azules y el texto dice diecisiete.



El mismo índice frente al coste por tarea. El área sombreada es el cuadrante más atractivo: puntuación alta y coste bajo, y los modelos abiertos ocupan la mayor parte de él.

para una delegación suele estar bien y para un proveedor que construye sobre uno de ellos no.

11.2 Cuánto cuesta

Una NVIDIA DGX Spark es una máquina de escritorio con 128GB de memoria unificada, actualmente por 4.699 dólares. Dos de ellas se conectan directamente entre sí sin conmutador, lo que da 256GB por alrededor de 9.400 dólares. Ese par ejecuta un modelo disperso de la clase DeepSeek a entre 55 y 60 tokens por segundo con un contexto de un millón de tokens, suficiente para atender a una delegación en funcionamiento.

Una sola unidad ejecuta un modelo de 27B a 35B con suficiente rapidez para sesenta usuarios simultáneos, lo que cubre la traducción, el resumen y la redacción para toda una oficina. Tres o cuatro unidades con 384 a 512GB alojan los mayores modelos abiertos sin podarlos.

Comparado con una suscripción para cada delegado, el hardware se amortiza en uno o dos ciclos, y sigue funcionando después de que termine la financiación.

11.3 Cómo configurarlo

Sirva el modelo con vLLM o SGLang. Ambos presentan la misma API que OpenAI, de modo que cualquier herramienta que se comunique con ChatGPT se comunicará con su propio servidor después de cambiar la dirección y la clave. Para una primera prueba en un portátil, Ollama o llama.cpp son más sencillos. NVIDIA publica guías de configuración para Spark, y las recetas comunitarias para conectar dos, tres y cuatro unidades son públicas y reproducibles.

Si ejecutar el suyo propio no es realista antes del próximo período de sesiones, la alternativa es un servicio alojado con un acuerdo escrito de conservación cero de datos. Trátelo como una solución provisional, no como un destino.

11.4 Una salvedad que conviene conocer

El índice anterior es un promedio de nueve evaluaciones, y los modelos abiertos no van a la zaga de manera uniforme en todas ellas. La brecha es mayor en las tareas de razonamiento más difíciles y en las alucinaciones, donde los modelos propietarios siguen estando claramente por delante. Para un negociador, ese segundo aspecto lo decide todo. Dirija el modelo a sus propios documentos, pídale citas y ábralas.

■ NOTA Nos complace ayudar

GainForest y la Youth Negotiators Academy asesoran a las delegaciones sobre esto sin costo alguno: dimensionar el hardware, elegir un modelo, ponerlo en servicio y formar a las personas que lo utilizarán. Escriba a team@gainforest.net.

La huella ambiental

Un negociador climático que utiliza una herramienta que consume energía y agua debería poder decir cuánto consume. La respuesta honesta tiene dos partes, y apuntan en direcciones distintas.

12.1 Su propio uso es reducido

Google midió una instrucción de texto mediana enviada a Gemini en 0,24 Wh de energía, 0,03 gramos de CO₂e y 0,26 mililitros de agua.¹ Eso equivale a unos pocos segundos de un ordenador portátil. Incluso una jornada de trabajo con uso intensivo sigue siendo menos que un breve trayecto en automóvil.

Polly comunica esto para cada cuenta, y el cálculo se publica en lugar de simplemente afirmarse.² Funciona a partir de los parámetros activos del modelo:



ENERGÍA POR TOKEN

2 FLOP por parámetro activo → dividir por el rendimiento del chip → multiplicar por la potencia del chip

AGUA Y CARBONO

1,1 litros por kWh para refrigeración → 480 gramos de CO₂e por kWh en una red promedio

Leer su instrucción cuesta aproximadamente una cuarta parte de lo que cuesta escribir una respuesta, y releer una instrucción almacenada en caché cuesta una décima parte, por lo que las tres operaciones se contabilizan por separado. En el modelo disperso que ejecuta Polly, que activa 13.000 millones de parámetros por token, un intercambio típico asciende a aproximadamente 0,13 Wh y 0,14 mililitros de agua. Esto queda apenas por debajo de la mediana de Google.

Trate cualquiera de estas cifras como aproximada dentro de un factor de dos. No incluyen el entrenamiento del modelo, el agua utilizada para generar la electricidad, su propio dispositivo ni el carbono incorporado en el hardware.

12.2 El uso de la industria no lo es

Las cifras relevantes están en la escala del sector, más que de la persona. La demanda de electricidad de los centros de datos aumenta más rápido de lo que las redes se descarbonizan, se está construyendo nueva capacidad en regiones con estrés hídrico, y las empresas que informan las cifras por instrucción son las mismas cuyas emisiones totales han aumentado desde que empezaron a construir estos sistemas.

Ambas partes son verdaderas al mismo tiempo. Su propio uso es un error de re-

¹ Google, *Measuring the environmental impact of delivering AI at Google scale*, arXiv:2508.15734, August 2025. La cifra incluye la capacidad inactiva y la sobrecarga del centro de datos, que la mayoría de las estimaciones publicadas omiten.

² El método y cada constante están en `footprint.ts` en `pi-village-core`. Una cifra que un negociador no puede interrogar es peor que no tener cifra alguna.

dondeo, y el agregado constituye una demanda seria y creciente sobre la energía, el agua y la tierra. Afirmar lo primero sin lo segundo es el tipo de contabilidad que un negociador no aceptaría de una Parte.

12.3 Qué hacer al respecto

Elija un modelo más pequeño cuando uno más pequeño sea suficiente. La traducción y la síntesis no necesitan un modelo de frontera, y un modelo disperso que activa unos pocos miles de millones de parámetros cuesta una fracción de lo que cuesta uno denso.

Pida a los proveedores sus cifras, incluida el agua. Pregunte desde qué red eléctrica se atiende la solicitud. Son preguntas razonables en materia de adquisiciones y rara vez se formulan.

Informe su propia cifra cuando utilice estas herramientas en su trabajo. Una delegación que publica la huella de su propio uso de la IA tiene autoridad para pedir a otros que hagan lo mismo.

Conclusión y recomendaciones

13.1 Fortalezas y debilidades

Fortalezas. Alta eficiencia en el procesamiento y la síntesis de los grandes volúmenes de texto en que se basa la política climática. La capacidad de generar informes, redactar políticas y simular negociaciones en lenguaje natural, lo que ahorra tiempo y recursos. La capacidad de asistir simultáneamente a muchas personas, desde responsables de políticas hasta personal investigador y activistas.

Debilidades. Posible sesgo en los modelos, que distorsiona la forma en que se procesa la información si no se supervisa y corrige. Dependencia de la calidad y amplitud de los datos de entrenamiento, lo que limita la capacidad de un modelo para abordar cuestiones climáticas matizadas. Dificultad para interpretar con precisión lenguaje jurídico y técnico complejo sin supervisión humana.

Oportunidades. Estrategias de comunicación personalizadas que implican a distintas partes interesadas en la acción climática. Integración con otras herramientas para el análisis de datos ambientales, lo que sustenta políticas mejor informadas. Mayor educación y difusión de información sobre política climática, haciéndola accesible para personas no expertas.

Amenazas. Desinformación, cuando un modelo genera contenido incorrecto o engañoso a partir de fuentes poco fiables. Dependencia excesiva, que frena el desarrollo de conocimientos especializados locales en política climática. Preocupaciones éticas sobre la transparencia y la rendición de cuentas en la toma automatizada de decisiones.

Los agentes añaden otras dos amenazas que conviene nombrar. Las delegaciones con recursos utilizarán sistemas bien gobernados, mientras que las delegaciones más pequeñas recurrirán a herramientas gratuitas con condiciones poco claras, lo que amplía la brecha bajo la apariencia de igualdad de acceso. Y, cuando varias personas y varias herramientas intervienen en una cadena, la responsabilidad por un texto resulta difícil de rastrear salvo que alguien decida de antemano dónde reside.

13.2 Recomendaciones

Hay tres cosas que merece la pena construir: herramientas para traducir el lenguaje técnico de la UNFCCC, plataformas para analizar datos históricos de negociación y entornos en los que las personas jóvenes negociadoras dispongan de la tecnología que necesitan para contribuir. Junto a ellas, seis más.

Alfabetización digital. Formación orientada específicamente a la navegación por bases de datos y plataformas internacionales de políticas.

Apoyo en tiempo real. Herramientas que proporcionen aclaraciones durante las negociaciones, no después de ellas.

Formación virtual en negociación. Entornos simulados para practicar escenarios de negociación.

Plataformas de apoyo financiero. Plataformas de financiación colectiva o de microsubvenciones que aborden las barreras financieras para asistir a las negociaciones, señaladas por negociadores de Liberia y Paraguay.

Formación en comunicación intercultural. Programas que aborden los malentendidos identificados por negociadores del Perú y Nigeria.

Programas de mentoría. Redes que conecten a personas jóvenes negociadoras con negociadores experimentados, de modo que el conocimiento se transfiera entre ciclos.

Para una persona negociadora, cuatro hábitos cubren la mayor parte de lo necesario. Utilizar estos sistemas para la preparación, el lenguaje y la memoria. Verificar cualquier cosa que se vaya a repetir en voz alta. No permitir nunca que un sistema hable en nombre de su Parte. Seguir realizando una parte suficiente de la investigación por cuenta propia como para poder hacerla también sin ayuda.

Para una delegación, conviene dejar por escrito qué herramientas están permitidas, qué puede pegarse en ellas y quién aprueba un texto asistido por IA. Una página basta, y es más barata que un incidente.



Al aprovechar el potencial de la IA, la juventud de hoy no solo hereda el futuro, sino que lo moldea activamente; que el próximo capítulo de la historia de nuestro planeta se escriba con los trazos audaces de la innovación y el liderazgo juvenil.



Las preguntas y comentarios son bienvenidos en team@gainforest.net.

Apéndice: resumen de entrevistas

Resumen consolidado de entrevistas con jóvenes negociadores climáticos, realizadas para comprender los desafíos que enfrentan y la manera en que los modelos de lenguaje pueden apoyar su labor.

Personas entrevistadas. Seis negociadores, procedentes de Liberia, Paraguay, Perú, Nigeria, Líbano e Indonesia. La experiencia variaba, y algunas personas eran nuevas en el proceso.

14.1 Desafíos y dificultades

En el desempeño de la función. Barreras lingüísticas cuando el inglés no es la lengua materna. La jerga técnica de la UNFCCC (Liberia, Paraguay). Brechas en los conocimientos técnicos y en la forma en que se comparte la información. Falta de conocimiento de las negociaciones históricas (Líbano). Dificultad para conversar con negociadores de alto nivel sobre temas complejos (Perú, Nigeria). Dificultad para adaptarse rápidamente a la dinámica de las negociaciones.

En situaciones específicas. Las primeras experiencias en la COP resultaron intimidantes debido a la falta de preparación (Paraguay, Perú). Hablar en público en un entorno de alto riesgo dificulta la articulación de temas complejos.

En la recopilación de información. El sitio web de la UNFCCC es la fuente principal y es complejo de navegar. No hay tiempo suficiente para procesar información crucial (Paraguay). La difusión depende de comunicaciones internas y redes personales.

En la comunicación. Expresar ideas complejas en inglés con rapidez y precisión es difícil. Los niveles de dominio del inglés varían y crean barreras (Indonesia).

14.2 Herramientas

En uso. Grammarly para apoyo a la redacción (Paraguay). Google Drive para la colaboración en documentos (Paraguay). No se utilizan herramientas de traducción, por considerar que el lenguaje es difícil de traducir (Paraguay).

Deseadas. Herramientas lingüísticas y gramaticales más sofisticadas para textos de la UNFCCC (Paraguay, Liberia). Plataformas para el examen eficiente de documentos y la extracción de información clave (Líbano, Nigeria). Recursos de aprendizaje adaptados a las necesidades (Liberia).

14.3 Procesamiento de información y toma de decisiones

Mantenerse al día. A través de colegas y recursos compartidos, así como mediante boletines y plataformas en línea (Liberia, Líbano).

Abordar información compleja. Trabajo en equipo y aprovechamiento de los conocimientos especializados de otras personas (Perú), e investigación exhaustiva sobre temas desconocidos.

Colaboración. Se fomentan activamente comunidades de apoyo entre negociadores (Perú, Nigeria). Están en marcha iniciativas para mejorar las competencias lingüísticas (Indonesia).

14.4 Ideas y soluciones

Mejoras deseadas. Simplificación de los documentos oficiales, en favor de la inclusividad (Líbano, Nigeria).

Tareas repetitivas que convendría automatizar. Recuperación de detalles históricos de las negociaciones (Líbano).

Visión de futuro. La IA y las herramientas digitales se consideran posibles apoyos para las negociaciones.

14.5 Oportunidades para los modelos de lenguaje

Beneficios. Traducción y resumen, que abordan las barreras lingüísticas y condensan la información.

Automatización. Síntesis a partir de documentos extensos.

Reservas. La dependencia de la tecnología podría disminuir las competencias de investigación crítica (Liberia, Paraguay).

Glosario de IA

Términos que oírás en la sala, definidos de forma sencilla.

15.1 Términos básicos

Inteligencia artificial (IA)	Sistemas diseñados para realizar tareas asociadas con la inteligencia humana, como la predicción, la clasificación, la generación o la planificación.
Aprendizaje automático	Método mediante el cual un sistema aprende patrones a partir de ejemplos, en lugar de recibir cada regla de forma explícita.
Modelo	Sistema matemático aprendido que asigna entradas a salidas.
Modelo lingüístico grande (LLM)	Modelo entrenado con grandes colecciones de texto para predecir y generar lenguaje.
IA generativa	IA que produce texto, imágenes, audio, código u otros resultados nuevos.
Sistema agéntico / agente de IA	Modelo conectado a herramientas, memoria, instrucciones y bucles de acción para poder completar varios pasos.
Automatización	Uso de un sistema para realizar una tarea con intervención humana directa limitada.
Aumento	Uso de la IA para apoyar el trabajo humano, no para sustituirlo.
Humano en el bucle	Una persona examina, aprueba o interviene en un proceso asistido por IA.
Humano sobre el bucle	Una persona supervisa un sistema sin examinar cada acción.
Humano fuera del bucle	Un sistema actúa sin que una persona examine cada decisión o acción.

15.2 Datos y conocimiento

Datos de entrenamiento	Ejemplos utilizados para ajustar los parámetros de un modelo durante el entrenamiento.
Conjunto de datos	Colección organizada de datos utilizada para entrenamiento, pruebas o análisis.
Procedencia de los datos	Información sobre el origen de los datos y sobre cómo se recogieron o modificaron.
Linaje de los datos	Historial de cómo los datos se desplazan a través de sistemas y transformaciones.
Datos estructurados	Datos organizados en campos, filas, categorías o esquemas definidos.
Datos no estructurados	Material como texto, imágenes, audio o vídeo sin una estructura tabular fija.
Datos sintéticos	Datos generados artificialmente y diseñados para asemejarse a datos reales.
Verdad de referencia	Etiqueta o resultado de referencia tratado como correcto para fines de evaluación.
Base de conocimientos	Colección curada de información que un sistema puede recuperar.
Conocimiento paramétrico	Información codificada en los parámetros aprendidos de un modelo.
Conocimiento fáctico	Información sobre hechos o relaciones que un modelo puede haber aprendido de sus datos de entrenamiento.

Generación aumentada por recuperación (RAG)	Sistema que recupera documentos o registros antes de generar una respuesta.
Conocimiento local	Conocimiento fundamentado en la experiencia de un lugar o una comunidad determinados.
Conocimientos indígenas	Sistemas de conocimiento desarrollados y mantenidos por los pueblos y las comunidades indígenas.
Soberanía de los datos	Principio según el cual los datos se rigen conforme a las leyes, los derechos y la autoridad de las personas o del lugar a que se refieren.
Minimización de datos	Recopilación y conservación únicamente de los datos necesarios para un propósito declarado.
Datos sensibles	Datos cuya divulgación o uso indebido podría causar daño.

15.3 Entradas y salidas de los modelos

Entrada / prompt	Instrucción o material proporcionado a un modelo.
Salida	Texto, predicción, clasificación, imagen o acción producidos por un sistema.
Inferencia	Ejecución de un modelo entrenado para producir una salida.
Ventana de contexto	Cantidad de texto u otro material que un modelo puede considerar a la vez.
Token	Unidad de texto u otros datos procesada por un modelo lingüístico.
Incrustación	Representación numérica de texto, imágenes u otros datos utilizada para comparar relaciones.
Clasificación	Asignación de un elemento a una o más categorías.
Predicción	Estimación de un resultado probable a partir de los datos disponibles.
Pronóstico	Predicción sobre un estado o acontecimiento futuro.
Recomendación	Salida que sugiere una línea de acción.
Optimización	Búsqueda del mejor resultado de acuerdo con un objetivo y unas restricciones definidos.
Función objetivo	Magnitud formal que un sistema intenta maximizar o minimizar.
Variable sustitutiva	Medida indirecta utilizada en lugar de un objetivo más difícil de medir.
Banco de pruebas	Prueba utilizada para comparar el desempeño de los modelos.
Evaluación	Medición de un modelo o sistema con respecto a criterios definidos.
Incertidumbre	Grado en que una salida puede ser errónea o incompleta.
Puntuación de confianza	Estimación numérica asociada con una predicción o clasificación.
Calibración	Relación entre la confianza predicha y la exactitud real.
Explicabilidad	Capacidad de proporcionar una explicación comprensible de cómo se produjo una salida.
Interpretabilidad	Grado en que puede comprenderse el comportamiento interno de un sistema.
Caja negra	Sistema cuyo funcionamiento interno es difícil de inspeccionar.
Alucinación	Salida de un modelo que es fluida pero no respaldada, falsa o fabricada.
Sesgo	Error sistemático o desempeño desigual asociado con los datos, el diseño o el uso.
Equidad	Enfoques para evaluar y abordar el trato o los resultados desiguales.
Robustez	Capacidad de funcionar de manera fiable en condiciones modificadas o imperfectas.
Ataque adversarial	Entrada diseñada deliberadamente para hacer que un sistema falle o se comporte de forma inesperada.

15.4 Modelos y entrenamiento

Red neuronal	Sistema computacional compuesto por capas conectadas que transforman entradas en salidas.
Perceptrón	Modelo matemático temprano de una sola célula cerebral, y base de investigaciones posteriores sobre redes neuronales.
Transformer	Arquitectura de red neuronal que utiliza la atención para procesar relaciones entre tokens.
Pesos	Números que un modelo aprendió durante el entrenamiento. En la práctica, los pesos son el modelo.
Pesos abiertos	Pesos publicados para que cualquiera pueda ejecutar, inspeccionar o alojar el modelo por sí mismo.
Pesos cerrados	Pesos conservados en los servidores del proveedor y accesibles únicamente a través de su servicio.
Preentrenamiento	Entrenamiento inicial con un gran corpus de datos, normalmente mediante la predicción de contenido faltante o subsiguiente.
Ajuste fino	Entrenamiento adicional con un conjunto de datos o una tarea más acotados.
Posentrenamiento	Procesos aplicados después del preentrenamiento para mejorar la utilidad, el comportamiento o la alineación.
Aprendizaje por refuerzo (RL)	Entrenamiento mediante retroalimentación que recompensa algunas acciones y penaliza otras.
Aprendizaje por refuerzo a partir de retroalimentación humana (RLHF)	Aprendizaje por refuerzo que utiliza preferencias o evaluaciones humanas.
Modelo de razonamiento	Modelo o sistema optimizado para dedicar computación adicional a problemas de varios pasos.
Cadena de pensamiento	Texto de razonamiento intermedio generado durante la resolución de problemas.
Leyes de escalamiento	Relaciones observadas entre el desempeño del modelo, el tamaño del modelo, los datos de entrenamiento y la capacidad de cómputo.
Mixture of experts (MoE)	Arquitectura que contiene múltiples subredes especializadas, de las cuales solo algunas se utilizan para cada entrada.

15.5 Gobernanza y responsabilidad

Alineación	Diseño de un sistema para que su comportamiento refleje objetivos o valores especificados.
Seguridad de la IA	Investigación y prácticas orientadas a reducir comportamientos dañinos o no controlados de los sistemas.
Gobernanza de la IA	Reglas, instituciones, procesos y prácticas que configuran el desarrollo y el uso de la IA.
IA responsable	Desarrollo y uso de la IA prestando atención a sus impactos sociales, éticos, jurídicos y ambientales.
Diseño sensible a los valores	Diseño de tecnología con atención explícita a los valores humanos y a los intereses afectados.
IA alineada con la soberanía	IA diseñada y gobernada para preservar la capacidad de acción de una comunidad o un lugar definidos.
Permisos granulares	Restricciones sobre las herramientas, archivos, datos o acciones a los que puede acceder un sistema.
Entorno aislado	Entorno separado en el que un sistema puede operar con acceso limitado.
Punto de aprobación humana	Autorización humana requerida antes de que una acción continúe.
Registro de auditoría	Registro de entradas, salidas, acciones y cambios del sistema.
Ficha del modelo	Documentación que describe el uso previsto, las limitaciones y la evaluación de un modelo.
Ficha del sistema	Documentación que abarca un sistema desplegado más amplio, incluidos los riesgos y las medidas de mitigación.
Deriva del modelo	Cambio en el desempeño a medida que los datos, el comportamiento o las condiciones cambian con el tiempo.

Control de versiones	Seguimiento de los cambios en modelos, datos, prompts y software.
Reversión	Restablecimiento de un sistema o despliegue a una versión anterior.
Autoalojamiento	Ejecución de un modelo en hardware que usted o su institución controlan.
Retención nula de datos	Promesa contractual de que un proveedor no almacena nada de lo que usted le envía.

Referencias

1. El perceptrón de Rosenblatt y la demostración de 1958, Cornell Chronicle
2. Krizhevsky, Sutskever, Hinton, *ImageNet Classification with Deep Convolutional Neural Networks*, NeurIPS 2012
3. Vaswani et al., *Attention Is All You Need*, Google Research 2017
4. Radford et al., *Language Models are Unsupervised Multitask Learners*, OpenAI 2019
5. Kaplan et al., *Scaling Laws for Neural Language Models*, OpenAI 2020
6. Hoffmann et al., *Training Compute-Optimal Large Language Models*, NeurIPS 2022
7. OpenAI, *Introducing ChatGPT*, 30 November 2022
8. OpenAI, *Learning to Reason with LLMs*, 12 September 2024
9. DeepSeek, *DeepSeek-R1*, 20 January 2025
10. METR, *Measuring AI Ability to Complete Long Tasks*, 2025
11. Anthropic, *Anthropic Economic Index*, 2025, uso por país y nivel de ingresos
12. Global Biodiversity Information Facility, datos de presencia y su sesgo de muestreo geográfico, gbif.org y el GBIF data blog
13. Artificial Analysis, *Intelligence Index by Open Weights and Proprietary*, 20 August 2026
14. Artificial Analysis, *Intelligence Index vs. Cost per Intelligence Index Task*, 20 August 2026
15. NVIDIA, guía de usuario de DGX Spark y descripción general del hardware
16. Google, *Measuring the environmental impact of delivering AI at Google scale*, arXiv:2508.15734, agosto de 2025
17. Epoch AI, *How much energy does ChatGPT use?*, 2025
18. GainForest, footprint.ts en pi-village-core, el cálculo en que se basan las cifras por cuenta de Polly
19. GainForest, *Generative AI + Climate Guide 2024*, preparado para el Climate Youth Negotiator Programme antes de la COP29