AI DAILY REPORT

Informe nocturno integral de IA del 17 de octubre de 2026

Condiciones operativas de modelos abiertos, validación en campo y juicio humano

2026. 10. 17

Se publican conjuntamente los pesos de modelos de IA y herramientas de búsqueda, mientras las empresas concretan las condiciones de uso y el equipamiento para ejecutarlos en servidores propios. En circuitos cuánticos y cálculos médicos, han avanzado las investigaciones que verifican respuestas de forma automática. En el ámbito industrial, se preparan instalaciones de demostración y herramientas de propuesta de procesos, al tiempo que se amplían el acceso para evaluadores externos y los procedimientos de verificación humana de resultados.

1 DeepSeek publica conjuntamente los pesos del modelo y herramientas de búsqueda

2 Step 5 Preview revela condiciones para ejecución en servidores propios

3 Claude reduce etapas de cálculo en el diseño de circuitos cuánticos

4 Aumenta el trabajo de agentes y cae la proporción de contrataciones de nivel inicial

5 Apertura de herramientas internas y derecho de publicación a evaluadores externos

6 Tencent conecta la memoria de múltiples herramientas de programación

7 BlenderWorld evalúa resultados 3D editables

8 Twelve Labs conecta la búsqueda de video con el premontaje

9 El robot de estacionamiento exterior Stan se exhibirá en Gangneung

10 Complejo siderúrgico de Pohang inicia construcción de centro piloto de IA

11 Mazak desarrolla IA para proponer procesos de mecanizado y cronogramas

12 Cálculo en registros clínicos: separar generación de texto y operaciones

13 Jiangsu realiza investigación de campo para legislar la promoción de la IA

14 Copilot para asuntos jurídicos conecta revisión de contratos y sugerencias de cambio

15 IA de defensa asiste en documentación y programación dentro de redes cerradas

Modelos

1 DeepSeek publica conjuntamente los pesos del modelo y herramientas de búsqueda

DeepSeek lanzó V4.1-Flash-1017 el 17 de octubre. Está disponible en su sitio web, aplicación y API, y también se pueden descargar los pesos, que son los valores de cálculo entrenados. En la evaluación de codificación interna de la empresa, registró un 77.0 % por menos de 0.20 dólares por tarea. En una configuración de inferencia baja, alcanzó un 72 % o más por alrededor de 0.10 dólares. Al tratarse de costos y puntuaciones de configuraciones de inferencia diferentes, no deben interpretarse conjuntamente como los resultados de una única condición.

También puso a disposición de los desarrolladores herramientas de búsqueda y extracción de contenido web. El precio es de 0.75 dólares por cada 1,000 llamadas y se ofrece de forma gratuita por tiempo limitado a los usuarios de la API que se conecten con funciones del lado del servidor a herramientas de desarrollo designadas. El rendimiento y los costos corresponden a anuncios propios de la empresa. Al incorporar el modelo al trabajo real, se deben verificar no solo la tasa de aciertos, sino también los fallos de búsqueda y las ejecuciones erróneas de herramientas.

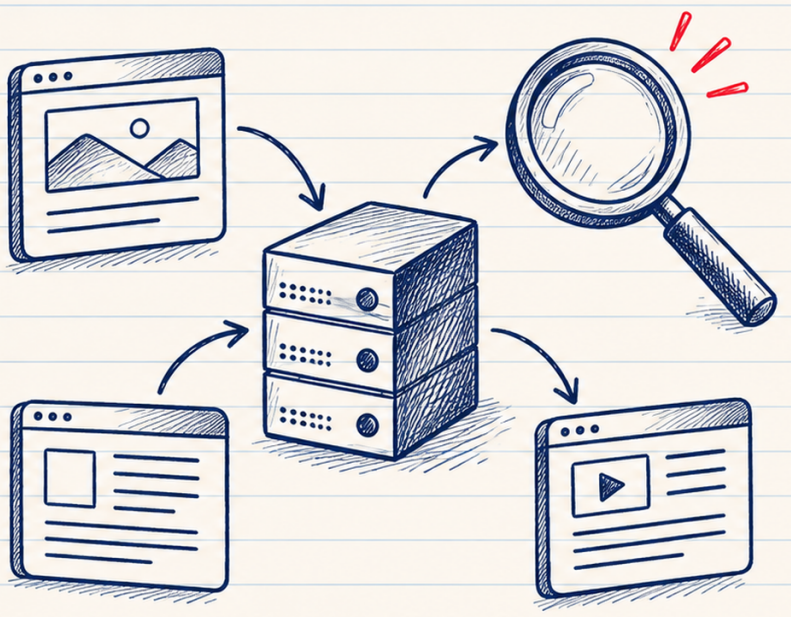
Cuaderno de IA de hoy

Conexión con búsqueda

Publicación de pesos

Herramientas de búsqueda

Rendimiento autoevaluado



Informe integral de IA de KIMKJ

Operación local

2 Step 5 Preview revela condiciones para ejecución en servidores propios

Step 5 Preview de StepFun es un modelo que activa 27 mil millones de un total de 600 mil millones de parámetros al procesar entradas. Admite un contexto de hasta 1 millón de tokens. Un token es la unidad en la que el modelo divide y procesa el texto. Con la publicación de los pesos, además de enviar solicitudes a servicios externos, las empresas ahora disponen de la opción de operarlo en sus propios servidores.

Sin embargo, que sea un modelo abierto no significa que pueda ejecutarse con pocos recursos de hardware. El ejemplo oficial de ejecución en FP8 utiliza 4 GPU B300 u 8 GPU H200. Los requisitos reales de memoria varían según la longitud del contexto y el volumen de procesamiento concurrente. También se aplican condiciones de licencia propias. Se requiere un contrato independiente para la incorporación de funciones en productos de fabricantes de dispositivos o para negocios de servicios de modelos que superen un determinado nivel de ingresos. Se debe distinguir entre la publicación de pesos y el uso comercial sin restricciones.

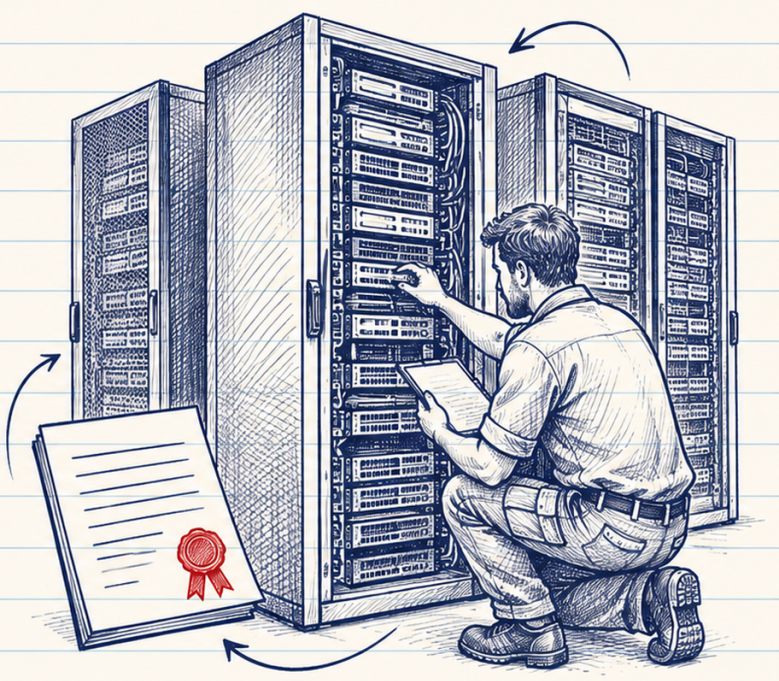
Cuaderno de IA de hoy

Condiciones de servidor propio

600 mil millones de parámetros

Verificar capacidad de hardware

Condiciones comerciales aparte



Informe integral de IA de KIMKJ

Ciencia

3 Claude reduce etapas de cálculo en el diseño de circuitos cuánticos

Anthropic anunció que Claude Mythos 5 mejoró diseños públicos existentes en un estudio enfocado en el diseño de 8 tipos de componentes de circuitos cuánticos. Cuando los investigadores describen las reglas de cálculo en texto, el modelo genera el código y un evaluador automático verifica la precisión y los recursos necesarios. El modelo recibe los resultados y corrige el diseño. El alcance del estudio se limitó a tareas cuyas respuestas correctas pueden verificarse mecánicamente.

Un diseño para sumar puntos de curvas elípticas utilizó un 12 % más de cúbits que el diseño mínimo existente, pero el número de compuertas, que representan las etapas de cálculo, se redujo a una decimoséptima parte (1/17). En la práctica, encontró un equilibrio diferente entre la cantidad de cálculo y los cúbits que actúan como memoria. Dicha exploración tomó alrededor de 53 horas, y los humanos proporcionaron las instrucciones iniciales y sugerencias intermitentes de dirección. Esto constituye un resultado de diseño de circuitos obtenido por los investigadores de la empresa y no demuestra la comercialización de computadoras cuánticas reales ni la aceleración de todos los cálculos.

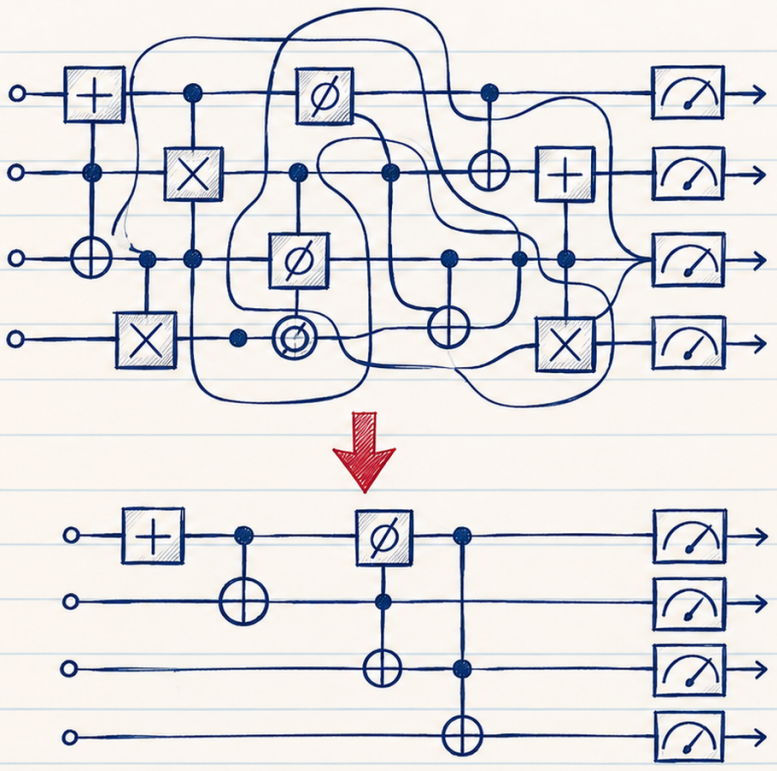
Cuaderno de IA de hoy

Rediseño de circuitos cuánticos

8 tareas de diseño

Verificación automática

Hardware terminado es aparte



Informe integral de IA de KIMKJ

Empleo

4 Aumenta el trabajo de agentes y cae la proporción de contrataciones de nivel inicial

Dentro de Anthropic, el tiempo de trabajo de los agentes de IA conversacionales alcanzó en agosto aproximadamente 6 veces el del personal. El tiempo humano es una estimación calculada multiplicando el número de empleados por 40 horas semanales. Los registros de cambios de código por ingeniero también aumentaron a cerca de 5 veces los de hace 6 meses. Sin embargo, el volumen de producción de código no equivale a la misma proporción de ingresos o rendimiento empresarial.

En las nuevas contrataciones de puestos técnicos, la proporción de empleados de nivel inicial se redujo del 27 % al 9 %. Aunque la empresa analizó datos de empleados y 17 entrevistas, no se trató de un experimento con asignación aleatoria de usuarios de IA. Dado que confluyen el crecimiento empresarial y diferencias previas a la adopción, no se puede concluir con certeza que la IA haya provocado la disminución de contrataciones iniciales. También surgió un análisis que indica que leer todo el contenido generado por los agentes requiere la mitad o más del tiempo de trabajo humano. La brecha entre la capacidad de generar mucho trabajo y la capacidad humana para verificar los resultados se está ampliando.

Cuaderno de IA de hoy

Carga laboral y contratación

- Aumento de horas de agentes
- Menor cuota de nivel inicial
- Causalidad no confirmada



Informe integral de IA de KIMKJ

Verificación de seguridad

5 Apertura de herramientas internas y derecho de publicación a evaluadores externos

Anthropic otorgó a 6 empleados de METR, una organización sin fines de lucro dedicada a evaluar riesgos de la IA, acceso a los equipos, espacios de oficina y herramientas internas de la empresa. El propósito es permitirles examinar los modelos y las prácticas operativas con herramientas similares a las del equipo interno de evaluación de riesgos, yendo más allá de pruebas consistentes únicamente en formular preguntas externas a los modelos. Esto no implica que se haya emitido un veredicto de que un modelo en particular sea seguro.

METR puede publicar de forma independiente análisis sobre niveles de riesgo, incidentes y prácticas de la empresa. Anthropic no puede rechazar la publicación en sí y solo puede solicitar la eliminación de una categoría restringida y predefinida de información, como datos de clientes o la arquitectura de los modelos. METR puede expresar públicamente su opinión sobre si dicha solicitud es justificada. El cambio concreto de este sistema radica en que amplía el alcance de acceso al tiempo que otorga el derecho de divulgar el juicio de los evaluadores fuera de la empresa.



Agentes

6 Tencent conecta la memoria de múltiples herramientas de programación

Hy Memory 2.0 de Tencent se compone de un modelo entrenado para la memoria, herramientas de desarrollo para almacenar y actualizar recuerdos, y MemHub, un administrador local. MemHub recopila el historial de trabajo de varias herramientas de codificación y lo organiza como memoria compartida. En las conversaciones, agrupa eventos que incluyen tiempo, lugar y personas, resuelve contenidos duplicados o contradictorios y los transfiere a la siguiente tarea.

En la evaluación LoCoMo presentada por Tencent, los tokens por consulta se redujeron de 1,159 a 703. No obstante, se trata de una autoevaluación de la empresa y los investigadores, y no significa que se obtenga la misma tasa de reducción en todas las tareas. También existe la condición de que se empleó un conjunto fijo de muestras para la comparación de modelos. La aplicación MemHub se proporciona para Mac.

Dado que el modelo de memoria, las herramientas de desarrollo y la aplicación de usuario tienen funciones distintas, el entorno requerido y el alcance del historial compartido varían según las partes que se instalen.



Evaluación

7 BlenderWorld evalúa resultados 3D editables

BlenderWorld, publicado por el equipo Qwen de Alibaba, evalúa si la IA puede completar tareas en el programa de creación 3D Blender. Consta de 223 tareas en 9 áreas, como productos, arquitectura, personajes y animación, y examina los resultados según 2,651 criterios. Además de tareas que solo proporcionan texto, incluye desafíos que suministran imágenes o videos junto con texto.

El resultado final no debe ser una sola imagen terminada, sino un proyecto editable de Blender equipado con materiales, geometría, iluminación y cámaras. Con base en el renderizado desde 8 puntos de vista y la inspección de escenas, se ponderan a partes iguales el cumplimiento de especificaciones y la calidad visual (50 % cada uno). La ejecución requiere Linux, Blender y un servicio de modelo que procese información visual. La publicación del entorno de evaluación no implica una clasificación definitiva de modelos. Las tareas se diseñaron para distinguir e inspeccionar entre imágenes estéticamente atractivas y archivos de producción modificables.



IA en Corea

8 Twelve Labs conecta la búsqueda de video con el premontaje

Jockey, el agente de IA para comprensión de video de Twelve Labs, localiza escenas en bibliotecas de video mediante solicitudes en lenguaje natural y expone los motivos por los que seleccionó dichas escenas. También ofrece funciones para elegir clips basándose en reacciones como visualizaciones y «me gusta», o para extraer partes necesarias del material original con el fin de generar un premontaje. Es un servicio que conecta un modelo de comprensión de contenido de video con tareas de producción reales.

La empresa explicó que Jockey puede invocarse desde las aplicaciones de ChatGPT y Claude, manteniendo el contexto de búsqueda incluso al alternar entre ambos servicios. Al introducir un enlace de video, también proporciona tendencias e indicadores de interacción asociados. Declaró haber incorporado los comentarios de pruebas preliminares realizadas con productoras que gestionan grandes archivos de video. Sin embargo, la mera presentación de estas funciones no permite confirmar la precisión de la búsqueda de video ni la tasa de ahorro de tiempo de edición. Queda pendiente el proceso en el que el usuario verifique los motivos de selección de escenas y el resultado del premontaje.



Robótica

9 El robot de estacionamiento exterior Stan se exhibirá en Gangneung

HL Robotics presentará por primera vez en Corea el robot autónomo de estacionamiento exterior Stan en el Congreso Mundial de Sistemas Inteligentes de Transporte de Gangneung, que se inaugura el 19 de octubre. Demostrará en un entorno real el proceso de entrada y salida de vehículos en el estacionamiento del Centro de Hockey de Gangneung. Al 17 de octubre se trata de un plan de demostración, por lo que no debe interpretarse como que el evento ya concluyó o que se completó su despliegue nacional.

La empresa explicó que Stan cuenta con 10 años o más de experiencia operativa en aeropuertos extranjeros y centros logísticos ferroviarios. También presentará un sistema de control integrado que traslada los espacios de estacionamiento reales a un modelo virtual para monitorizar simultáneamente la ubicación, la ruta de desplazamiento y el estado operativo de los robots y vehículos. El objetivo es automatizar la entrada y salida de vehículos tanto en interiores como en exteriores vinculándolo con Parkie, el robot de estacionamiento interior. La experiencia operativa en el extranjero, la demostración en el evento nacional y la futura expansión de adopción constituyen etapas distintas.

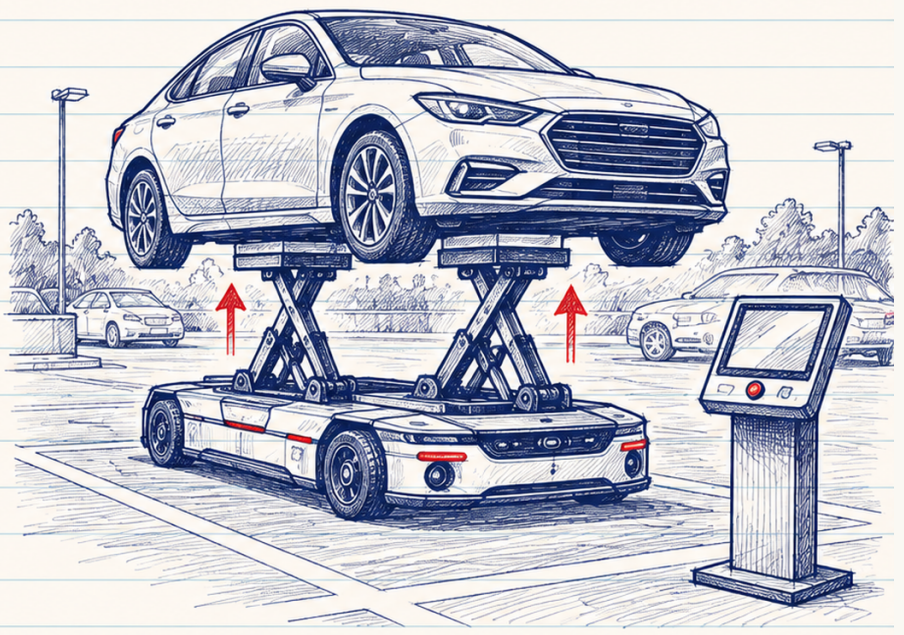
Cuaderno de IA de hoy

Demostración de aparcamiento

Evento del 19 de octubre

Prueba en estacionamiento real

Expansión nacional como meta



Informe integral de IA de KIMKJ

Industria

10 Complejo siderúrgico de Pohang inicia construcción de centro piloto de IA

El Instituto de Investigación de Ciencia y Tecnología Industriales de Pohang celebró del 15 al 16 de octubre en Gyeongju la reunión de inicio del proyecto de transformación con IA del complejo industrial siderúrgico de Pohang. El plan contempla una inversión total de 23,500 millones de wones durante 4 años, de 2026 a 2029, para construir un centro de apoyo integral e instalaciones de prueba que conecten procesos virtuales y reales. Se impulsará conjuntamente la fabricación de equipos y la verificación de su rendimiento in situ.

Con Dongkuk Steel y Jeil Technos como ejes centrales, se desarrollarán y validarán modelos representativos aplicables a los procesos siderúrgicos. Las universidades e institutos de investigación desarrollarán algoritmos adaptados al terreno, al tiempo que elaborarán planes de difusión y marcos de cooperación para todo el parque industrial. Actualmente se encuentra en la fase inicial del proyecto. El monto de financiamiento y el número de entidades participantes no deben interpretarse directamente como mejoras en la productividad real. La medida exacta en que se reducirán los defectos o los tiempos de trabajo en cada proceso se determinará mediante pruebas de campo tras la finalización de las instalaciones.

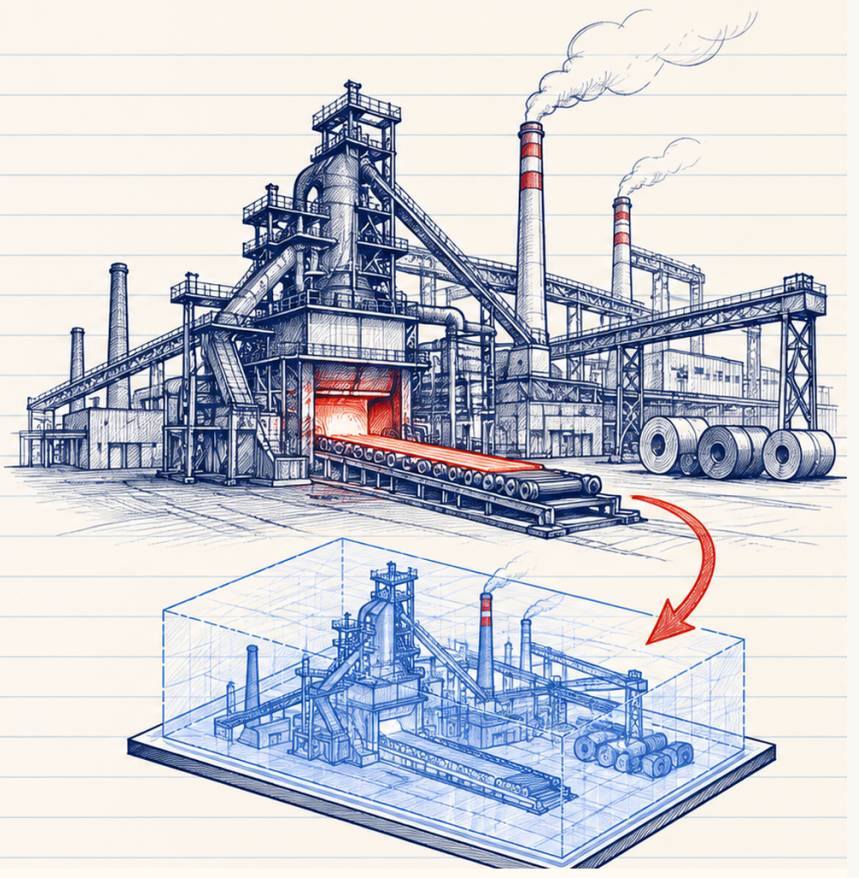
Cuaderno de IA de hoy

Validación siderúrgica in situ

23,500 millones de KRW en 4 años

Creación de centro de apoyo

Resultados tras la validación



Informe integral de IA de KIMKJ

Manufactura

11 Mazak desarrolla IA para proponer procesos de mecanizado y cronogramas

El fabricante japonés de máquinas herramienta Yamazaki Mazak tiene como objetivo desarrollar una IA que proponga procesos de mecanizado de piezas y cronogramas de producción para comercializarla dentro del año fiscal 2028. La IA para procesos de mecanizado analiza planos y modelos 3D para proponer métodos de mecanizado, secuencias y condiciones de corte. Emplea para el entrenamiento datos de intención de mecanizado que especialistas experimentados han plasmado en el lenguaje de programación MAZATROL.

Otra IA aborda la programación de la producción. Se diseña para vincularse con aplicaciones de gestión de equipos automatizados y ajustar prioridades ante paradas imprevistas de maquinaria o escasez de piezas. Adopta la plataforma operativa de ABEJA y se proyecta permitir el reentrenamiento tras la entrega mediante los datos operativos propios de la fábrica del cliente. Ambas funciones son planes de desarrollo y no representan productos terminados desplegados actualmente en fábricas ni cifras de productividad medidas en la práctica. Constituyen un intento de plasmar el criterio de técnicos expertos tanto en los procesos como en la planificación.

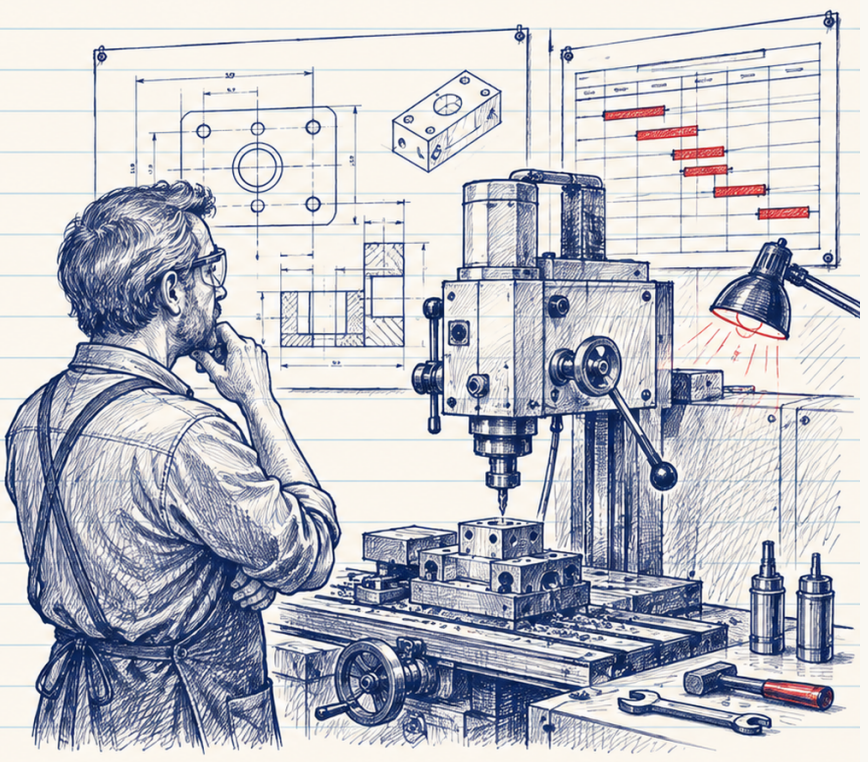
Cuaderno de IA de hoy

Propuesta de proceso y horario

De planos a procesos de corte

Respuesta a paradas imprevistas

Meta de producto para 2028



Informe integral de IA de KIMKJ

Investigación médica

12 Cálculo en registros clínicos: separar generación de texto y operaciones

Ha surgido una investigación de IA que separa la localización de valores necesarios en historiales clínicos del cálculo de puntuaciones médicas a partir de dichos valores. En lugar de hacer que el modelo memorice reglas de cálculo y redacte las respuestas directamente, busca las fórmulas de cálculo, crea herramientas de código ejecutables y extrae la información del paciente en un formato definido. La estructura delega el cómputo propiamente dicho a dicha herramienta.

En la evaluación MedCalc-Bench, la precisión final de respuestas correctas fue del 86.73 %. Superó en 24.0 puntos porcentuales al método de un solo ejemplo y en 17.27 puntos porcentuales al grupo de comparación con aprendizaje exclusivamente supervisado. Los errores residuales se concentraron en la interpretación semántica de los historiales y el formato de respuesta más que en la aritmética. El artículo fue aceptado en BIBM 2026 y publicado en arXiv. Esta evaluación corresponde a un conjunto de datos definido. Se requiere una validación independiente para determinar si se mantiene el mismo rendimiento ante registros reales de hospitales, complejos o con omisiones, y si mejora los resultados clínicos de los pacientes.



Legislación

13 Jiangsu realiza investigación de campo para legislar la promoción de la IA

Una delegación del Comité Permanente de la Asamblea Popular de la provincia china de Jiangsu realizó del 14 al 15 de octubre una investigación de campo en Nanjing y Suzhou para legislar sobre la promoción del desarrollo de la industria de la IA. Escuchó opiniones de asambleas locales, departamentos gubernamentales, asociaciones sectoriales y representantes de empresas, y visitó instalaciones corporativas. Esta actividad, divulgada el 16 de octubre, constituye una investigación preparatoria para redactar normativas y no significa que una nueva ley ya haya sido aprobada.

El ámbito de debate abarcó el suministro de recursos de computación, el flujo de datos, la innovación tecnológica, la integración intersectorial, la apertura de entornos de aplicación y la gestión de la seguridad. La delegación propuso reflejar las diferencias industriales y los niveles de desarrollo regional, y formular reglamentos locales eficaces a través de consultas, argumentación de expertos, deliberaciones y enmiendas. Aunque planteó una orientación para abordar conjuntamente el apoyo a la industria y la gestión de la seguridad, las obligaciones concretas aplicables a las empresas y los plazos de implementación no quedan determinados únicamente por esta investigación de campo.

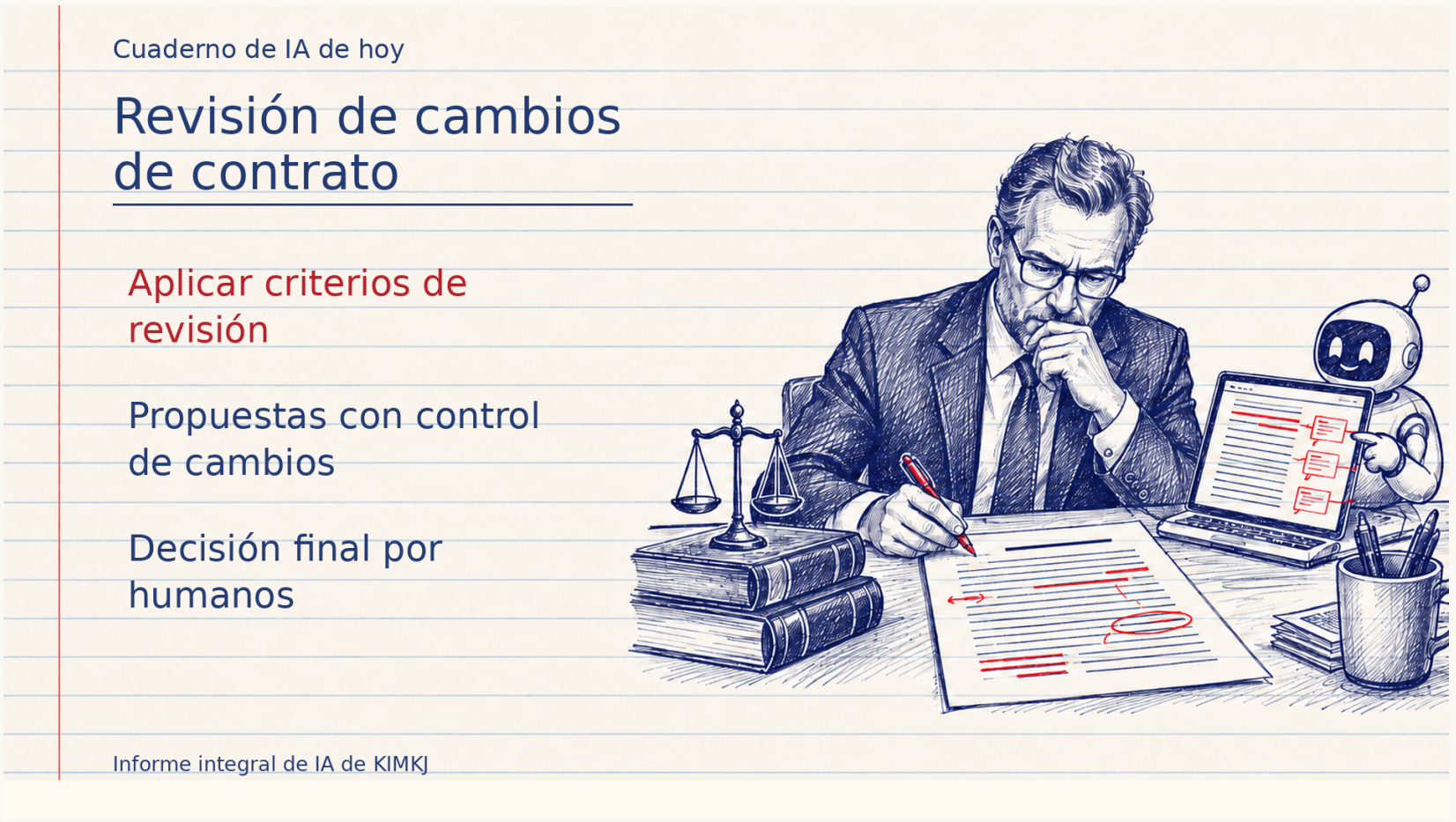


Herramientas de trabajo

14 Copilot para asuntos jurídicos conecta revisión de contratos y sugerencias de cambio

Copilot para el área legal de Microsoft agrupa una veintena de funciones que asisten en la redacción preliminar y revisión de contratos, así como en el seguimiento de obligaciones y plazos. En Word, resalta cláusulas que requieren atención según los criterios de revisión fijados por la empresa y propone modificaciones mediante control de cambios. También redacta borradores y notas jurídicas a partir de contratos anteriores y plantillas aprobadas de SharePoint. Su estructura requiere que el usuario examine las modificaciones una por una para decidir si las acepta.

En un espacio de trabajo que conecta múltiples aplicaciones, extrae obligaciones y plazos de contratos, correos electrónicos y actas de reuniones, y compara cláusulas entre distintos contratos. Requiere una licencia comercial de Microsoft 365 Copilot, y su disponibilidad varía según la región, el idioma y la configuración del administrador. La empresa especificó que se trata de una herramienta de asistencia para profesionales del derecho. Aunque el borrador resulte fluido, no finaliza automáticamente la verificación de citas, conclusiones ni los efectos jurídicos de las cláusulas.



Investigación en defensa

15 IA de defensa asiste en documentación y programación dentro de redes cerradas

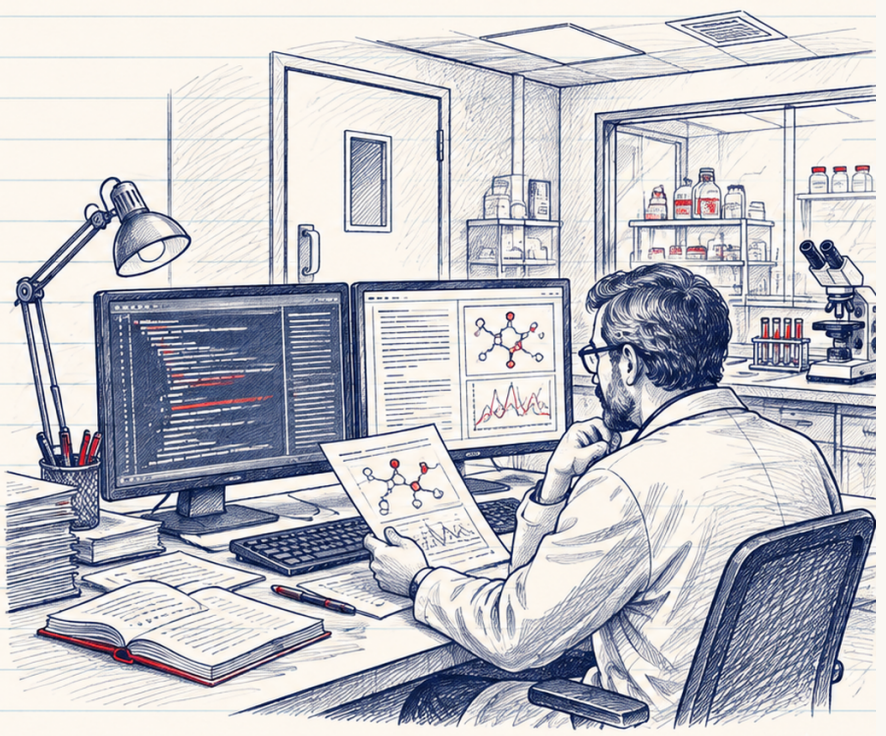
La Agencia para el Desarrollo de la Defensa de Corea comenzó el pasado mes de julio a operar para todo su personal ADDy, su propio servicio de IA generativa en un entorno aislado de redes de comunicación externas. Además de asistencia administrativa como la búsqueda de normativas y el resumen de documentos, ofrece funciones de ayuda para el desarrollo de software. Se trata de un caso de adopción previa confirmado en esta ocasión y no de un nuevo sistema de armas desplegado el 17 de octubre.

Los investigadores utilizan el servicio para redactar código en lenguaje natural y analizar errores. Se estructuró para realizar internamente el procesamiento de documentos y el soporte de desarrollo, incluso en entornos de investigación de defensa donde la conexión con servicios externos resulta difícil. Sin embargo, el mero hecho de su adopción no permite fijar en cifras la reducción del periodo de I+D ni el efecto en la prevención de incidentes de seguridad. La operación de un servicio que asiste en documentación y codificación tampoco debe extrapolarse como juicio autónomo en combate o autorización para el uso de armas.

Cuaderno de IA de hoy

Asistencia en red cerrada

- Inicio de operación en julio
- Búsqueda de normas y código
- Distinto de combate autónomo



Informe integral de IA de KIMKJ