



PUBLICACIONES DE LA
ACADEMIA NACIONAL DE
MEDICINA DE MÉXICO

ACTUALIDADES EN INTELIGENCIA ARTIFICIAL

Dr. Raúl Carrillo Esper
Dr. Rodolfo Palencia Díaz
Dr. Rodolfo de J. Palencia Vizcarra

Número 2

Evaluación de Modelos de Lenguaje de Gran Escala (LLM) para la Toma de Decisiones Clínicas Dinámicas y Multimodales.

Dr. Rodolfo Palencia Díaz

Dr. Rodolfo de J. Palencia Vizcarra

Médicos Internistas
Universidad de Guadalajara
Instituto Mexicano del Seguro Social (IMSS)
Colegiados y Certificados (CMMI)
Fundadores del TICC
28 de enero de 2026

Sintetizamos los hallazgos de una investigación avanzada sobre la capacidad de los modelos de lenguaje de gran escala (LLM) —específicamente agentes multimodales autónomos— para gestionar flujos de trabajo diagnósticos de extremo a extremo en entornos de simulación médica de alta fidelidad. A diferencia de las evaluaciones estáticas tradicionales, este análisis se centra en la toma de decisiones dinámica y multimodal en situaciones de cuidados agudos.

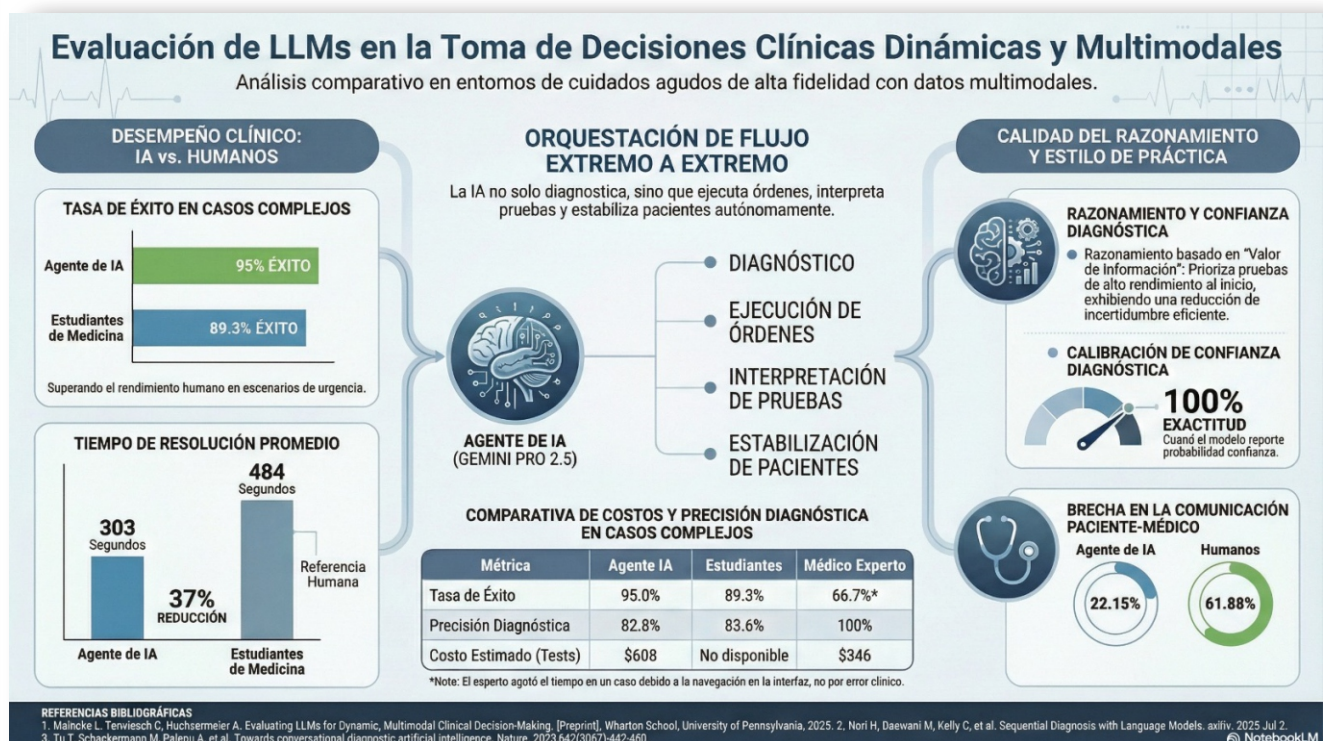
1. Marco Teórico: De Tareas Estáticas a Flujos de Trabajo Dinámicos

La medicina digital ha evolucionado de modelos especializados en tareas únicas (como la detección de sepsis o la interpretación de radiografías) hacia

sistemas universales. Sin embargo, la práctica clínica real no es estática; es un proceso dinámico de recopilación de información bajo incertidumbre.

La investigación adopta la perspectiva de un Proceso de Decisión de Markov Parcialmente Observable (POMDP) en tiempo continuo. En este modelo:

- Estado Latente (St): La condición fisiológica real del paciente.
- Observaciones (Ot): Signos vitales, resultados de laboratorio, imágenes y respuestas verbales.
- Acciones (At): Obtención de información (diálogo, examen físico) y tratamientos.
- Recompensa $^{\circ}$: Estabilización exitosa del paciente, precisión diagnóstica y eficiencia en tiempo y costo.



REFERENCIAS BIBLIOGRÁFICAS

1. Mainioke L, Tenwiesch C, Huchsermeier A. Evaluating LLMs for Dynamic, Multimodal Clinical Decision-Making. [Preprint]. Wharton School, University of Pennsylvania, 2025.
2. Nori H, Daewani M, Kelly C, et al. Sequential Diagnosis with Language Models. *arXiv*. 2025 Jul 2.
3. Tu T, Schackermann M, Palepu A, et al. Towards conversational diagnostic artificial intelligence. *Nature*. 2023;642(3057):442-460.

NotebookLM

2. Metodología y Arquitectura del Agente Autónomo

El estudio evaluó un agente basado en Gemini Pro 2.5 integrado a la plataforma de simulación BodyInteract.

Componentes del Sistema

- **Percepción Multimodal:** El agente procesa capturas de pantalla en tiempo real, señales de audio (auscultación con estetoscopio) y secuencias de video (tomografías computarizadas).
- **Arnés Técnico ("Harness"):** Una interfaz que permite al agente ejecutar comandos directamente sobre los elementos de la UI de la simulación, eliminando la latencia de la navegación manual y permitiendo un enfoque puro en el razonamiento clínico.
- **Ciclo de Control:** Un bucle de percepción-razonamiento-acción donde el agente evalúa el estado del paciente, emite un plan, ejecuta acciones y verifica los resultados.

3. Análisis de Competencia Clínica por Escenarios

Se evaluaron cuatro casos críticos que abarcan neurología, cardiología, neumología y endocrinología.

3.1. Caso de Baja Complejidad: Hipoglucemia Domiciliaria (Estudio 1)

El agente demostró una gestión impecable en un entorno de primer contacto:

- **Tasa de éxito:** 100% en 60 ejecuciones.
- **Precisión diagnóstica:** 97%.
- **Eficiencia temporal:** El agente fue significativamente más rápido (177 s) que los estudiantes de medicina (285 s).
- **Comportamiento:** Generó diagnósticos diferenciales apropiados, priorizó la medición de glucosa y escaló el tratamiento (de bebida azucarada a gel de glucosa) basándose en la respuesta del paciente.

3.2. Caso de Alta Complejidad: Neumonía y Sepsis en Emergencias (Estudio 2)

Este escenario requirió la interpretación de múltiples modalidades y pasos diagnósticos:

Resultados de éxito: 88.3% para el agente frente al 77.8% de los estudiantes.

Razonamiento clínico: Identificó correctamente leucocitosis con neutrofilia, consolidación radiográfica e hipoxemia.

Diagnóstico matizado: Aunque el estándar de la simulación era "neumonía", el agente diagnosticó "sepsis secundaria a neumonía" en muchos casos, demostrando una interpretación clínica sofisticada de la disfunción orgánica sistémica observada.

4. Comparativa: Agente IA vs. Estudiantes y Médicos Expertos

El estudio comparó el desempeño del agente contra un banco de datos de 14,691 sesiones de estudiantes y el benchmark de un médico de urgencias experto.

Métrica	Estudiantes de Medicina	Agente LLM (IA)	Médico Experto
Tasa de Éxito (Ajustada)	89.3%	95.0%	Alta (excepto por <i>timeouts</i>)
Tiempo Promedio	484 segundos	303 segundos	547 segundos
Precisión Diagnóstica	83.56%	82.78%	100% (en casos completados)
Comunicación con Paciente	61.88%	22.15%	Alta
Costo de Pruebas (Proxy)	Variable	Elevado (\$608 avg)	Bajo (\$346 avg)

Hallazgos Clave de la Comparativa (Estudio 3)

1. **Velocidad y Estabilización:** La IA actúa como un "estabilizador rápido". Prioriza acciones de alto impacto para evitar el deterioro del paciente, superando a los estudiantes en velocidad.
2. **Déficit de Comunicación:** El agente tiende a subinvertir en el diálogo con el paciente, centrándose casi exclusivamente en la logística diagnóstica y el tratamiento.
3. **Consistencia de Políticas:** La IA muestra una alta "consistencia de conjunto" (toma las mismas acciones necesarias), pero una "variabilidad de secuencia" (el orden de los pasos cambia entre ejecuciones).
5. El "Valor de la Información" y Calibración del Razonamiento (Estudio 4)

Al analizar el "estado de creencia" (bt) interno del agente, se identificaron propiedades fundamentales de su proceso de toma de decisiones:

- **Carga Frontal de Pruebas:** El agente aplica una estrategia de valor de información (VOI); las pruebas realizadas en los pasos iniciales producen los mayores cambios en la probabilidad diagnóstica (especialmente en el paso 2, tras obtener resultados de imagen o laboratorio).

- Reducción de Entropía: La incertidumbre diagnóstica (entropía de Shannon) disminuye drásticamente a lo largo de la sesión. En el caso de hipoglucemia, la entropía cayó de 1.78 a 0.31 bits.
- Calibración de Confianza: A diferencia de otros estudios sobre LLM, este agente mostró una calibración notable. Cuando el agente asignaba una probabilidad del 80-100% a un diagnóstico, su precisión era del 100%. Esto sugiere que el modelo "sabe lo que sabe".

6. Implicaciones para la Gestión Sanitaria y Colaboración Humano-IA

El estudio identifica tres roles críticos para estos agentes en el ecosistema de salud:

1. Motor de Triage Rápido: Capacidad para estabilizar casos sencillos y producir estimaciones de confianza calibradas para priorizar la atención humana en casos de alta incertidumbre.
2. Auditor en Tiempo Real: Actuar como un "segundo par de ojos" que procesa flujos multimodales en paralelo al médico, alertando sobre diagnósticos omitidos o pruebas críticas no solicitadas.
3. Extensión de Capacidad en Entornos Remotos: Proporcionar guía de estabilización en zonas sin médicos presenciales (zonas rurales, campos de batalla), donde un personal técnico puede ejecutar las acciones físicas bajo la dirección del agente.

7. Limitaciones Identificadas

- Fidelidad de la Simulación: Aunque BodyInteract es de alta fidelidad, sigue siendo un "gemelo digital" y no replica la complejidad total de un paciente humano real.
- Fricción de Interfaz: Parte de la ventaja de velocidad de la IA se debe a que no enfrenta la resistencia de navegación de menús que experimentan los humanos.
- Sesgo del Vendedor: Los resultados se limitan a la familia de modelos Gemini bajo una arquitectura específica.

Referencias Bibliográficas

1. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Intern Med.* 2023;183(6):589-596.
2. Bavafa H, Savin S, Terwiesch C. Customizing Primary Care Delivery Using E-Visits. *Prod Oper Manag.* 2021;30(11):4306-4327.
3. Chen TC, Multala E, Kearns P, Delashaw J, Dumont A, Maraganore D, et al. Assessment of ChatGPT's performance on neurology written board examination questions. *BMJ Neurol Open.* 2023;5(2):e000530.
4. Dai T, Singh S. Artificial Intelligence on Call: Physician's Decision of Whether to Use AI in Clinical Practice. *J Mark Res.* 2025;62(5):854-875.
5. Diaz-Navarro C, Armstrong R, Charnetski M, Freeman KJ, Koh S, Reedy G, et al. Global consensus statement on simulation-based practice in healthcare. *Adv Simul (Lond).* 2024;9(1):19.
6. Eisemann N, Bunk S, Mukama T, Baltus H, Elsner SA, Gomille T, et al. Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. *Nat Med.* 2025;31(3):917-924.
7. Geng J, Cai F, Wang Y, Koeppl H, Nakov P, Gurevych I. A Survey of Confidence Estimation and Calibration in Large Language Models. In: *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics*. Mexico City: Association for Computational Linguistics; 2024. p. 6577-6595.
8. Hopp WJ, Li J, Wang G. Big Data and the Precision Medicine Revolution. *Prod Oper Manag.* 2018;27(9):1647-1664.
9. Laker LF, Froehle CM, Windeler JB, Lindsell CJ. Quality and Efficiency of the Clinical Decision-Making Process: Information Overload and Emphasis Framing. *Prod Oper Manag.* 2018;27(12):2213-2225.
10. Su H, Sun Y, Li R, Zhang A, Yang Y, Xiao F, et al. Large Language Models in Medical Diagnostics: Scoping Review With Bibliometric Analysis. *J Med Internet Res.* 2025;27:e72062.



Sesión Clínica: Evaluación de LLM para la toma de decisiones clínicas dinámicas y multimodales

De la teoría estática a la práctica clínica simulada: Un análisis de Gemini Pro 2.5 en el entorno BodyInteract

Presentador: Dr. [Nombre del Experto]
Médico Internista | Doctorado en Medicina Digital e IA | Investigador Principal

Basado en la investigación de Meincke L, Terwiesch C, Huchzermeier A. (2025).



NotebookLM

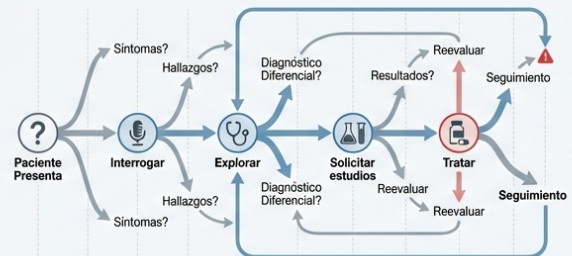
La medicina no es un examen de opción múltiple

El Paradigma Actual (Estático)



- Evaluación basada en preguntas aisladas (ej. USMLE, MedQA)
- Información completa de inicio ("Vignettes")
- Limitación: No refleja incertidumbre ni tiempo

La Realidad Clínica (Dinámica)

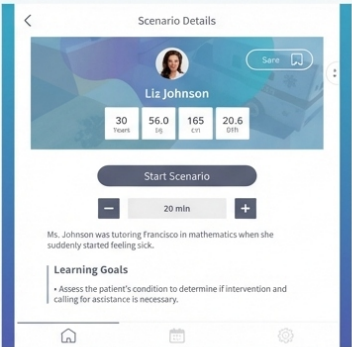
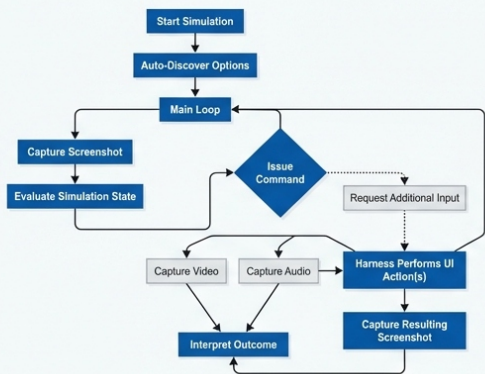


- Información incompleta al inicio (incertidumbre)
- Costos de información (tiempo y recursos)
- Orquestación secuencial: Interrogar → Explorar → Solicitar estudios → Tratar

"El verdadero reto no es interpretar una radiografía, sino saber cuándo pedirla y qué hacer con el resultado." [1]

NotebookLM

Metodología: El 'Agente Autónomo' en Simulación de Alta Fidelidad



Configuración

- Cerebro: Gemini Pro 2.5 (Multimodal)
- Cuerpo: Simulador BodyInteract
- Interfaz: Arnés técnico de 'loop' cerrado

Población Comparada

- IA (Agente): 60 corridas independientes
- Estudiantes: 14,000+ sesiones
- Gold Standard: Médico Experto

Escenarios Agudos

- Hipoglucemia
- Neumonía / Sepsis
- Ictus Isquémico
- Insuficiencia Cardíaca

NotebookLM

Caso 1: Hipoglucemia (Prueba de Concepto en Entorno Domiciliario)



Performance Dashboard

Éxito
100%
(Sin errores críticos)

Precisión Diagnóstica
97%

Tiempo de Resolución
177s (IA) vs. **285s** (Estudiantes)

Secuencia Lógica: Carbohidratos orales → Reevaluación → Gel de glucosa (escalada) → Llamada a emergencias

NotebookLM

Caso 2: Neumonía y Sepsis (El Reto de la Multimodalidad)



Escenario Clínico

Paciente de 58 años, fiebre, tos, desaturación.
Sala de Urgencias.

Manejo Escalado

Oxígeno (puntas → reservorio) + Antibióticos IV +
Fluidos.

Nota: La IA priorizó el diagnóstico de 'Sepsis' sobre
'Neumonía' debido al compromiso sistémico.

NotebookLM

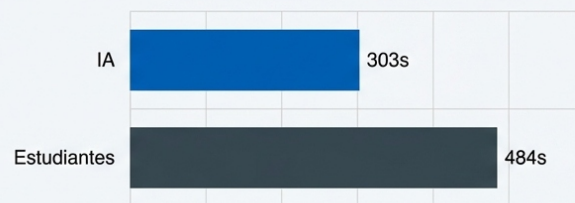
Comparativa de Desempeño: IA vs. Estudiantes de Medicina

Tasa de Éxito (Resolución)



La IA es más confiable para estabilizar.

Tiempo de Gestión (Segundos)



Reducción drástica del tiempo de decisión.

Precisión Diagnóstica: Comparable (IA 82.8% vs. Estudiantes 83.6%)

La IA actúa como un 'estabilizador rápido' en entornos agudos.

NotebookLM

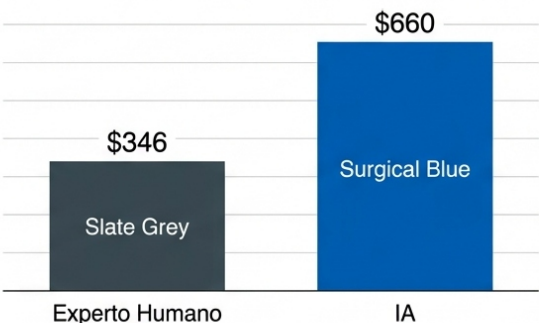
El Costo de la Eficiencia: Estilos de Práctica y ‘Bedside Manner’

Déficit de Comunicación



La IA prioriza la acción clínica sobre la empatía.

Uso de Recursos (Costos)

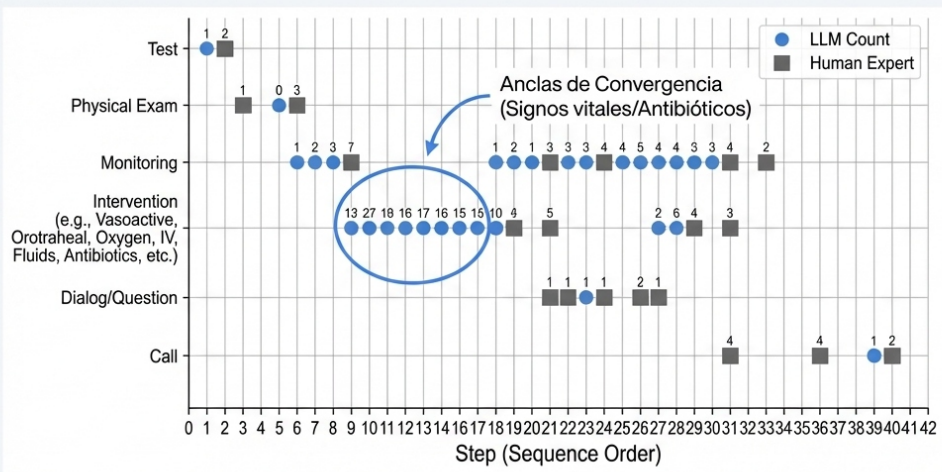


La IA gasta más en pruebas para reducir incertidumbre.

Conclusión: La IA actúa como un “Logista Clínico” eficiente pero distante.

NotebookLM

Anatomía de la Decisión: Variabilidad y Adaptabilidad

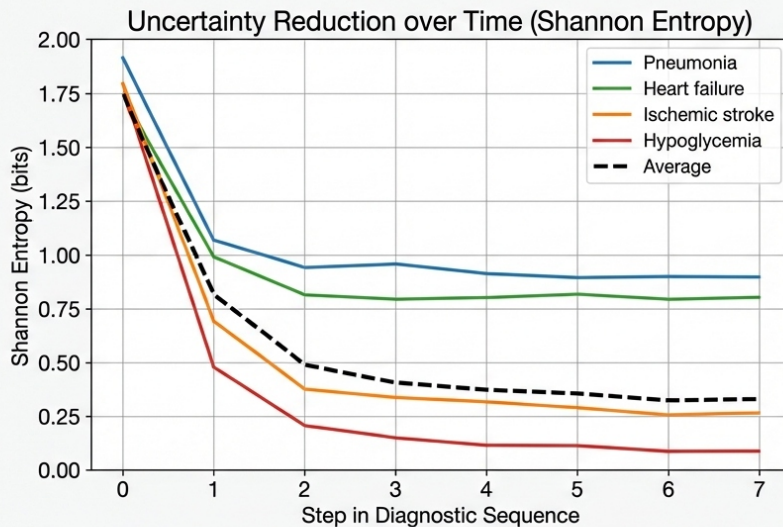


Solo el 3-8% de las corridas siguieron una secuencia idéntica.

La IA no es determinista. Muestra ‘flexibilidad de procedimiento’, adaptando su flujo de trabajo momento a momento similar a un clínico humano.

NotebookLM

Valor de la Información y Reducción de Entropía



Surgical Blue

Comportamiento racional (Value-of-Information):

La IA 'compra' información de alto valor en los pasos 1-2 (caída drástica de entropía) y refina marginalmente en los pasos 6-7.

NotebookLM

Calibración y Confianza: ¿Sabe la IA cuándo 'no sabe'?

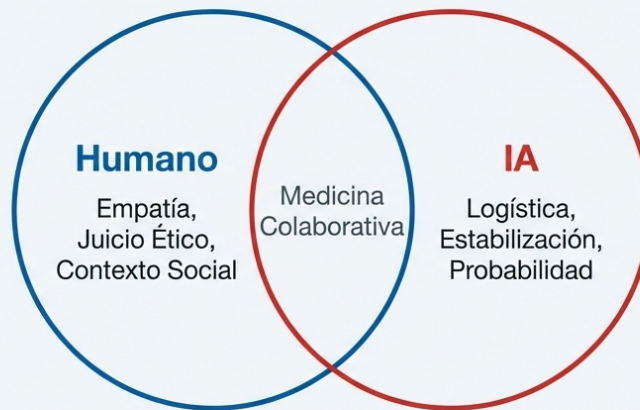
Confianza Reportada: 80-100% Precisión Real: 100%	
Confianza Reportada: 60-80% Precisión Real: 100%	
Confianza Reportada: 0-40% Precisión Real: 45%	

Hallazgo Crítico: La IA está bien calibrada. Su confianza reportada se alinea con su precisión real.

Implicación: Permite una supervisión efectiva. Si la IA duda, el humano debe intervenir.

NotebookLM

El Rol Futuro: El 'Logista Clínico' y la Red de Seguridad



- 1. **Triage Rápido:** Estabilización en entornos saturados.
- 2. **Segundo Par de Ojos:** Auditoría de flujos críticos (ej. Sepsis).
- 3. **Medicina Remota:** Apoyo a personal no experto.

NotebookLM

Limitaciones del Estudio y Cautela Académica



El 'Gemelo Digital' no es Realidad

La simulación es un modelo matemático, no fisiología humana real.



Comparador Limitado

Estudiantes vs. Un solo experto por caso (falta panel amplio de especialistas).



Sesgo de Interfaz

La velocidad de la IA puede deberse a la ausencia de fricción motora (UI).



Ventana Temporal

Solo casos agudos de 20 minutos; no evalúa manejo crónico.

NotebookLM

Conclusiones: Hacia una IA Clínica Dinámica

1. Competencia de Flujo de Trabajo

La IA supera las preguntas estáticas para gestionar pacientes completos de principio a fin.

2. Orquestación Multimodal

Integración exitosa de texto, audio e imagen en tiempo real para decisiones complejas.

3. Razonamiento Calibrado

Comportamiento racional de búsqueda de información y confianza alineada con la precisión.

La IA ha demostrado capacidad para actuar como un agente clínico competente en simulación, habilitando una colaboración Humano-IA segura.

NotebookLM

Referencias Bibliográficas

1. Meincke L, Terwiesch C, Huchzermeier A. Evaluating LLMs for Dynamic, Multimodal Clinical Decision-Making. [Preprint]. 2025.
2. Christensen CM, Cook S, Hall T. Marketing malpractice: the cause and the cure. Harv Bus Rev. 2005;83(12):74–83.
3. Tu T, et al. Towards conversational diagnostic artificial intelligence. Nature. 2025;642:442–450.
4. Nori H, et al. Sequential Diagnosis with Language Models. arXiv. 2025.

Nota: Todas las métricas y gráficos presentados provienen del estudio principal (Ref 1).

NotebookLM

Inteligencia Artificial Agente: El Poder de Cambiar la Medicina y Nuestro Mundo

Dr. Rodolfo Palencia Díaz
Dr. Rodolfo de J. Palencia Vizcarra

Médicos Internistas
Universidad de Guadalajara
Instituto Mexicano del Seguro Social (IMSS)
Colegiados y Certificados (CMMI)
Fundadores del TICC
28 de enero de 2026

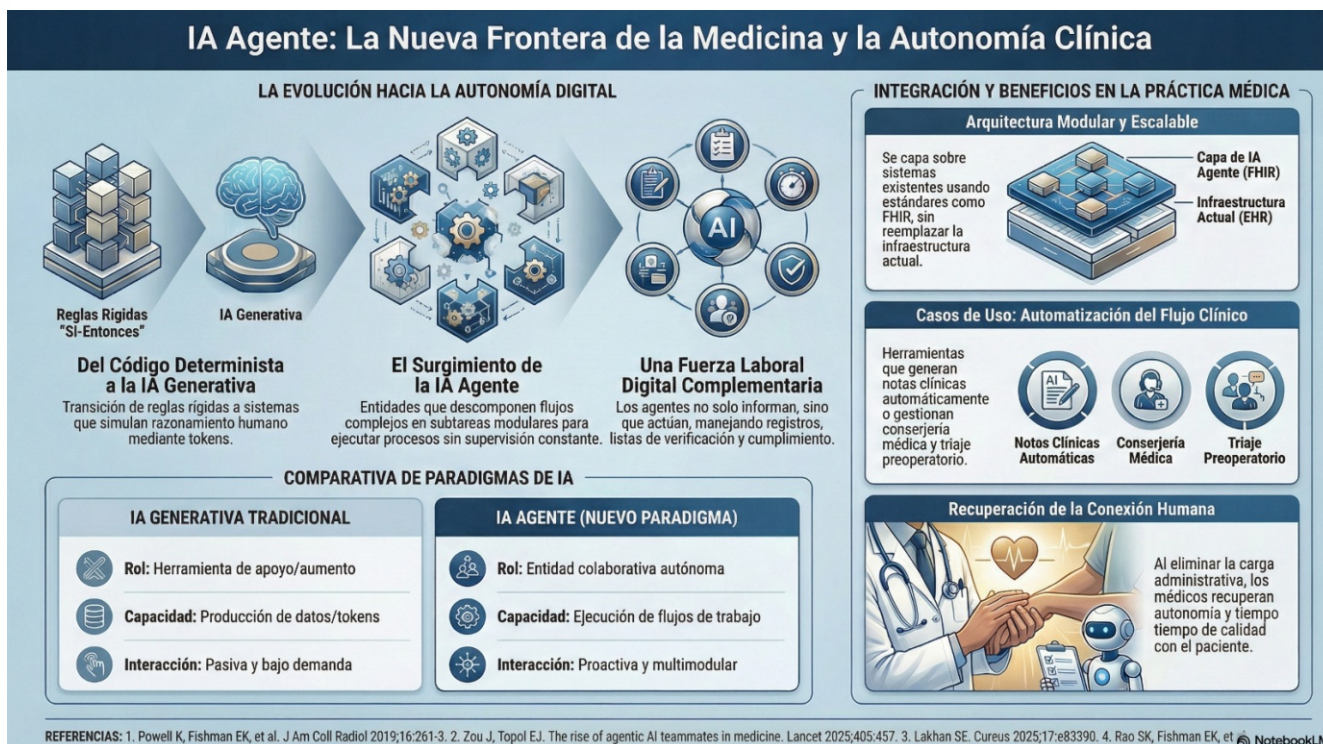
Resumen Ejecutivo

La industria de la salud se encuentra en el umbral de una era transformadora, impulsada por la transición de una inteligencia artificial (IA) generativa pasiva hacia una IA agente (o agentic AI). Este cambio representa una evolución desde herramientas que simplemente aumentan las capacidades individuales hacia entidades interactivas, colaborativas y autónomas capaces de ejecutar flujos de trabajo complejos. En un contexto donde los médicos están abrumados por la carga administrativa y el agotamiento profesional (burnout), la IA agente emerge no solo como una solución tecnológica, sino como un nuevo modelo operativo que permite reintegrar la conexión humana en el núcleo de la medicina [1-3].

Evolución del Paradigma Computacional en Medicina

La trayectoria de la computación en salud ha experimentado saltos cualitativos significativos:

1. Era Reglada: Dominada por códigos basados en estructuras simples de "si, entonces, de lo contrario" (if, then, else), ejecutados por unidades centrales de procesamiento (CPU) para resultados predecibles.
2. Revolución de las GPU: La llegada de las unidades de procesamiento gráfico (GPU) permitió el procesamiento en paralelo, fundamental para el desarrollo de redes neuronales y visión computacional.
3. IA Generativa: Capacidad de generar "tokens" (unidades discretas de lenguaje o datos), permi-



tiendo simular el razonamiento y el diálogo humano [1].

- 4. IA Agente: El estado actual, donde la IA trasciende la percepción y la generación para asumir la ejecución autónoma de tareas y flujos de trabajo multimodales [2,3].

Características Fundamentales de la IA Agente

A diferencia de la IA convencional, la IA agente se caracteriza por su modularidad y especialización. Funciona descomponiendo procesos de negocio y flujos de trabajo clínicos complejos en componentes modulares que imitan la forma en que los humanos realizan tareas [4,5].

Capacidades del Sistema de Agentes

- Autonomía: Ejecución de funciones complejas con mínima supervisión humana.
- Especialización por Subtareas: Un solo agente de salud puede integrar diversos servicios, tales como:
 - Recuperación y revisión de registros médicos electrónicos (EHR).
 - Análisis de literatura científica y tablas de referencia.
 - Gestión de listas de verificación (checklists).
 - Supervisión de cumplimiento normativo y privacidad.
- Interoperabilidad: Se integra sobre la infraestructura existente (como Epic) mediante estándares como FHIR (Fast Healthcare Interoperability Resources), actuando como un puente entre los datos del paciente y la información clínica accionable [5].

Aplicaciones Clínicas y Reducción de la Carga Administrativa

La IA agente aborda directamente el problema del tiempo limitado del personal de salud, ofreciendo soporte continuo y acceso a información bajo demanda.

Herramienta / Plataforma	Función Específica en el Flujo de Trabajo
Abridge	Escucha las visitas de los pacientes y genera automáticamente notas clínicas, permitiendo conversaciones naturales [5].
Deloitte's Frontline AI	Automatización de listas de verificación preoperatorias y conserjería médica [6].
Hippocratic.ai	Agentes de admisión digital y cuidados postoperatorios centrados en la seguridad [7].
Intrivo	Soporte de conserjería médica mediante agentes inteligentes [8].

Estos agentes han demostrado su capacidad para devolver la autonomía tanto a los proveedores como a los pacientes, integrando el poder de cómputo con los dominios intrincados de la atención directa.

Perspectiva Especializada: La Revolución en Radiología

La radiología se posiciona a la vanguardia de esta "revolución industrial de la inteligencia".

El impacto se prevé especialmente profundo en entornos colaborativos y multimodales:

- Toma de decisiones contextualizada: Integración en tiempo real de imágenes, genómica, patología, ensayos clínicos e historias longitudinales de los pacientes.
- Optimización de procesos: Mejora de la eficiencia en juntas de tumores (tumor boards) y la interacción con médicos referentes.
- Redescubrimiento de la práctica médica: Al optimizar los aspectos mundanos y administrativos, la IA permite al radiólogo enfocarse en el análisis complejo y el valor diagnóstico [5].

Conclusión: Un Sistema de Cuidado Equitativo y Compasivo

La emergencia de la IA agente trasciende las fronteras tradicionales de la IA perceptiva o generativa. No se trata simplemente de una mejora incremental de la eficiencia, sino de un cambio existencial en la funcionalidad de la industria de la salud. Al democratizar el acceso a la información y liberar a los profesionales de las cargas clericales, la IA agente establece las bases para un sistema de salud más eficiente, equitativo y, fundamentalmente, compasivo, donde la conexión humana vuelve a ser la prioridad [1,5].

Referencias Bibliográficas.

1. Powell K, Fishman EK, Horton KM, Johnson PT, Rowe SP. What health care can learn from self-driving vehicles. J Am Coll Radiol. 2019;16:261-3.
2. Zou J, Topol EJ. The rise of agentic AI teammates in medicine. Lancet. 2025;405:457.
3. Lakhan SE. The agentic era: why biopharma must embrace artificial intelligence that acts, not just informs. Cureus. 2025;17:e83390.
4. Dutta Roy TP. Bioethics artificial intelligence advisory (BAIA): an agentic artificial intelligence (AI) framework for bioethical clinical decision support. Cureus. 2025;17:e80494.
5. Rao SK, Fishman EK, Rizk RC, Chu LC, Rowe SP. Improving efficiencies while also delivering better

health care outcomes: a role for large language models. J Am Coll Radiol. 2024;21:1913-5.

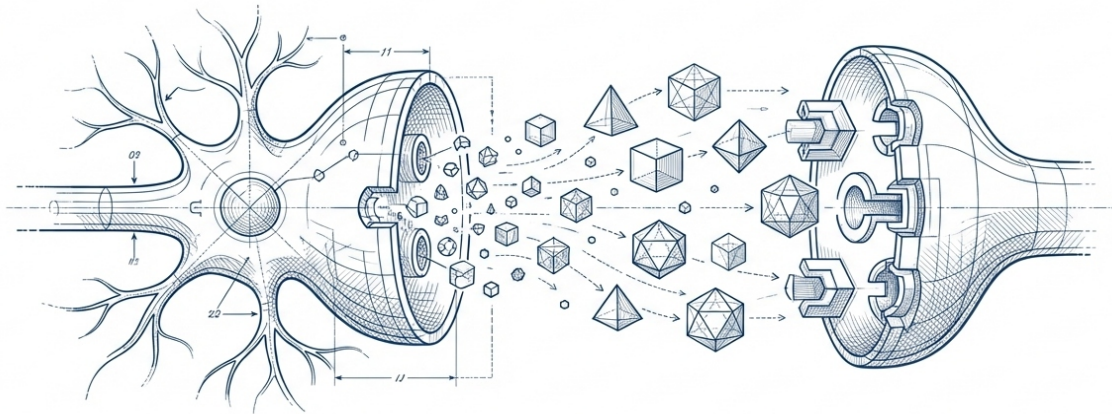
6. Vellody A, Sarker M. Agentic AI: autonomous generative AI agents. Deloitte [Internet]. 2025 [citado 21 de mayo de 2025]. Disponible en: <https://www.deloitte.com/in/en/services/consulting/services/core-businessoperations/agentic-ai.html>

7. Safety focused generative AI for healthcare. Hippocratic AI [Internet]. 2025 [citado 21 de mayo de 2025]. Disponible en: <https://www.hippocraticai.com>

8. Smart delightful AI. Intrivo [Internet]. 2025 [citado 21 de mayo de 2025]. Disponible en: <https://www.intrivo.com/ai-concierge>

Inteligencia Artificial Agente: El Poder de Cambiar la Medicina y Nuestro Mundo

De la generación de datos a la ejecución autónoma de tareas clínicas.

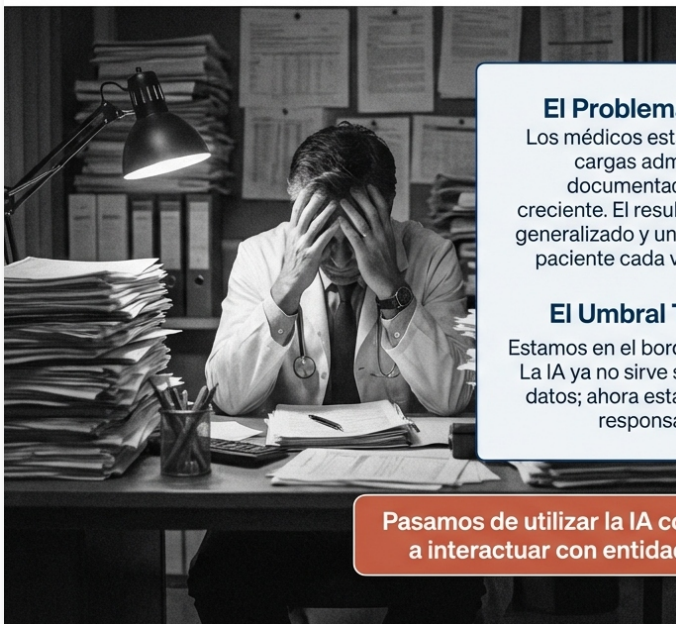


Presentado por: Dr. [Name], MD, PhD en Medicina Digital.

Basado en la investigación de: Kimberly Powell (NVIDIA), Elliot K. Fishman, MD (Johns Hopkins), Steven P. Rowe, MD, PhD (UNC), et al.

NotebookLM

La Paradoja Actual: Sobrecarga Clínica vs. Capacidad Tecnológica



El Problema Humanista

Los médicos están abrumados por cargas administrativas y documentación creciente. creciente. El resultado es el 'burnout' generalizado y una interacción con el paciente cada vez más reducida.

El Umbral Tecnológico

Estamos en el borde de una nueva era. La IA ya no sirve solo para aumentar datos; ahora estamos transfiriendo responsabilidades.

Pasamos de utilizar la IA como una herramienta pasiva a interactuar con entidades colaborativas activas.

Revisión de Imágenes



Actualización de Historial



Coordinación de Citas

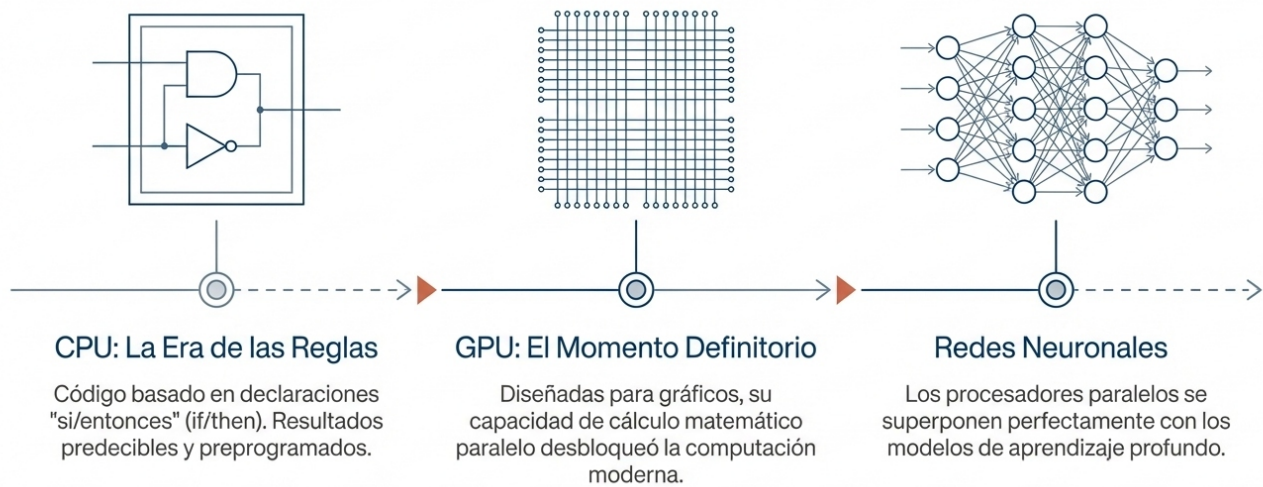


Análisis de Datos



NotebookLM

La Evolución del Cómputo: De Reglas a Razonamiento



Este descubrimiento catalizó un cambio fundamental: no solo mejoramos el software; cambiamos la arquitectura del razonamiento.

NotebookLM

Más Allá de los 'Tokens': Definiendo la IA Agente



IA Generativa (El Presente)

Genera "tokens" (unidades de lenguaje). Simula el razonamiento y el diálogo. Es perceptiva y creativa.

IA Agente (El Futuro Inmediato)

Capacidad de ejecución autónoma de tareas. No solo crea contenido; maneja flujos de trabajo complejos y procesos comerciales.

Una fuerza laboral digital complementaria.

NotebookLM

Anatomía de un Agente Digital: La Estructura Modular



Al igual que los humanos dividen tareas complejas, la IA Agente utiliza componentes modulares especializados que, integrados, manejan flujos de trabajo con mínima supervisión.

NotebookLM

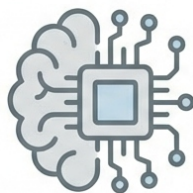
Recuperando el Tiempo: El Agente Administrativo

El Desafío:

El trabajo administrativo distrae a los proveedores y debilita la relación médico-paciente.



Voz / Conversación Natural



Notas Clínicas Estructuradas

La Solución (Caso: Abridge):

Agentes que escuchan, sintetizan y documentan automáticamente notas clínicas estructuradas.

Impacto: Elimina la pantalla como barrera.

El médico y el paciente pueden conversar, el agente recuerda los detalles.

NotebookLM

Extendiendo el Cuidado: El Agente de Servicio 24/7

Concepto:

Los agentes digitales no están limitados por el tiempo. Ofrecen acceso a información y triaje bajo demanda.



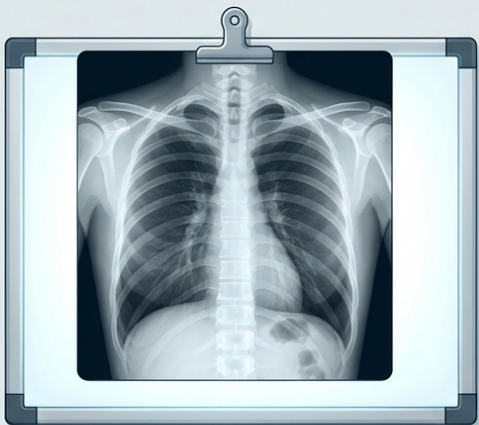
Ejemplos Reales:

- **Hippocratic.ai & Intrivo:** Admisión y listas preoperatorias.
- **Deloitte Frontline AI:** Conserjería médica.

Resultado: Miles de agentes desplegados globalmente complementan la estrategia de atención, permitiendo al personal humano centrarse en casos críticos.

NotebookLM

Lecciones desde la Radiología: La Vanguardia del Cambio



La radiología lidera la "Revolución Industrial de la Inteligencia".



El cambio: De "leer píxeles" a "optimizar flujos de trabajo".

La oportunidad: Automatizar lo mundano para redescubrir la alegría de la medicina diagnóstica.

NotebookLM

El Consejo de Tumores del Futuro: Integración Multimodal



Estado Futuro:
Convergencia perfecta de **datos dispares en la sala de conferencias o la clínica**, superando los flujos de trabajo fragmentados de la última década.

NotebookLM

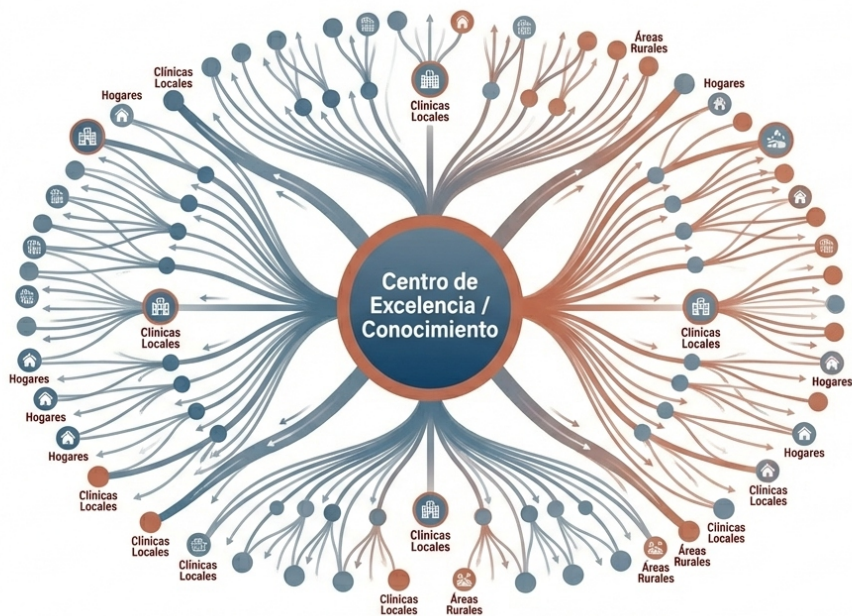
Implementación: Integración sin Fricción



Estrategia: No reemplazar ('Rip and replace'), sino construir encima. La IA Agente actúa como un puente entre los datos rígidos y la información procesable, integrándose directamente en los flujos de trabajo clínicos.

NotebookLM

Escalabilidad y Democratización del Conocimiento



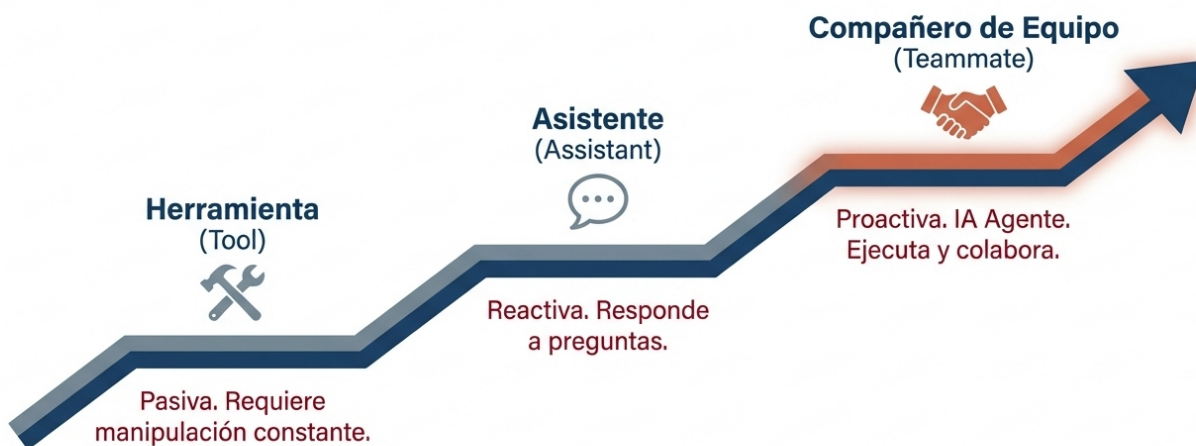
Personalización: La naturaleza modular permite adaptar los agentes a la carga de trabajo y necesidades locales.

Democratización: El software es escalable infinitamente. Llevamos el conocimiento médico de alto nivel a cualquier lugar con conexión.

Visión: Un sistema operativo de salud nuevo, modular y equitativo.

NotebookLM

Una Revolución Industrial de la Inteligencia



Estamos presenciando un cambio existencial que fusiona la Inteligencia con el Servicio, transformando la funcionalidad estándar de la industria de la salud.

NotebookLM

Conclusión: El Retorno a la Conexión Humana

El objetivo final no es la tecnología.

Al devolver la autonomía y el tiempo a los proveedores a través de la IA Agente, reinstauramos la disponibilidad de la conexión humana que yace en el corazón de la medicina.

La promesa de la IA Agente es permitirnos redescubrir la alegría de la práctica médica.



NotebookLM

Referencias Bibliográficas

1. Powell K, Fishman EK, Horton KM, Johnson PT, Rowe SP. What health care can learn from self-driving vehicles. *J Am Coll Radiol* 2019;16:261-3.
2. Zou J, Topol EJ. The rise of agentic AI teammates in medicine. *Lancet* 2025;405: 457.
3. Lakhan SE. The agentic era: why biopharma must embrace artificial intelligence that acts, not just informs. *Cureus* 2025;17:e83390.
4. Dutta Roy TP. Bioethics artificial intelligence advisory (BAIA): an agentic artificial intelligence (AI) framework for bioethical clinical decision support. *Cureus* 2025;17: e80494.
5. Rao SK, Fishman EK, Rizk RC, Chu LC, Rowe SP. Improving efficiencies while also delivering better health care outcomes: a role for large language models. *J Am Coll Radiol* 2024;21:1913-5.
6. Vellody A, Sarker M. Agentic AI: autonomous generative AI agents. Deloitte. Accessed May 21, 2025.
7. Hippocratic AI. Safety focused generative AI for healthcare. Accessed May 21, 2025.
8. Intrivo. Smart delightful AI. Accessed May 21, 2025.

NotebookLM

Sobre el Autor



Dr. [Nombre]

Médico Internista | PhD en Medicina Digital e IA

Investigador y Educador especializado en la intersección de la informática clínica y el cuidado del paciente. Experto en el desarrollo de estrategias de IA para sistemas de salud y autor frecuente en publicaciones científicas.

Misión: Integrar tecnologías emergentes para restaurar la relación médico-paciente.

 NotebookLM

Redefiniendo la atención a las personas mayores con IA agencial: Desafíos y oportunidades

Dr. Rodolfo Palencia Díaz

Dr. Rodolfo de J. Palencia Vizcarra

Médicos Internistas
Universidad de Guadalajara
Instituto Mexicano del Seguro Social (IMSS)
Colegiados y Certificados (CMMI)
Fundadores del TICC
28 de enero de 2026

El análisis exhaustivo sobre la integración de la Inteligencia Artificial (IA) agencial, potenciada por Modelos de Lenguaje Extensos (LLMs), en el cuidado de la población adulta mayor. Frente a una crisis global del cuidado y un cambio demográfico sin precedentes, la IA agencial emerge no solo como una herramienta de automatización, sino como un paradigma de asistencia proactiva, personalizada y autónoma.

1. Contexto demográfico y la urgencia de innovación

El envejecimiento global de la población plantea desafíos críticos para los sistemas de salud y protección social. Según proyecciones de las Naciones Unidas, para el año 2030, una de cada seis personas en el mundo tendrá más de 65 años, cifra que se espera que se duplique para 2050 [1, 2].

El panorama del cuidado actual:

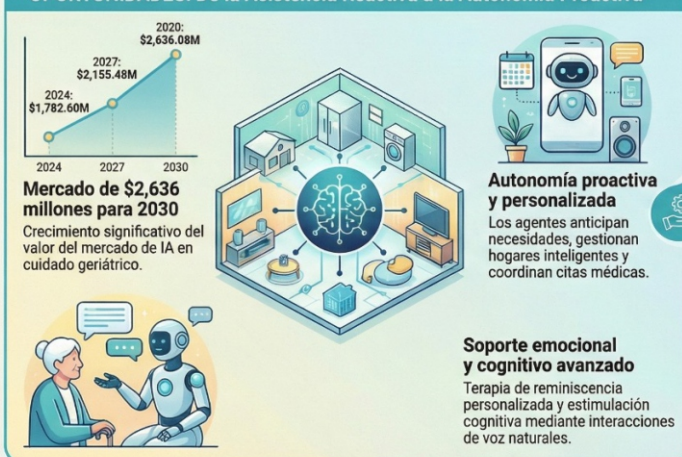
- **Gasto y Financiamiento:** En Inglaterra, el presupuesto de 2024-25 para el cuidado social de adultos alcanzó los £24.5 mil millones, de los cuales el 50% se destina a mayores de 65 años [3].
- **Escasez de Personal:** En el Reino Unido, las solicitudes de visa para trabajadores de cuidado cayeron de 18,300 en agosto de 2023 a solo 2,300 en agosto de 2024 [3].
- **Brechas Globales:** Para 2050, el 80% de los adultos mayores residirá en países de ingresos bajos y medios, donde la infraestructura sanitaria es limitada [4]. Menos del 60% de los países han integrado el cuidado a largo plazo en sus marcos nacionales [5].

IA Agencial: Redefiniendo el Cuidado Geriátrico

Una transformación en la atención de personas mayores, equilibrando la innovación con los desafíos éticos y técnicos fundamentales.

Ante el envejecimiento global y la escasez de cuidadores, la IA Agencial evoluciona los sistemas reactivos. Basada en Modelos de Lenguaje Extensos (LLM), actúa de forma proactiva y autónoma para gestionar la salud y la independencia de los adultos mayores.

OPORTUNIDADES: De la Asistencia Reactiva a la Autonomía Proactiva



DESAFÍOS: Seguridad, Ética y Fiabilidad Clínica



Referencias (Estilo Vancouver)

1. Khalil RA, Ahmad K, Ali H. Redefining Elderly Care With Agentic AI: Challenges and Opportunities. IEEE OJCS. 2026;7:1-25. doi:10.1109/OJCS.2026.3650842. 2. World Health Organization. WHO calls for urgent transformation of care and support systems for older people. Geneva: WHO; 2024. 3. Chul M, et al. The economic potential of generative AI: The next productivity frontier. McKinsey & Company; 2024.

NotebookLM

2. La revolución de la IA agencial y los LLMs.

A diferencia de la IA tradicional que sigue instrucciones predefinidas, la IA agencial actúa de manera proactiva para alcanzar metas específicas [14]. Su auge actual se debe al desarrollo de LLMs avanzados (como ChatGPT, Gemini, Claude y LLaMA) que permiten una comprensión matizada y diálogos sofisticados [16].

Capacidades clave:

- **Autonomía Proactiva:** Puede aprender, adaptarse y tomar decisiones autónomas, como ajustar la temperatura de un hogar basándose en señales conductuales o programar citas médicas al detectar anomalías en signos vitales [39, 40].
- **Procesamiento de Datos Masivos:** Para 2025, se espera que la generación global de datos supere los 180 zettabytes, con la salud contribuyendo en más de un tercio [13]. La IA agencial puede integrar datos de dispositivos vestibles (wearables), registros médicos electrónicos y sistemas de hogar inteligente [13].
- **Estructuración de Tareas:** Mediante la "encadenación" (chaining), descompone problemas complejos en pasos accionables, automatizando flujos de trabajo y triajes de alto riesgo [14, 15].

3. Crecimiento del mercado de asistencia tecnológica.

La adopción de estas tecnologías refleja una tendencia creciente impulsada por la escasez de cuidadores y la demanda de soluciones personalizadas.

Tabla 1: Proyecciones de mercado (2024-2030).

Indicador	Valor en 2024	Valor proyectado (2030)	Tasa de Crecimiento (CAGR)
Mercado de robots asistenciales	—	—	~14.8% [9]
Soluciones de IA en cuidado de mayores	—	USD 2,250 millones	~9.73% [10]
Valor Total del Mercado	USD 1,782.60 millones	USD 2,636.08 millones	Proyección Figura 1

4. Áreas de aplicación en el cuidado del adulto mayor.

La IA agencial se organiza en cinco dominios principales para mejorar la calidad de vida y la independencia:

1. **Compañía y Soporte Emocional:** Combate la soledad mediante agentes conversacionales que recuerdan interacciones pasadas y ofrecen terapia de reminiscencia personalizada (curando música o fotos del pasado) [31, 23].

2. **Asistencia Sanitaria Personalizada:** Monitoreo de signos vitales y gestión de la polifarmacia. Los modelos de Generación Aumentada por Recuperación (RAG) permiten crear Modelos de Lenguaje de Cuidado (CaLMs) específicos para demencia [28, 30].
3. **Compromiso Cognitivo:** Sesiones de narración interactiva, rompecabezas colaborativos y módulos de aprendizaje que ajustan su dificultad según el rendimiento cognitivo del usuario [31, 23].
4. **Habilitación de la Independencia:** Control ambiental dinámico (IoT), recordatorios de medicación y asistencia en tareas digitales complejas como la banca en línea [38, 39].
5. **Inclusión y Accesibilidad:** Los agentes de voz eliminan las barreras de la alfabetización digital, permitiendo interacciones naturales para personas con visión reducida o limitaciones motoras [42, 47].

5. Desafíos críticos y soluciones potenciales

La integración de sistemas autónomos en poblaciones vulnerables presenta riesgos que deben ser mitigados mediante un enfoque de múltiples capas.

Tabla 2: Desafíos y estrategias de mitigación.

Desafío	Descripción	Soluciones Propuestas
Privacidad y Seguridad	Riesgo de brechas en datos sensibles de salud [52].	Encriptación de extremo a extremo, aprendizaje federado (<i>federated learning</i>) y blockchain para auditoría [74, 76].
Exactitud (Alucinaciones)	Tasas de error de LLMs estimadas entre 5% y 30% [61].	Mecanismos de consenso multi-agente y validación con "humano en el bucle" (<i>human-in-the-loop</i>) [17, 63].
Sesgo Algorítmico	Amplificación de inequidades históricas en salud [55].	Auditorías continuas de sesgo, métricas de equidad y conjuntos de datos representativos [55, 77].
Vulnerabilidades Técnicas	Susceptibilidad a inyección de comandos (<i>prompt injection</i>) y "jail-breaking" [83].	Validación robusta de entradas, monitoreo de comportamientos adversarios y controles de acceso por roles [95, 63].
Interoperabilidad	Sistemas heredados fragmentados (EMR/EHR) [55].	APIs compatibles con estándares FHIR/HL7 y middleware de mapeo automático [10, 79].

6. Perspectivas futuras y prioridades de investigación.

Para lograr una integración responsable, la comunidad científica y tecnológica debe priorizar las siguientes áreas:

- **Estandarización:** Desarrollo de protocolos de validación y guías de diseño de comandos (prompts) específicos para geriatría [96, 97].
- **Integración Multimodal:** Combinar datos de voz, visión y sensores para interacciones más ricas y una detección de caídas o emergencias más precisa [40, 102].

- Robustez Adversaria: Diseñar modelos que resistan intentos de manipulación maliciosa de información médica [84].
- Evaluación Longitudinal: Estudios que midan no solo la precisión técnica, sino el impacto a largo plazo en el bienestar emocional y la inclusión digital [98, 107].

Referencias Bibliográficas

1. Guillemot JR, Zhang X, Warner ME. Population aging and decline will happen sooner than we think. *Soc Sci.* 2024;13(4):190.
2. United Nations. World population prospects 2022. Summary of results [Internet]. 2022 [citado 16 Dic 2024]. Disponible en: www.un.org/development/desa/pd/files/wpp2022_summary_of_results.pdf
3. Bancalari A, Zaranko B. Adult social care in England: What next? [Internet]. Institute for Fiscal Studies; 2024 [citado 15 Ene 2025]. Disponible en: <https://ifs.org.uk/publications/adult-social-care-england-what-next>
4. World Economic Forum. Could these old and new ideas be the future of social care for the elderly [Internet]. 2023 [citado 15 May 2025]. Disponible en: <https://www.weforum.org/stories/2023/08/elderly-social-care-dementia-villages/s>
5. Hasan H (WHO). WHO calls for urgent transformation of care and support systems for older people [Internet]. 2024 [citado 11 Jul 2025]. Disponible en: <https://tinyurl.com/498ytc2x>
6. Khalil RA, Ahmad K, Ali H. Redefining Elderly Care With Agentic AI: Challenges and Opportunities. *IEEE OJCS.* 2026;7:10.1109/OJCS.2026.3650842.
7. Sun Q, Xie J, Ye N, Gu Q, Guo S. Enhancing nursing and elderly care with large language models: An AI-driven framework. En: *Proc 31st Int Conf Comput Linguistics*; 2025. p. 10083-10090.
8. Adams SJ, Acosta JN, Rajpurkar P. How generative AI voice agents will transform medicine. *npj Digit Med.* 2025;8(1):1-4.
9. Ma B, et al. Artificial intelligence in elderly healthcare: A scoping review. *Ageing Res Rev.* 2023;83:101808.
10. Han T, et al. Medical large language models are susceptible to targeted misinformation attacks. *npj Digit Med.* 2024;7(1):288.

Redefiniendo el Cuidado del Adulto Mayor con IA Agéntica

Desafíos, Oportunidades y el Futuro de la Longevidad Autónoma



Basado en el estudio de Ruhul Amin Khalil, Kashif Ahmad y Hazrat Ali (2026). Survey Paper: IEEE Open Journal of the Computer Society.

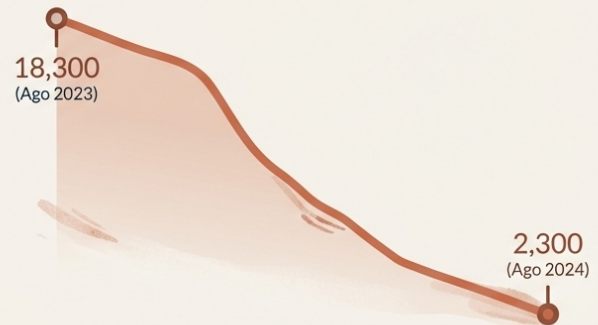
NotebookLM

La Crisis Silenciosa: El Imperativo Demográfico

Población 65+ (Global)



Solicitudes de Visa para Cuidadores (Reino Unido)



“El mundo está envejeciendo más rápido de lo que podemos cuidar a nuestra población.”

Fuente: Naciones Unidas & Home Office UK Statistics.

NotebookLM

Una Oportunidad de Mercado de \$2,600 Millones



Key Insights

- CAGR 9.73%: Soluciones de IA
- CAGR 14.8%: Robots de Asistencia

NotebookLM

Más Allá del Chatbot: La Era de la IA Agéntica

IA Tradicional



IA Tradicional

Reactiva.
Espera comandos.
Instrucciones fijas.

IA Agéntica (LLM-based)



Autónoma.
Razonamiento cognitivo.
Toma de decisiones proactiva.

Evolución impulsada por LLMs
y capacidad de 'Chaining'

NotebookLM

Un Ecosistema Holístico de Cuidado



Integración de voz, texto y sensores para una vida sin fricciones.

NotebookLM

Salud Personalizada: De la Reacción a la Predicción

Inputs



Registros Médicos
(EMR)



Wearables



Sensores IoT



Outputs



Predicción de Riesgos
(Caídas, Hospitalización)



Dietas a Medida
(DASH, Índice Glucémico)



Recomendaciones vía
RAG (Retrieval-
Augmented Generation)

NotebookLM

Compañía Inteligente y Resiliencia Cognitiva



Terapia de Reminiscencia

Curación de música, fotos
e historias del pasado
para mejorar el estado de
ánimo y la memoria.



Conversación Empática

Modelos (ChatGPT,
LaMDA) que ajustan tono
y estilo para combatir el
aislamiento social.



Estimulación Cognitiva

Juegos y narración
interactiva para retrasar
el deterioro mental.

“Tecnología que entiende, recuerda y acompaña.”

NotebookLM

Autonomía Cotidiana e Inclusión Digital



- **Hogar Inteligente:** Control automático de luces y temperatura.



- **Accesibilidad Universal:** Interfaces de voz para superar barreras visuales.



- **Gestión de Tareas:** Automatización de citas y compras.



- **Objetivo:** Eliminar la brecha digital.

NotebookLM

El Delicado Equilibrio: Innovación vs. Ética

Autonomía
& Eficiencia



Privacidad
& Seguridad

A medida que los sistemas se vuelven más autónomos, los riesgos de seguridad y los dilemas éticos aumentan exponencialmente.

NotebookLM

Desafío 1: Blindando la Privacidad del Paciente



- - **Riesgo:** Superficie de ataque ampliada en datos sensibles.
- - **Solución:** Aprendizaje Federado (Federated Learning).
- - **Solución:** Privacidad Diferencial (Ruido estadístico).
- - **Cumplimiento:** HIPAA / GDPR & Cifrado de extremo a extremo.

NotebookLM

Desafío 2: Precisión y Confianza

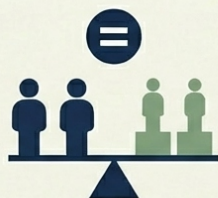
El Problema de las Alucinaciones (Tasas del 5% - 30%)



Uso de RAG para anclar respuestas en datos médicos reales.

NotebookLM

Desafío 3: Equidad Algorítmica y Seguridad Adversaria



Sesgo Técnico

- Auditorías en tiempo real para evitar discriminación en el cuidado.



Seguridad ante "Jail-breaking"

- Protección contra "Prompt Injection" y manipulación maliciosa.
- Validación robusta de entradas (*Input Validation*).

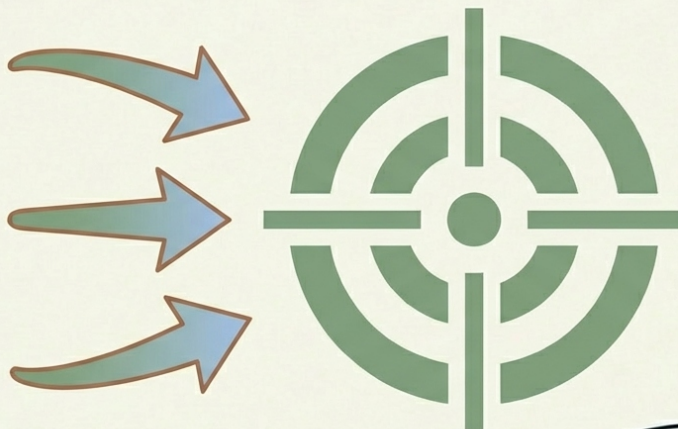
NotebookLM

Hoja de Ruta: Prioridades de Investigación

1. **Estandarización:** Guías para prompts médicos.

2. **Multimodalidad:** Integración de voz, visión y sensores.

3. **Robustez:** Sistemas inmunes a ataques adversarios.



NotebookLM

Recomendaciones Estratégicas para la Implementación



Establecer **protocolos éticos** claros y **consentimiento** dinámico.



Asegurar **interoperabilidad** con sistemas heredados (Legacy Systems).



Invertir en **integración multimodal** (física + cognitiva).



Priorizar el **diseño centrado** en el humano y la capacitación.

NotebookLM

Hacia una Longevidad Digna y Autónoma



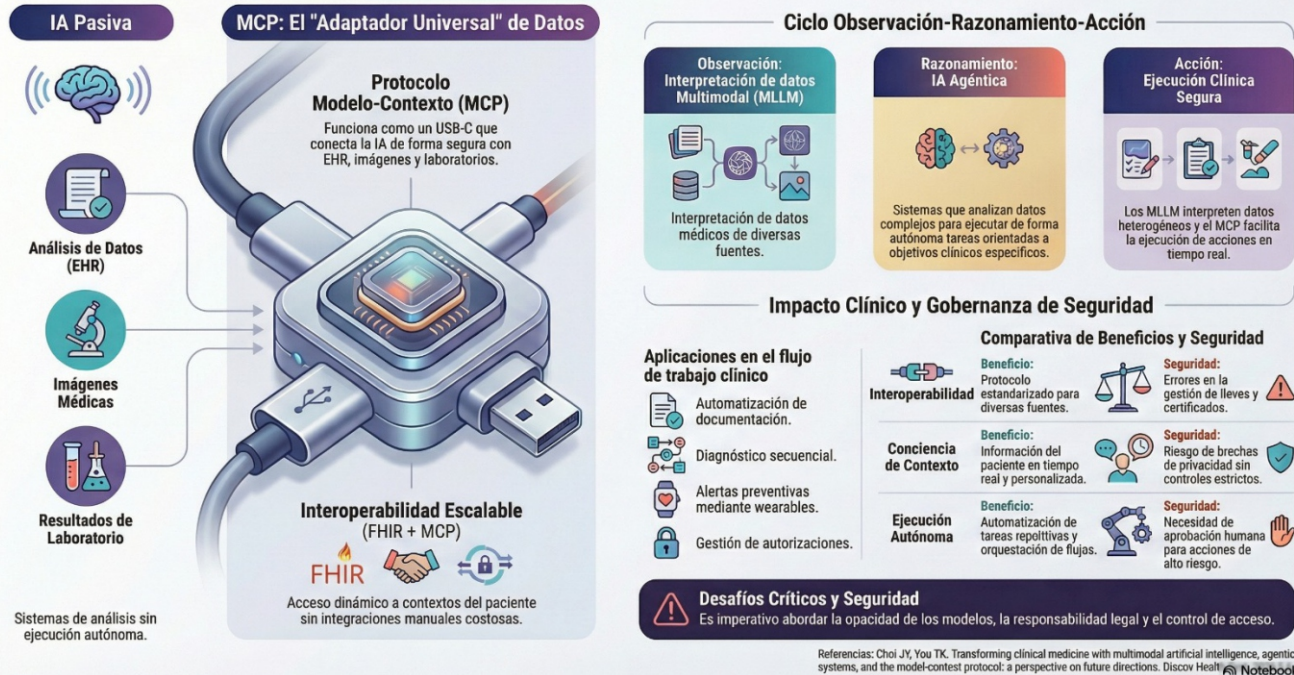
La IA Agéntica no llega para reemplazar el cuidado humano, sino para escalarlo, personalizarlo y hacerlo sostenible.

Construyamos un futuro donde el cuidado sea compasivo, equitativo y receptivo.

Gracias.

NotebookLM

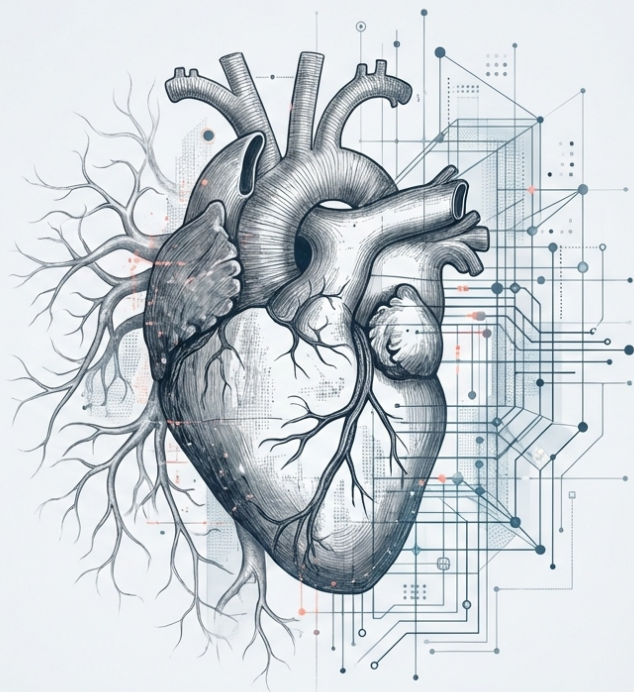
El Futuro de la Medicina Clínica: IA Multimodal y el Protocolo Modelo-Contexto (MCP)



Transformando la Medicina Clínica: IA Multimodal, Sistemas Agénticos y el Protocolo Modelo-Contexto

Una perspectiva experta sobre las futuras direcciones de la infraestructura en salud digital

Dr. [Name]
Médico Internista e Investigador Clínico
PhD en Medicina Digital e Inteligencia Artificial



De oráculos pasivos a compañeros de equipo activos

Resumen Ejecutivo



La convergencia de estas tecnologías permite sistemas que no solo ven y leen, sino que actúan dentro de la infraestructura clínica existente.

NotebookLM

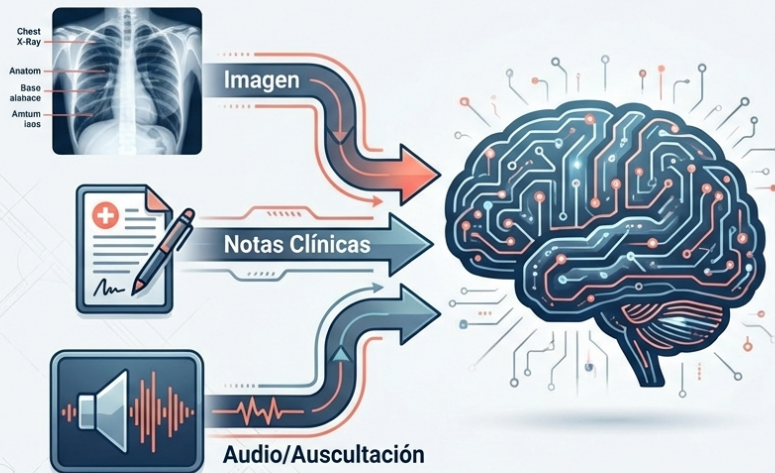
La realidad clínica actual: Fragmentación de datos y carga cognitiva

- **Interfaces Aisladas:** Jardines vallados de datos (EHR, PACS, LIS).
- **IA Desconectada:** Herramientas de propósito único sin contexto.
- **Fricción:** Alta carga cognitiva y errores de transcripción.



NotebookLM

El cerebro: Modelos de Lenguaje Grande Multimodales (MLLMs)



Definición: Sistemas capaces de procesar texto, imágenes, audio y señales simultáneamente.

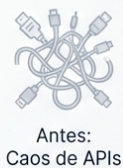
Por qué importa:

- La medicina es intrínsecamente multimodal.
- Rendimiento a nivel humano en benchmarks (USMLE, MedQA) con modelos como GPT-4o y Med-Gemini.
- Capacidad de interpretar una radiografía, leer la historia y escuchar sonidos pulmonares en un solo contexto.

NotebookLM

El conector: Protocolo Modelo-Contexto (MCP)

La metáfora del USB-C para la Inteligencia Artificial



- **Estándar Abierto:** Reemplaza conexiones frágiles y APIs ad hoc.
- **Plug-and-Play:** Conexión universal a servidores de datos.
- **Interoperabilidad:** Escalabilidad sin reescribir código.
- **Contexto en Tiempo Real:** Recuperación de datos vivos del paciente.



NotebookLM

El actor: Inteligencia Artificial Agéntica

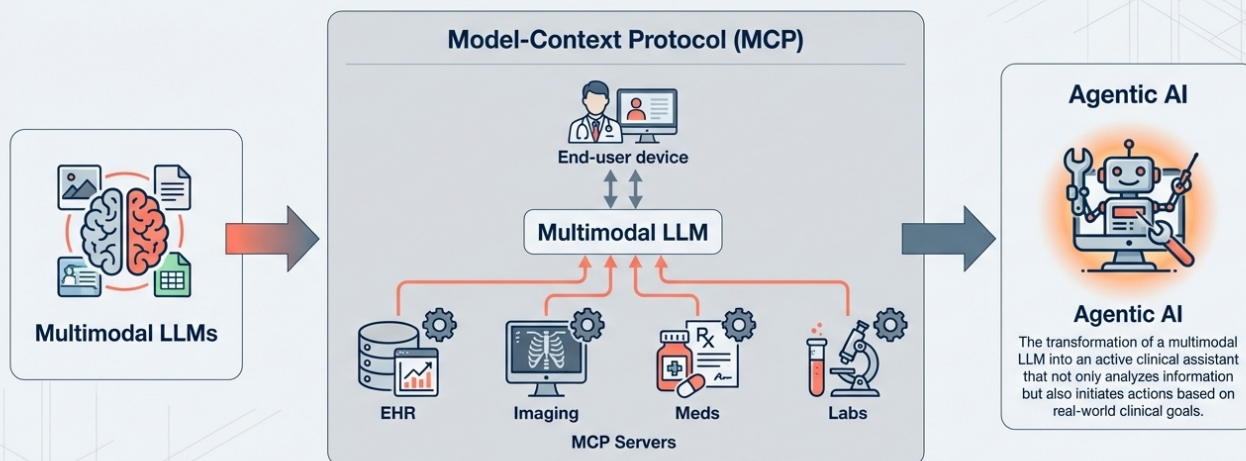


Definición: Sistemas diseñados para perseguir objetivos, usar herramientas y ejecutar tareas de múltiples pasos.

Cambio clave: De "¿Qué piensas?" a "Por favor, encárgate de esto".

NotebookLM

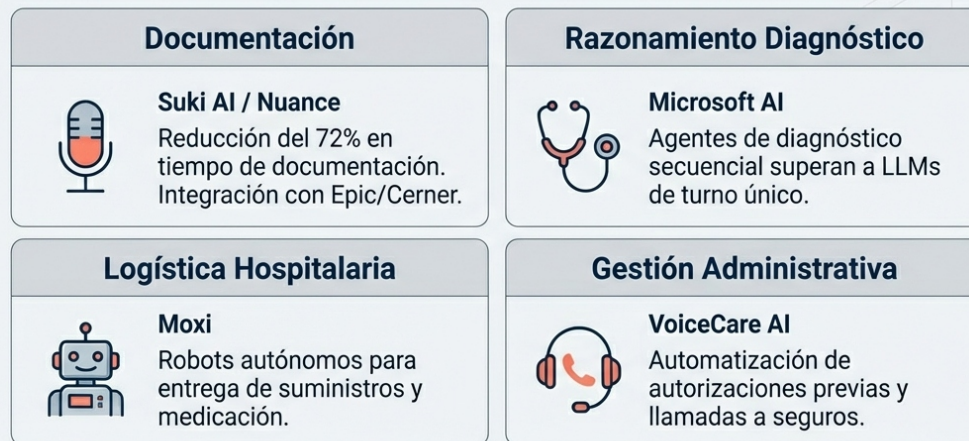
Arquitectura de convergencia para la salud digital



Interoperabilidad semántica + Ejecución autónoma = Asistencia clínica escalable

NotebookLM

Estado del arte: Fase Semi-Agéntica Actual

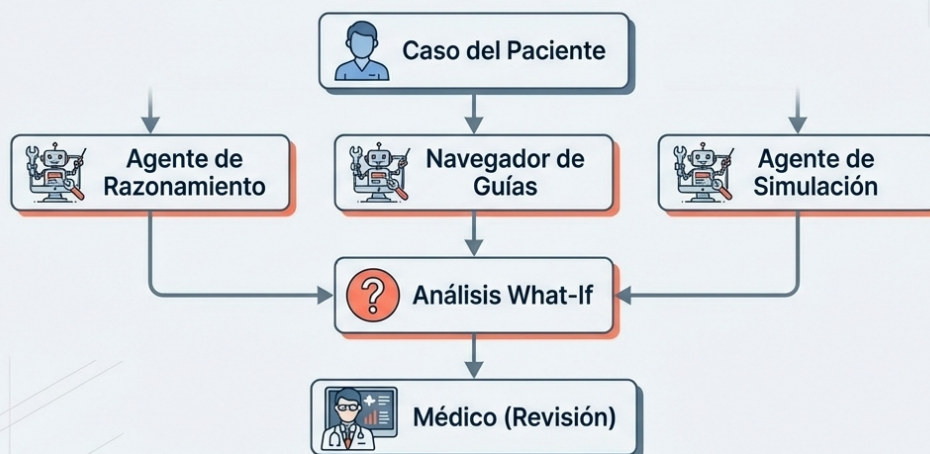


Veredicto: Fase Semi-Agéntica (Supervisión humana obligatoria).

NotebookLM

El futuro clínico: El agente de diagnóstico y navegación

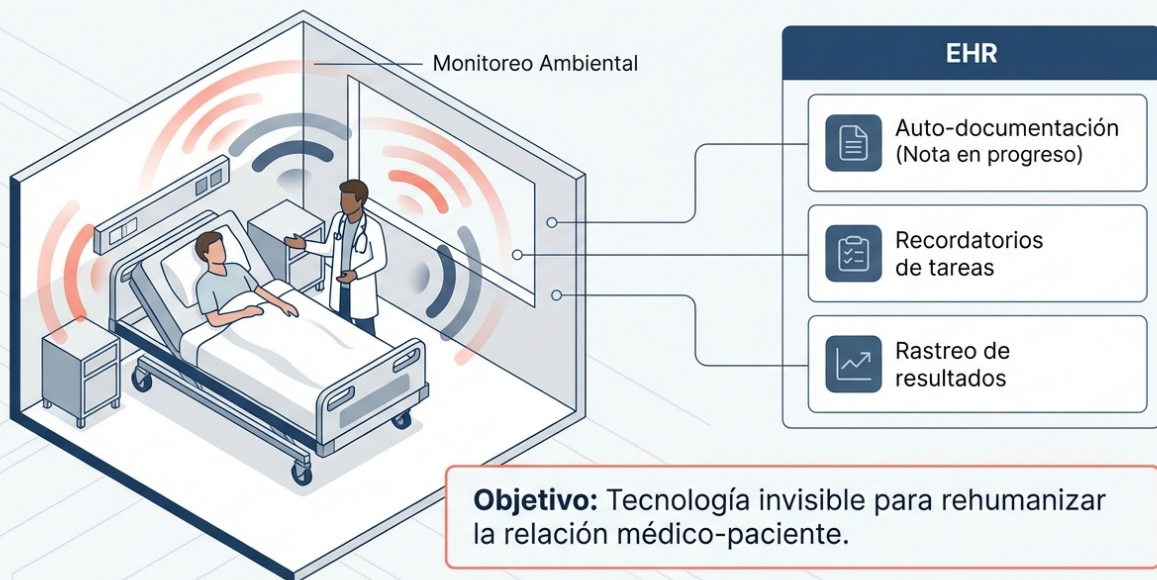
Soporte a la Decisión Clínica Dinámico



Propuesta de plan validada vs. Predicción probabilística.

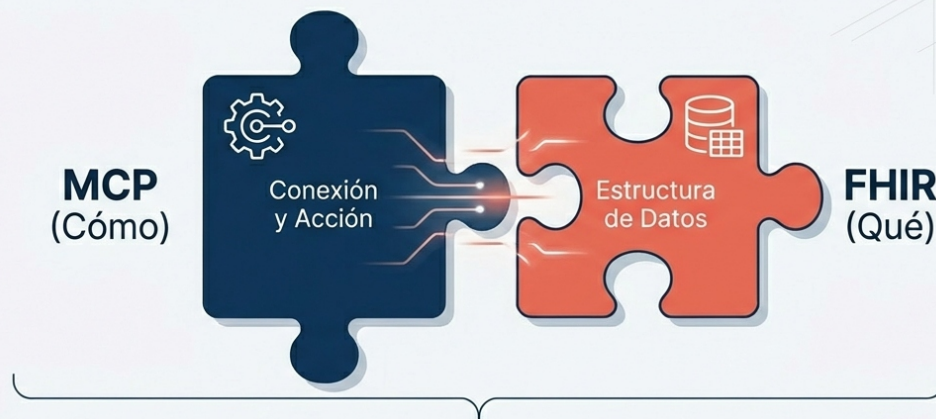
NotebookLM

El flujo de trabajo ambiental y la asistencia ubicua



NotebookLM

El desafío de la interoperabilidad: Alineación con FHIR



Estrategia: Endpoints MCP que envuelven APIs FHIR existentes.

- **Desafío:** Infraestructuras heredadas y sistemas 'legacy'.
- **Solución:** Integración de estándares para facilitar aprobación regulatoria.

NotebookLM

Riesgos críticos y limitaciones tecnológicas



Inmadurez de Herramientas

Ecosistema experimental.
Falta de interfaces estandarizadas críticas.



Alucinaciones y Errores

Riesgo de acciones autónomas incorrectas.
Validación obligatoria.



Responsabilidad Legal

¿Quién es responsable?
(Médico vs. Desarrollador vs. Vendor).



Latencia Computacional

Costos elevados de MLLMs
y retrasos en tiempo real.



NotebookLM

Un marco para el despliegue seguro

Lista de Verificación de Gobernanza

- ✓ **Control de Acceso**
Autenticación mutua y principio de mínimo privilegio.
- ✓ **Seguridad de Datos**
Aislamiento de contexto (Pacientes A vs B).
- ✓ **Human-in-the-Loop**
Aprobación obligatoria para acciones de alto riesgo (ej. medicación).
- ✓ **Auditoría**
Registros inmutables de razonamiento y acción.

NotebookLM

Hoja de ruta estratégica (Roadmap)



NotebookLM

Conclusión: Hacia una colaboración aumentada



El objetivo no es la autonomía total, sino una asistencia fiable que **reduzca la carga administrativa** y permita a los médicos centrarse en el **juicio clínico** y la **empatía**.

Referencias Seleccionadas:

1. Choi JY, Yoo TK. Transforming clinical medicine with multimodal artificial intelligence... Discover Health Systems (2026).
2. Hou X, et al. Model Context Protocol (MCP): Landscape... (2025).
3. Singhal K, et al. Toward expert-level medical question answering... Nat Med (2025).
4. Acharya DB, et al. Agentic AI: Autonomous intelligence... IEEE Access (2025).

NotebookLM