

Del texto clínico al valor institucional: sistema de terminología médica basado en Inteligencia Artificial (IA) y Procesamiento de Lenguaje Natural (NLP).

La historia clínica electrónica deja de ser solo un registro asistencial y se convierte en una fuente estratégica de calidad, educación, investigación y toma de decisiones basada en evidencia.

Reto institucional



¿Qué problema queríamos resolver?

Imposibilidad de revisar manualmente grandes volúmenes de documentación clínica

- Auditoría documental**
Identificar similitud textual y posibles eventos de copy-paste en las historias clínicas.
- Detección de actividades educativas**
Reconocer automáticamente actividades de educación al paciente registradas por los profesionales de salud.
- Reclutamiento para investigación**
Buscar de forma automatizada pacientes elegibles para estudios y ensayos clínicos mediante criterios multicampo.
- Análítica avanzada**
Transformar texto clínico no estructurado en información útil para gestión, investigación, seguridad del paciente y calidad.

Solución implementada

Se desarrolló un sistema web modular basado en Inteligencia Artificial (IA) y Procesamiento de Lenguaje Natural (NLP) integrado al ecosistema GHIPS, para analizar grandes volúmenes de información clínica no estructurada y convertirla en evidencia accionable.

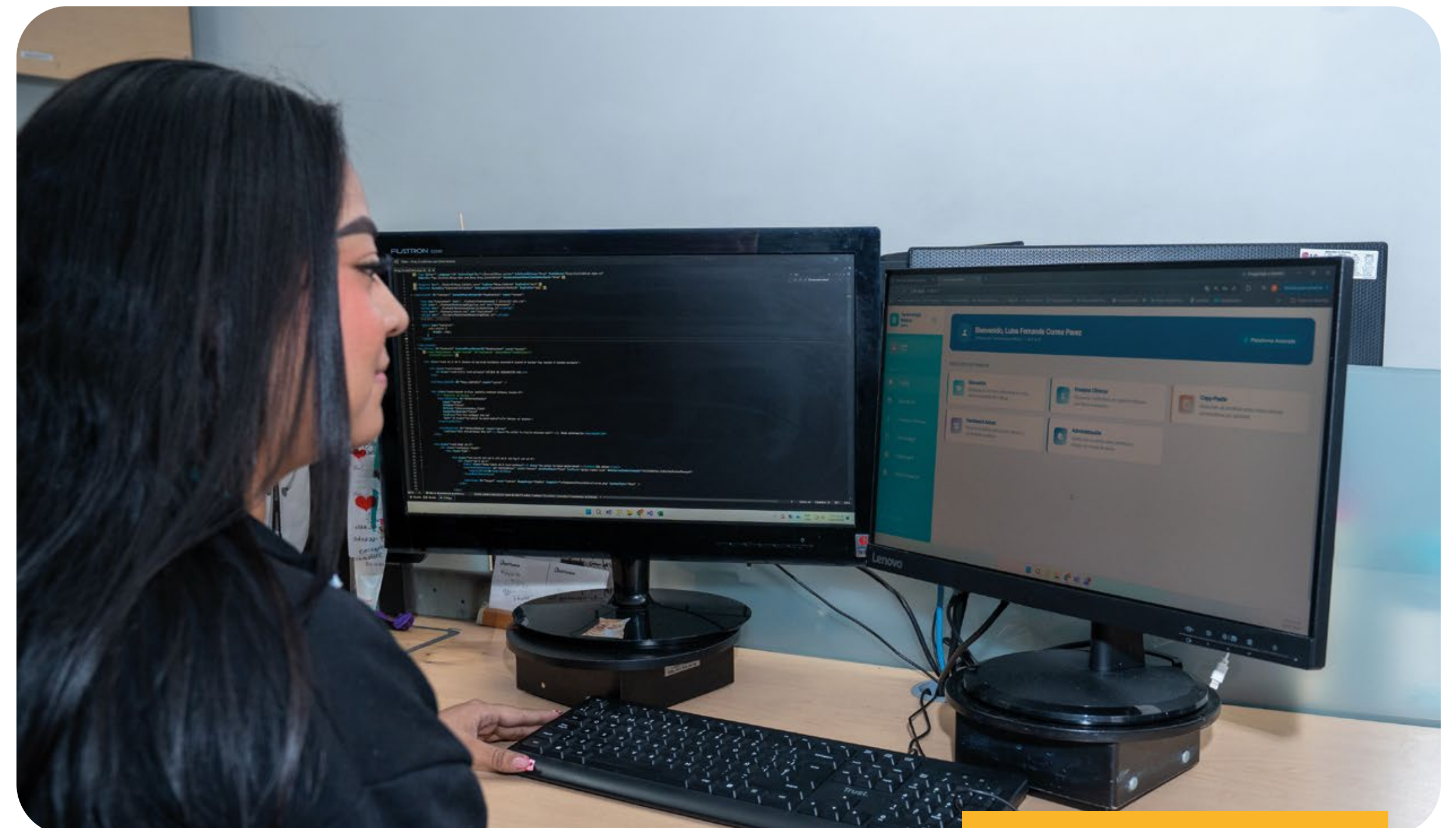
El proceso incluyó la curaduría de un catálogo de 73 términos educativos, la definición de umbrales de similitud, y la integración de paneles de resultados interactivos. Se siguieron recomendaciones de la literatura internacional sobre NLP clínico [4, 5, 1] y se adaptaron metodologías de análisis de copy-paste validadas en estudios multicéntricos [1, 2, 3].



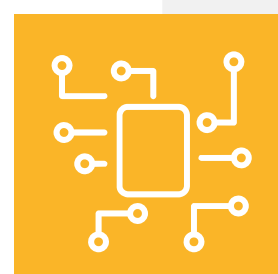
Figura 1. Vista del modelo.

Referencias

- [1] Justin M. Weis y Paul C. Levy. «Copy, paste, and cloned notes in electronic health records: prevalence, benefits, risks, and best practice recommendations». En: Chest 145.3 (2014), págs. 632-638. doi: 10.1378/chest.13-0886.
- [2] Robert E. Hirschtick. «A piece of my mind. Copy-and-paste». En: JAMA 295.20 (2006), págs. 2335-2336. doi: 10.1001/jama.295.20.2335.
- [3] Eugenia L. Siegler y Ronald Adelman. «Copy and paste: a remediable Hazard of electronic health records». En: American Journal of Medicine 122.6 (2009), págs. 495-496. doi: 10.1016/j.amjmed.2009.02.010.
- [4] Yanshan Wang et al. «Clinical information extraction applications: A literatura review». En: Journal of Biomedical Informatics 77 (2018), págs. 34-49. doi: 10.1016/j.jbi.2017.11.011.
- [5] Yanshan Wang et al. «The 2019 n2c2/OHNL Track on Clinical Semantic Textual Similarity: Overview». En: JMIR Medical Informatics 8.11 (2020), e23375. doi:10.2196/23375.



Tecnología empleada



Frontend: Angular 17
Backend: FastAPI (Python) con aceleración GPU (CUDA)
Base de Datos: SQL Server 2022
Motor NLP: RapidFuzz
Seguridad: Single Sign-On GHIPS y control de acceso por roles

Resultados destacados

Módulo copy paste

Similitud media **41,81%** | Copy-paste literal **≥95%: 2,96%** | Tiempo de procesamiento **108 segundos**

Módulo de educación al paciente

75,72% de las notas hospitalarias contenían al menos un término educativo

Módulo de ensayos clínicos

5.972 pacientes identificados | Tiempo promedio de **95,48 segundos** por búsqueda

Notas clínicas procesadas por año (datos reales)

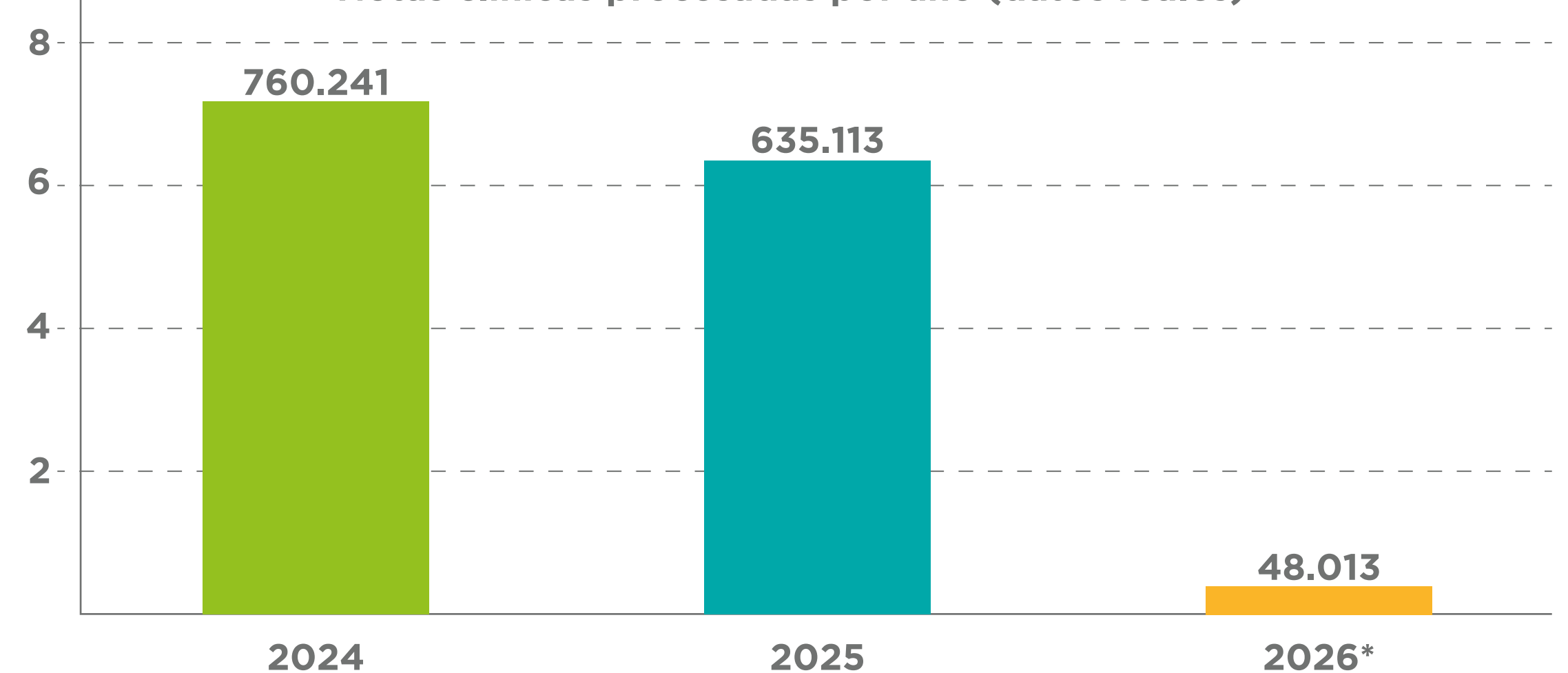


Tabla 1. Distribución de notas clínicas por año en la base de datos institucional. *2026: datos parciales (enero - mayo). Fuente: tabla clinica_notes, consulta directa a SQL Server.

Lecciones claves

- La similitud media permite priorizar auditorías por servicio, incluso cuando la copia literal es baja con umbral estricto.
- La detección de educación evidencia cultura docente y soporta procesos de calidad y acreditación.
- La búsqueda multicriterio acelera el cribado de pacientes para investigación clínica.

Mensaje final

La transformación digital en salud es posible y medible. El éxito radica en combinar tecnología, métricas claras y trabajo colaborativo. **Cada hospital puede convertir sus datos en conocimiento y calidad asistencial.** La historia clínica es mucho más que un registro asistencial: es una fuente estratégica de conocimiento. Cuando la Inteligencia Artificial transforma el texto clínico en información accionable, se fortalecen la calidad, la investigación, la educación y la toma de decisiones basadas en evidencia.

