

INTEGRACIÓN DE IMAGEN MÉDICA Y DATOS CLÍNICOS EN OMOP PARA INVESTIGACIÓN E INTELIGENCIA ARTIFICIAL

Navarro Ventura, Beatriz (Veratech)
De La Iglesia Vaya, Mariam (Conselleria de Sanitat)

MOTIVACIÓN:

- El **dato clínico** y la **imagen médica** **NO** hablan el mismo idioma en muchos hospitales.
- La imagen médica y sus metadatos son esenciales en el diagnóstico, seguimiento y evaluación técnica y han sido escasamente aprovechados en su **reúso para la investigación clínica y aplicaciones de IA**.
- Desde 2024, OMOP ya propone una **metodología estandarizada** para integrar imágenes y sus metadatos.
- La Fundación **Fisabio** y el Instituto de Investigación Sanitaria La Fe (**IIS La Fe**) como proveedores clave de datos.

METODOLOGÍA:

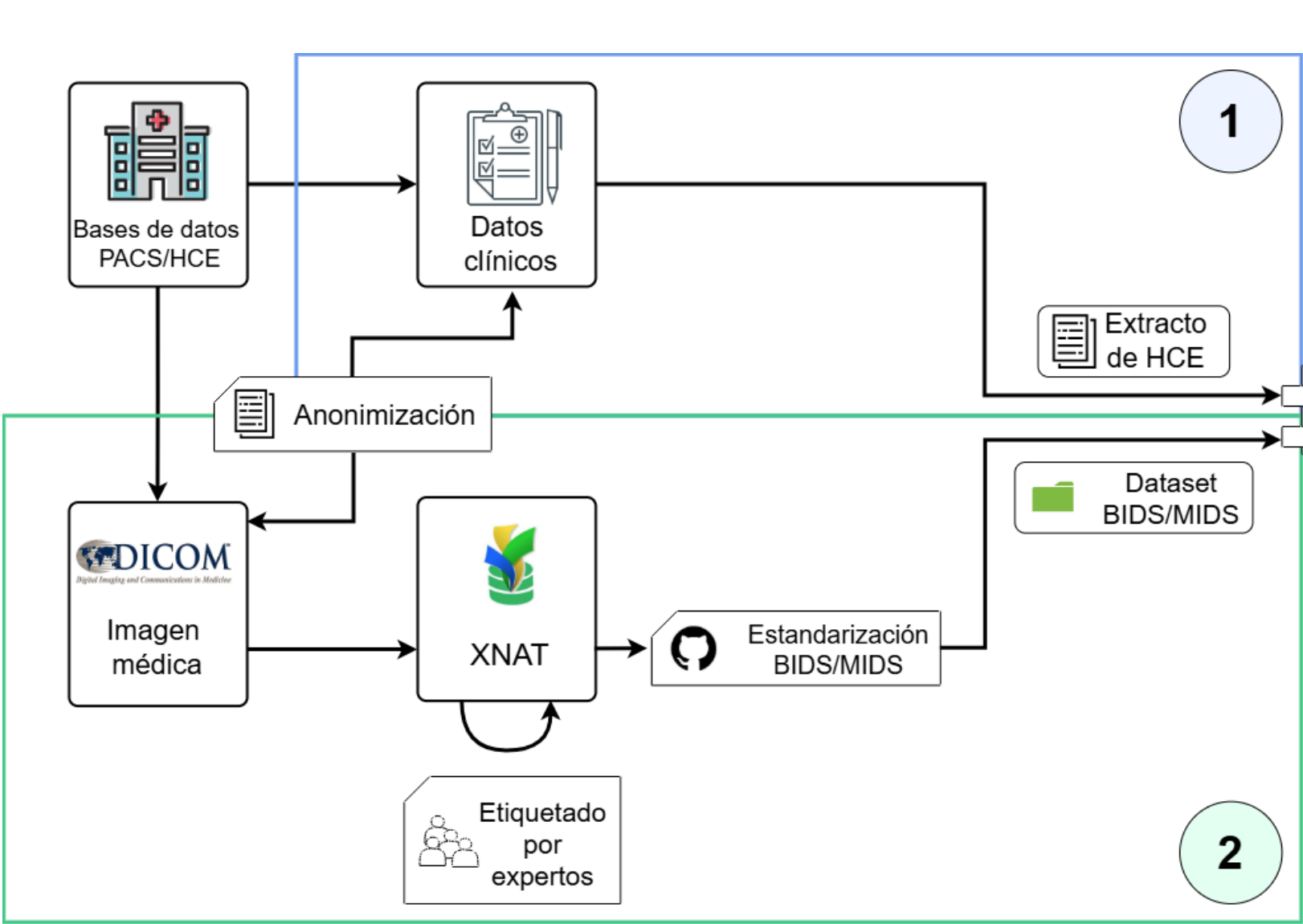
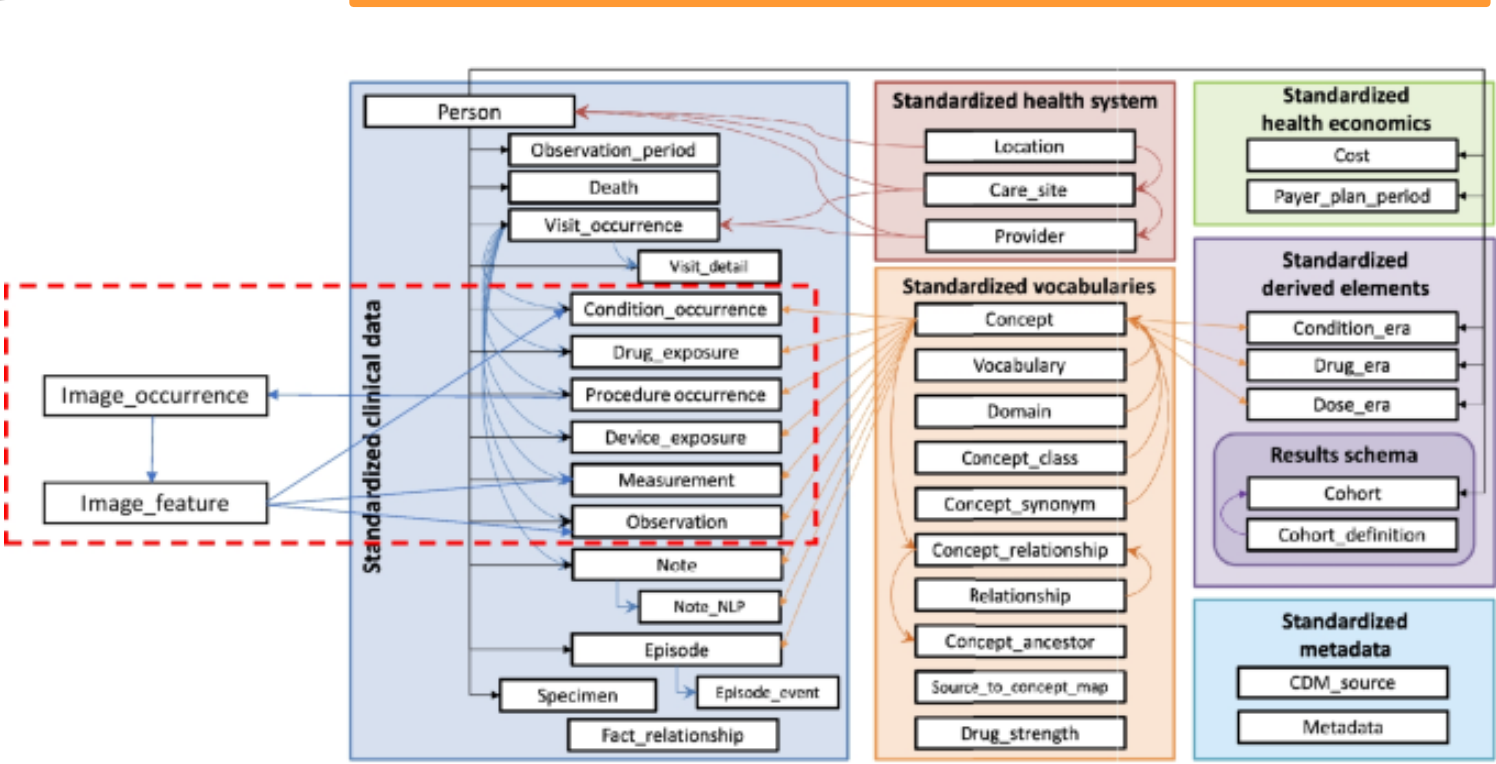


Figura 1: Esquema completo del flujo de datos

OBJETIVOS:

- Interoperabilidad:** un solo modelo para datos clínicos e imagen.
- Datasets listos para IA:** entrenamiento de modelos robustos.
- Investigación multicéntrica y federada,** esencial para la reproducibilidad.
- Contexto clínico completo:** relación entre hallazgos en imagen, laboratorio, antecedentes...
- Medicina de precisión** y seguimiento longitudinal.

OMOP CDM es un estándar que normaliza la estructura y el contenido de datos de salud y facilita la realización de estudios clínicos de manera eficiente.



Origen: PACS con imágenes DICOM y fuentes de datos de HCE

Flujo 1: Datos clínicos → Tablas clínicas de OMOP CDM

Flujo 2: Imágenes DICOM y metadatos → OMOP CDM Extendido con las tablas: **Image_occurrence** e **Image_feature**.

Pasos intermedios: - XNAT: Repositorio de imágenes médica compatible con DICOM.
- BIDS/MIDS: Formato estandarizado de los archivos DICOM.

RESULTADOS:

Esta metodología se ha aplicado en dos casos de estudio: **retinopatía diabética y cáncer de próstata**.

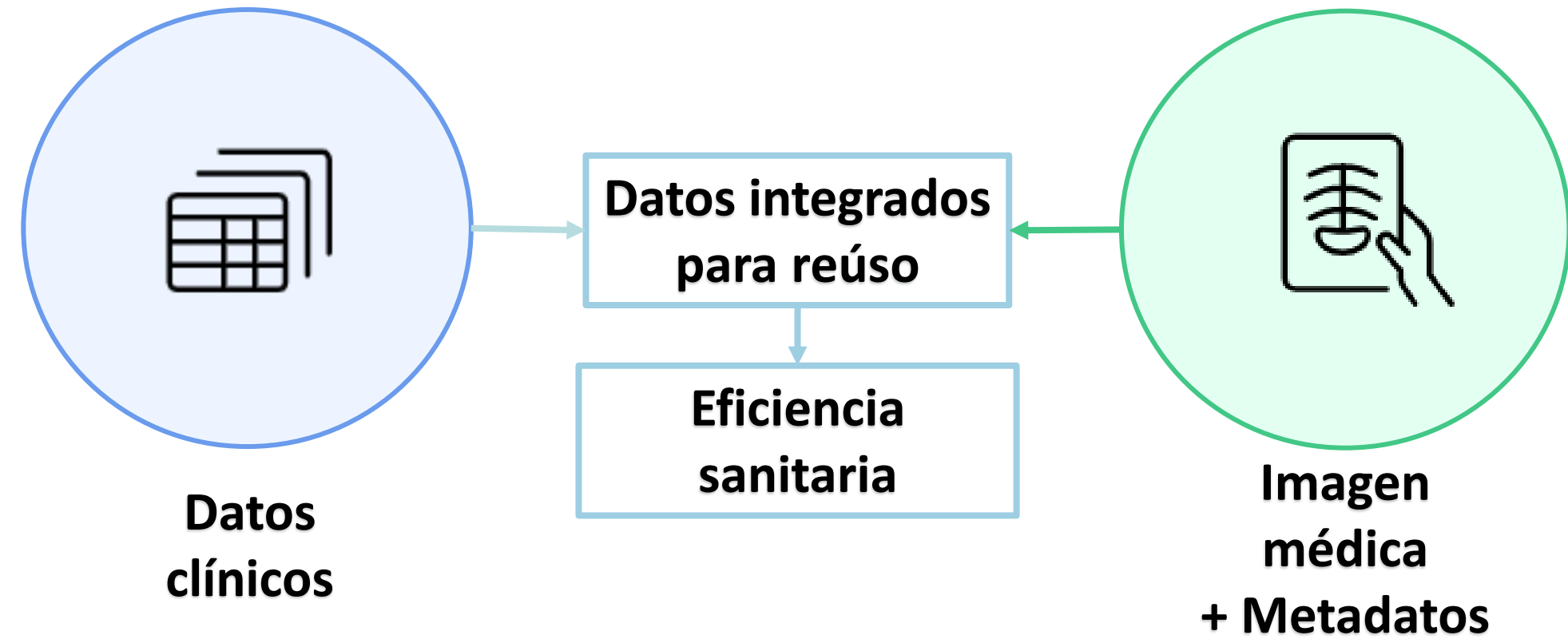
	Retinopatía diabética	Cáncer de próstata
Número de imágenes	4852 imágenes de fondo de ojo	47832 imágenes de RM
Número de pacientes	518 pacientes	8380 pacientes
Variables clínicas	-Tipo DM -Tratamiento para DM -Hemoglobina glicosilada -Tensión arterial -Filtrado glomerular -Ratio albúmina/creatinina en orina	-Antecedentes familiares de cáncer de próstata -PI-RADS -Antígeno prostático específico -Tacto rectal -Tipo de biopsia -Volumen prostático
Objetivo del modelo de IA	- Cribado automático de la imagen por grado de retinopatía diabética	- Diagnóstico precoz de cáncer de próstata en la imagen
Comentarios	- Importante registrar la lateralidad	- También se empezó a trabajar con imagen de anatomía patológica

El resultado ha sido una **base de datos OMOP cargada** con:

- Variables clínicas de los pacientes
- La **integración de la extensión de imagen médica**:
 - Metadatos de las imágenes
 - La ruta a estas imágenes

CONCLUSIONES:

- OMOP + DICOM estandarizado = **datasets listos para IA, redes federadas y con contexto clínico completo**.
- En algunos hospitales, la imagen y el dato clínico **YA** son **interoperables**.
- Reúso de imagen médica en **investigación clínica e inteligencia artificial**.
- Eficiencia sanitaria y calidad asistencial**.



Esta carga de datos clínicos, imágenes y sus metadatos a OMOP CDM permite que este actúe como **origen para cualquier aplicación de IA** que requiera el uso combinado de imágenes y de variables clínicas.

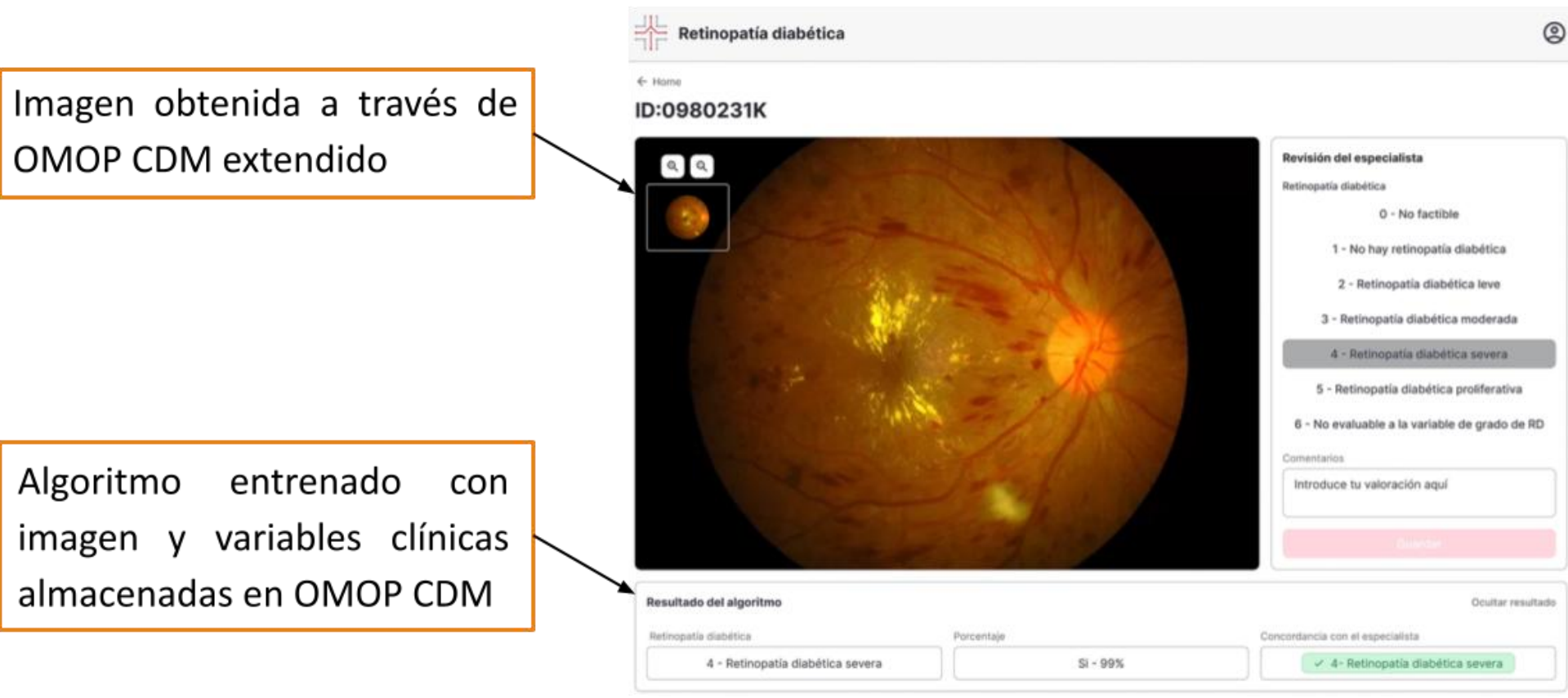


Figura 2: Herramienta de humanización para el cribado automático de retinopatía diabética desarrollada por Opinno dentro del marco del Proyecto Tartaglia.