

SYNTHIA eCRF: Standardized Multimodal Data Platform with Agentic AI

- Adrian Galiana-Bordera, Cristina Abad-Rodríguez, Maria Algarra-Deltoro, Leonor Cerda-Alberich
Health Research Institute La Fe, Spain (SYTHIA Project Coordinator)

FACTS & FIGURES

Start date:	1/09/2024
End date:	31/08/2029
Call	IHI – Call 5
EU funding	12,438,775 €
Industry contribution	10,738,015 €
Contributing partners	268,750 €
Project website	www.ihl-synthia.eu
Social media	 ihi-synthia   

CHALLENGE

The General Healthcare AI & Data Challenge:

Training clinically validated AI and generating high-fidelity synthetic data requires diverse multimodal records (clinical, imaging, genomics). However, European health data remains locked in legacy hospital silos with non-standard vocabularies. Furthermore, GDPR and EHDS regulations increasingly discourage centralized raw data pooling, widening the "AI Chasm" between lab models and clinical adoption.

The Concrete SYNTHIA Research Challenge:

Within WP4, our goal was to deploy a unified, multi-tenant electronic Case Report Form (SYNTHIA eCRF) that standardizes multimodal data directly at capture into OMOP-CDM v5.4 across 6 disease domains. The challenge: empowering clinicians with point-of-care OMOP harmonization, interactive deep-learning image segmentation, and on-premise AI agents without sensitive data ever leaving hospital premises.

APPROACH & METHODOLOGY

- Modern Full-Stack Architecture (Next.js):** High-performance platform in TypeScript/React with a Backend-for-Frontend (BFF) layer, dynamic clinical dashboards, form validation, and client-side DICOM parsing.
- Operational Relational Core (NocoDB + PostgreSQL):** Agile relational store with autogenerated REST APIs, audit trails, and native Model Context Protocol (MCP) allowing AI agents to query and curate patient tables.
- Medical Imaging PACS (Orthanc DICOMweb):** Orthanc server with full DICOMweb support (WADO/QIDO/STOW-RS), linking image series directly to clinical records via OMOP concept IDs.
- AI-Powered Viewer (Custom OHIF + nnInteractive):** Web viewer for all image modalities, featuring live promptable lesion segmentation via nnInteractive with positive/negative point clicks.
- Point-of-Capture OMOP Standardization:** Eliminates post-hoc ETL losses by mapping data upon entry. Automated bootstrap scripts translate domain metafiles into NocoDB, binding fields directly to SNOMED-CT, LOINC, and RxNorm/ATC.
- Enterprise Security & RBAC (Keycloak):** Centralized OAuth 2.0 / OpenID Connect SSO with short-lived JWT tokens, enforcing role-based access per hospital and use case, unified with Orthanc via orthanc-auth-service.
- On-Premise Privacy AI Agents (n8n + Ollama):** Local workflows running open medical LLMs (MedGemma 1.5 / Gemma 4) to parse unstructured reports and propose verified OMOP mappings with Human-in-the-Loop validation and zero data egress.

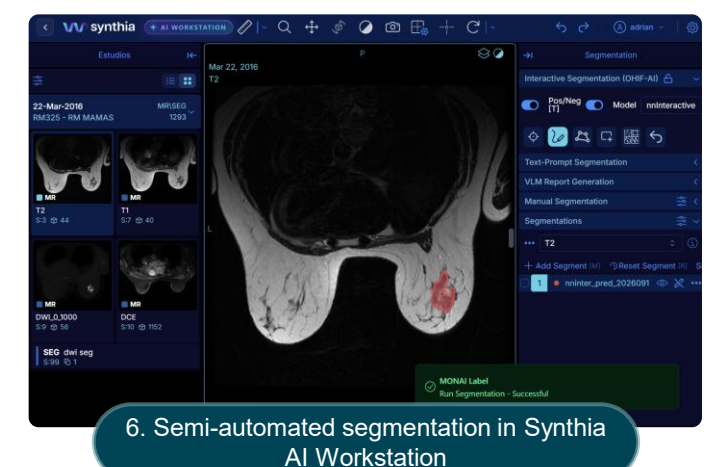
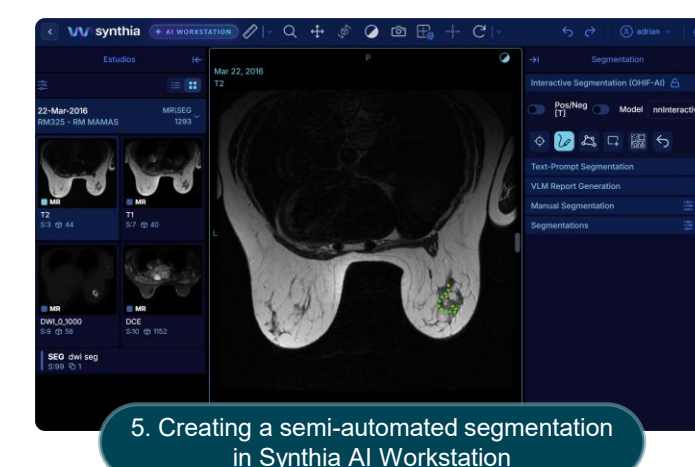
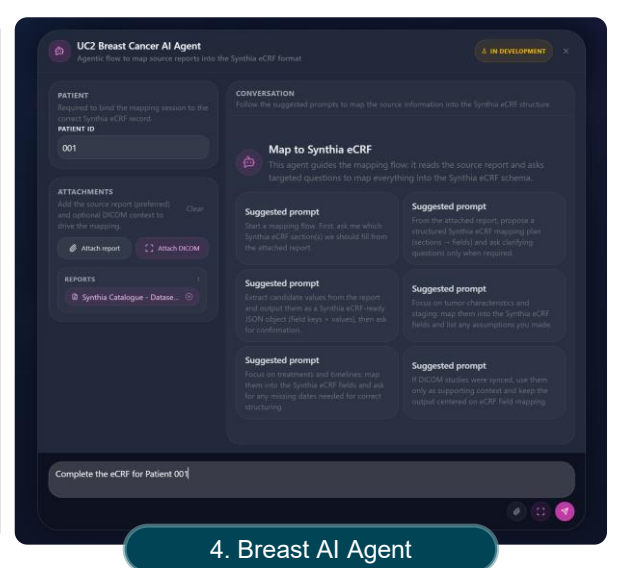
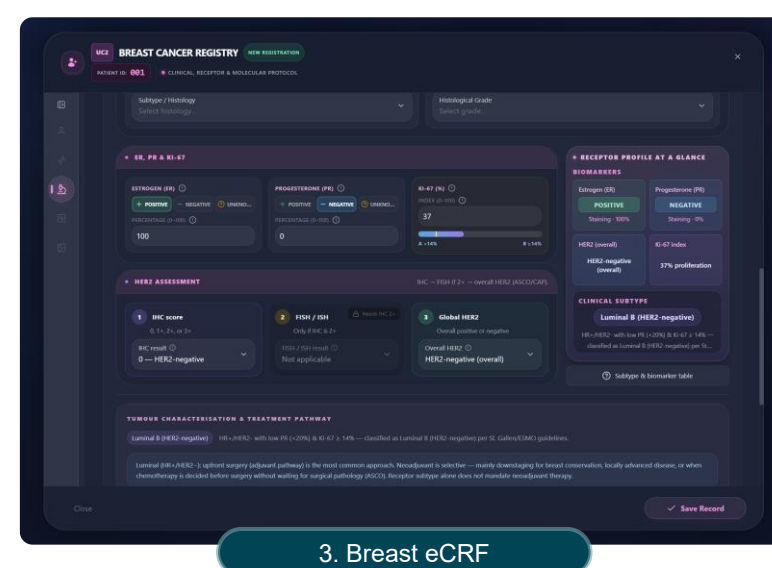
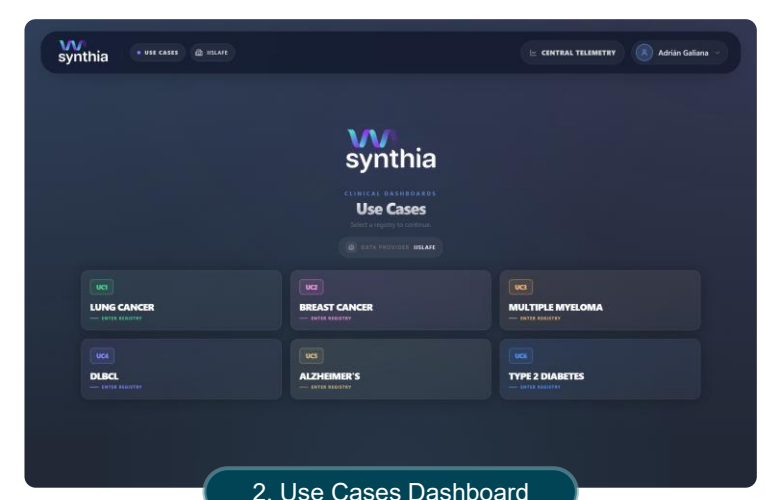
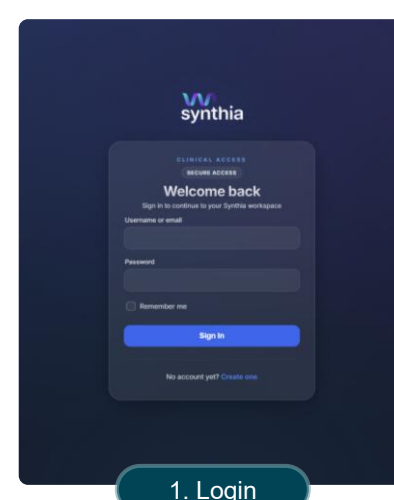
RESULTS

End-to-end multimodal implementation across 6 Use Cases:

The platform was successfully deployed across 6 distinct disease domains (lung cancer, breast cancer, multiple myeloma, diffuse large B-cell lymphoma, Alzheimer's disease, type-2 diabetes).

The workflow seamlessly integrates:

(1) Keycloak SSO secure access, (2) multi-disease registry selection, (3) structured OMOP-CDM capture with automated clinical subtyping (e.g., Luminal B breast cancer via ER, PR, Ki-67 & HER2), (4) local agentic AI mapping from raw reports, and (5–6) zero-latency interactive MRI lesion segmentation via nnInteractive in the custom OHIF-AI workstation.



VALUE OF IHI COLLABORATION

- Public-private partnership:** IHI enables close collaboration between healthcare organisations, academy and industry, combining clinical needs with technological expertise and implementation perspectives.
- Access to complementary data and expertise:** The consortium brings together diverse clinical datasets, technical capabilities and domain knowledge across six use cases, creating opportunities that individual partners could not achieve alone.
- European-scale interoperability and reuse:** Collaboration across institutions and countries supports harmonized and federated approaches that can be applied across different healthcare settings while respecting European data-protection requirements.
- Supporting future implementation:** The IHI ecosystem connects technological development with clinical needs, data governance and regulatory considerations, supporting the future uptake of SYNTHIA solutions.

CAREER IMPACT

- Advanced Technical Expertise:** Developed hands-on expertise in secure health data architectures, OMOP-CDM harmonization, multimodal data integration and agentic AI
- Cross-sector Collaboration:** Strengthened experience working across clinical, technical and research teams within a large European public-private partnership
- Research to Implementation Skills:** Enhanced the ability to translate clinical and research requirements into reusable technical solutions designed for deployment across different healthcare environments.