Digital Twin : modelli artificiali delle patologie umane a supporto della pratica clinica.

Saverio D'Amico

Firenze | 4-6 marzo 2026
Palazzo degli Affari

Disclosures of Name Surname

Company name	Research support	Employee	Consultant	Stockholder	Speakers bureau	Advisory board	Other

A wide range of definitions and applications…

Digital twins are replicas of real-world objects or systems created through high-resolution models from real-world data to better **understand how the system works** and to see what would happen in **simulation scenarios**.

Including mechanistic simulations that produce behavior based on explicitly defined biological hypotheses and **multiscale mechanisms** may generate new clinical evidence and supports **more efficient and dynamic clinical decision-making** through insights from network science, quantitative biology, and digital medicine.

Utilizing digital twins in healthcare requires simplifying complex biological data and using strategic abstractions due to current computational limits.

Defining digital twin in healthcare

01 **Drug discovery**
Digital twins of cell types and disease pathways

02 **Preventative medicine and care**
Digital twins of organs

03 **Healthcare workflow optimization**
Digital twins of healthcare systems and processes

04 **Personalized medicine**
Digital twins of patients and diseases

05 **Clinical trials**
Digital twins of trial partecipans

06 **Medical education and devices**
Digital twin of medical imaging and diagnostic equipment

Defining digital twin in healthcare

01 **Drug discovery**
Digital twins of cell types and disease pathways

02 **Preventative medicine and care**
Digital twins of organs

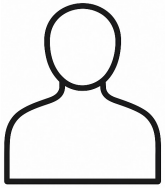
03 **Healthcare workflow optimization**
Digital twins of healthcare systems and processes

04 **Personalized medicine**
Digital twins of patients and diseases

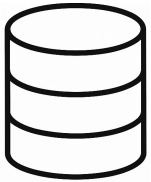
05 **Clinical trials**
Digital twins of trial partecipans

06 **Medical education and devices**
Digital twin of medical imaging and diagnostic equipment

The 5 blocks of an MDT Architecture:



Patient

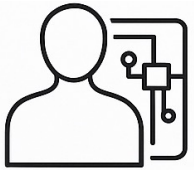


Data Availability

Collection

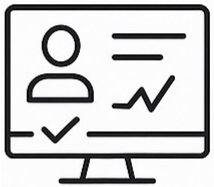
Harmonization

Fusion



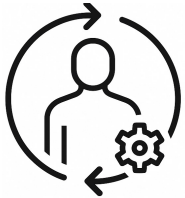
Patient-in-silico

AI + mechanistic models



Interface

Interpretation and querying of the patient-in-silico



Synchronization

Iterative updates

2021 WHO guidance on ethics and governance of AI for health

We have to address three important topics for a **right deployment of AI in hematology**:

- **Transparency of models.** We have to provide a good understanding of the models (interpretability and explainability)

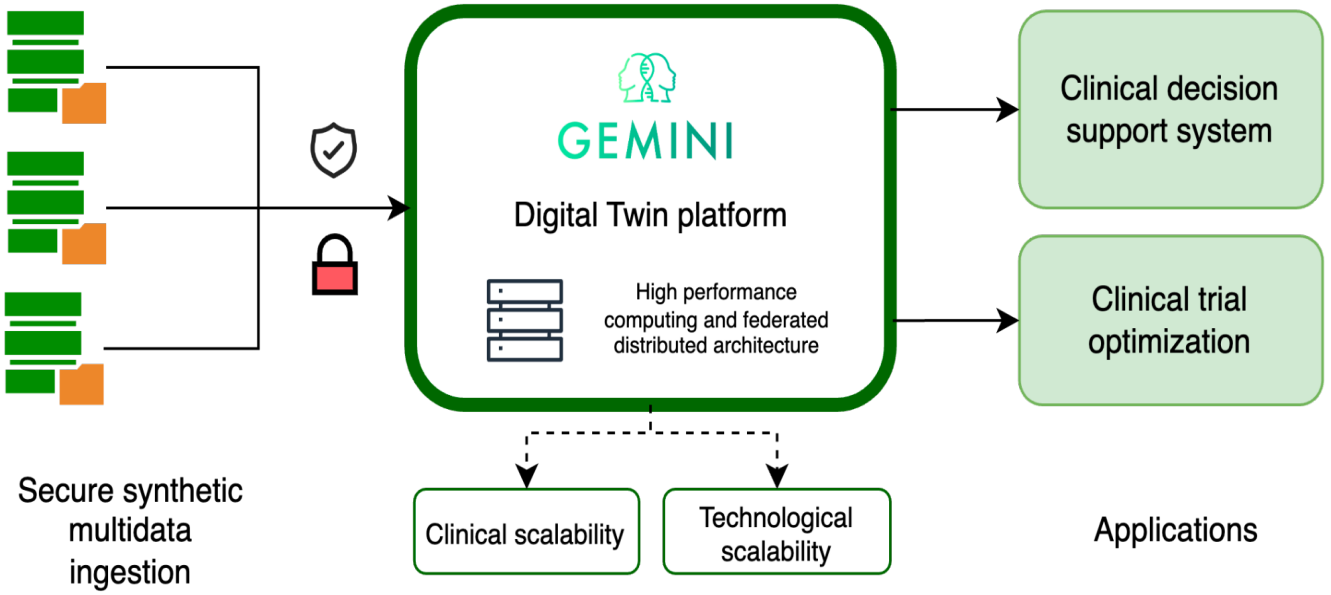
- **Reliability of models.** The main vulnerabilities of AI models are related to lack of generalizability. Therefore, extensive, independent validation of generated AI-models is required.

- **Protection of data and data sharing**. Innovative technologies such as federated learning procedures for data collection and analysis (without moving sensitive medical data from their original locations) are required to facilitate clinical implementability of AI solutions

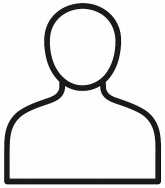
The World Health Organization. 2021 WHO guidance on ethics and governance of artificial intelligence for health. https://www.who.int/publications/i/item/9789240029200

The GEMINI AI Project, an overview

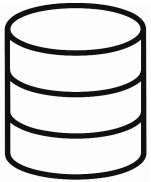
A Digital Twin by generative Artificial Intelligence to boost personalized medicine in Hematology



The 5 blocks of an MDT Architecture:



Patient

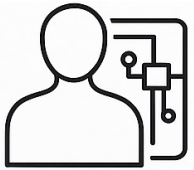


Data Availability

Collection

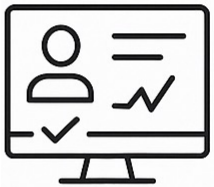
Harmonization

Fusion



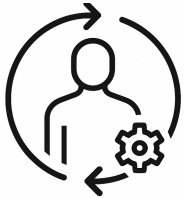
Patient-in-silico

AI + mechanistic models



Interface

Interpretation and querying of the patient-in-silico



Synchronization

Iterative updates

Core technologies for digital twin implementation: Data Availability

Technological solutions

Synthetic data. Synthetic data are artificial data generated by an algorithm trained to learn all the essential clinical and statistical characteristics of a real dataset. Since they are not real data, they are not regulated by particular limitations so they can be easily accessed and shared.

Federated learning Platform. Federated Learning is a learning paradigm seeking to address the problem of data governance and privacy by training algorithms collaboratively without exchanging the data itself

Voice AI for data collection.

Large Language Models (LLMs) for data extraction and structuring.

For clinical problems

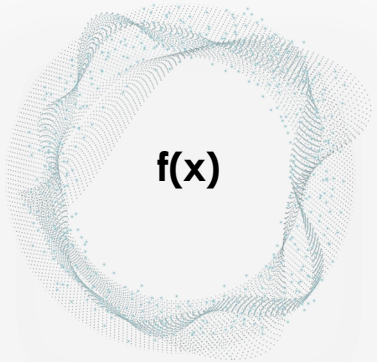
97%

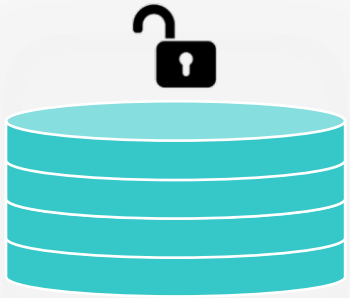
of the data produced by hospitals remains unused

Source: Deloitte, Health Data, 2023

Chen R et al. Nature Biomedical Engineering 2021;5:493–497; Savage N et al. Nature 2023; doi: 10.1038/d41586-023-01445-8; D'Amico S et al. Journal of Clin Oncol CCI 2023 Jun;7:e2300021; Jacobs F et al. Journal of Clin Oncol CCI 2023 Aug;7:e2300045

Synthetic Data Generation





REAL SOURCE DATA

FIT A MODEL

USE FITTED MODEL TO GENERATE SYNTHETIC DATA

Synthema & Synthia: synthetic hematological over federated learning frameworks

Synthema, is a European consortium founded in 2023 (www.synthema.eu) consisting of academic and industrial partners, is building a Synthetic Validation Framework (SVF) for privacy, utility and clinical evaluation of synthetic data in healthcare. SVF is being submitted to the EMA for approval and design improvement to expedite and extend the use of synthetic data as control arms in clinical trials.

Synthia is a European consortia under the HORIZON EUROPE program (www.ihi-synthia.eu). It is an ambitious collaboration between public and private institutions to facilitate the responsible use of Generative AI in healthcare applications such as Synthetic Data (SD) generation and advanced technology for precision medicine.

Synthetic data in the AI Act:

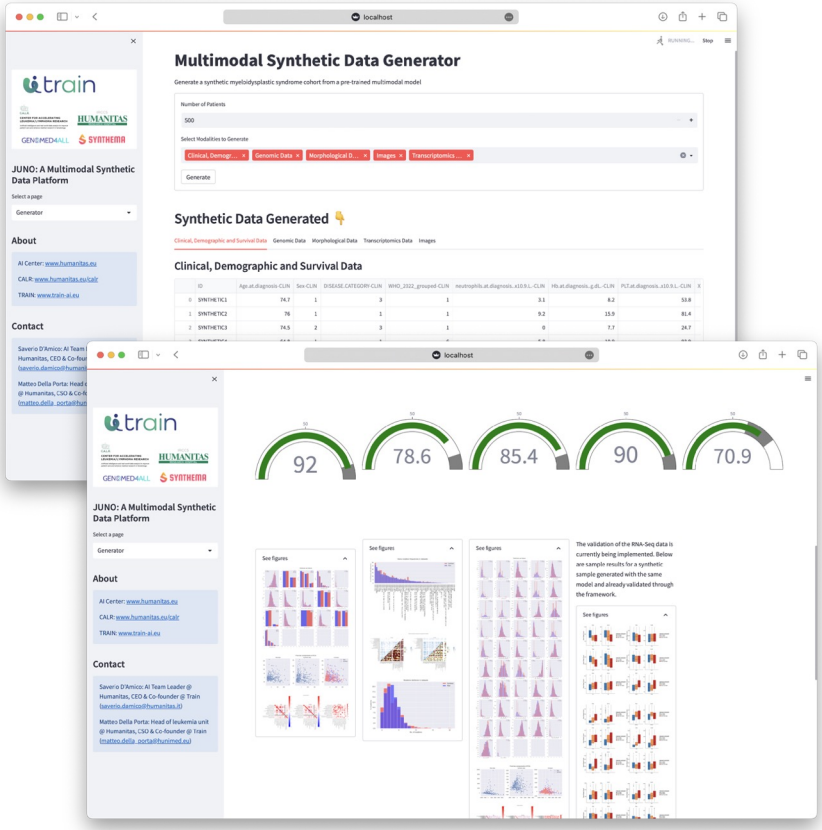
Data Governance obligations for High Risks Systems:
Art. 10(5) lett. a: "(...) the bias detection and correction cannot be effectively fulfilled by processing other data, including synthetic or anonymised data;"

Further Processing of Personal Data for Developing Certain AI Systems in the Public Interest in the AI Regulatory Sandbox:
Art. 59(1) lett. b: " the data processed are necessary for complying with one or more of the requirements referred to in Chapter III, Section 2 where those requirements cannot be effectively fulfilled by processing anonymised, synthetic or other non-personal data;"

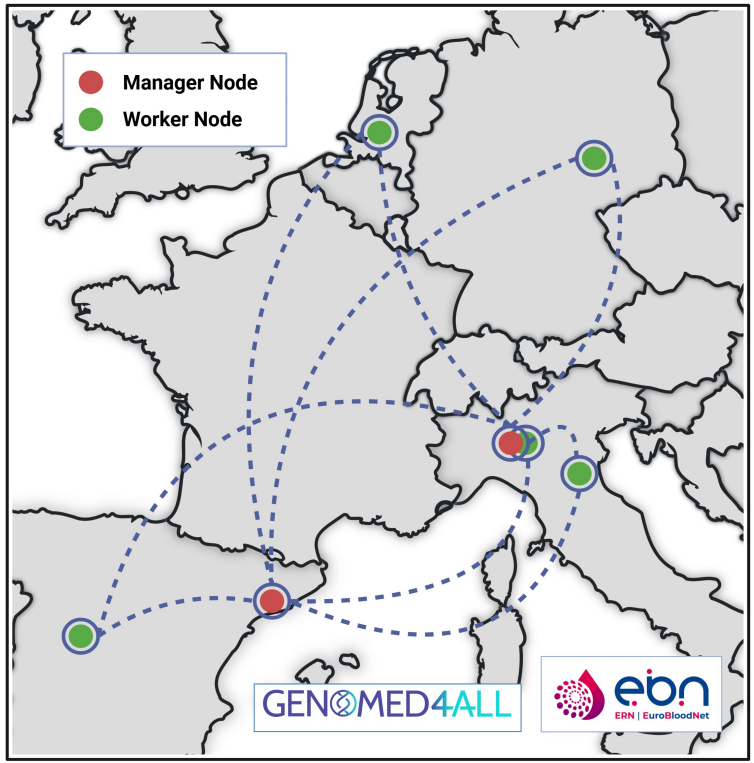
Real case study

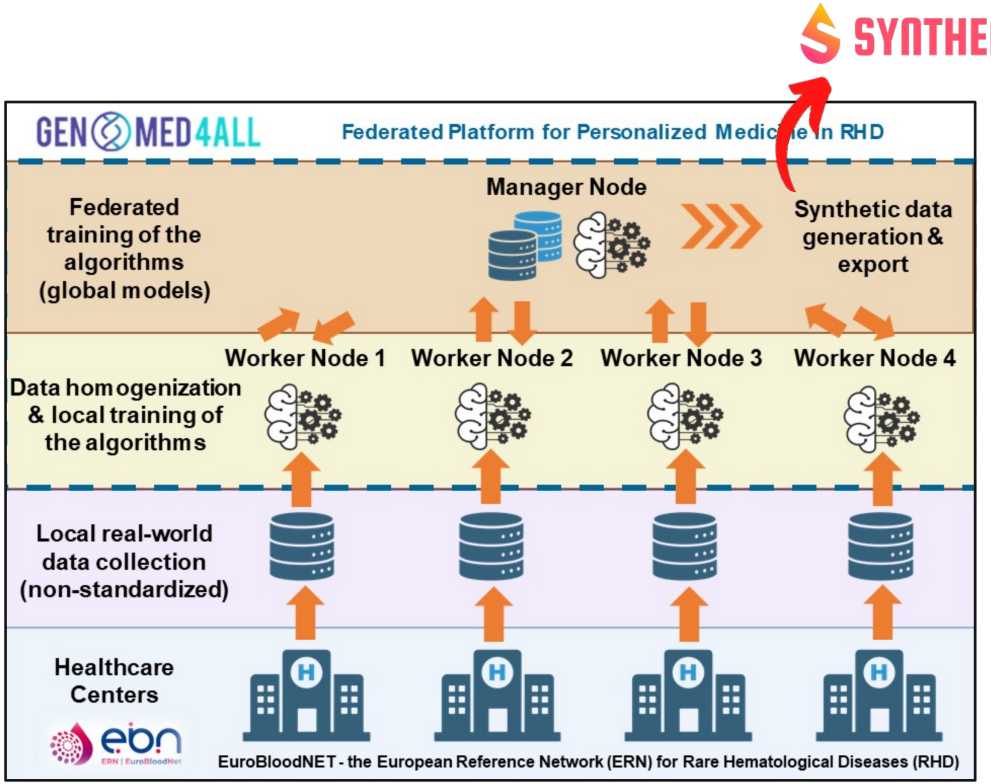
Generation of Multimodal Longitudinal Synthetic Data By Artificial Intelligence to Improve Personalized Medicine in Hematology

To address data scarcity in hematology, this project used Generative AI to create high-fidelity Synthetic Data (SD) for patients with myeloid neoplasms (MN), specifically training on 605 MDS and 877 AML patients. The goal was to generate reliable, privacy-compliant datasets spanning a rich multimodal spectrum, including clinical information, cytogenetics, somatic mutations, bulk RNA-seq (transcriptomics), longitudinal data, and bone marrow pathology images. A sophisticated AI pipeline (using GANs, VAEs, LLMs, and Stable Diffusion) was developed and rigorously assessed with a Synthetic Validation Framework (SVF). The resulting synthetic data accurately replicated the complex statistical properties, biological features, and longitudinal outcomes (like Overall Survival) of the real patients, demonstrating high fidelity (87-96%) and low privacy risk. This approach was validated by showing that augmenting real data with SD improved the performance of predictive models for disease classification and prognosis, proving its value in advancing personalized medicine.

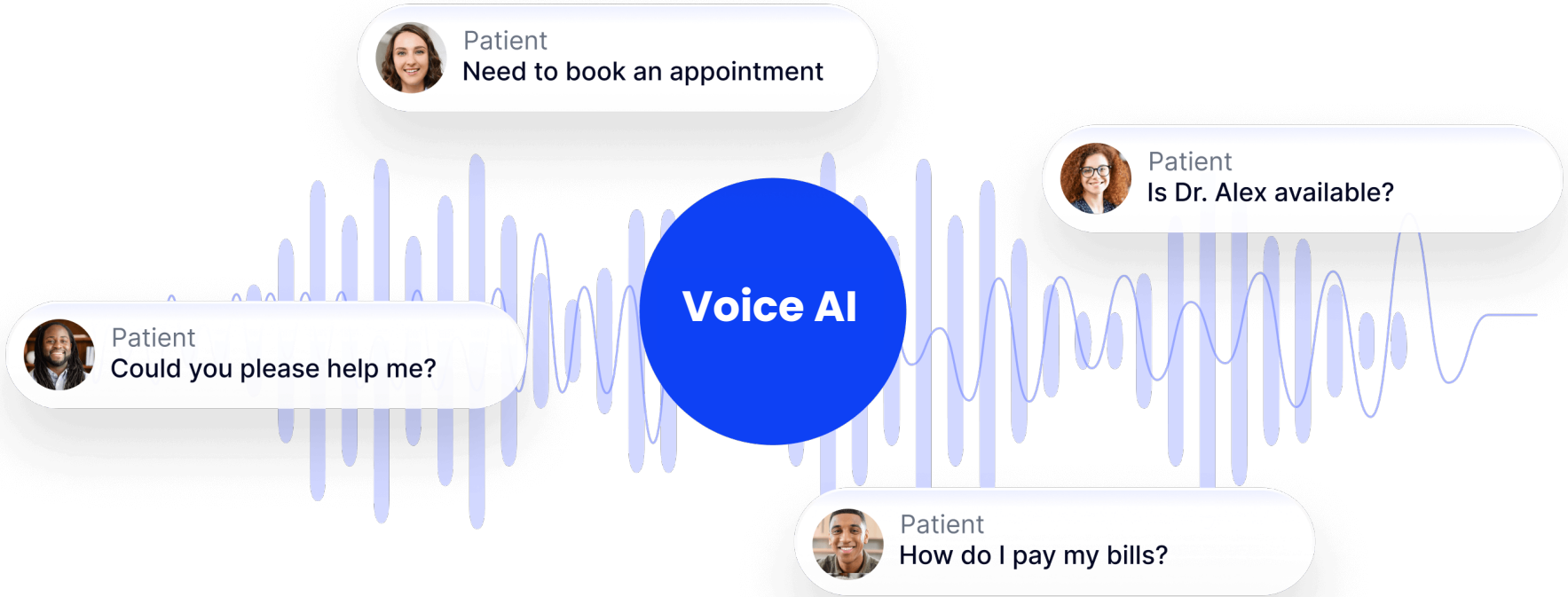


Federated Learning: GenoMed4All example for Onco-Hematology





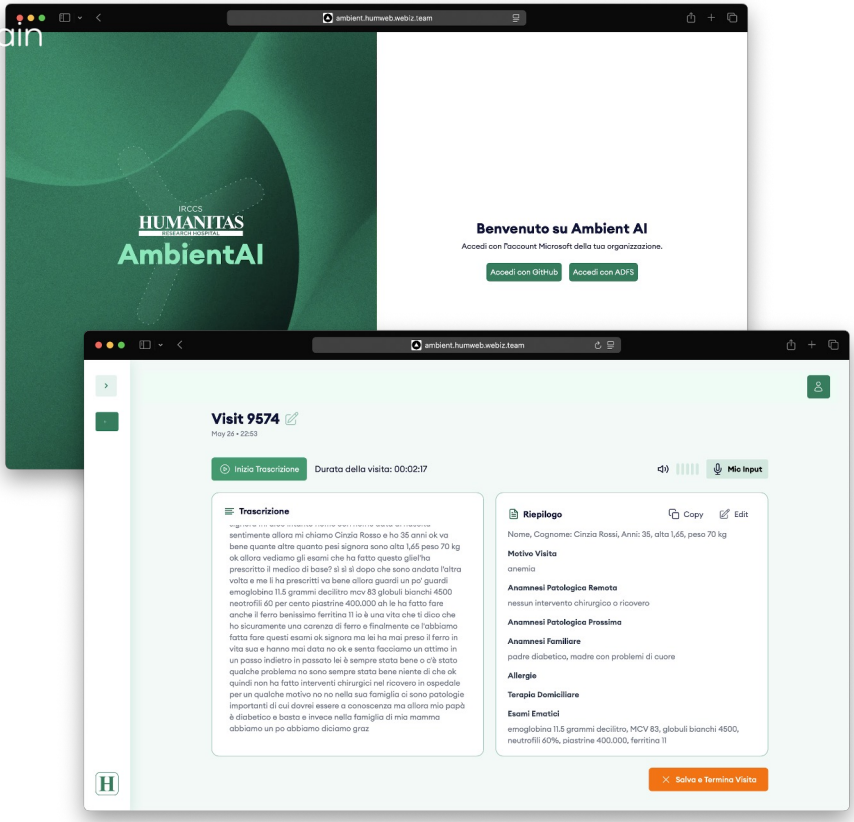
Voice AI for data collection



Ambient AI Platform for Hematological Visits in Humanitas Hospital

AI-generated Report Analysis:
A preliminary analysis of 100 simulated reports demonstrated that AI-assisted transcription significantly improves documentation efficiency by reducing manual data entry. AI-generated reports exhibited high clinical relevance, requiring minimal edits from physicians. Strong transcription accuracy (Jaccard Similarity - JS: 0.85) and effective contextual interpretation (Word Mover Distance - WMD: <1) were observed. Moreover, the Ambient AI platform leverages LLMs to automate data retrieval and extraction from medical records, enabling the generation of structured datasets for research purposes.

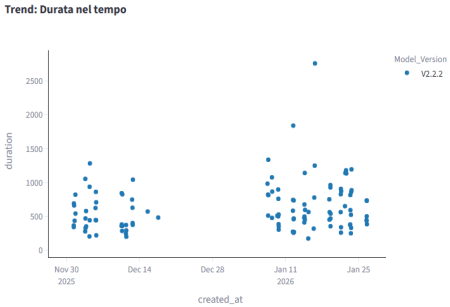
Clinical Validation:
The Ambient AI platform is currently undergoing validation at IRCCS Humanitas Research Hospital, Italy, involving 1,000 patients. Early results from this ongoing validation indicate increased physician satisfaction, improved workflow efficiency, and enhanced patient engagement.

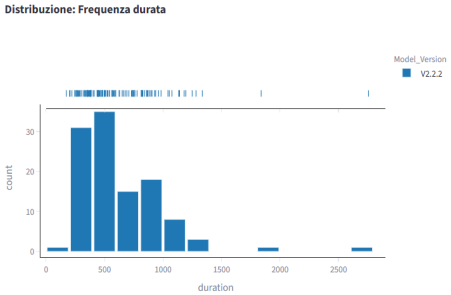


Performance Metrics | User Metrics

Metriche Medie

Approfondimento Durata

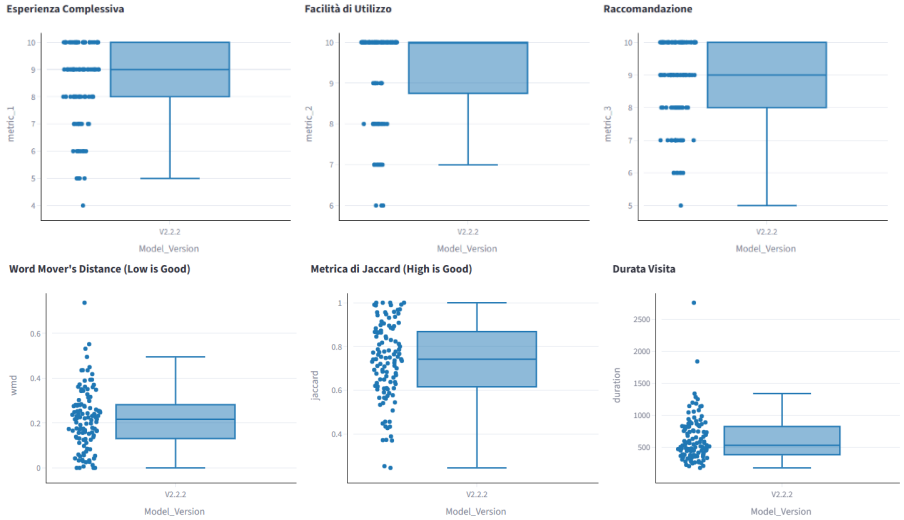




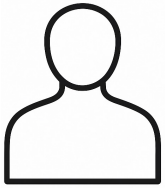
Number of patients already tested: 301

Edit rate (accuracy of transcription): 75-80%

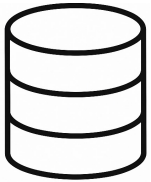
Median Duration 13.48 min (8.20-18.5)



The 5 blocks of an MDT Architecture:



Patient

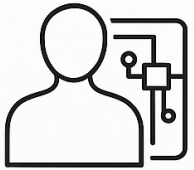


Data Availability

Collection

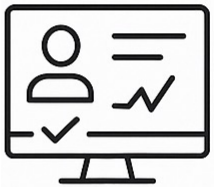
Harmonization

Fusion



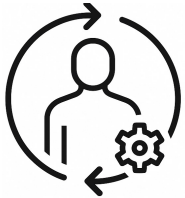
Patient-in-silico

AI + mechanistic models



Interface

Interpretation and querying of the patient-in-silico



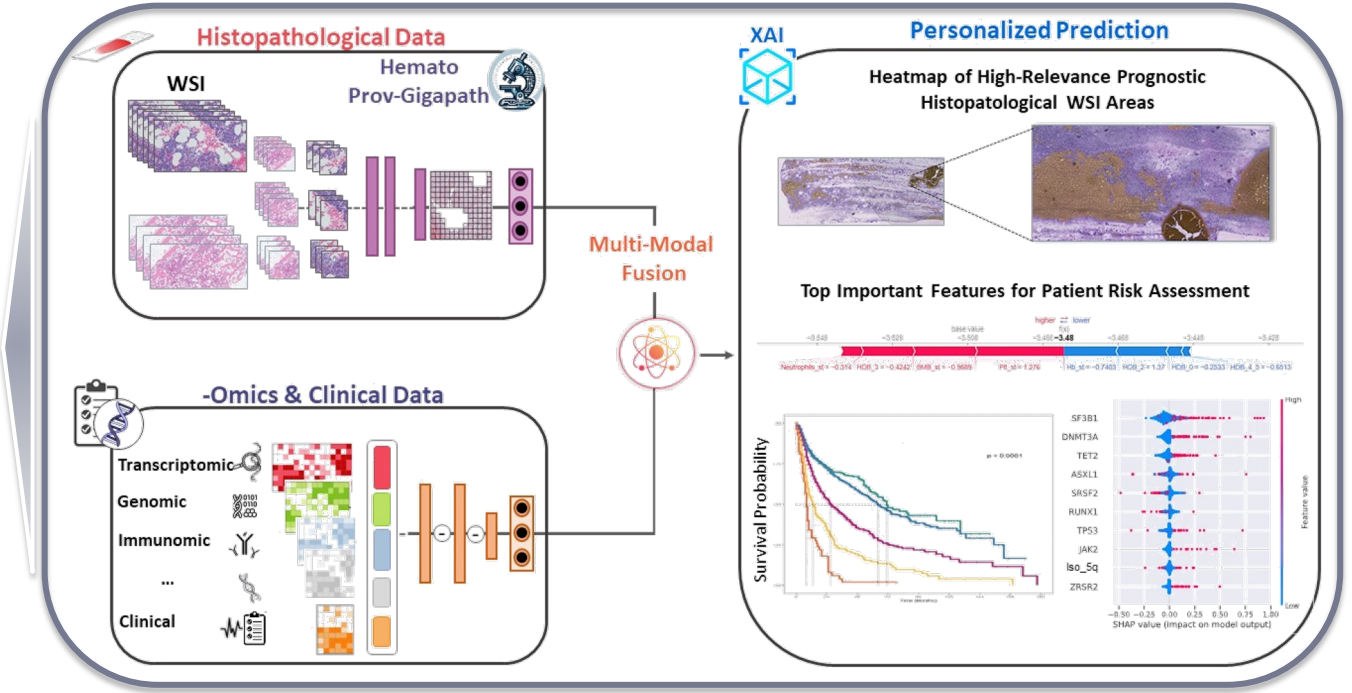
Synchronization

Iterative updates

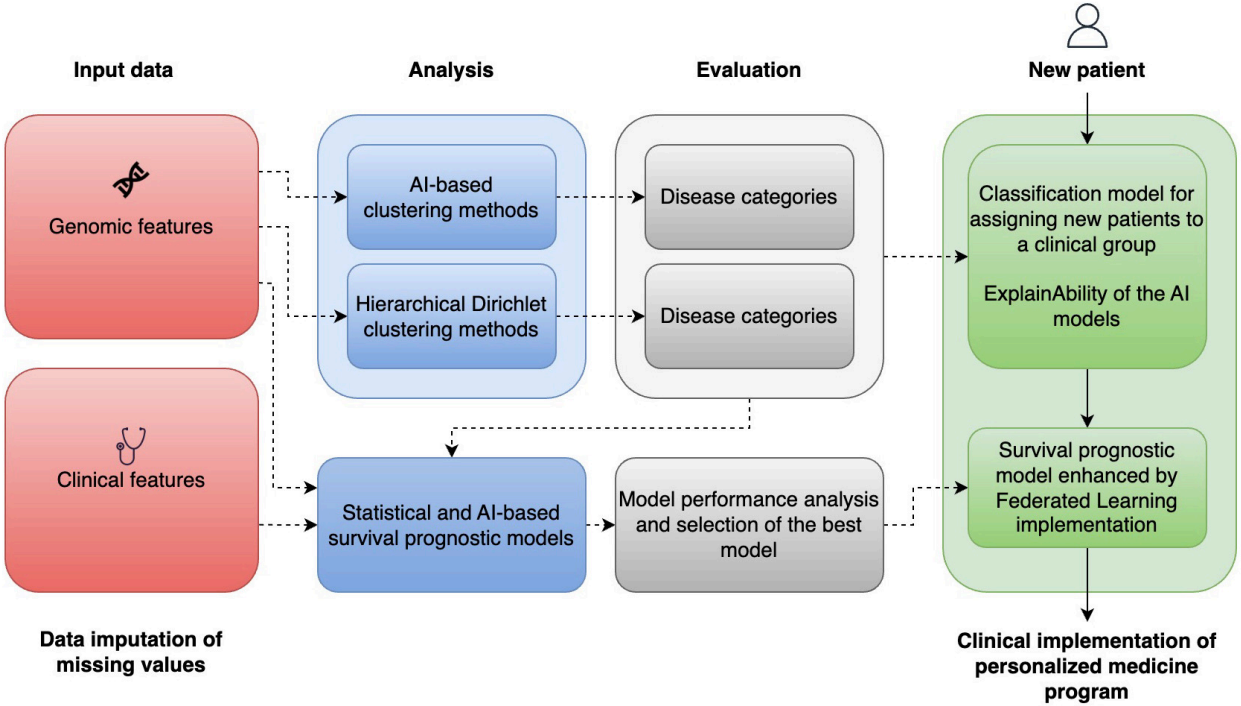
Core technologies for digital twin implementation: Multimodal data integration

MEGAERA: Multi-modal Explainable and Grounded AI-based Engine for Research Advancements in personalized care in hematology

MEGAERA



MOSAIC - Artificial Intelligence-based Framework for Multi-Modal Analysis of Classification and Personalized Prognostic Assessment in Rare cancers



GENOMED4A

D'Amico. et al. JCO CCI 2024

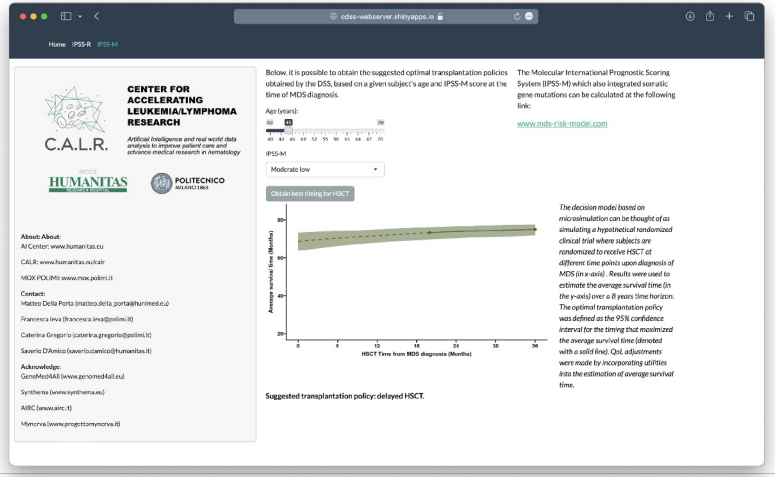
Core technologies for digital twin implementation: Clinical decision support system

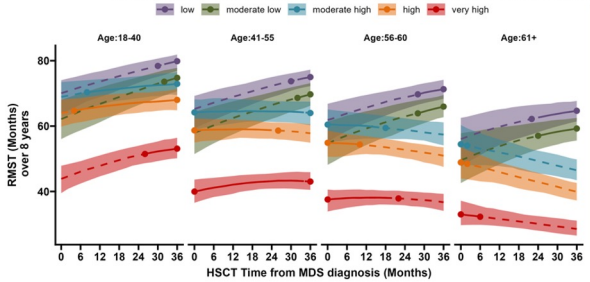
Real case study

Clinical and Genomic-Based Decision Support System to Define the Optimal Timing of Allogeneic Hematopoietic Stem Cell Transplantation in Patients with Myelodysplastic Syndromes (MDS)

Allogeneic hematopoietic stem cell transplantation (HSCT) is the only potentially curative treatment for patients with myelodysplastic syndromes (MDS).

We aimed to develop and validate a Decision Support System to define the optimal timing of HSCT in MDS patients based on clinical and genomic information as provided by IPSS-M vs conventional IPSS-R.





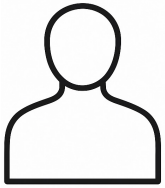
Very high: mean survival increases with HSCT waiting time in 18-40 years old subjects and the optimal timing is after 24 months from MDS diagnosis. In 41-55 years old, mean survival is quite stable across HSCT timing. In older patients, HSCT in the first 24 months (56-60) and in the first 6 months (61+) is important in order to increase the mean survival.

High/Moderate High: in younger subjects a slight increase is observed in the mean survival time. RMST tend to stabilize across HSCT timing in 41-55 years old, while in older patients a delayed intervention seems to be more favorable.

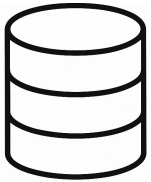
Moderate low/low: life expectancy increases the more the HSCT is delayed; the rate of increase is greater for younger subjects leading to shorter optimal windows.

Source: Tentori et al, J Clin Oncol 2024
Gregorio et al, J Clin Oncol CCI, 2024

The 5 blocks of an MDT Architecture:



Patient

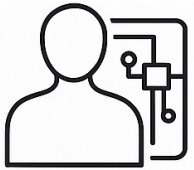


Data Availability

Collection

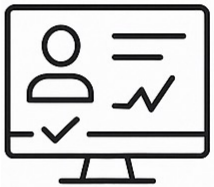
Harmonization

Fusion



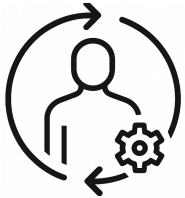
Patient-in-silico

AI + mechanistic models



Interface

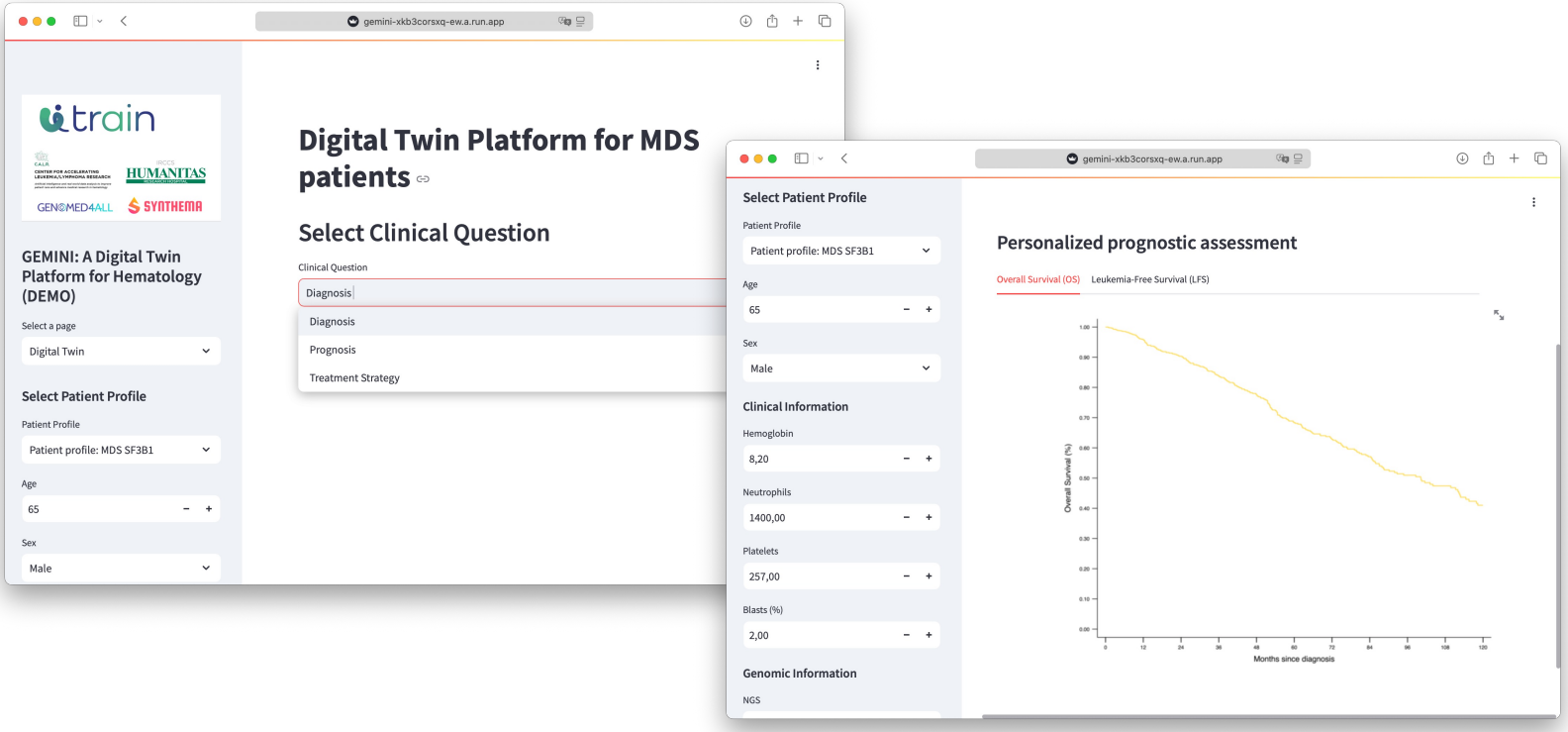
Interpretation and querying of the patient-in-silico



Synchronization

Iterative updates

Defining digital twin in healthcare



D'Amico et al. Blood 2024

GEMINI: A Digital Twin Platform for Hematology (DEMO)

Select a page

Digital Twin

Select Patient Profile

Patient Profile

Patient profile: MDS SF3B1

Age

65

Sex

Male

Clinical Information

Hemoglobin

8,20

Neutrophils

1400,00

Platelets

257,00

Digital Twin Platform for MDS patients

Select Clinical Question

Clinical Question

Diagnosis

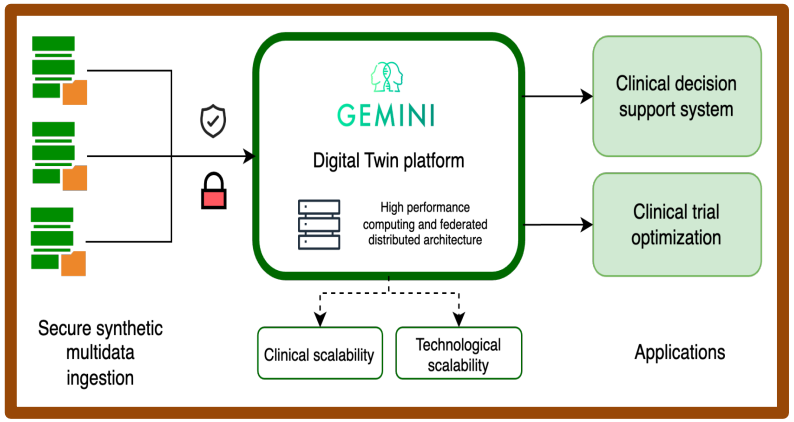
Execute Model

Data source for GEMINI Project

N=7,117

N>2, 000

Including QoL



Canadian MDS registry N>1, 000

Including frailty, comorbidity

VALIDATE dataset

N=2,135

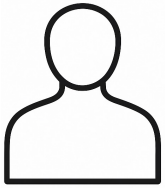
>2,000

N=8, 050

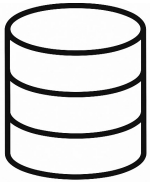
Lombardia region

EHR from >10 M people

The 5 blocks of an MDT Architecture:



Patient

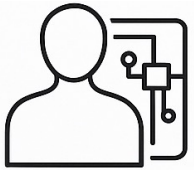


Data Availability

Collection

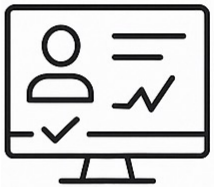
Harmonization

Fusion



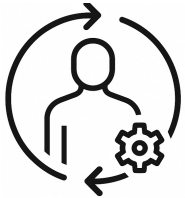
Patient-in-silico

AI + mechanistic models



Interface

Interpretation and querying of the patient-in-silico



Synchronization

Iterative updates
Virtual Assistant

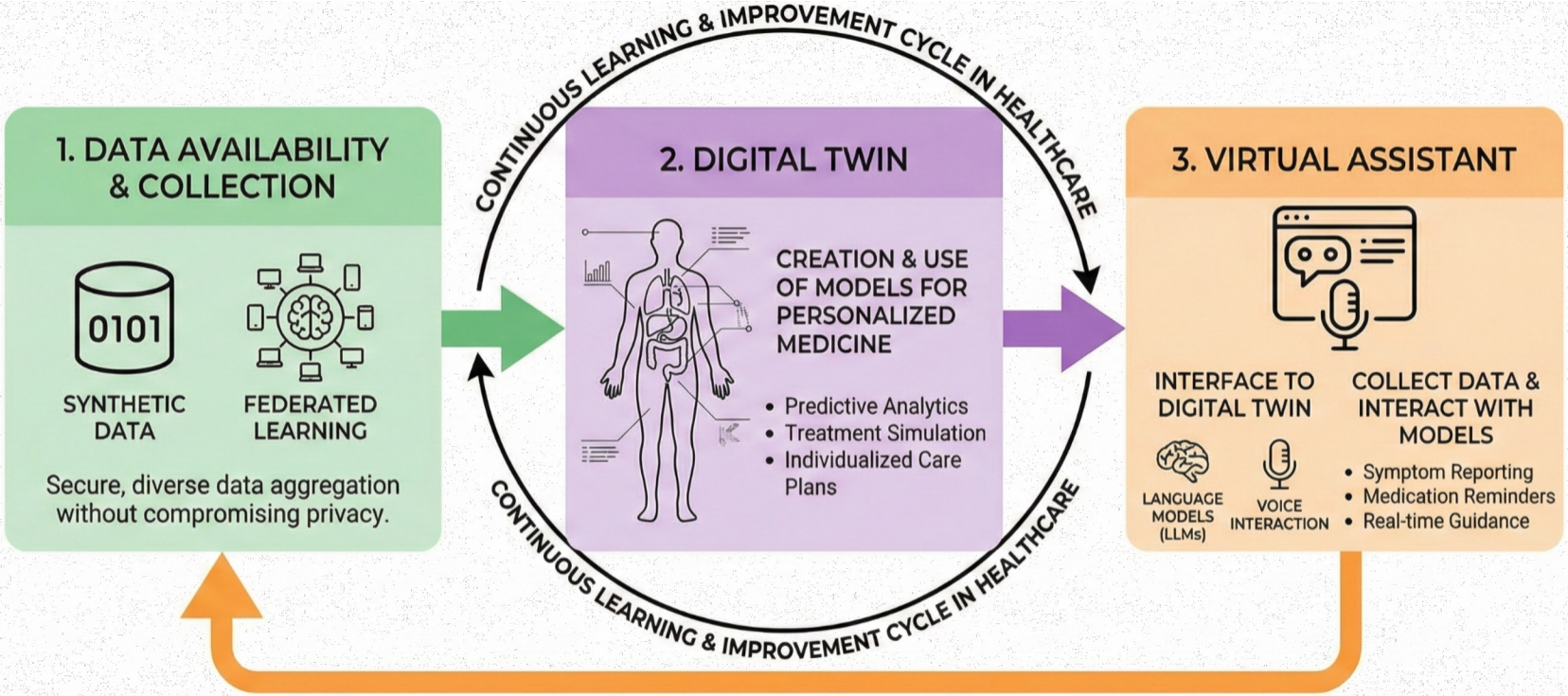
AI as a technological solution

Intelligent Support

The answer to healthcare challenges lies in **Advanced AI tools** acting as an extension of medical capabilities.

A **Virtual Assistant** can support healthcare with strong clinical validation to automate routine tasks, monitor patients, and provide decision support.

Goal: technology serving *humans*, not replacing them.


CONTINUOUS LEARNING & IMPROVEMENT CYCLE IN HEALTHCARE
1. DATA AVAILABILITY & COLLECTION
0101
SYNTHETIC DATA
FEDERATED LEARNING
Secure, diverse data aggregation without compromising privacy.
2. DIGITAL TWIN
CREATION & USE OF MODELS FOR PERSONALIZED MEDICINE
• Predictive Analytics
• Treatment Simulation
• Individualized Care Plans
3. VIRTUAL ASSISTANT
INTERFACE TO DIGITAL TWIN
LANGUAGE MODELS (LLMs)
VOICE INTERACTION
COLLECT DATA & INTERACT WITH MODELS
• Symptom Reporting
• Medication Reminders
• Real-time Guidance
CONTINUOUS LEARNING & IMPROVEMENT CYCLE IN HEALTHCARE

A.N.I.T.A. Virtual Assistant

FEATURES

- **24/7 Accessibility:** Continuous availability for patients and staff, no fixed hours.
- **Appointment Scheduling:** Automatic booking, modification, and cancellation of visits.
- **Initial Triage & Symptoms:** Preliminary assessment of symptoms and routing to the appropriate level of care.
- **Medical Information & FAQs:** Instant answers to common questions about procedures, services, and preparation.
- **Reminders & Follow-up:** Alerts for medications, appointments, and post-discharge check-ins.
- **Multilingual Support:** Communication in different languages to overcome linguistic barriers.
- **Hospital Navigation:** Interactive guide to find departments, clinics, and services within the facility.

A.N.I.T.A.
(Hospital Virtual Assistant)

BENEFITS

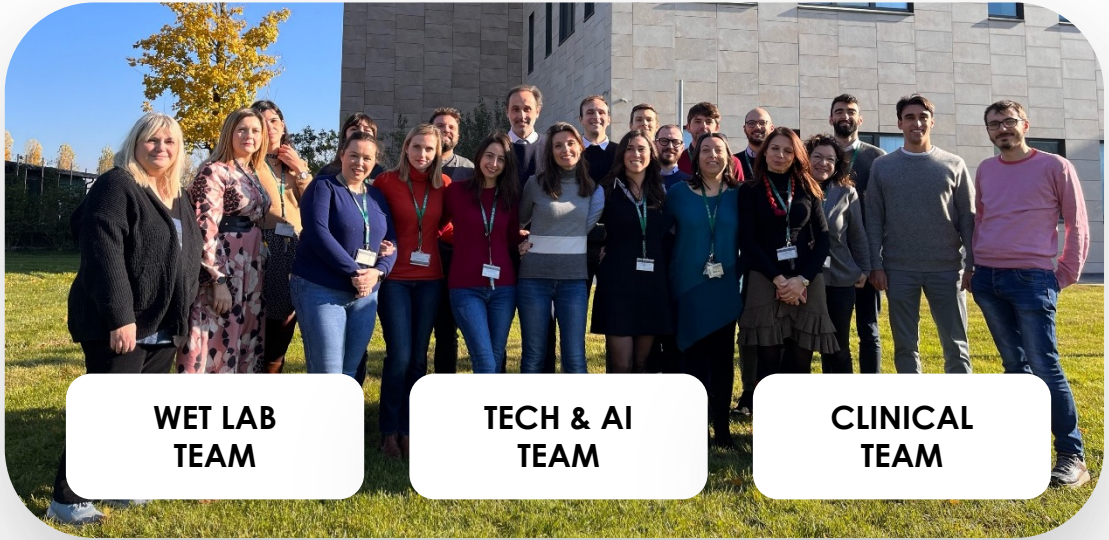
- **Improved Patient Experience:** Quick access, reduced anxiety, and greater overall satisfaction.
- **Operational Efficiency:** Automation of repetitive administrative tasks, freeing up valuable time for staff.
- **Reduced Waiting Times:** Instant responses and faster handling of requests, decreasing queues.
- **Cost Savings:** Decrease in operational expenses and optimization of resource allocation.
- **Continuous Monitoring & Prevention:** Early detection of health problems and improvement of therapy adherence.
- **Easy Data Access:** Secure and quick consultation of personal health information and medical records.
- **Support for Healthcare Staff:** Reduction of workload and focus on more complex clinical care and human interactions.

Thank you!

IRCCS HUMANITAS RESEARCH HOSPITAL

CENTER FOR ACCELERATING LEUKEMIA/LYMPHOMA RESEARCH
C.A.L.R.

Saverio D'Amico

saverio.damico@humanitas.it

train