

# AI Sym4MED. Synthetic and scalable data platform for medical empowered AI

**Consortio:** Associação Fraunhofer Portugal Research (Pt); Imperial College of Science Technology And Medi (Uk); Instrumentacion y Componentes Sa (Es); Consorci Sanitari de L'alt Penedès I Garraf (Es); Tiga Bilgi Teknolojileri Anonim Sirketi (Tr); Zabala Brussels (Be); Asociacion Instituto de Investigacion Sanitaria Bi (Es) ; Servicio Vasco de Salud Osakidetza (Es) ; Time.Lex (Be); Universidade do Porto (Pt) ; Nova Id Fct - Associação para a Inovacao E Desenv (Pt) ; Ibermatica (Es) ; Saidot Oy (Fi) ; Universitair Medisch Centrum Utrecht (NI) ; Universitat Zurich (Ch);

**Tecnología:** Administración Digital & Salud; Inteligencia Artificial

## Descripción general:

AI Sym4Med aims at developing a platform that will provide healthcare data engineers, practitioners, and researchers access to a trustworthy dataset system augmented with controlled data synthesis for experimentation and modeling purposes.

This platform will address data privacy and security by combining new anonymization techniques, attribute-based privacy measures, and trustworthy tracking systems. Moreover, data quality controlling measures, such as unbiased data and respect to ethical norms, context-aware search, and human-centered design for validation purposes will also be implemented to guarantee the representativeness of the synthetic data generated. Indeed, an augmentation module will be responsible for exploring and developing further the techniques of creating synthetic data, also dynamically on demand for specific use cases. Furthermore, this platform will exploit federated technologies for reproducing un-identifiable data from closed borders, promoting the indirect assessment of a broader number of databases, while respecting the privacy, security, and GDPR-compliant guideline

The proposed framework will support the development of innovative unbiased AI-based and distributed tools, technologies, and digital solutions for the benefit of researchers, patients, and providers of health services, while maintaining a high level of data privacy and ethical usage. AI Sym4Med will help in the creation of more robust machine learning (ML) algorithms for real-world readiness, while considering the most effective computation configuration.

Furthermore, a machine-learning meta-engine will provide information on the quality of the generalized model by analyzing its limits and breaking points, contributing to the creation of a more robust system by supplying on-demand real and/or synthetic data. This platform will be validated against local, national, and cross-border use-cases for both data engineers, ML developers, and aid for clinicians' operations.

**Programa:** HORIZON-HLTH-2022-IND-13 (101095387)

**Duración:** 48 meses (2019 – 2022)

**Presupuesto global proyecto:** 6.949.765,00 €

**Presupuesto Grupo Ayesa:** 212.500,00 €



# AI Sym4MED. Synthetic and scalable data platform for medical empowered AI

**Consortio:** **Associação Fraunhofer Portugal Research (Pt)**; Imperial College of Science Technology And Medi (Uk); Instrumentación y Componentes Sa (Es); Consorci Sanitari de L'alt Penedès I Garraf (Es); Tiga Bilgi Teknolojileri Anonim Sirketi (Tr); Zabala Brussels (Be); Asociación Instituto de Investigación Sanitaria Bi (Es) ; Servicio Vasco de Salud Osakidetza (Es) ; Time.Lex (Be); Universidade do Porto (Pt) ; Nova Id Fct - Associação para a Inovação E Desenv (Pt) ; Ibermática (Es) ; Saidot Oy (Fi) ; Universitair Medisch Centrum Utrecht (NI) ; Universitat Zurich (Ch);

**Tecnología:** Administración Digital & Salud; Inteligencia Artificial

## Rol de Ayesa:

Ibermática will collaborate on the Short-term prediction of Advanced Parkinson Disease (APD) by using unsupervised classification by Artificial intelligence applied to multimodal digital data from Electronic Health Records and open sources/datasets for augmentation of the data .

1. Definition of functional requirements: Definition of care processes for the identification, evaluation and follow-up of patients with APD in the Basque Health Care System (OSAK) (primary care, general neurology, specialized consultation in PD). Based on this study, and with the aim of having a reference model with which to compare the potential benefit of applying AISYM4MED in APD care processes, a quantitative descriptive analysis of the current situation in OSAK will be carried out for the processes of diagnosis and monitoring of APD.
2. Selection of criteria and parameters to define APD: considering two types of classification will be used to define APD patients in the project: 1) Simplified retrospective APD classification: based on retrospective data from the OSAK(Osabide) electronic medical records (tabulated and free text data); 2) Comprehensive APD classification: based on additional synthetic data created starting from existing DB and open sources.
3. Implementation: Retrospective identification of APD and data extraction according to the simplified retrospective APD classification patients with PD: will be identified according to the ICD 10-CM Diagnosis Codes registered Osabide to prove the capability of the software to correctly recognize them.

4. Synthesis of data from open sources and auditing of their quality: based on the approaches and models defined in objective 3 and 4, the consortium will create a synthetic dataset, thus augmenting the available population for running the classification algorithm.

5. Performance Evaluation of AISYM4MED APD Classification Algorithms Compared to Standard Clinical Practice Procedures of Oskidetza to Identify APD

