

Bioberri. Nuevas herramientas para la interoperabilidad semántica y analítica de datos en Biobancos



Consortio: Noray Bioinformatics; Ibermatica; i3B

Tecnología: Administración Digital & Salud; Inteligencia Artificial

Descripción general:

Desarrollo de nuevos productos y servicios que permitan la interoperabilidad, una mejor explotación de los datos y la integración de la información que actualmente se desperdicia, lo que conducirá a Biobancos más eficientes y rentables e impulsará su papel como elemento clave en el ecosistema de investigación en el ámbito de la medicina personalizada.

Ibermática, además de participar en todos los paquetes de trabajo, desarrollará tanto una herramienta como una metodología que de soporte a la codificación automatizada de textos escritos en lenguaje natural a SNOMED-CT.

La primera línea de trabajo está orientada a crear una base de conocimiento en la que partiendo de un Corpus de información procedente de la información textual actualmente almacenada, y mediante análisis del corpus a través de analítica de datos, se va a generar un modelo que almacene todo el conocimiento relativo a los textos utilizados en un Biobanco. Se obtendrá un texto origen de alta calidad, ya que uno de los problemas existentes en cualquier codificación automática de textos es la falta de calidad en el texto origen, con fallos ortográficos, expresiones ambiguas, anómalas para un contexto y se trabajará en el desarrollo de un sistema de codificación que, basado en técnicas de NLP y machine learning, automaticen la codificación de los textos en Snomed-CT. Ambas líneas de trabajo son independientes y darán como fruto dos productos complementarios que puedan ser usados de manera consecutiva o por separado.

Programa: Hazitek Competitivo (ZL-2019/00575; ZL-2020/00285)

Duración: 20 meses (2019 – 2020)

Presupuesto global proyecto: 297.205,32 €

Presupuesto Grupo Ayesa: 76.603,91 €

ACTUACIÓN COFINANCIADA POR EL DEPARTAMENTO DE DESARROLLO ECONÓMICO, SOSTENIBILIDAD Y MEDIO AMBIENTE DE GOBIERNO VASCO Y DE LA UNIÓN EUROPEA A TRAVÉS DEL FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)



Europar Batasuna
Unión Europea

Fondo Europeo de
Desarrollo Regional (FEDER)
"Una manera de hacer Europa"

Eskualde Garapenerako
Europar Funtsa (EGEF)
"Europa ogitiko modu bat"



EUSKO JAURLARITZA
GOBIERNO VASCO

Bioberri. Nuevas herramientas para la interoperabilidad semántica y analítica de datos en Biobancos

Consorcio: Noray Bioinformatics; Ibermatica; i3B

Tecnología: Administración Digital & Salud; Inteligencia Artificial

Rol de Ayesa:

Ibermática desempeña un papel clave en el proyecto BIOBERRI, centrando su contribución en el desarrollo de soluciones avanzadas de procesamiento del lenguaje natural y web semántica para mejorar la interoperabilidad de datos en biobancos.

Participa en la definición de requisitos, arquitectura y escenarios de validación, asegurando la alineación con necesidades reales del sector. Su principal aportación es el desarrollo de un sistema de codificación automática de textos médicos a SNOMED-CT, basado en una base de conocimiento construida con ontologías y enriquecida mediante técnicas de análisis lingüístico y aprendizaje automático.

Además, desarrolla algoritmos para mejorar la calidad de los datos (corrección de errores, desambiguación y detección de inconsistencias) y servicios inteligentes que asisten en la introducción de información clínica. Paralelamente, implementa herramientas de validación que permiten a expertos supervisar y mejorar los resultados.

Finalmente, contribuye a la integración de estas capacidades en la plataforma global del proyecto y a su validación en entornos reales, asegurando una solución robusta, escalable y orientada a su aplicación práctica.

