

## **PROGRAMA 46RF**

### **C18.I06 DATA LAKE SANITARIO**

#### **1. DENOMINACIÓN DEL COMPONENTE**

Renovación y ampliación de capacidades del Sistema Nacional de Salud.

#### **2. DESCRIPCIÓN GENERAL DEL COMPONENTE**

La crisis sanitaria ha dejado patente la fortaleza del Sistema Nacional de Salud español, pero también ha puesto de manifiesto las dificultades a las que se enfrenta a la hora de abordar situaciones que requieren anticipación, respuesta rápida y coordinación. También ha evidenciado la necesidad de corregir los problemas estructurales existentes relacionados con las tendencias demográficas, sociales, tecnológicas o económicas. Este componente del Plan de Recuperación, Transformación y Resiliencia español abordará los siguientes retos:

- la vulnerabilidad ante la crisis sanitaria mundial;
- la transformación del sistema de salud ante el envejecimiento de la población;
- la igualdad de género;
- y la sostenibilidad y la resiliencia a largo plazo del sistema.

#### **3. PRINCIPALES OBJETIVOS DEL COMPONENTE**

Los objetivos de este componente son:

- Preparar el sistema de salud para prevenir y hacer frente a posibles amenazas sanitarias globales como la actual, derivada de la pandemia de COVID-19, incrementando las capacidades de salud pública y los sistemas de vigilancia epidemiológica.
- Prestar servicios de salud con la máxima rapidez, calidad y seguridad, independientemente de los recursos económicos de los pacientes, su lugar de residencia, género, origen o edad.
- Mantener a las personas como centro del sistema sanitario, mejorando su participación y rediseñando la atención sanitaria para orientarla hacia las necesidades de las personas y de las comunidades.

- Garantizar sistemas de información que no solo midan la actividad, sino también los resultados finales en materia de salud.
- Promover la salud y el bienestar de forma activa y prevenir la enfermedad y la dependencia a lo largo de toda la vida.
- Atraer y retener a los y las mejores profesionales ofreciéndoles oportunidades de desarrollo individual y colectivo.
- Avanzar hacia un Sistema Nacional de Salud generador de información y conocimiento, digitalizado, y en el que se potencie la investigación y la innovación en el ámbito de la salud, como motor del empleo, el crecimiento, la productividad y la innovación.
- Asegurar una financiación suficiente y sostenible para abordar los nuevos retos en salud propios de una sociedad moderna y desarrollada, y garantizar asimismo la eficiencia del gasto.
- Fortalecer y desarrollar la coordinación y la gobernanza multinivel en la gestión del Sistema Nacional de Salud y potenciar la cohesión territorial. Promover activamente estrategias para alcanzar la igualdad de género en el sistema sanitario.

#### **4. DESCRIPCIÓN DE LA INVERSIÓN**

Esta inversión consiste en la generación de un Data Lake sanitario, que recoja la información de los sistemas de información, con el objetivo de proporcionar un análisis masivo con capacidad de respuesta en tiempo real para la identificación y mejora del diagnóstico y tratamiento, identificación de factores de riesgo, análisis de tendencias, identificación de patrones, predicción de situaciones de riesgo sanitario y programación de recursos para su atención incluyendo algoritmos de inteligencia artificial, y utilizando nuevas arquitecturas de sistemas escalables y nuevas herramientas de procesamiento y descubrimiento de modelos. La definición de usos se establecerá en fases posteriores del proyecto y se prevé la posibilidad de incorporar a los proveedores sanitarios privados. La Secretaría de Estado de Digitalización e Inteligencia Artificial es la responsable del proyecto y se llevará a cabo en colaboración con el Ministerio de Sanidad.

Se compone de diversas medidas:

- La adquisición de infraestructura tecnológica por las Comunidades Autónomas para la construcción del Data Lake sanitario.
- La creación de un repositorio de datos que permita incorporar datos provenientes de múltiples fuentes y dispositivos, tanto de las administraciones públicas como de redes abiertas y, en general, lo que se conoce como open data, en múltiples formatos

(imagen, texto, sensores) requiere un dimensionamiento específico de los recursos de almacenamiento y tratamiento de la información.

➤ Además, es preciso contar con herramientas de última tecnología, que permitan usos analíticos sofisticados, incluyendo algoritmos de inteligencia artificial. También es preciso contar con las necesidades de proporcionar respuesta en tiempo real.

➤ Es preciso una nueva arquitectura de sistemas escalables y nuevas herramientas de procesamiento y análisis de los datos, que den apoyo a la exploración de datos y al descubrimiento de modelos, al análisis ad hoc de diferente nivel de complejidad, a los reportes de gestión y seguimiento, y a los análisis de datos a tiempo real.

➤ Por ello, se utilizará una combinación de conceptos y tecnologías de data warehouse (almacén de datos) y data lake (repositorio de datos raíz) que incluirá una gran variedad de procesos y herramientas analíticas.

– Incorporación de Comunidades Autónomas al Data Lake.

➤ Definición de estructura, gobernanza y estrategia de coordinación común a través de grupos de trabajo específicos, y aprobado en el seno del Consejo Interterritorial del Sistema Nacional de Salud.

➤ Identificación de colecciones de datos y servicios a proporcionar por las Comunidades Autónomas, así como las condiciones de seguridad y conservación aplicables en un modelo de federación de espacios de datos.

➤ Definición de modelos de consumo de información por las Comunidades Autónomas, usos permitidos de datos, teniendo en cuenta particularidades en la organización y modelos de gestión de sus servicios de salud y de las necesidades derivadas de las diferentes estrategias específicas (salud pública, medicina personalizada y otras).

➤ Establecimiento de derechos y obligaciones para agente en relación a la aplicación de tecnologías disruptivas sobre el Data Lake.

➤ Implementación de sistemas, plataformas y procesos tecnológicos necesarios para la incorporación y explotación de la información, en cada una de las Consejerías de Salud participantes, con una certificación técnica y funcional común previa.

➤ Definición y puesta en marcha de proyectos de tratamiento masivo de datos por las Comunidades Autónomas, intentando primar la búsqueda de colaboración y sinergias entre organizaciones del sector privado y centros de investigación.

Esta arquitectura deberá tener un dimensionamiento en seguridad y capacidad ajustado a los niveles de servicio definidos en cada uno de los escenarios de uso.

## 5. COSTE DE LA INVERSIÓN Y DISTRIBUCIÓN ANUALIZADA

(millones de euros)

Periodificación	2020	2021	2022	2023	2024	2025	2026	Total
Coste del Mecanismo	--	--	35,00	65,00	--	--	--	100,00
Otra financiación	--	--	--	--	--	--	--	--
<b>Total</b>	--	--	35,00	65,00	--	--	--	100,00

## 6. HITOS Y OBJETIVOS DE LA INVERSIÓN

Nº	Medida	Hito/ objetivo	Nombre	Indicadores cualitativos para los hitos	Indicadores cuantitativos para los objetivos			Tiempo		Descripción de cada hito y objetivo
					Unidad	Valor de referencia	Meta	Tr	Año	
284	C18.I6	Objetivo	Operatividad del Data Lake sanitario		Nº	0	17	T4	2023	Operatividad del Data Lake sanitario a escala estatal e incorporación a él de las 17 Comunidades Autónomas y de las Ciudades Autónomas con el objetivo de permitir el análisis masivo de datos para la identificación y mejora del diagnóstico y los tratamientos.

## 7. CENTRO GESTOR RESPONSABLE

- Secretaría de Estado de Digitalización e Inteligencia Artificial.