

Nº Expediente: PA1/2025

**MEMORIA JUSTIFICATIVA DE LA NECESIDAD E IDONEIDAD DEL CONTRATO MIXTO DE SUMINISTRO DESTINADO A LA ADQUISICIÓN DE LICENCIAS DE PRODUCTOS DE SOFTWARE PARA EL PROCESAMIENTO DE LENGUAJE NATURAL, EN FORMATO DE SUSCRIPCIÓN, JUNTO CON LOS SERVICIOS ASOCIADOS DE INSTALACIÓN Y SOPORTE, CON EL OBJETIVO DE DESARROLLAR CAPACIDADES ADICIONALES E INTEGRARLAS EN LA PLATAFORMA TECNOLÓGICA DEL ESPACIO DE DATOS SANITARIOS DE USO SECUNDARIO DE LA COMUNIDAD DE MADRID CON FINES DE INVESTIGACIÓN E INNOVACIÓN “HEALTHDATA@MAD-R&I”, EN CONSONANCIA CON EL ESPACIO EUROPEO DE DATOS EN SALUD, Y CONFORME A LA CONVOCATORIA DEL MINISTERIO DE TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA, EN EL MARCO DEL PLAN DE RECUPERACIÓN, TRANSFORMACIÓN Y RESILIENCIA – NEXT GENERATION EU (PROGRAMA ESPACIOS DE DATOS SECTORIALES), MEDIANTE PROCEDIMIENTO ABIERTO SIMPLIFICADO Y TRAMITACIÓN URGENTE.**

En las últimas décadas, la proporción de información de salud contenida en los sistemas de registros de salud electrónicos (RSE) está creciendo constantemente, dando lugar a un recurso valioso pero fragmentado de "datos del mundo real", a partir del cual todos los agentes interesados podrían obtener conocimientos significativos. Los sistemas sanitarios necesitan urgentemente mejorar su capacidad para aprender de los datos que poseen, profundizando en el conocimiento del proceso salud-enfermedad, con el fin de optimizar las vías de atención, lograr los mejores resultados posibles para los pacientes, hacer un uso más adecuado de los recursos y mejorar la seguridad de los pacientes. Esta necesidad de una mejora en la atención sanitaria basada en la evidencia, que puede proporcionar el uso de los datos, cumple con la creciente expectativa social de equidad en los estándares de atención entre los sistemas de salud y entre diferentes grupos de población, incluyendo la perspectiva de género y de edad.

Durante la última década, las administraciones públicas y otros agentes financiadores han realizado grandes inversiones en infraestructuras de transformación digital en salud, generando RSE cada vez más completos y complejos, cuyo contenido, en una alta proporción, necesita ser estructurado y normalizado. Por otra parte, existen grandes infraestructuras de datos multinacionales, como la Iniciativa de Medicamentos Innovadores (IMI) de la Unión Europea (<https://www.medicamentosinnovadores.org/sites/medicamentosinnovadores/HomeIMI.html>), la Red Europea de Datos y Evidencias de Salud (EHDEN, <https://www.ehden.eu/>), la iniciativa DARWIN EU® (<https://www.ema.europa.eu/en/aboutus/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu>) de la Agencia Europea de Medicamentos (EMA) y las propuestas de la Comisión Europea para un Espacio Europeo de Datos Sanitarios (EHDS). A nivel mundial, se ha establecido la red colaborativa de ciencia abierta *Observational Health Data Sciences and Informatics* (OHDSI, <https://www.ohdsi.org/>) para respaldar estudios rápidos de redes a nivel internacional.

La Estrategia Europea de Datos, lanzada en 2020 por la Comisión Europea (<https://digitalstrategy.ec.europa.eu/en/policies/strategy-data>), sienta las bases para la creación de espacios comunes europeos de datos en sectores clave, como la movilidad, financiero, energía, agricultura o salud.



La autenticidad de este documento se puede comprobar en  
<https://gestiona.comunidad.madrid/csv>  
mediante el siguiente código seguro de verificación: **0945554439685526598867**

Con fecha 18 de enero de 2024, el Ministerio para la Transformación Digital y de la Función Pública publicó la Orden TDF/1461/2023, de 29 de diciembre, por la que se aprueban las bases reguladoras y se efectúa la convocatoria para la concesión de ayudas, en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de Compartición de Datos, en el marco del Plan de Recuperación, Transformación y Resiliencia- Next Generation EU (Programa Espacios de Datos Sectoriales).

A la citada convocatoria se presentó la Fundación para la Investigación e Innovación Biosanitaria de Atención Primaria de la Comunidad de Madrid (en adelante, FIIBAP o La Fundación), con el proyecto Espacio de datos de salud de uso secundario con fines de investigación científica e innovación HealthData@MAD-R&I.

Con fecha 20 de diciembre de 2024, el Ministerio para la Transformación Digital y de la Función Pública notificó resolución de concesión definitiva para el proyecto HealthData@MAD-R&I con un presupuesto de 2.337.568 €.

Por tanto, el proyecto HealthData@MAD-R&I está financiado, en el marco de la convocatoria para la concesión de ayudas, en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de Compartición de Datos, por el Ministerio para la Transformación Digital y de la Función Pública y por la Unión Europea a través de los recursos financieros derivados del Plan de Recuperación, Transformación y Resiliencia-Next Generation EU (Programa Espacios de Datos Sectoriales).

### **Cumplimiento del Plan de Recuperación, Transformación y Resiliencia (PRTR)**

El presente contrato se enmarca en el desarrollo del proyecto *Espacio de datos de salud de uso secundario con fines de investigación científica e innovación HealthData@MAD-R&I*, financiado en el marco del **Plan de Recuperación, Transformación y Resiliencia (PRTR)**, aprobado por la Decisión de Ejecución del Consejo ECOFIN de 13 de julio de 2021 [COM (2021) 322 final].

Más concretamente, la actuación se integra en:

- Componente 12: Política Industrial España 2030
- Inversión 1: Espacios de datos sectoriales
- Programa: Espacios de Datos Sectoriales, promovido por el Ministerio para la Transformación Digital y de la Función Pública
- Objetivo CID 180, con vencimiento el 30 de junio de 2026:

*“Grandes espacios de datos sectoriales e interoperables de alto valor creados en sectores estratégicos. Al menos cuatro en los sectores agroalimentario, movilidad sostenible, salud y comercio.”*

El proyecto HealthData@MAD-R&I contribuye directamente al cumplimiento de este objetivo mediante la creación y despliegue de un espacio de datos interoperable en el **sector salud**, que responde a los criterios establecidos en las bases reguladoras del programa, entre ellos:



- la compartición soberana, segura y conforme a derecho de datos sanitarios,
- la interoperabilidad y alineación con los estándares europeos,
- la generación de casos de uso con impacto real en la investigación, la innovación y la mejora de la atención sanitaria,
- y la incorporación de mecanismos de gobernanza y seguimiento alineados con la Estrategia Europea de Datos y el Reglamento (UE) 2021/241.

Este contrato en concreto —de adquisición de licencias y servicios asociados para el procesamiento de lenguaje natural, detección de entidades clínicas y codificación a terminología estándar— tiene por finalidad dotar al Espacio de Datos Sanitarios de capacidades avanzadas para el tratamiento de información no estructurada procedente de historias clínicas, lo que resulta esencial para alcanzar los objetivos científicos, técnicos y de interoperabilidad del proyecto.

La Fundación para la Investigación e Innovación Biosanitaria de Atención Primaria de la Comunidad de Madrid se constituye de conformidad con el Decreto 53/2015, de 21 de mayo, por el que se autoriza la constitución de la misma, como organización sin ánimo de lucro de las previstas en el título I de la Ley 1/1998, de 2 de marzo, de Fundaciones de la Comunidad de Madrid, teniendo la consideración de Entidad de Derecho Privado integrada en el sector público autonómico. Asimismo, las Fundaciones de Investigación Biomédica, de acuerdo con lo dispuesto en el artículo 113 bis, de la Ley 12/2001, de 21 de diciembre, de Ordenación Sanitaria de la Comunidad de Madrid, están configuradas como Organismos Públicos de Investigación de la Comunidad de Madrid.

La Fundación tiene atribuida como finalidad, de acuerdo con lo previsto, en el apartado I, del artículo primero, de sus Estatutos, promover la investigación científico-técnica, y la innovación, así como la formación y docencia en el área de ciencias de la salud, con el objetivo de potenciar la calidad asistencial en los centros sanitarios dependientes del órgano competente en materia de atención primaria de la Comunidad de Madrid.

Entre las actividades de la Fundación destacan las de promocionar y coordinar la realización y desarrollo de programas de investigación científica e innovación aplicadas a las ciencias de la salud; facilitar la financiación y la gestión de los procesos de investigación e innovación de atención primaria, salud pública, urgencias y emergencias y proyectar a la sociedad y al entorno sanitario los avances de la investigación y la innovación.

En lo relativo a la Comunidad de Madrid, existen diversas iniciativas ya consolidadas, en concreto, nuestra región dispone de una plataforma que incluye el repositorio de todos los datos clínicos generados desde las historias clínicas de hospitales, atención primaria y el servicio de urgencias y emergencias con más de 11 millones de pacientes, y con recursos muy amplios para la gestión, estandarización y normalización de los datos. A esta plataforma se une la experiencia de Infobanco, un proyecto de Compra Pública Innovadora liderado por la Consejería de Sanidad de la Comunidad de Madrid y financiado por el Ministerio de Ciencia e Innovación en el marco de la Línea de fomento de Innovación desde la Demanda (Línea FID). Este proyecto constituye un importante ejemplo de colaboración público-privada de gestión, estandarización, armonización y normalización de datos, así como de desarrollo de casos de uso, que incluyen diferentes usos secundarios y, de forma preferente, casos de uso relacionados con la investigación y la innovación. Esta situación de partida, desde el punto de vista tecnológico, nos ubica,



por tanto, en una situación ventajosa para afrontar el reto planteado al amparo del proyecto HealthData@MAD-R&I.

El proyecto HealthData@MAD-R&I tiene como objetivo general, definir, desarrollar, validar y documentar casos de uso relacionados con la investigación y la innovación que permitan validar los servicios del Espacio de datos sanitarios electrónicos de uso secundario y como objetivos específicos y transversales los siguientes:

- **OE1.** Desarrollar y poner en producción un entorno de investigación e innovación que incluya un Espacio de Datos Sanitarios Electrónicos de Uso Secundario en Investigación e Innovación.
- **OE2.** Definir, validar y poner en marcha un modelo de gobernanza para el acceso a datos de uso secundario que tenga en cuenta a los titulares, usuarios, representantes de pacientes y organismos de acceso a datos.
- **OE3.** Integrar y converger las iniciativas innovadoras y de colaboración público-privada existentes con los sistemas de acceso a datos en producción.
- **OE4.** Proporcionar herramientas que faciliten la detección de patologías reumatológicas inflamatorias o de mala evolución impactando en el curso de patologías de alta prevalencia e impacto negativo en la calidad de vida.
- **OE5.** Contribuir a la mejora de la salud y de la atención sociosanitaria en mujeres con larga supervivencia de cáncer de mama.
- **OE6.** Generar sistemas predictivos que ayuden a detectar el riesgo de reingreso y a orientar la provisión de servicios sociosanitarios para evitarlos.
- **OE7.** Contribuir a la mejora de la calidad de vida de las personas mayores de 75 años.
- **OT1.** Contribuir a la economía del dato: el proyecto va a generar un marco adecuado para la generación de la economía del dato en el sector sanitario en la Comunidad de Madrid.
- **OT2.** Adaptar las iniciativas y plataformas disponibles a los requisitos definidos por el Reglamento del Espacio Europeo de Datos para su uso secundario
- **OT3.** Incorporar nuevos actores al ecosistema del espacio de datos sanitario de uso secundario donde estén representados el sistema sanitario, los pacientes y la industria.

En el marco del proyecto HealthData@MAD-R&I se han previsto cuatro casos de uso en diferentes áreas de investigación que sirven como demostradores del potencial científico-tecnológico y comercial que los datos de salud aportan a la economía del dato en el sector sanitario:

- **CASO DE USO 1:** Optimización de las derivaciones de sujetos con enfermedades reumáticas y musculoesqueléticas entre atención primaria y reumatología
- **CASO DE USO 2:** Trayectorias asistenciales de las mujeres largas supervivientes de cáncer de mama SURBCANMADRID
- **CASO DE USO 3:** Hospitalización No Programada (MAD-HNP.eSTRATA)
- **CASO DE USO 4:** Modelo explicativo en vida real sobre la efectividad de las estatinas en la reducción de eventos cardiovasculares y la mortalidad en la población anciana sin antecedentes de enfermedad cardiovascular de la Comunidad de Madrid.



Finalmente, también se considera un plan de sostenibilidad y escalabilidad para que puedan servir como referente para otros potenciales casos de uso.

El resultado final buscado es permitir que España y la industria nacional en su conjunto, se consolide como un líder y referente en Europa. Todo ello tendrá un efecto tractor sobre la cadena de valor del uso secundario de datos sanitarios existente en nuestro país y estimularán el despliegue de nuevos modelos de negocio, actividad y empleo en línea con la Estrategia Europea de Datos.

Para lo anterior, se precisa ampliar las capacidades de la Plataforma HealthData@MAD-R&I que dará soporte a los casos de uso, por lo que **resulta necesario la adquisición de licencias** que cubran todas las necesidades y requerimiento técnicos.

Por todo lo cual, se precisa el suministro consistente en la adquisición de licencias y la implementación de los servicios asociados en el entorno tecnológico del *DataLake* de HealthData@MAD-R&I para fortalecer y ampliar la capacidad de análisis de datos sanitarios, con el objetivo final de mejorar la capacidad para utilizar datos existentes y nuevos para la toma de decisiones en tiempo real y el desarrollo de modelos predictivos en el ámbito de la salud. La plataforma HealthData@MAD-R&I se alinea con las directrices y especificaciones técnicas para facilitar el uso ininterrumpido de los datos sanitarios en toda Europa en el marco del Espacio Europeo de Datos Sanitarios (EHDS).

En concreto, **el objetivo de este contrato es implantar una solución de Procesamiento de Lenguaje Natural** para extraer valor de los datos sanitarios existentes en la organización para tomar decisiones informadas en el ámbito de la salud, incluyendo:

- Facilitar la anonimización de datos de carácter personal en documentación clínica
- Desarrollar herramientas de Detección de Entidades (*Named Entity Recognition – NER* – por sus siglas en inglés)
- Codificar a terminología estándar las entidades reconocidas

El contrato propuesto es un **contrato mixto** de suministro destinado a la adquisición de licencias de productos de software para el procesamiento de lenguaje natural, en formato de suscripción, junto con los servicios asociados de instalación y soporte, con el objetivo de desarrollar capacidades adicionales e integrarlas en la plataforma tecnológica del Espacio de Datos Sanitarios de Uso Secundario de la Comunidad de Madrid con fines de investigación e innovación “Healthdata@MAD-R&I”, de los previstos en el artículo 18 de la Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público. En este sentido, para la determinación de las normas que regirán la adjudicación del contrato acudimos a la **prestación principal, en el presente caso, suministro**. Así, la presente contratación se califica principalmente como contrato de suministro, atendiendo al sentido o la finalidad que se pretende alcanzar con el contrato en cuestión, así como a la magnitud del suministro que lleva aparejado el servicio accesorio de instalación y soporte de la licencia. De esta forma, la prestación de los servicios accesorios de instalación y soporte son consecuencia de la realización de un suministro de licencias que resulta necesario, relevante y de gran envergadura sin la cual esos servicios no tendrían ninguna trascendencia.

Todo ello de acuerdo con lo establecido en el art. 18 de la LCSP y lo contemplado por la





doctrina y jurisprudencia, entre otras la Resolución 1122/2018, de 7 de diciembre, del Tribunal Administrativo Central de Recursos Contractuales, toda vez que el fin último perseguido mediante el contrato no es la prestación de los servicios de soporte y su instalación, sino el resultado propio de un suministro de licencias siendo el servicio accesorio o complementario del mismo.

De acuerdo con todo lo anteriormente expuesto, y con el fin de dar respuesta a las necesidades a las que responde el proyecto "Espacio de datos de salud de uso secundario con fines de investigación científica e innovación HealthData@MAD-R&I", así como para el cumplimiento a la concesión de la ayuda del Ministerio para la Transformación Digital y de la Función Pública, para el desarrollo del mismo, se hace necesaria y se propone la contratación por un **importe máximo de 150.000 € sin IVA**, y 181.500 € IVA incluido.

El contrato tendrá una duración estimada de once (11) meses, puesto que se prevé su formalización el 1 de julio de 2025. No obstante, con independencia de la fecha efectiva de formalización, **el contrato deberá finalizar, en todo caso, el 15 de junio de 2026**, por lo que la duración definitiva del mismo se ajustará a dicho plazo límite. **La eventual reducción del plazo efectivo de ejecución no conllevará una disminución del importe de adjudicación, puesto que el presupuesto del contrato se ha calculado en función del cumplimiento de los resultados esperados, con independencia del tiempo requerido para su consecución.** En este sentido, la necesidad del presente contrato "llave en mano" radica en la necesidad de obtener un trabajo completo y funcional, incluyendo el licenciamiento, la integración y puesta en marcha del software, así como su soporte hasta pleno funcionamiento, todo dentro de un precio y plazo definidos. Así, en caso de que el plazo de ejecución del contrato sea menor, el contratista deberá adscribir cuantos medios materiales y personales adicionales sean necesarios a efectos de asegurar la correcta ejecución del contrato.

Debido a la duración del contrato propuesto resulta necesario tramitar el presente expediente de contratación como plurianual, 2025-2026, con la imputación de costes por anualidad que figura en la memoria económica y en la propuesta de contratación.

Finalmente, cabe reseñar que, la naturaleza y extensión de las necesidades que pretenden cubrirse mediante el contrato proyectado quedan determinadas en la documentación preparatoria del expediente, resultando su objeto y contenido idóneos para la satisfacción de las mismas.

Para dar cumplimiento a los objetivos del proyecto, se requiere la contratación del presente objeto contractual enmarcado dentro de las **siguientes actuaciones conforme a la resolución de concesión de la ayuda:**

Contrato propuesto	Licencia relacionada	Importe en la resolución, IVA no incluido	Actividad en la resolución	PT relacionado	Caso de uso relacionado
<b>C1 Licencias Procesado Lenguaje Natural.</b>	Licencia 1	30.000,00 €	A3 - Normalización, estandarización y armonización de los datos	PT4	1



<b>Detección de Entidades y Codificación a Estándar</b>	Licencia 3	30.000,00 €	A2 - Armonización y carga de datos. Validación	PT5	2
	Licencia 5	30.000,00 €	A2 - Armonización y validación de datos (MADHNP. eARM)	PT6	3
	Licencia 2	30.000,00 €	A4 - Análisis estadístico y generación de modelos	PT4	1
	Licencia 4	30.000,00 €	A3 - Análisis de datos. Validación de trayectorias y modelos	PT5	2
		<b>150.000,00 €</b>			

Todo ello de conformidad con lo dispuesto en el artículo 28 de la Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público.

En Madrid, a fecha de última firma electrónica  
La Presidencia del Patronato

