

### III. OTRAS DISPOSICIONES

#### MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

**12053** *Resolución de 3 de junio de 2025, del Consorcio Barcelona Supercomputing Center-Centro Nacional de Supercomputación, por la que se publica el Convenio con las entidades participantes para la realización de actuaciones ligadas al desarrollo del Plan Nacional de Tecnologías del Lenguaje y del Ecosistema Empresarial de las Tecnologías del Lenguaje, en el marco de la Estrategia de Inteligencia Artificial 2024 y el Plan de Recuperación, Transformación y Resiliencia -financiado por la Unión Europea- Next Generation EU.*

Con fecha 2 de junio de 2025, el Consorcio Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS) ha suscrito el convenio referenciado arriba con las entidades firmantes de dicho convenio.

Por ello y en cumplimiento de lo dispuesto en el artículo 48.8 de la Ley 40/2015, de 1 de octubre de Régimen Jurídico del Sector Público, procede la publicación en el «Boletín Oficial del Estado» de dicho convenio, que figura como anexo a esta resolución.

Barcelona, 3 de junio de 2025.–El Director del Consorcio Barcelona Supercomputing Center-Centro Nacional de Supercomputación, Mateo Valero Cortés.

#### ANEXO

**Convenio entre el Barcelona Supercomputing Center-Centro Nacional de Supercomputación y demás entidades participantes para la realización de actuaciones ligadas al desarrollo del Plan Nacional de Tecnologías del Lenguaje y del Ecosistema Empresarial de las Tecnologías del Lenguaje en el marco de la Estrategia de Inteligencia Artificial 2024 y el Plan de Recuperación, Transformación y Resiliencia –financiado por la Unión Europea– Nextgeneration EU**

Barcelona, a 2 de junio de 2025.

#### INTERVIENEN

De una parte, don Mateo Valero Cortés, actuando en nombre y representación del Barcelona Supercomputing Center-Centro Nacional De Supercomputación (en adelante BSC-CNS), provisto del CIF S0800099D y con domicilio social sito en C/Jordi Girona 31, 08034, Barcelona, con el cargo de Director que ostenta en virtud de la renovación de su nombramiento por el Consejo Rector del BSC-CNS en la reunión celebrada el día 25 de febrero de 2025, en virtud de las competencias atribuidas al director en el artículo 15.i) de los Estatutos de este Consorcio.

De otra parte, don José Ramón Bengoetxea Caballero, Rector Magnífico de la Universidad del País Vasco/Euskal Herriko Unibertsitatea (en adelante, UPV/EHU), nombrado por el Decreto 1/2025, de 20 de enero, (BOPV, núm. 17, lunes 27 de enero de 2025), en nombre y representación de dicha Universidad, en virtud de las atribuciones que tiene conferidas por razón de su cargo, conforme a lo dispuesto en el artículo 179 de los Estatutos de la Universidad del País Vasco/Euskal Herriko Unibertsitatea.

De otra parte, don Antonio López Díaz, Rector magnífico da Universidad de Santiago de Compostela (en adelante, USC), en virtud de lo establecido en sus estatutos y del nombramiento como Rector por el Decreto 35/2022, del 17 de marzo, (DOG núm. 56 do 22 de marzo de 2022), actuando en nombre y representación de la mencionada entidad.

De otra parte, doña Amparo Navarro Faure, Rectora Magnífica de la Universidad de Alicante, nombramiento que acredita mediante Decreto 186/2024 de 17 de diciembre del Consell de la Generalitat, y en nombre y representación de este organismo, en virtud de las atribuciones que tiene conferidas en razón de su cargo conforme a lo dispuesto en el artículo 64, apartado h del Estatuto de la Universidad de Alicante, situado en el Campus de San Vicente del Raspeig, Ap. 99, I-03080 Alicante.

Y, de otra parte, don Nicolás Ruiz Reyes Rector Magnífico de la Universidad de Jaén, BOJA, nombramiento que acredita a través del Decreto 132/2023, de 12 de junio («Boletín Oficial de la Junta de Andalucía» número 113, jueves, 15 de junio de 2023 página 10352/1) (en adelante, UJA).

BSC-CNS y las entidades participantes arriba descritas (en adelante, las entidades participantes o colaboradoras) serán denominadas conjuntamente como las «partes» o individualmente como «parte».

Las partes declaran hallarse debidamente facultadas y con la capacidad legal suficiente para suscribir el presente convenio, a cuyo efecto.

## EXPONEN

Primero.

Que el Plan de Recuperación, Transformación y Resiliencia establece la hoja de ruta para modernizar la economía española, reactivar el crecimiento económico y generar empleo. Además, pretende conseguir una reconstrucción económica sólida, inclusiva y resiliente tras la crisis provocada por la COVID-19, así como responder a los desafíos de la próxima década. Este Plan está financiado por los fondos Next Generation EU, incluyendo el Mecanismo de Recuperación y Resiliencia.

El Plan de Recuperación, que consta de treinta y un componentes tras su modificación mediante adenda aprobada por la Comisión Europea el 2 de octubre de 2023 y ratificada por el ECOFIN (Consejo de Ministros de Finanzas de la UE) el 17 de octubre de 2023, abarca un conjunto de reformas e inversiones destinadas a transformar el modelo productivo de España. Sus principales áreas de acción son la transición energética, la digitalización, la cohesión territorial y social, y la igualdad.

Segundo.

Que el Componente 16, Reforma 1 del Plan de Recuperación comprende la Estrategia Nacional de Inteligencia Artificial (ENIA), lanzada el 23 de julio de 2020 por el Gobierno de España como parte de la Agenda España Digital. Esta estrategia está diseñada para guiar el proceso de transformación digital del país. En el tercer eje estratégico de ENIA, «Desarrollar plataformas de datos e infraestructuras tecnológicas que den soporte a la IA», se incluye la medida 14: «Impulso al Plan Nacional de Tecnologías del Lenguaje».

El objetivo principal del Plan Nacional de Tecnologías del Lenguaje es promover la investigación y el desarrollo industrial en áreas como el procesamiento del lenguaje natural, la traducción automática y los sistemas conversacionales en España, tanto en español como en las lenguas cooficiales. Con ello se pretende fortalecer el desarrollo de la IA de manera transparente y responsable, facilitando su expansión tanto en el sector público como en el privado.

Tercero.

Que la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA) ejerce, entre otras funciones, la elaboración, formulación, coordinación y evaluación de la Estrategia Nacional de Inteligencia Artificial (ENIA). Esta actividad se enmarca dentro del Componente 16, Reforma 1, del Plan de Recuperación, Transformación y Resiliencia y

se realiza a través de la Dirección General de Digitalización e Inteligencia Artificial, específicamente por la Subdirección General de Inteligencia Artificial y Tecnologías Habilitadoras Digitales, conforme al artículo 3.a) del Real Decreto 2014/2024, de 27 de febrero, que establece la estructura orgánica básica del Ministerio para la Transformación Digital y de la Función Pública.

Cuarto.

Que desde la publicación de la Estrategia Nacional de Inteligencia Artificial se ha experimentado un notable avance en la capacidad computacional y en la IA generativa. Esta evolución resalta la necesidad de acelerar la adopción de estas tecnologías en la Administración General del Estado, actuando como motor para su adopción en el tejido empresarial español.

Por este motivo, el Consejo de Ministros aprobó la «Estrategia de Inteligencia Artificial 2024» la cual se basa en tres ejes fundamentales: i) fortalecer los elementos clave para el desarrollo de la IA, ii) facilitar la expansión de la IA en los sectores público y privado promoviendo la innovación y la ciberseguridad, y iii) desarrollar una IA transparente, responsable y humanística.

Que en el contexto actual es esencial incluir el castellano y las lenguas cooficiales en futuras aplicaciones de inteligencia artificial. El objetivo es garantizar un mayor porcentaje de datos de entrenamiento en español y en las lenguas cooficiales, promoviendo así una mayor soberanía tecnológica y transparencia. Para lograr este objetivo, el BSC-CNS desempeña un papel fundamental en la generación de estos modelos ya que cuenta con una avanzada infraestructura de supercomputación. Esta infraestructura permite manejar grandes volúmenes de datos y realizar cálculos complejos a alta velocidad, esenciales para el desarrollo y entrenamiento de modelos de inteligencia artificial y procesamiento de lenguaje natural. Es por ello que la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA) otorga al BSC-CNS la financiación objeto de esta resolución, tal y como se indica en el expositivo siguiente, en la medida que esta aportación potenciará significativamente la capacidad de desarrollar tecnología lingüística avanzada, posicionando España como un referente en el ámbito de la inteligencia artificial multilingüe.

Quinto.

Que la Resolución de fecha 19 de agosto de 2024, la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA) autoriza al BSC-CNS una aportación de ocho millones cuatrocientos treinta mil seiscientos cinco euros con setenta y cuatro céntimos (8.430.605,74 €) con cargo a la aplicación presupuestaria 33.50.46PR.74054 para la realización de actuaciones ligadas al desarrollo del Plan Nacional de Tecnologías del Lenguaje y del Ecosistema Empresarial de las Tecnologías del Lenguaje en el marco de la estrategia de Inteligencia Artificial 2024 y el Plan de Recuperación, Transformación y Resiliencia financiado por la Unión Europea-Next Generación EU. Que, en virtud de dicha resolución, el BSC-CNS actuará como entidad instrumental del Ministerio bajo la responsabilidad del Ministerio para la Transformación Digital y de la Función Pública y éste como entidad ejecutora del Proyecto.

Sexto.

Que el BSC-CNS es un consorcio formado por la Administración General del Estado, representada por el Ministerio de Ciencia, Innovación y Universidades, la Generalitat de Catalunya, representada por el Departament de Recerca i Universitats, y por la Universitat Politècnica de Catalunya (UPC). El Consorcio BSC-CNS está adscrito al Ministerio de Ciencia, Innovación y Universidades en representación de la Administración General del Estado. El presupuesto del BSC-CNS, como entidad dependiente de dicha Administración, se integra con carácter anual en los Presupuestos Generales del Estado.

El BSC-CNS tiene como misión científica la computación de alto rendimiento (HPC, por sus siglas en inglés) y desarrolla un papel activo promoviendo el HPC en España y Europa como herramienta esencial para la competitividad internacional en ciencia e ingeniería. Asimismo, el BSC-CNS está al servicio de la comunidad científica internacional y de la industria que requiere recursos HPC, configurándose como un centro internacional de excelencia en ciencias electrónicas. Entre sus principales líneas de investigación se incluye el procesamiento de lenguaje natural, un área que demanda una gran capacidad computacional para el desarrollo, entrenamiento y ejecución de modelos de lenguaje natural. El BSC-CNS integra diferentes disciplinas científicas y tecnológicas, lo que permite un enfoque holístico y multifacético en el desarrollo de modelos, abarcando desde la teoría hasta la aplicación práctica. Por último, la colaboración del BSC-CNS con otras instituciones en este campo permite el intercambio de conocimientos y el acceso a las últimas innovaciones y técnicas en IA.

Séptimo.

Que, de acuerdo con la resolución de la SEDIA, el BSC-CNS contará con la colaboración de las entidades que se detallan en los siguientes exponendos:

La UPV/EHU es la universidad pública vasca, comprometida con su cultura, con los retos sociales y con el desarrollo sostenible a través de la formación, la investigación, la transferencia de conocimiento y las alianzas locales y globales, en todas las ramas del saber.

La Universidad de Santiago de Compostela es una institución pública de enseñanza superior que, mediante la docencia, la investigación y la innovación se dedica a la generación, difusión y transferencia de conocimiento de forma socialmente responsable y con una conexión directa con las demandas y retos de la sociedad gallega.

La Universidad de Alicante tiene como misión esencial procurar una formación integral de sus estudiantes, potenciar una investigación científica puntera y promover su transferencia eficaz a nuestro entorno económico y social.

La Universidad de Jaén es una institución pública de educación superior que, a través de la docencia, la investigación y la transferencia del conocimiento, aúna la formación integral de una ciudadanía crítica, profesional y comprometida con la sostenibilidad de su entorno y adaptada a las necesidades de una sociedad cambiante mediante la generación, desarrollo y transmisión de la ciencia, de la técnica y de la cultura; presenta como rasgos singulares un personal comprometido con la orientación a la mejora continua en el desarrollo de todas sus actividades, un alto grado de compromiso social, y un equilibrio entre la proyección universal y la internacionalización transversal de toda su actividad ejerciendo, a su vez, el liderazgo en el desarrollo socioeconómico y cultural de la provincia de Jaén.

Que las partes arriba mencionadas, que no tienen vocación de mercado, tienen por objetivo perseguir el interés público mediante el desarrollo conjunto de las actividades científicas descritas en el anexo 1. En este sentido, la colaboración entre el BSC-CNS y las universidades en el desarrollo de un modelo de lenguaje de IA generará beneficios científicos y académicos para cada una de ellas. Este esfuerzo conjunto impulsará avances significativos en diversas disciplinas, ya que cada parte intercambiará conocimientos en áreas de inteligencia artificial, lingüística, matemáticas, ciencias de la computación y procesamiento de lenguaje natural. La integración y la puesta en común de estas metodologías, know-how y tecnologías de dichas áreas en los resultados contribuirá al avance del conocimiento y al beneficio del interés general.

Así pues, las universidades aportarán su capacidad para realizar investigación básica y exploración teórica, mientras que el BSC-CNS, con su experiencia en infraestructura y tecnologías avanzadas, llevará a cabo la coordinación e implementación eficiente y escalable de estos desarrollos. Esta sinergia acelerará la innovación y permitirá que estudiantes e investigadores apliquen sus conocimientos adquiridos en un entorno

práctico. El impacto de esta colaboración va más allá de la infraestructura científica y tecnológica, refuerza un modelo inclusivo de innovación y crea una red de talento interdisciplinario.

En resumen, la colaboración del BSC-CNS y las universidades para el desarrollo de un modelo de lenguaje de IA no solo eleva el estado del arte en inteligencia artificial, sino que también facilita el aprendizaje práctico, promueve la aplicación directa de los conocimientos adquiridos y genera beneficios tangibles en términos de formación y soluciones tecnológicas con impacto real en la sociedad.

Octavo.

Que el artículo 34.1 de la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación establece que los agentes públicos de financiación o ejecución del Sistema Español de Ciencia, Tecnología e Innovación, incluidas las Administraciones Públicas, las universidades públicas, los organismos públicos de investigación de la Administración General del Estado, los consorcios y fundaciones participadas por las administraciones públicas, los organismos de investigación de otras administraciones públicas, y los centros e instituciones del Sistema Nacional de Salud, podrán suscribir convenios sujetos al derecho administrativo. Podrán celebrar estos convenios los propios agentes públicos entre sí, o con agentes privados que realicen actividades de investigación científica y técnica, nacionales, supranacionales o extranjeros, para, entre otras actividades, la realización conjunta de proyectos y actuaciones de investigación científica, desarrollo e innovación, así como la creación o financiación de centros, institutos, consorcios o unidades de investigación, e infraestructuras científicas.

Noveno.

Que según lo dispuesto en el artículo 59.1 del Real Decreto-ley 36/2020, de 30 de diciembre, por el que se aprueban medidas urgentes para la modernización de la Administración Pública y para la ejecución del Plan de Recuperación, Transformación y Resiliencia, la tramitación de los convenios que celebre la Administración General del Estado, sus organismos públicos y entidades de derecho público, vinculados o dependientes, para la ejecución de los proyectos con cargo a fondos europeos previstos en el Plan de Recuperación, Transformación y Resiliencia, se regirá por lo previsto en el capítulo VI del título preliminar de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, con las especialidades recogidas en dicho artículo 59.

Que, en atención a todo lo expuesto, se estima necesaria y justificada la suscripción de un convenio que suponga la puesta en común, la distribución y utilización conjunta y/o coordinada de medios, servicios y recursos pertenecientes a cada una de las entidades participantes y su ordenación para un fin compartido bajo las premisas de interdisciplinariedad y eficiencia en la gestión de los recursos públicos.

En consecuencia y de conformidad con lo dispuesto en la Ley 40/2015, de 1 de octubre, en el Real Decreto-ley 36/2020, de 30 de diciembre, y en la Ley 14/2011, de 1 de junio, y con lo anteriormente expuesto, las partes acuerdan suscribir el presente convenio que se regirá por las siguientes:

## CLÁUSULAS

Primera. *Objeto del convenio.*

1.1 El objeto de este convenio consiste en recoger las condiciones de la colaboración de las distintas instituciones que participan en el proyecto para la realización de las actuaciones ligadas al desarrollo del Plan Nacional de Tecnologías del Lenguaje y del Ecosistema Empresarial de las Tecnologías del Lenguaje en el marco de

la Estrategia de Inteligencia Artificial de 2024 en el marco del Componente 16 del Plan de Recuperación, Transformación y Resiliencia (en adelante, Proyecto ALIA).

1.2 Este proyecto tiene como objetivo principal el desarrollo de actuaciones definidas en distintas iniciativas de las Palancas 3 y 6 de la Estrategia de Inteligencia Artificial 2024, actuaciones desarrolladas con mayor detalle en el anexo I de la Resolución de SEDIA del 19 de agosto de 2024 y que se listan a continuación:

- Desarrollo de un repositorio de corpus de elevada calidad en español y lenguas cooficiales.
- Actividades adicionales para el desarrollo de modelos fundacionales de lenguaje en español y lenguas cooficiales.
- Desarrollo y puesta a disposición pública de un marco de evaluación de modelos.
- Creación de una plataforma de entrenamiento de modelos de inteligencia artificial y tecnologías de lenguaje (IA/TL) en entorno seguro pre-comercial.
- Red territorial de equipos de colaboración técnica en plataforma de entrenamiento de modelos de inteligencia artificial y tecnologías de lenguaje (IA/TL) y acceso al entorno seguro pre-comercial.

1.3 En el anexo 1 del convenio se recoge la distribución de fondos entre las partes firmantes de este convenio y en el anexo 2 se recoge la memoria técnica, así como la participación de las partes en las distintas tareas científico-técnicas.

1.4 BSC-CNS es la «entidad beneficiaria y coordinadora de este proyecto», mientras que las demás entidades públicas Universidad del País Vasco (UPV/EHU), Universidad de Santiago de Compostela (USC), Universidad de Alicante (UA), Universidad de Jaén (UJA) colaboran en el citado convenio como «entidades participantes». Las Partes declaran que conocen el contenido de la Resolución de la Secretaría de Estado de Digitalización e Inteligencia Artificial, de fecha 19 de agosto de 2024, (en adelante, la «Resolución») y se comprometen a cumplir con los términos establecidos en la misma.

1.5 La entidad beneficiaria y las entidades participantes manifiestan no estar incurso en ninguna de las circunstancias recogidas en el artículo 13 ap. 2 de la Ley 38/2003, General de Subvenciones.

1.6 El objeto de este convenio no reúne el carácter de prestación contractual de la Ley de Contratos del Sector Público. En este sentido, no estamos ante una contraprestación de obligaciones, en la que una parte se obliga frente a la otra, a realizar una prestación contractual a cambio de un precio. En el presente convenio no existe una contraposición de intereses, ni beneficio económico, ni enriquecimiento patrimonial. Las partes persiguen objetivos ajenos al mercado y su participación tiene como fin la consecución de un fin común y la satisfacción del interés general. Es una colaboración científica entre entidades del sector público.

Segunda. *Obligaciones de las partes.*

2.1 Las partes firmantes de este convenio se comprometen a:

2.1.1 Desarrollar las actividades científico-técnicas descritas en la memoria técnica recogida en el anexo 2. Estas actividades se realizarán con cargo a los fondos transferidos por la resolución de la SEDIA al BSC-CNS y desde el BSC-CNS, en su función de coordinador, al resto de partes de este convenio según el reparto establecido en el anexo 1. Con arreglo a la citada resolución, el BSC-CNS actúa como entidad instrumental y el Ministerio de Transformación Digital y de la Función Pública, anteriormente denominado como el Ministerio de Asuntos Económicos y Transformación Digital, como entidad ejecutora.

2.1.2 Llevar a cabo las actuaciones objeto de la resolución en el periodo indicado en la resolución, esto es, en el primer trimestre de 2026.

2.1.3 A realizar las acciones que se les asigne para que la entidad ejecutora pueda desarrollar y aplicar las medidas contenidas en el Plan de Recuperación y, en concreto, a cumplir con los hitos y objetivos asignados en la forma y plazo previstos. Asimismo, las partes facilitarán a la entidad ejecutora toda la información de seguimiento que ésta precise, y en particular proporcionarán la información que tiene que suministrarse al sistema de información de gestión y seguimiento del Plan de Recuperación, Transformación y Resiliencia, conforme a lo dispuesto en la Orden HFP/1031/2021, de 29 de septiembre, y según se establezca en el manual de uso y funcionamiento del sistema de información, Manual de usuario CoFFEE-MRR.

2.1.4 A velar por los principios transversales del Mecanismo de Recuperación y Resiliencia de acuerdo con el Reglamento (UE) 2021/241 del Parlamento Europeo y del Consejo de 12 de febrero de 2021 y con los principios de gestión específicos del Plan de Recuperación, Transformación y Resiliencia, previstos en el artículo 2 de la Orden HFP/1030/2021, de 29 de septiembre y cumplir las siguientes obligaciones:

1) Disponer de un Plan de medidas antifraude.  
2) Cumplir con las normas de etiquetado verde y etiquetado digital de acuerdo con el Plan de Recuperación, Transformación y Resiliencia y el Mecanismo de Recuperación y Resiliencia, establecido por el Reglamento (UE) 2021/241 del Parlamento Europeo y del Consejo, de 12 de febrero de 2021. En este caso, al tratarse de una actuación de la medida C16.R1 contribuye en un 100 % a la consecución del objetivo digital (etiqueta 09.bis) y en un 0 % al objetivo de transición verde.

3) Realizar un análisis de riesgos en relación con posibles impactos negativos significativos en el medioambiente (do no significant harm, DNSH), así como el seguimiento y verificación de los resultados respecto a la evaluación inicial. Para garantizar este cumplimiento, se deberán cumplir las siguientes condiciones:

- Mitigación del cambio climático.
- Adaptación al cambio climático.
- Uso sostenible y protección de los recursos hídricos y marinos.
- Transición a una economía circular.
- Protección y restauración de la biodiversidad y los ecosistemas.

Por lo tanto, para realizar el análisis de riesgos conforme a las condiciones mencionadas anteriormente y se deberá consultar el apartado 10 de la resolución.

2.1.5 Firmar los modelos del anexo IV de la Orden HFP/1030/2021, de 29 de septiembre:

- Modelo de declaración de ausencia de conflicto de intereses (DACI).
- Modelo de declaración de cesión y tratamiento de datos en relación con la ejecución de actuaciones del plan de recuperación, transformación y resiliencia (PRTR).
- Compromiso de cumplimiento de principios transversales mediante la firma del Modelo declaración de compromiso en relación con la ejecución de actuaciones del plan de recuperación, transformación y resiliencia (PRTR).

2.1.6 En caso de subcontratar parte o toda la actividad financiada con cargo al Mecanismo de Recuperación y Resiliencia, la entidad beneficiaria, las entidades participantes y los subcontratistas, tendrán que cumplir con las obligaciones del apartado 2.1.4 de este convenio y con las relativas al Reglamento del Mecanismo de Recuperación y Resiliencia: etiquetado digital y verde, principio de no causar daño significativo, evitar conflictos de intereses, fraude, corrupción, comunicación, transferencia de datos y publicación obligatoria, cuando corresponda, en la Plataforma de Contratos del Sector Público («PLACSP») o en la herramienta de contratación aplicable en cada caso.

2.1.7 Llevar un inventario de los bienes, materiales o inmateriales adquiridos con cargo a las aportaciones previstas. El inventario deberá registrar, para cada elemento que se incluya en los apartados correspondientes, información sobre cada uno de los

gastos imputados, concepto, importes abonados, fechas de los gastos y de los pagos, y denominación y NIF o equivalente de los proveedores o destinatarios, de acuerdo con lo recogido en el artículo 22, letra d) del Reglamento (UE) 2021/241.

2.1.8 Cumplir con las obligaciones en materia de información, comunicación y publicidad contempladas en el artículo 34 del Reglamento (UE) 2021/241, relativo al Mecanismo de Recuperación y Resiliencia, sobre la financiación europea en las medidas incluidas en el Plan de Recuperación, Transformación y Resiliencia.

2.1.9 Remitir con la periodicidad, procedimiento y formato, según se establece en los Órdenes HFP 1030/2021 y HFP 1031/2021, y con referencia a las fechas contables que establezca la Autoridad Responsable del Mecanismo de Recuperación y Resiliencia (MRR), información sobre el seguimiento de la ejecución contable de los gastos financiados con estos fondos. La información se emitirá de conformidad con lo dispuesto en el artículo 46.2 del Real Decreto-ley 36/2020, de 30 de diciembre. Las partes remitirán esta información al BSC-CNS con tiempo suficiente para que, en su calidad de entidad instrumental del Proyecto ALIA, a efectos de la ejecución de los fondos MRR, el BSC-CNS pueda tener la información exigida completa para que pueda enviársela a la entidad ejecutora en los plazos establecidos.

2.1.10 Cumplir los requisitos de pista de auditoría: obligación de guardar la trazabilidad de cada una de las actuaciones y cada uno de los perceptores finales de los fondos, en los términos definidos por el artículo 22.2.d) del Reglamento (UE) 2021/241 y las Órdenes Ministeriales HFP/1031/2021 y HFP/1030/2021, de 29 de septiembre. Con la finalidad de dar adecuado cumplimiento a esta obligación se deberá incluir la información del artículo 8.1. de la Orden HFP/1030/2021.

2.1.11 A conservar la documentación vinculada a la actividad financiada por el MRR de conformidad con el artículo 132 del Reglamento Financiero, así como de disponer de un sistema que permita calcular y seguir el cumplimiento de objetivos y medición de indicadores.

2.1.12 A someterse a los controles preceptivos por las autoridades de control que se establezcan, entre otras, el órgano gestor y responsable del seguimiento de las actuaciones financiadas, al control financiero de la Intervención General de la Administración del Estado y al control fiscalizador del Tribunal de Cuentas, así como a los controles señalados en el Reglamento (UE) 2021/241 del Parlamento Europeo y del Consejo de 12 de febrero de 2021. Las partes deberán también someterse a las actuaciones de control de las instituciones comunitarias, en virtud del artículo 22.2.e) del Reglamento (UE) 2021/241 del Parlamento Europeo y del Consejo de 12 de febrero, cuando sea preciso.

2.1.13 A prestar colaboración y proporcionar cualquier documentación o información que se estime necesaria, así como a facilitar las inspecciones y comprobaciones que, en su caso, se efectúen por parte de las autoridades nacionales o de la Unión. La oposición a la realización de estas comprobaciones e inspecciones podrá constituir causa de reintegro de la ayuda, sin perjuicio de la tramitación del correspondiente procedimiento sancionador. Cualquier irregularidad que afecte al control desarrollado por las autoridades nacionales o las de la Unión que impida el retorno de fondos al Tesoro, implicará el reintegro por las partes del importe del retorno no percibido en relación con las actuaciones responsabilidad de las partes.

2.1.14 A tomar las medidas necesarias para garantizar el cumplimiento de lo dispuesto en el artículo 3 de la Orden HFP/55/2023, de 24 de enero, relativa al análisis sistemático del riesgo de conflicto de interés en los procedimientos que ejecutan el Plan de Recuperación, Transformación y Resiliencia. Asimismo, las entidades participantes deberán realizar dicho análisis conforme al artículo 3 de la citada orden y en observancia a los requerimientos de prevención, detección y corrección del conflicto de interés que la Comisión Europea exige a los Estados Miembros beneficiarios del MRR sobre la base del artículo 22 del Reglamento 2021/241 del Parlamento Europeo y del Consejo de 12 de febrero de 2021, por el que se establece el Mecanismo de Recuperación y Resiliencia y de las normas financieras aplicables al presupuesto de la Unión Europea. En este

sentido, todas las entidades ejecutoras e instrumentales encargadas de la ejecución de las obligaciones adquiridas en este convenio deberán realizar el análisis ex ante del riesgo de conflicto de interés sobre sus instrumentos jurídicos de ejecución (generalmente en contratos).

2.2 Además, el BSC-CNS se compromete a:

2.2.1 De acuerdo con la resolución de la SEDIA, el BSC-CNS como entidad instrumental y bajo la responsabilidad y dirección del Ministerio de Transformación Digital y de la Función Pública se compromete a actuar como coordinador de la participación de los grupos de investigación de las instituciones españolas en el marco de este convenio. Además, BSC-CNS como receptor de los fondos del Plan de Recuperación, Transformación y Resiliencia por parte de la SEDIA se compromete a transferir a las otras partes firmantes de este convenio los fondos correspondientes a la financiación de las actividades a realizar por las mismas según de describe en la cláusula Sexta de este convenio.

2.2.2 Las cantidades a transferir por el BSC-CNS al resto de partes firmantes de este convenio son las provenientes de la aportación realizada por la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA) de la partida presupuestaria 33.50.46PR. 74054 que se imputarán en el proyecto pn072900 Desarrollo de Modelos ALIA del presupuesto anual del BSC-CNS.

*Tercera. Compatibilidad del régimen de ayudas de Estado y prevención de la doble financiación.*

3.1 La aportación dineraria mencionada en la cláusula sexta será compatible con otras subvenciones, ayudas, ingresos o recursos de una finalidad similar, siempre y cuando se garantice el cumplimiento con el artículo 9 del Reglamento (UE) 2021/241 del Parlamento Europeo y del Consejo de 12 de febrero de 2021.

3.2 En todo caso, las partes deberán comunicar al BSC-CNS y éste a la SEDIA, la obtención de otra subvención, ayuda, ingreso o recurso que financie las actividades subvencionadas tan pronto esté en conocimiento de dicha obtención, y en todo caso, con anterioridad a la justificación de la aplicación dada a los fondos percibidos.

3.3 El receptor final de los fondos tiene la obligación de informar sobre cualesquiera otros fondos, no solo europeos, que hayan contribuido a la financiación de los mismos costes.

*Cuarta. Uso de los fondos a cargo del Mecanismo de Recuperación y Resiliencia.*

4.1 El uso de los fondos del MRR para la cofinanciación de actividades y proyectos estará sujeto a los siguientes principios: el marco temporal de las actuaciones financiadas por los fondos del MRR, de acuerdo con el artículo 17.2 y 18.4 (i) del reglamento del MRR, y los principios de adicionalidad y complementariedad a los que se refiere el artículo 9 del citado reglamento y la sección 2 de la parte 3 de la guía de Estados Miembros publicada el 22 de enero de 2021 (SWD (2021)12 part 1/2).

4.2 No podrán financiarse con cargo a este instrumento gastos ordinarios y recurrentes. Todos los gastos deberán responder de manera indubitada a la naturaleza de la actividad seleccionada para su financiación por el Mecanismo de Recuperación y Resiliencia que, en todo caso y circunstancias, deberán contribuir a los hitos y objetivos de la medida. Tampoco podrá financiarse el Impuesto sobre el Valor Añadido o impuesto equivalente.

4.3 En ningún caso, los fondos del MRR o cualquier otro fondo europeo podrán utilizarse para financiar dos veces el mismo coste.

## Quinta. *Verificación y justificación de los fondos.*

5.1 Las actuaciones que se financien deben haber concluido materialmente en el primer trimestre de 2026.

5.2 La verificación de la actividad realizada consistirá en la comprobación de que la financiación del MRR empleada se ha utilizado para los fines previstos y que se ha gestionado de conformidad con todas las normas aplicables y, en particular, si se han cumplido los hitos 251 y 253 recogidos en el anexo a la Decisión de Ejecución del Consejo (*Council Implementing Decision* o CID) vinculados al Componente 16, Reforma 1. Concretamente, los fondos recogidos en este convenio contribuyen al cumplimiento del hito 251 «Apoyo a proyectos en inteligencia artificial», que requiere haber comprometido, antes del cuarto trimestre de 2023, al menos quinientos (500.000.000,00 €) millones de euros de presupuesto en las medidas recogidas como parte de la Estrategia Nacional de Inteligencia Artificial (ENIA), y entre las que se encuentra el Plan Nacional de Tecnologías del Lenguaje, en el que se circunscribe esta actuación, y del hito 253 «Finalización de proyectos en inteligencia artificial», que establece el compromiso de finalización de los trabajos asociados a las medidas de la Estrategia Nacional de Inteligencia Artificial (ENIA) en el primer trimestre del 2026.

5.3 Con objeto de justificar la ejecución presupuestaria, las partes deberán remitir al BSC-CNS como máximo en los siguientes diez (10) días a finalizar cada trimestre de ejecución del proyecto, un informe justificativo de la plena aplicación de fondos para la actuación regulada por este convenio. Por su parte, el BSC-CNS deberá remitir, el correspondiente informe justificativo a la SEDIA, como máximo en los siguientes treinta (30) días desde la fecha de finalización del plazo marcado en el apartado 2.1.2.

5.4 La justificación económica de gastos se realizará mediante cuenta justificativa, con aportación de justificantes de gasto, de acuerdo con el artículo 72 del Real Decreto 887/2006, de 21 de julio, por el que se aprueba el Reglamento de la Ley 38/2003, de 17 de noviembre, General de Subvenciones. Todas las partes deberán acreditar los costes reales incurridos, sin incluir el IVA u otros impuestos análogos.

## Sexta. *Financiación.*

6.1 El presupuesto total que se establece para esta colaboración científico-técnica es de siete millones seiscientos sesenta y cuatro mil seiscientos euros (7.664.600,00 €) según se recoge en el anexo 1 de este convenio.

6.2 El BSC-CNS, como parte integrante de este convenio y entidad beneficiaria de la ayuda, transferirá las asignaciones correspondientes a las entidades participantes conforme a la siguiente distribución:

1) El 70% de la aportación respectiva a cada entidad participante en un plazo máximo de cuarenta y cinco (45) días tras la inscripción del presente convenio en REOICO.

2) El restante 30% de la aportación respectiva a cada entidad participante en el mes de noviembre del año 2025. Esta cantidad se abonará siempre y cuando la entidad participante haya cumplido con las tareas asignadas en el anexo 2 y haya entregado la información técnica-económica solicitada por la entidad beneficiaria para hacer frente a las justificaciones anuales y final del proyecto ante el organismo concedente.

6.3 Las cantidades indicadas se harán efectivas mediante transferencia bancaria realizada por la entidad beneficiaria en la cuenta abierta a nombre de las entidades participantes que éstas indiquen y certifiquen (indicando en el concepto de la transferencia el acrónimo o código identificativo del proyecto) y previa conformidad del ordenante del pago con el cumplimiento hasta ese momento del Plan de trabajo.

6.4 Los importes a abonar no incluirán IVA puesto que dichas cantidades no son en concepto de contraprestación directa y equivalente a las actividades que se realizarán. Estas cantidades son compensación de los costes generados en la ejecución del

proyecto y carecen de onerosidad, al tener como finalidad la compensación de costes y coadyuvar a la consecución de un interés común y satisfacer el interés general, tal y como se describe en la cláusula 1 de este convenio. En este sentido, la ausencia de onerosidad del objeto, fines y actuaciones es clara y patente a la luz de los preceptos contenidos en el artículo 4.1 (sujeción de las entregas de bienes y prestaciones al impuesto) de la Ley 37/1992, de 28 de diciembre, del Impuesto sobre el Valor Añadido.

6.5 En ningún caso, las entidades participantes podrán reclamar más cantidad de la establecida en el presupuesto acordado en el anexo 1.

#### Séptima. *Comisión de Seguimiento.*

7.1 Para garantizar la correcta ejecución y el seguimiento de lo establecido en el presente convenio se constituirá una Comisión de Seguimiento que adecuará su funcionamiento a lo previsto en los artículos 15 y siguientes de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (en adelante LRJSP).

7.2 La Comisión de Seguimiento será mixta y paritaria estará integrada por un o una representante de cada una de las partes firmantes de este convenio.

7.3 Las partes designarán nominalmente a las personas titulares que les representen en la Comisión de Seguimiento y se lo comunicarán al BSC-CNS en el plazo de un (1) mes desde la entrada en vigor del presente convenio y designarán igualmente a las personas que sustituyan a los titulares en caso de ausencia de estos.

7.4 Las funciones de la Comisión de Seguimiento serán las siguientes:

- 1) Realizar el seguimiento y la vigilancia de las actuaciones recogidas en este convenio.
- 2) Resolver los problemas de interpretación y cumplimiento que pudieran surgir en la aplicación e interpretación de las cláusulas y en el cumplimiento de este convenio.
- 3) Proponer modificaciones y la prórroga.
- 4) Establecer las definiciones, contenidos y entregables de las tareas recogidas en el anexo 2, de acuerdo con el criterio conjunto de las instituciones involucradas en cada una de ellas. En caso de discrepancia, se tendrá en cuenta el criterio de la institución coordinadora en cada caso.

7.5 La Comisión de Seguimiento se reunirá de forma ordinaria, al menos una vez al trimestre (ya sea de forma presencial o virtualmente) y, de forma extraordinaria, cuando lo solicite una de las partes.

1) La Presidencia de dicha Comisión la asumirá el BSC-CNS, que tendrá funciones ejecutivas y decisorias sobre los asuntos que se debatan en el seno de dicha Comisión, en la medida que es el beneficiario de esta ayuda. El BSC-CNS informará a las entidades participantes de las decisiones técnicas, económicas y jurídicas más relevantes y las debatirá en el seno de dicha comisión a los efectos de recabar la opinión y parecer de las entidades participantes.

2) En lo no previsto en el presente convenio, el régimen de organización y funcionamiento de la Comisión de Seguimiento será, de forma supletoria, el previsto para los órganos colegiados en la sección 3.<sup>a</sup> del capítulo II del título preliminar de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público.

#### Octava. *Comunicación y publicidad de la financiación europea.*

8.1 En la documentación administrativa para la gestión de expedientes correspondientes a las actuaciones recogidas en este convenio se deberá incluir una mención a que el expediente está financiado por el Mecanismo de Recuperación y Resiliencia.

8.2 Las partes firmantes de este convenio deberán cumplir con las obligaciones de información y publicidad que las autoridades competentes establezcan relativas a la financiación del Mecanismo de Recuperación y Resiliencia.

8.3 En las medidas de información y comunicación del convenio, así como en las actuaciones que desarrollen las inversiones recogidas en el mismo (carteles informativos, placas, publicaciones impresas y electrónicas, material audiovisual, páginas web, anuncios e inserciones en prensa, certificados, etc.), se deberá cumplir con lo establecido en la Orden HFP/1030/2021, de 29 de septiembre. En todo caso, se deberán incluir los siguientes logos:

- El emblema de la Unión;
- Junto con el emblema de la Unión, se incluirá el texto «Financiado por la Unión Europea –NextGenerationEU–»;
- Se usará también el logo del Plan de Recuperación, Transformación y Resiliencia disponible en el enlace: <https://planderecuperacion.gob.es/identidad-visual>.
- Se tendrán en cuenta las normas gráficas del emblema de la Unión y los colores normalizados establecidos en el anexo 2 del Reglamento de Ejecución 821/2014. También se puede consultar la siguiente página web: <http://publications.europa.eu/code/es/es-5000100.htm> y descargar distintos ejemplos del emblema en: [https://europa.eu/european-union/about-eu/symbols/flag\\_es#download](https://europa.eu/european-union/about-eu/symbols/flag_es#download).

8.4 Los carteles informativos y placas deberán colocarse en un lugar bien visible y de acceso al público. Toda la publicidad deberá cumplir con la normativa vigente en materia de fondos europeos.

8.5 Asimismo, en las publicaciones, infraestructuras y otros resultados a los que puedan dar lugar las actividades financiadas, deberá también mencionarse que las mismas han sido financiadas por el Ministerio para la Transformación Digital y de la Función Pública y coordinadas por el Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS).

#### Novena. *Protección de datos.*

9.1 De conformidad con el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016 (Reglamento General de Protección de Datos o RGPD), aplicable desde el 25 de mayo de 2018, las partes hacen constar de manera expresa que se abstendrán de cualquier tipo de tratamiento de los datos personales de que dispongan como consecuencia de este convenio, exceptuando aquel que sea estrictamente necesario para las finalidades del mismo. En este sentido se comprometen, especialmente, a no ceder a terceros los datos mencionados o los archivos que los contengan, así como a guardar estricta confidencialidad sobre los mismos. Asimismo, las partes quedan sometidas a lo dispuesto en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, y demás normativa aplicable en materia de protección de datos.

9.2 El tratamiento de datos de los firmantes y personas de contacto de cada entidad se realizará en calidad de responsable por cada una de las partes, y de acuerdo a las políticas de privacidad accesible en la página web de cada entidad.

9.3 Sin perjuicio de lo anterior, en el caso que sea necesario para la finalidad y la ejecución del Proyecto transmitir o almacenar datos personales en el BSC-CNS – distintos a los relativos a las personas de contacto y de ejecución regulados en los párrafos anteriores–, éstos deberán venir anonimizados en origen por la entidad responsable de los datos. Si los datos no pudieran ser anonimizables, éstos no se transmitirán hasta que las partes correspondientes formalicen la herramienta jurídica aplicable al caso concreto. En todo caso, la entidad que vaya a ceder los datos ostentará la condición de responsable de tratamiento y será el encargado de redactar el acuerdo de tratamiento y el protocolo técnico de acceso, tratamiento, almacenamiento de los datos seudonimizados con el receptor de los datos. En el caso de datos personales anonimizados, el responsable de los datos deberá redactar el protocolo técnico de acceso, tratamiento, almacenamiento de datos anonimizados para garantizar que se han

aplicado las medidas técnicas y organizativas necesarias para la anonimización irreversible de esos datos.

Se indican a continuación los responsables del tratamiento de datos de cada entidad:

BSC-CNS: [dpo@bsc.es](mailto:dpo@bsc.es)

Universidad del País Vasco, UPV/EHU:

<https://www.ehu.es/es/web/idazkaritza-nagusia/babestu-harremanetarako>

Contacto del delegado de protección de datos: [dpd@ehu.es](mailto:dpd@ehu.es)

Universidad de Alicante, UA, CENID:

Gerencia de la Universidad de Alicante.

[Gerencia@ua.es](mailto:Gerencia@ua.es)

Tel.: 96.590.34.00 Fax: 96.590.34.64.

Apartado de correos, 99.

Email: [dpd@ua.es](mailto:dpd@ua.es)

Universidad de Santiago de Compostela:

El Delegado de Protección de Datos Marcos Almeida Cerredá, [dpd@usc.es](mailto:dpd@usc.es)

La responsabilidad genérica del tratamiento de datos recae en la Universidad de Santiago de Compostela, Rectorado de la USC, Praza do Obradoiro s/n, 15782 Santiago de Compostela.

Tel: 881811000.

Contacto a través de: <https://www.usc.gal/es/contacto>.

Universidad de Jaén:

Delegado de Protección de Datos (DPO).

Telefónica, SAU.

Email: [dpo@ujaen.es](mailto:dpo@ujaen.es) Tel: 953 212121.

## Décima. *Confidencialidad.*

10.1 Toda información intercambiada entre las partes en el marco de este convenio tendrá la consideración de reservada y confidencial, y no podrá ser divulgada a terceros sin el consentimiento previo por escrito de las partes afectadas ni utilizada para tareas que no estén específicamente acordadas.

10.2 La confidencialidad no será de aplicación cuando:

- 1) La parte receptora pueda demostrar que conocía previamente la información recibida.
- 2) La información recibida sea o pase a ser de dominio público.
- 3) La parte receptora obtenga autorización previa y por escrito para su revelación, o la información sea requerida judicialmente o por orden de autoridad administrativa o gubernamental competente.
- 4) Sea recibida de manera lícita de un tercero.
- 5) Se haya generado de forma independiente y de buena fe por los miembros de su institución sin ninguna conexión con la información confidencial.

10.3 Las partes adoptarán cuantas medidas sean necesarias para que su personal cumpla con lo dispuesto en esta cláusula. Las obligaciones en materia de confidencialidad subsistirán hasta los cinco años después de la pérdida de vigencia del presente convenio.

Undécima. *Publicación de resultados.*

11.1 En caso de que existan resultados publicables, las partes aplicarán las normas de publicación de resultados recogidas en el modelo DESCA de Consortium Agreement de la Unión Europea para Horizonte Europa.

1. **Publicación conjunta.** Las partes podrán cooperar para publicar los resultados de la investigación en un artículo conjunto y la revisión se hará de forma conjunta por las partes. En tal caso, la autoría de las publicaciones se basará en las normas recogidas en el modelo DESCA. Las partes deberán reconocer todas las contribuciones científicas o de otro tipo en cualquier publicación o presentación relacionada con la investigación realizada en virtud del presente convenio.

2. **Publicación por separado.** Durante el proyecto y por un período de un (1) año tras la finalización del mismo, cada parte podrá hacer difusión de los propios resultados, incluyendo, pero no limitándose a publicaciones y presentaciones, con sujeción a las siguientes disposiciones:

1) Cualquier publicación prevista se notificará previamente a las otras partes al menos cuarenta y cinco (45) días naturales antes de la publicación. Cualquier objeción a la publicación deberá notificarse por escrito al coordinador y a la parte o partes que proponen la difusión en un plazo de treinta (30) días naturales de la recepción de la notificación. Si no se presenta ninguna objeción en el plazo antes indicado, se autoriza la publicación.

2) Una objeción estará justificada si: (a) afecta negativamente a la protección de los resultados o de los conocimientos previos de la parte que presenta la objeción, o (b) perjudica significativamente los intereses legítimos de la parte que presenta la objeción en relación con sus resultados o conocimientos previos, o (c) la publicación propuesta incluye información confidencial de la parte que presenta la objeción.

3) La objeción debe incluir una solicitud precisa de las modificaciones necesarias y debe estar motivada.

4) Si se presenta una objeción, las partes implicadas deberán discutir la forma adecuada de superar los motivos de la objeción (por ejemplo, modificando la publicación prevista y/o protegiendo la información). La parte objetora no deberá continuar la oposición de forma injustificada si se han alcanzado medidas apropiadas tras la discusión.

5) La parte objetora puede solicitar un aplazamiento de la publicación de no más de noventa (90) días naturales a partir del momento en que plantee dicha objeción. Transcurridos los noventa (90) días naturales, la publicación será permitida, siempre que se hayan tenido en cuenta las objeciones de la parte objetora.

6) Una parte no incluirá en ninguna actividad de difusión los resultados o conocimientos previos de otra Parte sin obtener la previa aprobación por escrito de la parte propietaria, a menos que ya estén ya publicados.

11.2 Titularidad de los Conocimientos Previos.

11.2.1 Cada parte seguirá siendo propietaria de los Conocimientos Previos aportados al proyecto. En virtud del presente convenio no se entienden cedidos a las otras partes ninguno de los Conocimientos Previos aportados al proyecto. Se entiende por «Conocimientos Previos» cualquier dato, información, conocimiento técnico, metodología, modelo, código, software o cualquier otro activo tangible o intangible, independientemente de su formato o naturaleza, incluidos todos los derechos asociados, como derechos de propiedad intelectual e industrial, que cumpla con los siguientes criterios: a) Pertenezca a alguna de las Partes con anterioridad a la entrada en vigor del convenio, y b) sea necesario para la ejecución del proyecto.

11.2.2 Cada una de las Partes deberá aportar los Conocimientos Previos de la manera más permisiva posible, es decir, con licencias opensource permisivas y Creative

Commons Comerciales. Las Partes podrán utilizar dicho Conocimiento Previo bajo las condiciones de las licencias Open Source permisivas y Creative Commons comerciales.

11.2.3 No se podrán aportar Conocimientos Previos sujetos a licencias open source restrictivas o de carácter viral, como la GPLv3, si dichos Conocimientos generan dependencias con los Resultados y/o impiden la cesión de los Resultados de explotación a la SEDIA. Asimismo, no se podrá acoplar, ni integrar software sujeto a licencias virales o restrictivas en los Resultados.

11.2.4 Las partes acuerdan que, de aportarse tecnologías, conocimientos, metodologías, modelos, bases de datos, código y/u otros elementos no pertenecientes a alguna de las partes, sea ésta una tecnología al que una parte tenga acceso bajo una licencia de uso con un tercero no vinculado con el convenio u otra casuística, deberán respetar los derechos de propiedad intelectual e industrial de terceros. BSC-CNS y las demás partes no se responsabilizan de posibles infracciones de derechos de propiedad intelectual e industrial de terceros provocados por la aportación de dicho elemento por la respectiva parte.

11.2.5 Las Partes declaran y garantizan que los Conocimientos Previos que aporten están libres de cargas o restricciones que puedan limitar su uso en el proyecto o la explotación de los resultados.

Duodécima. *Propiedad de los resultados.*

12.1 Los resultados generados en el marco del presente convenio serán propiedad de las Partes que los desarrollen o co-desarrollen, conforme a su contribución respectiva. No obstante, las Partes ceden en exclusiva a la SEDIA los derechos de explotación de los Resultados propietarios, mundialmente y por tiempo indefinido, siempre que dichos resultados no infrinjan derechos de terceros o estén afectados por licencias opensource, en cuyo caso, aplicarán las reglas de las licencias opensource. Asimismo, tendrán la autorización del uso del correspondiente resultado, los entes, organismos y entidades pertenecientes al sector público, a semejanza de lo que se establece en el artículo 308 de la Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público, por la que se transponen al ordenamiento jurídico español las Directivas del Parlamento Europeo y del Consejo 2014/23/UE y 2014/24/UE, de 26 de febrero de 2014.

12.2 Cuando los Resultados hayan sido generados conjuntamente por varias Partes, éstas compartirán la titularidad en proporción a sus respectivas contribuciones. Las Partes co-titulares colaborarán para establecer los mecanismos necesarios para ceder los Resultados de explotación a la SEDIA, garantizando el cumplimiento de las obligaciones derivadas del presente convenio y los requisitos establecidos por el organismo financiador para la cesión de los derechos de explotación.

12.3 Para garantizar la cesión en exclusiva de los Resultados a la SEDIA, las Partes acuerdan generar Resultados no sujetos a:

1. Licencias restrictivas que impongan obligaciones adicionales o incompatibilidades con la explotación del proyecto.

2. Dependencias de tecnologías o elementos sujetos a licencias virales o no comerciales que limiten la transferencia de derechos al organismo financiador. Cada Parte garantiza que los Resultados que aporte estarán libres de restricciones y/o limitaciones que impidan su uso y/o explotación conforme a los términos del presente convenio.

12.4 La SEDIA será responsable de determinar el modelo de licenciamiento aplicable a los Resultados. Las Partes se reservan para sí mismas una licencia no exclusiva, irrevocable, mundial y de duración indefinida sobre los Resultados para actividades académicas, investigación no comerciales.

12.5 Las entidades participantes garantizan al BSC-CNS que los Resultados que aporten estarán libres de restricciones o limitaciones que puedan impedir su uso o

explotación conforme a los términos del presente convenio y a lo dispuesto en la resolución correspondiente.

Decimotercera. *Vigencia, modificación y extinción del convenio.*

13.1 De conformidad con el artículo 48.8 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, el presente convenio se perfecciona con el consentimiento de las partes y resultará eficaz una vez inscrito en el Registro Estatal de Órganos e Instrumentos de Cooperación (REOICO) del sector público estatal, debiendo publicarse, a continuación, en el «Boletín Oficial del Estado».

13.2 Este convenio permanecerá en vigor desde la fecha de su firma y hasta el primer trimestre de 2026 y podrá ser prorrogado, de mutuo acuerdo entre las partes, por un período máximo de un (1) año mediante adenda al mismo, adoptada antes de la finalización del periodo inicial de vigencia, todo ello con los límites que establece la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. Entre las condiciones para su prórroga se encuentra la ampliación del plazo de ejecución de los fondos del Plan de Recuperación, Transformación y Resiliencia. Las cláusulas de este convenio podrán ser modificadas por acuerdo unánime entre las partes.

13.3 De acuerdo con lo previsto en el artículo 51.1 de la LRJSP, este convenio se extinguirá por el cumplimiento de las actuaciones que constituyen su objeto o por incurrir en alguna de las siguientes causas de resolución:

- 1) El transcurso del plazo de vigencia del convenio sin haberse acordado la prórroga del mismo.
- 2) Por acuerdo unánime de las partes.
- 3) Por incumplimiento de las condiciones y compromisos asumidos por alguna de las partes. En este caso, cualquiera de las partes podrá notificar a la parte incumplidora un requerimiento para que cumpla en un determinado plazo con las obligaciones o compromisos que se consideran incumplidos. Este requerimiento será comunicado a través la Comisión de Seguimiento. Si transcurrido el plazo indicado en el requerimiento persistiera el incumplimiento, la parte que lo dirigió notificará a las partes firmantes la concurrencia de la causa de resolución y se entenderá resuelto el convenio.
- 4) Por decisión judicial declaratoria de la nulidad del convenio.
- 5) Si se produjesen circunstancias que hicieran imposible o innecesaria la colaboración.
- 6) Por cualquier otra causa distinta de las anteriores prevista en el convenio o en otras leyes.

13.4 En caso de resolución del convenio, las partes quedan obligadas al cumplimiento de sus respectivos compromisos hasta la fecha en que ésta se produzca, y dará lugar a la liquidación del mismo con el objeto de determinar las obligaciones y compromisos de cada una de las partes en los términos establecidos en el artículo 52 de la Ley 40/2015.

13.5 Asimismo, de conformidad con el artículo 84 de la Ley 47/2003, la responsabilidad financiera derivados de la gestión de los fondos procedentes de la Unión Europea se someterá a lo previsto en la Ley 38/2003, de 17 de noviembre, y en la normativa comunitaria. Serán causas de reintegro total de la aportación dineraria:

- 1) El incumplimiento de los hitos y objetivos establecido para esta actuación, que afecten al cumplimiento de los hitos y objetivos establecidos para el Plan de Recuperación, Transformación y Resiliencia. De este modo, la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA) podrá exigir al BSC-CNS el reintegro total por anticipado, según lo dispuesto en la Orden HFP/1030/2021, de 29 de septiembre, a efectos de garantizar el cumplimiento de los hitos y objetivos del Plan de Recuperación, Transformación y Resiliencia.
- 2) La falta de presentación de la documentación justificativa.

3) La constatación, durante las actuaciones de comprobación, que los objetivos, servicios o gastos justificados no se corresponden con lo establecido en esta resolución o que estos gastos hubieran sido desarrollados con anterioridad a este convenio, lo que evidenciaría la no aplicación de los fondos aportados a los fines perseguidos en la resolución.

4) El incumplimiento de las obligaciones establecidas por la Orden HFP/1030/2021, de 29 de septiembre: a) el incumplimiento del principio de «no causar un perjuicio significativo al medio ambiente» (Do No Significant Harm-DNSH), b) el incumplimiento de las obligaciones de información y publicidad. c) el incumplimiento de las obligaciones en materia de etiquetado digital y verde y d) el incumplimiento de cualquiera de obligaciones establecidas en la citada orden.

13.6 Será causa de reintegro parcial de la aportación dineraria la existencia de fondos transferidos que no se hayan destinado a financiar la actividad descrita en el anexo 2 o que no estén convenientemente justificados de acuerdo con lo establecido en esta resolución, y deberán ser reintegrados al Tesoro. En caso de reintegro, la Parte afectada deberá reintegrar total o parcialmente los fondos recibidos por el BSC-CNS, como institución coordinadora, para que ésta a su vez los devuelva al órgano competente.

13.7 Los eventuales reintegros a que hubiera lugar por no realizarse el gasto o, en su caso, por no poder cumplir total o parcialmente con los objetivos previstos, se tramitarán conforme a lo establecido en el artículo 7 ap.7 de la Ley 31/2022, de 23 de diciembre, de Presupuestos Generales del Estado para el año 2023.

#### Decimocuarta. *Incumplimiento.*

14.1 En el supuesto que el BSC-CNS considere que: a) una entidad participante en el Proyecto incumple su actividad al calendario y/o tareas asignadas, b) que la calidad de las mismas no es la necesaria, c) que no se provee del soporte documental que justifique debidamente los costes o tareas técnicas correspondientes o d) que los gastos se apartan de los previstos en los anexos 1 y 2, podrá requerir mediante preaviso a la entidad participante que repare y ajuste su actividad a lo previsto y comprometido en los anexos 1 y 2 del este convenio.

14.2 En el supuesto que la entidad participante no rectifique, el BSC-CNS podrá poner término a la participación de dicha entidad participante, reclamar la devolución de los importes de la ayuda, los intereses de demora y los daños y perjuicios ocasionados. En todo caso, la responsabilidad por daños y perjuicios queda limitada al importe de la ayuda concedida.

#### Decimoquinta. *Régimen jurídico y resolución de controversias.*

15.1 Este convenio tiene naturaleza interadministrativa de acuerdo con el artículo 47 de la Ley 40/2015, de 1 de octubre, y se regirá por sus propios términos y condiciones y por el ordenamiento jurídico administrativo, con las especialidades previstas en el artículo 59 del Real Decreto-ley 36/2020, de 30 de diciembre, por el que se aprueban medidas urgentes para la modernización de la Administración Pública y para la ejecución del Plan de Recuperación, Transformación y Resiliencia.

15.2 Las cuestiones litigiosas que puedan surgir en la interpretación y cumplimiento de este convenio que no cuenten con acuerdo en el seno de la Comisión de Seguimiento, serán de conocimiento y competencia del orden jurisdiccional de lo contencioso-administrativo.

Y en prueba de conformidad, las partes firman este convenio, en la fecha indicada en la última de ellas, el 2 de junio de 2025.–Por el Barcelona Supercomputing Center, el Director, Mateo Valero Cortés.–Por la Universidad del País Vasco/Euskal Herriko Unibertsitatea, el Rector, José Ramón Bengoetxea Caballero.–Por la Universidad de Santiago de Compostela, el Rector, Antonio López Díaz.–Por la Universidad de Alicante,

la Rectora, Amparo Navarro Faure.–Por la Universidad de Jaén, el Rector, Nicolás Ruiz Reyes.

## ANEXO 1

### Presupuesto

#### *Presupuesto personal por entidad participante*

	BSC* – Euros	UPV-EHU – Euros	USC – Euros	CENID UA – Euros	UJA – Euros	Total – Euros
Personal.	1.440.000,00	840.000,00	840.000,00	840.000,00	840.000,00	4.800.000,00
Viajes.	38.500,00	36.000,00	36.000,00	36.000,00	36.000,00	182.500,00
Subcontratación.	180.000,00	400.000,00	400.000,00	400.000,00	400.000,00	1.780.000,00
Comunicación/Diseminación.	182.100,00	–	–	–	–	182.100,00
Costes Ind. de personal (15 %).	216.000,00	126.000,00	126.000,00	126.000,00	126.000,00	720.000,00
Total por entidad*.	2.056.600,00	1.402.000,00	1.402.000,00	1.402.000,00	1.402.000,00	7.664.600,00

## ANEXO 2

### Memoria técnica

Se muestra el detalle de las tareas y/o paquetes de trabajo del proyecto en el que deben cooperar las distintas partes. Este anexo 2 es conforme a la memoria técnica del Proyecto presentada por el BSC-CNS a la SEDIA.

#### *Objeto*

El proyecto ALIA ha sido financiado por la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA), Ministerio de Transformación Digital y de la Función Pública, entre las medidas urgentes para la modernización de la Administración Pública y para la ejecución del Plan de Recuperación, Transformación y Resiliencia, en el marco temporal definido por la Estrategia de Inteligencia Artificial 2024 y los hitos CID del Componente 16, Reforma 1.

La Estrategia de Inteligencia Artificial 2024 se articula en torno a tres ejes fundamentales: i) reforzar las palancas clave para el desarrollo de la Inteligencia Artificial, ii) facilitar la expansión de la Inteligencia Artificial en el sector público y privado fomentando la innovación y la ciberseguridad y iii) desarrollar una Inteligencia Artificial transparente, responsable y humanística. El objetivo del proyecto se centra en los dos primeros ejes. Concretamente, se persigue financiar actuaciones ligadas a la generación de modelos y corpus para desarrollar una infraestructura pública, así como ayudar a la expansión de esta infraestructura en el sector privado, singularmente en pymes, a través de la puesta a disposición de estas de una infraestructura segura de entrenamiento de modelos de propósito específico.

En un contexto actual, el desarrollo de la familia de modelos ALIA se hace necesario para mejorar la inclusión del español y las lenguas cooficiales en las futuras aplicaciones de inteligencia artificial, garantizando un mayor porcentaje de datos de entrenamiento en español y lenguas cooficiales, y por tanto una mayor soberanía tecnológica y transparencia. El desarrollo de esta familia de modelos, así como el de las infraestructuras de datos y de entrenamiento de modelos, es parte esencial de los ejes habilitadores para desarrollar la inteligencia artificial generativa en España, además de

favorecer su adopción por parte tanto de la Administración Pública como del Sector Privado.

Adicionalmente, se persigue el desarrollo y colaboración del ecosistema investigador de las tecnologías del lenguaje en español, para lo que se prevé la colaboración científica y técnica de otros centros de investigación (Universidad del País Vasco-HiTZ, Centro de Inteligencia Digital Alicante-CENID, Universidad de Santiago de Compostela y Universidad de Jaén) y de empresas especializadas, con el Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS).

### *Plan de proyecto*

Las distintas entidades participan en el proyecto cumpliendo con los objetivos y desarrollos marcados en el plan de que se detalla en adelante. El BSC-CNS participará en la coordinación y supervisión de las tareas llevadas a cabo en el proyecto y, por extensión, las detalladas en este convenio. Constituirá el enlace entre las instituciones firmantes de este convenio, informando de progresos que puedan ser de interés general, así como coordinando las actividades que involucren a todo el ecosistema (las presentes en este convenio). Además, asegurará la correcta realización de las tareas descritas en este convenio transfiriendo los fondos necesarios a estas instituciones.

Los objetivos del proyecto incluyen desarrollos a nivel técnico que se especifican en una serie de iniciativas:

- 1) Desarrollo de un repositorio de corpus de elevada calidad en español y lenguas cooficiales.
- 2) Desarrollo de modelos de propósito general (LLM).
- 3) Desarrollo y puesta a disposición pública de un marco de evaluación de modelos.
- 4) Actividades transversales y desarrollo de casos de uso piloto.
- 5) Red territorial de equipos de colaboración técnica en IA/TL y acceso al entorno seguro pre-comercial.

Para cada iniciativa se especifican los resultados esperados y el papel de cada institución.

Iniciativa 1). Desarrollo de un repositorio de corpus de elevada calidad en español y lenguas cooficiales.

– Actividad 1.1 Identificación, análisis y contacto con potenciales repositorios de contenidos en todas las modalidades, no restrictivos y susceptibles de formar parte del corpus general del proyecto ALIA para todas las lenguas de España BSC: Coordinará y hará el seguimiento de la actividad.

– UPV/EHU, USC, CENID y UJA: Contribuirán al entregable común.

Actividad 1.3 Operacionalización del acceso a corpus de texto para el entrenamiento de modelos en todas las lenguas de España y Portugal. Tareas de desarrollo de scripts/protocolos de acceso y descarga a las diferentes fuentes de datos, publicados en github y debidamente documentados.

– BSC: Coordinará y hará el seguimiento de la actividad.

– UPV/EHU, USC, CENID y UJA: Contribuirán al entregable común.

Actividad 1.5 Operacionalización del acceso a corpus de voz para el entrenamiento de modelos en todas las lenguas de España.

Se desarrollan los sistemas para la captura de los repositorios de datos de voz no restrictivos y susceptibles de formar parte del corpus general de ALIA en cada uno de los

idiomas, a partir de sus scripts/protocolos de acceso y descarga a las diferentes fuentes de datos (repositorios de datos de voz):

- BSC: castellano y catalán.
- UPV/EHU: castellano y euskera.
- USC: castellano y gallego.
- CENID: castellano y valenciano y para los dominios turístico y el de la propiedad intelectual.

Actividad 1.7 *Pipeline* de preproceso de datos de voz.

Se desarrollan herramientas de preproceso de voz con el objetivo de disponer de datos de entrenamiento, especialmente datos alineados texto/voz para:

- BSC: castellano y catalán.
- UPV/EHU: euskera.
- USC: gallego.

Actividad 1.8 Dataset de alineamiento. Los datos de alineamiento son la última fase después de los datos de pre-entrenamiento y de instrucción.

Corpus de alineamiento con anotaciones de preferencias sobre las respuestas para acercarlas a las preferencias humanas en:

- BSC: castellano y catalán.
- UPV/EHU: castellano y euskera.
- USC: castellano y gallego.
- CENID: castellano y valenciano.
- UJA: castellano.

Actividad 1.9 Generación de datos (sintéticos) para corpus paralelos (Traducción Automática) incluyendo adaptación a dominio.

Desarrollo de corpus paralelos con datos sintéticos en dominios concretos para entrenamiento de modelos de TA en:

- BSC: inglés, castellano, chino y catalán.
- UPV/EHU: inglés, castellano y euskera
- USC: inglés, castellano y gallego.
- CENID: inglés, castellano y valenciano
- UJA: inglés y castellano.

Actividad 1.10 Generación de corpus de dominio (legal, médico, científico...). Estos corpus sirven como base para el entrenamiento de modelos discriminativos de dominio contemplados en la resolución anterior (21 de diciembre de 2023) así como para los nuevos SLMs generativos de dominio contemplados en la actividad 2.5.

Generación y proceso de datos de dominio en cada una de las lenguas:

- BSC: corpus de dominio legal, médico, y científico en castellano.
- UPV/EHU: corpus de dominio de noticias, legal/administrativo y científico procesado y con sus correspondientes metadatos
- USC: corpus de noticias, cultura y legal/administrativo.
- CENID: corpus de dominio turístico y de la propiedad intelectual procesado y con sus correspondientes metadatos
- UJA: dominio legal-administrativo, patrimonio-cultural y biosanitario procesado y con sus correspondientes metadatos.

Iniciativa 2). Desarrollo de modelos de propósito general (LLM).

Actividad 2.2 Modelo autosupervisado de voz tipo wavLM (o equivalente) Ibérico en mínimo 4 idiomas. Entrenado en MN5 por el BSC con la aportación de datos de UPV/EHU, USC y CENID en las respectivas lenguas.

Creación y desarrollo del modelo autosupervisado de reconocimiento del habla, tipo wavLM (o equivalente), Ibérico.

- BSC: castellano y catalán.
- UPV/EHU: euskera.
- USC: gallego.
- CENID: valenciano.

Actividad 2.3 Un encoder fonético ibérico (para tareas de síntesis de voz).

Diseño y desarrollo de un encoder fonético ibérico necesario para tareas de síntesis de voz, en particular para:

- BSC: castellano y catalán.
- UPV/EHU: euskera.
- USC: gallego.
- CENID: valenciano.

Actividad 2.4 Sintetizador de voz ibérico (StyleTTS2).

Diseño y desarrollo de un modelo de síntesis de voz ibérico a partir del modelo StyleTTS2 (o equivalente) en:

- BSC: castellano y catalán.
- UPV/EHU: euskera.
- USC: gallego.
- CENID: valenciano.

Actividad 2.5 Generación modelos SLMs generativos de dominio a partir de los modelos fundacionales.

El BSC según la actividad 2.6 coordinará y hará el seguimiento de las tareas de desarrollo de modelos SLMs generativos (solo decoders) en diferentes dominios a partir de los modelos fundacionales.

Diseño y desarrollo de modelos SLMs generativos (solo decoders con técnicas de quantization o similares para reducción de tamaño) en los dominios de interés (por ejemplo, legal y médico) a partir de los modelos fundacionales en las siguientes lenguas:

- UPV/EHU: castellano y euskera.
- USC: castellano y gallego.
- CENID: castellano y valenciano.
- UJA: castellano.

Actividad 2.7 Generación de modelos discriminativos (encoders con arquitectura deBERTA o similar) de dominio y tarea.

El BSC según la actividad 2.8 coordinará y hará el seguimiento de las tareas de desarrollo de modelos discriminativos auto-encoders con arquitectura estado-del-arte en diferentes dominios.

Diseño y desarrollo de modelos multilingües discriminativos auto-encoders con arquitectura estado-del-arte en diferentes dominios (administrativo, científico, legal, propiedad intelectual, médico, turístico...) para tareas discriminativas que al menos incluya:

- UPV/EHU: inglés, castellano y euskera.
- USC: castellano y gallego.
- CENID: inglés, castellano y valenciano.
- UJA: inglés y castellano.

Actividad 2.9 Generación de datasets para tareas discriminativas.

Diseño y desarrollo de datasets de tareas discriminativas para las diferentes lenguas:

- BSC: al menos tres dataset para tareas discriminativas de impacto en castellano o catalán.
- UPV/EHU: al menos tres dataset para tareas discriminativas de impacto en castellano y euskera.
- USC: al menos tres dataset para tareas discriminativas de impacto en castellano y gallego.
- CENID: al menos tres dataset para tareas discriminativas de impacto en castellano y valenciano, tales como detección de fraude, detección de marcas falsas y detección de lenguaje claro.
- UJA: al menos tres dataset para tareas discriminativas de impacto en castellano, por ejemplo: detección lenguaje ofensivo, detección de trastornos mentales, tratamiento de subjetividad.

Iniciativa 3). Desarrollo y puesta a disposición pública de un marco de evaluación de modelos.

Actividad 3.1 Creación datasets de evaluación e integración en frameworks de referencia, así como en la plataforma de benchmarking. La evaluación incluye no sólo performance, sino también aspectos de sesgo y toxicidad entre otros.

El BSC según la actividad 3.2 coordinará y hará el seguimiento de la creación datasets de evaluación e integración en frameworks de referencia, así como en la plataforma de benchmarking.

El resto de las entidades, llevarán a cabo el diseño y desarrollo de datasets de evaluación que permita comparar las capacidades y rendimiento (benchmarking) de los distintos modelos de lenguaje, TA y de voz (síntesis y reconocimiento). Los datasets deben contemplar distintas métricas de evaluación de modelos de lenguaje, modelos de TA y también de reconocimiento y síntesis de voz. No sólo medirán la capacidad de los modelos sino también aspectos de eficiencia temporal y energética, sesgo, veracidad y toxicidad entre otros. Integración en marcos de referencia (frameworks) así como en la plataforma pública de evaluación del rendimiento (benchmarking) de los distintos modelos de lenguaje, en las distintas lenguas:

- UPV/EHU: al menos tres datasets para tareas de evaluación en castellano y euskera.
- USC: al menos tres datasets para tareas de evaluación en castellano y gallego.
- CENID: al menos tres datasets para tareas de evaluación en castellano y valenciano.
- UJA: al menos tres datasets para tareas de evaluación en castellano.

Iniciativa 5). Red territorial de equipos de colaboración técnica en IA/TL y acceso al entorno seguro pre-comercial.

El BSC será el encargado de la coordinación de esta iniciativa. Además de las actividades que se describen a continuación, estará en continuo contacto con el personal responsable a cargo de cada nodo territorial. Realizará reuniones regulares con el objetivo de realizar el seguimiento técnico de las actividades y aplicar las medidas adecuadas, en caso necesario, para mitigar cualquier riesgo que pueda surgir durante la ejecución de la actuación.

Actividad 5.1 Convocatorias de acceso a recursos computacionales para el desarrollo de pruebas de concepto en IA/TL.

– Descripción: Definición y gestión de las convocatorias de acceso gratuito al entorno seguro para las empresas y organizaciones para la evaluación de modelos. La sugerencia inicial será elaborar convocatorias específicas en base el sistema de convocatorias de la RES como punto de partida conocido por las entidades participantes.

- Entregable: 2 convocatorias de acceso al entorno seguro.
- El BSC según la actividad 5.5 coordinará y hará el seguimiento de esta actividad.

Actividad 5.2 Acompañamiento y aprendizaje colaborativo para la realización de pruebas de concepto y evaluación en entorno real de modelos de propósito específico, incluyendo el apoyo al desarrollo de los casos de uso piloto de la actividad 4.5 (Desarrollo de casos de uso piloto en ámbitos estratégicos definidos por la SEDIA a partir de la familia de modelos ALIA desarrollada) de la resolución de la SEDIA del 19 de agosto de 2024 para la realización de actuaciones ligadas al desarrollo del PlanTL y del ecosistema empresarial de las TL en el marco de la Estrategia de IA 2024 y el PRTR financiado por la UE-NextGeneration EU.

– Descripción: Servicio de apoyo a la estructuración de proyectos de IA y tecnologías del lenguaje para facilitar actividades como la inferencia de modelos, el ajuste de modelos a tareas, la evaluación o la alineación de los mismos. Este apoyo en las pruebas y evaluación de modelos lo realizarán las diferentes entidades participantes con personal técnico específico.

- Entregable: Informe de servicios prestados.
- El BSC según la actividad 5.5 coordinará y hará el seguimiento de esta actividad.

Actividad 5.3 Tareas de difusión y divulgación de IA/TL.

– Descripción: Campañas comunicativas/jornadas y eventos para promoción del uso de los modelos de la familia ALIA. Creación y/o adaptación de material formativo acerca de la IA/TL y su aplicación alrededor de los recursos de ALIA. Contribución al ALIA Kit con los recursos generados por los diferentes nodos.

- Entregable: 2 jornadas organizadas en el BSC acerca de los modelos de la familia ALIA y su aplicabilidad. 3 campañas comunicativas. 4 actividades formativas.
- El BSC según la actividad 5.5 coordinará y hará el seguimiento de esta actividad.

Actividad 5.4 Promoción de las actividades de la red y encuentros.

– Se realizará la promoción y difusión de la cartera de servicios, en campañas comunicativas y mediante la organización de jornadas y eventos para mostrar los recursos de ALIA. Además, se prevé organizar un encuentro anual en el BSC de las entidades participantes en el convenio.

- Entregable: Informe del encuentro.
- El BSC según la actividad 5.5 coordinará y hará el seguimiento de esta actividad.

En la tabla resumen, a continuación, se especifican las iniciativas, sus actividades y entregables/resultados, sí como las fechas de inicio y fin y qué entidades se hacen cargo.

La numeración de las actividades indicadas en la tabla corresponde con la resolución de la SEDIA firmada el 19 de agosto de 2024 para la realización de actuaciones ligadas al desarrollo del PlanTL y del ecosistema empresarial de las TL en el marco de la Estrategia de IA 2024 y el PRTR financiado por la UE-NextGeneration EU.

Iniciativas	Actividades	Fecha inicio	Fecha fin	Entidades participantes	Entregables y/o resultados
1) Desarrollo de un repositorio de corpus de elevada calidad en español y lenguas cooficiales.	1.1 Identificación, análisis y contacto con potenciales repositorios de contenidos en todas las modalidades, no restrictivos y susceptibles de formar parte del corpus.	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, UJA, CENID	E1.1 Informe de actividad y acuerdos establecidos para la cesión/compartición de datos (v1 M6; v2 M12). Con la contribución de UPV/EHU, USC, UJA, CENID y la coordinación del BSC. Responsable compilación del entregable: BSC.
	1.3 Operacionalización del acceso a corpus de texto para el entrenamiento de modelos en todas las lenguas de España y Portugal.	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, UJA, CENID	E1.3 Scripts/protocolos de acceso y descarga a las diferentes fuentes de datos, publicados en github y debidamente documentados (v1 M6; v2 M12). Con la contribución de UPV/EHU, USC, UJA, CENID y la coordinación del BSC. Responsable compilación del entregable: BSC.
	1.5 Operacionalización del acceso a corpus de voz para el entrenamiento de modelos en todas las lenguas de España.	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, CENID	E1.5 Scripts/protocolos de acceso y descarga de las diferentes fuentes de datos de voz (incluyendo alineación texto/voz cuando sea posible), publicados en github y debidamente documentados (M12). Con la contribución de todas las entidades participantes. Responsable compilación del entregable: UPV/EHU.
	1.7 Pipeline de preproceso de datos de voz.	Fecha publicación en el BOE.	M9	BSC, UPV/EHU, USC	E1.7 Pipeline de preproceso de datos de voz para todas las lenguas oficiales debidamente documentada (v1 6M; v2 9M). Con la contribución de UPV/EHU, USC para el desarrollo de los componentes monolingües de gallego y vasco. Y del BSC para el castellano. Responsable compilación del entregable: BSC.

Iniciativas	Actividades	Fecha inicio	Fecha fin	Entidades participantes	Entregables y/o resultados
1) Desarrollo de un repositorio de corpus de elevada calidad en español y lenguas cooficiales.	1.8 Dataset de alineamiento. Los datos de alineamiento son la última fase después de los datos de pre-entrenamiento y de instrucción	Fecha publicación en el BOE.	M11	BSC, UPV/EHU, USC, UJA, CENID	E1.8 Corpus de alineamiento de modelos en todas las lenguas de España (M11). Con la contribución de todas las entidades participantes. Responsable compilación del entregable: UJA.
	1.9 Generación de datos (sintéticos) para corpus paralelos (Traducción Automática) incluyendo adaptación a dominio.	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, UJA, CENID	E1.9.1 BSC: Corpus paralelos inglés, castellano, catalán, chino para entrenamiento de modelos de TA (v1 M6; v2 M12). E1.9.2 UPV/EHU: Corpus paralelos inglés, castellano y euskera para entrenamiento de modelos de TA (v1 M6; v2 M12). E1.9.3 USC Corpus paralelos inglés, castellano y gallego para entrenamiento de modelos de TA (v1 M6; v2 M12). E1.9.4 UJA Corpus paralelos inglés, y castellano para entrenamiento de modelos de TA (v1 M6; v2 M12). E1.9.5 CENID Corpus paralelos inglés, castellano y valenciano (v1 M6; v2 M12).
	1.10 Generación de corpus de dominio (legal, médico,...).	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, UJA, CENID	E1.10.1 BSC: Corpus de dominio: jurídico/legal y científico en castellano; procesado y debidamente metadatos (V1 M6; v2 M12). E1.10.2 UPV/EHU: Corpus de dominio de noticias, legal/administrativo y científico procesado y con sus correspondientes metadatos (V1 M6; v2 M12). E1.10.3 USC: Corpus de dominio de noticias, cultura y legal/administrativo (V1 M6; v2 M12). E1.10.4 UJA: Corpus de dominio legal-administrativo, patrimonio-cultural y biosanitario procesado y con sus correspondientes metadatos. (V1 M6; v2 M12). E1.10.5 CENID: Corpus de dominio turístico y de la propiedad intelectual procesado y con sus correspondientes metadatos (V1 M6; v2 M12).

Iniciativas	Actividades	Fecha inicio	Fecha fin	Entidades participantes	Entregables y/o resultados
2) Desarrollo de modelos de propósito general (LLM).	2.2 Modelo autosupervisado de voz tipo wavLM (o equivalente) Ibérico en mínimo 4 idiomas.	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, CENID	E2.2 Modelo autosupervisado de voz tipo wavLM (o equivalente) Ibérico (v1 M6; v2 M12). Entrenado en MN5 por el BSC con la aportación de datos de UPV/EHU, USC y CENID en las respectivas lenguas. Responsable del entregable: BSC.
	2.3 Un encoder fonético ibérico (para tareas de síntesis de voz).	Fecha publicación en el BOE.	M6	BSC, UPV/EHU, USC, CENID	E2.3.1 BSC: Encoder fonético para el castellano (M6). E2.3.2 USC: Encoder fonético para el gallego (M6). E2.3.3 UPV/EHU Encoder fonético para el euskera (M6). E2.3.4 CENID Encoder fonético para la variedad valenciana (M6).
	2.4 Sintetizador de voz ibérico (StyleTTS2)	Fecha publicación en el BOE.	M9	BSC, UPV/EHU, USC, CENID	E2.4.1 BSC: Sintetizador de voz ibérico (StyleTTS2) para el castellano y catalán (M9). E2.4.2 USC: Sintetizador de voz ibérico (StyleTTS2) para el gallego (M9). E2.4.3 UPV/EHU: Sintetizador de voz ibérico (StyleTTS2) para el euskera (M9). E2.4.4 CENID: Sintetizador de voz ibérico (StyleTTS2) para la variedad valenciana (M9).
	2.5 Generación modelos SLMs generativos de dominio y/o lengua a partir de los modelos fundacionales	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, UJA, CENID	E2.5.1 UPV/EHU: Modelos SLMs generativos de dominio y/o lengua a partir de los modelos fundacionales (v1 M9; v2 M12). E2.5.2 USC: Modelos SLMs generativos de dominio y/o lengua a partir de los modelos fundacionales (v1 M9; v2 M12). E2.5.3 UJA: Modelos SLMs generativos de dominio a partir de los modelos fundacionales (v1 M9; v2 M12). E2.5.4 CENID: Modelos SLMs generativos de dominio y/o lengua a partir de los modelos fundacionales (v1 M9; v2 M12). Coordinación de la actividad: BSC.
	2.7 Generación de modelos discriminativos (encoders con arquitectura deBERTA o equivalente) de dominio y tarea.	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, UJA, CENID	E2.7.1 UPV/EHU: Modelos discriminativos adaptados a cada dominio y/o tarea (v1 M6, v2 M12). E2.7.2 USC: Modelos discriminativos adaptados a dominio y/o tarea (v1 M6, v2 M12). E2.7.3 UJA: Modelos discriminativos adaptados a dominio y/o tarea (v1 M6, v2 M12). E2.7.4 CENID: Modelos discriminativos adaptados a dominio y/o tarea (v1 M6, v2 M12). Coordinación de la actividad: UJA.
	2.9 Generación de datasets para tareas discriminativas	Fecha publicación en el BOE.	M9	BSC, UPV/EHU, USC, UJA, CENID	E2.9.1 BSC: al menos tres dataset para tareas discriminativas de impacto (M9). E2.9.2 UPV/EHU: al menos tres dataset para tareas discriminativas de impacto (M9). E2.9.3 USC: al menos tres dataset para tareas discriminativas de impacto (M9). E2.9.4 UJA: al menos tres dataset para tareas discriminativas de impacto: por ejemplo: detección lenguaje ofensivo, detección de trastornos mentales, tratamiento de subjetividad (M9). E2.9.5 CENID: al menos tres dataset para tareas discriminativas de impacto, tales como detección de fraude, detección de marcas falsas y detección de lenguaje claro (M9). Coordinación de la actividad: UJA.

Iniciativas	Actividades	Fecha inicio	Fecha fin	Entidades participantes	Entregables y/o resultados
3) Desarrollo y puesta a disposición pública de un marco de evaluación de modelos.	3.1 Creación de datasets de evaluación e integración en frameworks de referencia así como en la plataforma de benchmarking. No sólo performance, incluye también aspectos de sesgo y toxicidad entre otros.	Fecha publicación en el BOE.	M9	BSC, UPV/EHU, USC, UJA, CENID	E3.9.1 UPV/EHU: al menos 3 dataset para tareas de evaluación (M9). E3.9.2 USC: al menos 3 dataset para tareas de evaluación (M9). E3.9.3 UJA: al menos 3 dataset para tareas de evaluación (M9). E3.9.4 CENID: al menos 3 datasets para tareas de evaluación(M9). Coordinación de la actividad: BSC.
5) Red territorial de equipos de colaboración técnica en IA/TL y acceso al entorno seguro pre-comercial.	5.1 Convocatorias de acceso recursos computacionales para el desarrollo de pruebas de concepto en IA/TL.	Fecha publicación en el BOE.	M9	BSC, UPV/EHU, USC, UJA, CENID	E5.1 Informes de evaluación de las propuestas (v1 M6; v2 M11). Con la contribución de todas las entidades participantes. Responsable compilación del entregable: BSC.
	5.2 Acompañamiento y aprendizaje colaborativo para el desarrollo de pruebas de concepto de modelos de propósito específico, incluyendo el apoyo al desarrollo de los casos de uso.	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, UJA, CENID	E5.2.1 Informe de apoyo y soporte prestados (M12). Con la contribución de todas las entidades participantes. Responsable compilación del entregable: BSC.
	5.3 Tareas de difusión y divulgación de IA/TL.	Fecha publicación en el BOE.	M12	BSC, UPV/EHU, USC, UJA, CENID	E5.3.1 Resumen sobre jornadas de difusión acerca de los modelos de la familia ALIA y su aplicabilidad. E5.3.2 Resumen de campañas comunicativas. E5.3.3 Resumen de actividades formativas. Con la contribución de todas las entidades participantes. Responsable compilación del entregable: BSC.
	5.4 Promoción de las actividades de la red y encuentros.	Fecha publicación en el BOE.		BSC, UPV/EHU, USC, UJA, CENID	E5.4 Informe del encuentro (v1- M6; v2 M12).

## ANEXO 3

## Conocimientos previos

*Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS)*

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
alvocat-vocos-22khz	Modelo	BSC	Apache 2.0
whisper-large-v3-ca-3catparla	Modelo	BSC	Apache 2.0
faster-whisper-large-v3-ca-3catparla	Modelo	BSC	Apache 2.0
aina-translator-ca-es	Modelo	BSC	Apache 2.0
aina-translator-es-ca	Modelo	BSC	Apache 2.0
aina-translator-ca-en	Modelo	BSC	Apache 2.0
aina-translator-en-ca	Modelo	BSC	Apache 2.0
aina-translator-ca-de	Modelo	BSC	Apache 2.0

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
aina-Translator-de-ca	Modelo	BSC	Apache 2.0
aina-Translator-ca-fr	Modelo	BSC	Apache 2.0
aina-Translator-fr-ca	Modelo	BSC	Apache 2.0
aina-Translator-ca-it	Modelo	BSC	Apache 2.0
aina-Translator-it-ca	Modelo	BSC	Apache 2.0
aina-Translator-ca-pt	Modelo	BSC	Apache 2.0
aina-Translator-pt-ca	Modelo	BSC	Apache 2.0
aina-Translator-ca-zh	Modelo	BSC	Apache 2.0
aina-Translator-zh-ca	Modelo	BSC	Apache 2.0
aina-Translator-eu-ca	Modelo	BSC	Apache 2.0
aina-Translator-gl-ca	Modelo	BSC	Apache 2.0
Plume32k	Modelo	BSC	Apache 2.0
Plume128k	Modelo	BSC	Apache 2.0
Plume256k	Modelo	BSC	Apache 2.0
SalamandraTA-2B	Modelo	BSC	Apache 2.0
roberta-base-ca-v2	Modelo	BSC	Apache 2.0
roberta-large-ca-v2	Modelo	BSC	Apache 2.0
salamandra-2b	Modelo	BSC	Apache 2.0
salamandra-2b-instruct	Modelo	BSC	Apache 2.0
salamandra-7b	Modelo	BSC	Apache 2.0
salamandra-7b-instruct	Modelo	BSC	Apache 2.0
annotated_catalan_common_voice_v17	Dataset	BSC	CC BY
Common Voice Benchmark Catalan Accents	Dataset	BSC	CC BY
ParlamentParla	Dataset	BSC	CC BY
ParlamentParla v3	Dataset	BSC	CC BY
cv17_es_other_automatically_verified	Dataset	BSC	CC BY
CA-EN_Parallel_Corpus	Dataset	BSC	CC BY
magma	Software	BSC	Apache 2.0

*Universidad del País Vasco/Euskal Herriko Unibertsitatea (HiTZ/UPV/EHU)*

El declarante manifiesta que ha realizado el análisis del correspondiente background y afirma que la información listada está revisada, es veraz y correcta:

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
HiTZ/Medical-mT5-large	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/Medical-mT5-xl	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/Medical-mT5-large-multitask	LLM	HiTZ/UPV/EHU	apache-2.0

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
HiTZ/Medical-mT5-xl-multitask	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/mdeberta-base-metaphor-detection-es	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-metaphor-detection-es	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/deberta-large-metaphor-detection-en	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/deberta-base-metaphor-detection-en	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-metaphor-detection-en	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt5-counter-narrative-en	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt5-counter-narrative-es	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt5-counter-narrative-eu	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/lmloss-opt-rm-1.3b	LLM	HiTZ/UPV/EHU	MIT
HiTZ/rmloss-opt-rm-13b	LLM	HiTZ/UPV/EHU	MIT
HiTZ/gpt2-eus-euscrawl	LLM	HiTZ/UPV/EHU	CC-BY
HiTZ/stt_eu_conformer_ctc_large	STT	HiTZ/UPV/EHU	apache-2.0
HiTZ/stt_eu_conformer_transducer_large	STT	HiTZ/UPV/EHU	apache-2.0
HiTZ/EriBERTa-base	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt-hitz-gl-eu	MT	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt-hitz-eu-es	MT	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt-hitz-es-eu	MT	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt-hitz-eu-en	MT	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt-hitz-en-eu	MT	HiTZ/UPV/EHU	apache-2.0
HiTZ/mt-hitz-ca-eu	MT	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-xnli-es	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-lemma-eu	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-lemma-ru	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-lemma-es	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-lemma-en	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-lemma-cz	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-lemma-pl	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-lemma-tr	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/judge-eus	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/mdeberta-expl-extraction-multi	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/xlm-roberta-large-expl-extraction-multi	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/mbert-argmining-abstrct-multilingual	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/Mistral-7B-MedExpQA-EN	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/mbert-argmining-abstrct-en-es	LLM	HiTZ/UPV/EHU	apache-2.0
HiTZ/A2T_RoBERTa_SMFA_ACE-arg	LLM	HiTZ/UPV/EHU	MIT

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
HiTZ/A2T_RoBERTa_SMFA_WikiEvents-arg_ACE-arg	LLM	HiTZ/UPV/EHU	MIT
HiTZ/A2T_RoBERTa_SMFA_WikiEvents-arg	LLM	HiTZ/UPV/EHU	MIT
HiTZ/A2T_RoBERTa_SMFA_ACE-arg_WikiEvents-arg	LLM	HiTZ/UPV/EHU	MIT
HiTZ/A2T_RoBERTa_SMFA_TACRED-re	LLM	HiTZ/UPV/EHU	MIT
ixa-ehu/roberta-eus-mc4-base-cased	LLM	HiTZ/UPV/EHU	MIT
ixa-ehu/roberta-eus-euscrawl-base-cased	LLM	HiTZ/UPV/EHU	MIT
ixa-ehu/ixambert-base-cased	LLM	HiTZ/UPV/EHU	apache-2.0
ixa-ehu/berteus-base-cased	LLM	HiTZ/UPV/EHU	apache-2.0
ixa-ehu/SciBERT-SQuAD-QuAC	LLM	HiTZ/UPV/EHU	apache-2.0
ixa-ehu/roberta-eus-cc100-base-cased	LLM	HiTZ/UPV/EHU	MIT
ixa-ehu/roberta-eus-euscrawl-large-cased	LLM	HiTZ/UPV/EHU	MIT
EDBL	Lexicon	HiTZ/UPV/EHU	CC-BY
EUSLEM	Lexicon	HiTZ/UPV/EHU	CC-BY
ElhuyarWord	Lexicon	HiTZ/UPV/EHU	CC-BY
BasqueWordNet/EuskalWordNet	Lexicon	HiTZ/UPV/EHU	CC-BY
Lexkit	Lexicon	HiTZ/UPV/EHU	CC-BY
EuSemCor	Corpora	HiTZ/UPV/EHU	CC-BY
EPEC-ROLSEM	Corpora	HiTZ/UPV/EHU	CC-BY
EPEC-DEP (BDT)	Corpora	HiTZ/UPV/EHU	CC-BY
IXA-pipes	Software	HiTZ/UPV/EHU	apache-2.0
HiTZ/Euscrawl	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/This-is-not-a-dataset	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/Multilingual-Opinion-Target-Extraction	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/CONAN-EUS	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/basqueparl	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/EusProficiency	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/casimedicos-exp	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/EusReading	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/EusTrivia	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/latxa-corpus-v1.1	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/Multilingual-Medical-Corpus	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/EusIE	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/cometa	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/casimedicos-squad	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/Multilingual-BioASQ-6B	Corpora	HiTZ/UPV/EHU	apache-2.0

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
HiTZ/meta4xnli	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/XCOPA-eu	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/wnli-eu	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/BertaQA	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/PixT3	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/MedExpQA	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/EusParallel	Corpora	HiTZ/UPV/EHU	apache-2.0
HiTZ/casimedicos-arg	Corpora	HiTZ/UPV/EHU	CC-BY
HiTZ/composite_corpus_eu_v2.1	Corpora	HiTZ/UPV/EHU	CC-BY
MCR	Lexicon	HiTZ/UPV/EHU	CC-BY
AdimenSUMO	Lexicon	HiTZ/UPV/EHU	CC-BY
BLC	Lexicon	HiTZ/UPV/EHU	CC-BY
Predicate Matrix	Lexicon	HiTZ/UPV/EHU	CC-BY
QWN-PPV	Lexicon	HiTZ/UPV/EHU	CC-BY
WordNet2TO	Lexicon	HiTZ/UPV/EHU	CC-BY
KnowNet	Lexicon	HiTZ/UPV/EHU	CC-BY
DeepKnowNet	Lexicon	HiTZ/UPV/EHU	CC-BY
WordFrameNet	Lexicon	HiTZ/UPV/EHU	CC-BY
XWND	Lexicon	HiTZ/UPV/EHU	CC-BY
ImpPar	Software	HiTZ/UPV/EHU	CC-BY
CrossTimeLines	Corpora	HiTZ/UPV/EHU	CC-BY
WSD-WN-Glosses	Corpora	HiTZ/UPV/EHU	CC-BY
Ask2Transformers Domains	Lexicon	HiTZ/UPV/EHU	CC-BY

*Universidad de Santiago de Compostela (CiTIUS-ILG-USC)*

El declarante manifiesta que ha realizado el análisis del correspondiente background y afirma que la información listada está revisada, es veraz y correcta:

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
Nos_TTS-sabela-vits-phonemes	Model	USC	Apache License, Version 2.0
Nos_TTS-celtia-vits-graphemes	Model	USC	Apache License, Version 2.0
Llama-3.1-Carballo	Model	USC	MIT License
Carballo-bloom-1.3B	Model	USC	MIT License
Carvalho_pt-gl-1.3B	Model	USC	MIT License
Carballo-cerebras-1.3B	Model	USC	MIT License
Nos_D2T-gl	Model	USC	MIT License
Nos_MT-OpenNMT-en-gl	Model	USC	MIT License

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
Nos_MT-OpenNMT-es-gl	Model	USC	MIT License
Nos_MT-OpenNMT-eu-gl	Model	USC	MIT License
Nos_MT-OpenNMT-ca-gl	Model	USC	MIT License
Nos_MT-OpenNMT-gl-en	Model	USC	MIT License
Nos_MT-OpenNMT-gl-es	Model	USC	MIT License
Nos_MT-OpenNMT-multilingual	Model	USC	MIT License
belebele_gl	Dataset	USC	CC BY 4.0
mgs_m_gl	Dataset	USC	CC BY 4.0
truthfulqa_gl	Dataset	USC	Apache License 2.0
summarization_gl	Dataset	USC	CC BY 4.0
pipeline	Software	USC	MIT License
Nos_CorpusNOS-GL: Galician Macrocorpus for LLM training	Dataset	USC	Creative Commons Attribution 4.0 International, Apache License 2.0, Creative Commons Zero v1.0 Universal, Creative Commons Attribution 4.0 International
Nos_Celtia-GL	Dataset	USC	Creative Commons Attribution 4.0 International
Nos_ParlaSpeech-GL	Dataset	USC	Creative Commons Attribution 4.0 International
Nos_TranscriSpeech-GL	Dataset	USC	Creative Commons Attribution 4.0 International

*Universidad de Alicante-Grupo de investigación de Procesamiento de Lenguaje y Sistemas de Información (UA-GPLSI)*

El declarante manifiesta que ha realizado el análisis del correspondiente background y afirma que la información listada está revisada, es veraz y correcta:

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
eHealth-KD	corpus	UA-GPLSI, UH	MIT License
ehealthkd-2019-ensembled-corpus	corpus	UA-GPLSI, UH	Creative Commons BY 4.0
ClearText Cas v1.0	corpus	UA-GPLSI	Creative Commons BY 4.0
ClearTextSim Cas v1.0	corpus	UA-GPLSI	Creative Commons BY 4.0
ClearText Val v1.0	corpus	UA-GPLSI	Creative Commons BY 4.0
COCOTEROS	corpus	UA-GPLSI	Creative Commons BY 4.0
FNDDeepML	corpus	UA-GPLSI	Apache 2.0
RUN-AS	corpus	UA-GPLSI	Apache 2.0
RUN-AS-SFN	corpus	UA-GPLSI	Apache 2.0
VIL	corpus	UA-GPLSI	Creative Commons BY 4.0

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
Twin-AnnotatES	corpus	UA-GPLSI	Creative Commons By 4.0
LIPCOL	corpus	UA-GPLSI	Creative Commons BY 4.0
SSAID	corpus	UA-GPLSI	Creative Common BY 4.0
reviews-gender-corpus	corpus	UA-GPLSI, UH	Creative Commons BY 4.0
AutoGOAL	librería	UA-GPLSI, UH	MIT License
BFair	librería	UA-GPLSI, UH	MIT License
ClearSim	modelo	UA-GPLSI	Creative Commons BY 4.0
LenOfen	modelo	UA-GPLSI	Vreative Commons BY 4.0

*Universidad de Jaén. Grupo Sistemas Inteligentes de Acceso a la Información (SINA)*

El declarante manifiesta que ha realizado el análisis del correspondiente background y afirma que la información listada está revisada, es veraz y correcta:

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
CRiSOL	Lexicón	SINAI-UJA	Creative Commons BY 4.0
emoti-sp	Lexicón	SINAI-UJA	Creative Commons BY 4.0
eSOL	Lexicón	SINAI-UJA	Creative Commons BY 4.0
eSOLdomainGlobal	Lexicón	SINAI-UJA	Creative Commons BY 4.0
Hashtags-sp	Lexicón	SINAI-UJA	Creative Commons BY 4.0
iSOL	Lexicón	SINAI-UJA	Creative Commons BY 4.0
COAH	Corpus	SINAI-UJA	Creative Commons BY 4.0
COAR	Corpus	SINAI-UJA	Creative Commons BY 4.0
COST	Corpus	SINAI-UJA	Creative Commons BY 4.0
EVOCA Corpus	Corpus	SINAI-UJA	Creative Commons BY 4.0
MCE Corpus	Corpus	SINAI-UJA	Creative Commons BY 4.0
OCA Corpus	Corpus	SINAI-UJA	Creative Commons BY 4.0
Colección HEP	Corpus	SINAI-UJA	Creative Commons BY 4.0
COPOD	Corpus	SINAI-UJA	Creative Commons BY 4.0
COPOS	Corpus	SINAI-UJA	Creative Commons BY 4.0
DOS	Corpus	SINAI-UJA	Creative Commons BY 4.0
EmoEvent	Corpus	SINAI-UJA	Creative Commons BY 4.0
CARES	Corpus	SINAI-UJA	Creative Commons BY 4.0
OffendES	Corpus	SINAI-UJA	Creative Commons BY 4.0
SAD: Spanish Anorexia Detection	Corpus	SINAI-UJA	Creative Commons BY 4.0
SHARE: Lexicón de expresiones ofensivas de la población española	Lexicón	SINAI-UJA	Apache 2.0

Conocimientos previos	Tipo	Titularidad	Condiciones de uso y explotación (Licencia)
SOL	Lexicón	SINAI-UJA	Creative Commons BY 4.0
TASS-Politics	Corpus	SINAI-UJA	Creative Commons BY 4.0
TASS-Social-TV Corpus	Corpus	SINAI-UJA	Creative Commons BY 4.0
TASS-STOMPOL Corpus	Corpus	SINAI-UJA	Creative Commons BY 4.0
Inter TASS Corpus	Corpus	SINAI-UJA	Creative Commons BY 4.0
MentalRiskES	Corpus	SINAI-UJA	Creative Commons BY 4.0
PRECOM	Corpus	SINAI-UJA	Creative Commons BY 4.0
DataBench	Corpus	SINAI-UJA	MIT
DataBenchSPA	Corpus	SINAI-UJA	MIT