

III. OTRAS DISPOSICIONES

MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA

24153 *Resolución de 14 de noviembre de 2025, de la Secretaría de Estado de Digitalización e Inteligencia Artificial, por la que se publica el Convenio con la Comunidad Autónoma del País Vasco, para la puesta en marcha de actuaciones en el ámbito de las tecnologías del lenguaje en Euskera, en el marco de la iniciativa Euskorpus del Gobierno Vasco.*

Con fecha 23 de octubre de 2025 se ha suscrito el Convenio entre el Ministerio para la Transformación Digital y de la Función Pública y la Consejería de Industria, Transición Energética y Sostenibilidad de la Comunidad Autónoma de Euskadi. Se trata de impulsar el desarrollo del euskera en el ámbito digital, especialmente en tecnologías del lenguaje e inteligencia artificial en el marco de la iniciativa Euskorpus del Gobierno Vasco.

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» del citado convenio.

Madrid, 14 de noviembre de 2025.–La Secretaria de Estado de Digitalización e Inteligencia Artificial, María González Veracruz.

ANEXO

Convenio entre el Ministerio para la Transformación Digital y de la Función Pública y la Consejería de Industria, Transición Energética y Sostenibilidad de la Comunidad Autónoma de Euskadi para la puesta en marcha de actuaciones en el ámbito de las tecnologías del lenguaje en euskera, en el marco de la iniciativa Euskorpus del Gobierno Vasco

REUNIDOS

De una parte, don Óscar López Águeda, Ministro para la Transformación Digital y de la Función Pública, en virtud del RD 890/2024, de 5 de septiembre, actuando conforme a lo dispuesto en el artículo 61 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (LRJSP); y el Real Decreto 1185/2024, de 28 de noviembre, por el que se modifican el Real Decreto 210/2024, de 27 de febrero, por el que se desarrolla la estructura orgánica básica del Ministerio para la Transformación Digital y de la Función Pública, y el Real Decreto 1009/2023, de 5 de diciembre, por el que se establece la estructura orgánica básica de los departamentos ministeriales, y cuantas disposiciones le atribuyen competencia.

De otra parte, don Mikel Jauregi Letemendia, Consejero de Industria, Transición Energética y Sostenibilidad, en virtud del Decreto 24/2024, de 24 de junio, del Lehendakari, de nombramiento del Consejero de Industria, Transición Energética y Sostenibilidad, actuando mediante la autorización del Consejo de Gobierno, de fecha 7 de octubre de 2025, para manifestar el consentimiento y suscribir el presente convenio en nombre de la Administración General de la Comunidad Autónoma de Euskadi, conforme al artículo 62.1 del Decreto 144/2017, de 25 de abril, del Servicio Jurídico del Gobierno Vasco.

Todos intervienen en uso de las facultades que, de conformidad con la normativa vigente, les confieren los cargos que desempeñan y se reconocen mutuamente

legitimidad y la capacidad jurídica necesaria para la firma del presente convenio y, a tal efecto,

EXPONEN

Primero.

Que la finalidad de este convenio es establecer los términos de la colaboración entre el Ministerio para la Transformación Digital y de la Función Pública y el Departamento de Industria, Transición Energética y Sostenibilidad del Gobierno Vasco, para desarrollar y apoyar iniciativas en el ámbito de las tecnologías del lenguaje, convirtiendo al euskera en una herramienta clave para el desarrollo en un escenario digital y globalizado, y contribuyendo a la creación de corpus digital en euskera.

El corpus –escrito y hablado–, el léxico computacional, las bases de datos terminológicas, la recopilación de voz, etc. son recursos lingüísticos, entendidos como un conjunto de datos y descripciones de voz o lenguaje en formato legible por máquina, utilizados para crear, mejorar o evaluar sistemas o algoritmos de lenguaje natural y de voz.

No existe en este momento ningún modelo base de inteligencia artificial que posea un conocimiento significativo del euskera, efectivo para tareas de procesamiento del habla. Los modelos de traducción automática disponibles para este idioma son de calidad inferior a los modelos líderes actuales, y, además, los grandes modelos de lenguaje entrenados específicamente en euskera tienen una cobertura limitada, al estar basados en corpus de noticias web, lo que reduce su eficacia en otros dominios, por todo lo cual se hace necesario potenciar el desarrollo del euskera en la generación de corpus digitales lingüísticos.

Segundo.

La Estrategia para la Transformación Digital de Euskadi 2025 («Euskadi 2025 Euskadiko Eraldaketa Digitalerako Estrategia», ETDE2025) marca como objetivo general «acelerar la adopción de las palancas tecnológicas incipientes, fortalecer el desarrollo y aprovechar el potencial demostrado de los habilitadores, activar y potenciar su rápida incorporación a los ámbitos de aplicación esenciales, contribuyendo así a la transición tecnológica digital, la transición energética medioambiental y la transición social y sanitaria que debe afrontar Euskadi para el 2025».

A principios de año, el Gobierno de Euskadi anunció la puesta en marcha de la iniciativa, Euskorpus, con el fin de llevar al euskera a la revolución de la IA, garantizando de esta manera el futuro de la lengua vasca en una sociedad cada vez más digital.

El objetivo de esta iniciativa es promocionar y asegurar la presencia digital del euskera, patrimonio cultural inmaterial nacional en peligro de extinción digital, mediante la recopilación, generación y desarrollo activo y metódico del corpus digital del euskera; y la promoción de modelos base de código abierto para el euskera, para su conservación, mantenimiento, y generación de impacto tanto en el tejido industrial como social, conectado al marco europeo de recursos lingüísticos digitales. En definitiva, gracias a Euskorpus se desarrollarán paquetes de datos en euskera, de forma que los fabricantes puedan importarlos en sus dispositivos conectados a internet dotándoles de capacidad de funcionar en este idioma. De esta forma, teléfonos móviles, tabletas, relojes inteligentes, aplicaciones o asistentes digitales tendrán la capacidad de interactuar con los usuarios en euskera, y los usuarios de las herramientas de IA podrán utilizarlas en esta lengua.

Así mismo, el Gobierno Vasco y la Real Academia de la Lengua Vasca disponen de múltiples proyectos que buscan liderar a nivel mundial el desarrollo de herramientas y tecnologías que permitan la proyección y uso del euskera en la Inteligencia Artificial, impulsar la transformación digital en el sector industrial y de servicios y promover un cambio estructural hacia un modelo económico más innovador y emprendedor en el País

Vasco, basado en el desarrollo de capacidades en el ámbito de la economía de la lengua.

Tercero.

Que a la Secretaría de Estado de Digitalización e Inteligencia Artificial le corresponden las funciones dispuestas en el artículo 2, del Real Decreto 210/2024, de 27 de febrero, por el que se establece la estructura orgánica básica del Ministerio para la Transformación Digital y de la Función Pública, modificado por el Real Decreto 1185/2024, de 28 de noviembre, esto es, las funciones previstas en el artículo 62 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, en el ámbito de sus competencias, relativas a la política de impulso a la digitalización de la sociedad y economía de forma respetuosa con los derechos individuales y colectivos, así como con los valores del ordenamiento jurídico español.

En el año 2015 la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital puso en marcha el Plan de Impulso de las Tecnologías del Lenguaje, cuyo objetivo fue fomentar el desarrollo del procesamiento del lenguaje natural, la traducción automática y los sistemas conversacionales en lengua española y lenguas cooficiales.

Posteriormente, en el año 2022 el Ministerio de Asuntos Económicos y Transformación Digital (actual Ministerio para la Transformación Digital y de la Función Pública) puso en marcha el PERTE Nueva Economía de la Lengua, iniciativa que buscaba aprovechar el potencial económico y competitivo del español y las lenguas cooficiales en áreas como la inteligencia artificial, la traducción, la divulgación cultural y la ciencia. En este marco destaca el Proyecto ILENIA, así como la iniciativa ALIA impulsada por el Ministerio, que cuenta con los objetivos de: (i) avanzar en el desarrollo de su propio modelo fundacional de lenguaje, en castellano y lenguas cooficiales, como infraestructura fundamental para el desarrollo de aplicaciones que puedan ser de utilidad para el sector productivo; (ii) impulsar las infraestructuras necesarias para el fomento de la IA, a través de una familia de modelos de lenguaje desarrollados bajo criterios de transparencia y confiabilidad, con un aumento de la participación del castellano y las lenguas cooficiales; y (iii) generar estándares de IA ética y confiable, con modelos abiertos y transparentes, garantizando la protección de los derechos fundamentales, la protección de los derechos de propiedad intelectual y la protección de datos personales, desarrollando un marco de mejores prácticas en este terreno. De esta forma se hará pública la información que permitirá conocer desde la composición de los corpus hasta los entrenamientos realizados, la estructura de los modelos y los recursos utilizados para ello.

En el desarrollo de la familia de modelos ALIA se ha constatado que el rendimiento de los mismos es menor en euskera que en el resto de idiomas cooficiales, debido a la menor disposición de corpus en estas lenguas.

Cuarto.

Que este convenio se firma al amparo de la competencia exclusiva del Estado en materia de fijación de las bases y coordinación de la planificación general de la actividad económica y de fomento y coordinación general de la investigación científica y técnica, de acuerdo con las reglas 13.^a y 15.^a, del artículo 149, apartado 1, de la Constitución Española, por cuanto las actividades financiadas generarán un impacto de alto valor y retorno económico para el tejido productivo y el sector público, así como para el fomento de la investigación en materia de corpus digitales.

Por ello, a la Secretaría de Estado de Digitalización e Inteligencia Artificial le corresponde el diseño, liderazgo y coordinación de proyectos estratégicos clave para la transformación digital de España, así como la interlocución con actores estratégicos, incluyendo industriales, académicos, del sector público y de la sociedad civil, de acuerdo con lo dispuesto en el Real Decreto 829/2023, de 20 de noviembre, por el que se

reestructuran los departamentos ministeriales y se crea el Ministerio para la Transformación Digital y de la Función Pública. El mencionado RD confiere a éste las políticas de transformación digital y el desarrollo y fomento de la inteligencia artificial, y, más concretamente, en el Real Decreto 1185/2024, de 28 de noviembre, por el que se modifican el Real Decreto 210/2024, de 27 de febrero, por el que se desarrolla la estructura orgánica básica del Ministerio para la Transformación Digital y de la Función Pública, y en el Real Decreto 1009/2023, de 5 de diciembre, por el que se establece la estructura orgánica básica de los departamentos ministeriales.

Quinto.

Que la Comunidad Autónoma del País Vasco tiene competencia exclusiva en materia de promoción, desarrollo económico y planificación de la actividad económica del País Vasco, de acuerdo con la ordenación general de la economía, así como en materia de investigación científica y técnica, en coordinación con la Administración General del Estado, prestando especial atención al euskera, por constituir parte esencial de su cultura, de acuerdo con el artículo 6, apartado uno, de la Ley Orgánica 3/1979, de 18 de diciembre, por la que se aprueba el Estatuto de Autonomía para el País Vasco.

Y cuenta con la Real Academia de la Lengua Vasca (Euskaltzaindia) como institución consultiva oficial en materia de euskera de acuerdo con lo establecido en el artículo 6 del Estatuto de Autonomía del País Vasco y reconocida por Decreto 573/1976, de 26 de febrero (BOE 73, de 25 de marzo de 1976).

Sexto.

Que las instituciones firmantes entienden de común interés la conveniencia de impulsar lazos de colaboración y cooperación con el fin de aunar esfuerzos que permitan situar el euskera entre las lenguas europeas con elevado nivel tecnológico, y para ello es importante que esté integrado en el desarrollo de la Inteligencia Artificial, de forma que se puedan ofrecer servicios de calidad en euskera, tanto a las Administraciones públicas, como al conjunto de la sociedad.

El fin común que se pretende satisfacer con la celebración del presente convenio es el interés general de las partes en desarrollar actuaciones que potencien la promoción, generación y desarrollo activo y metódico de corpus digital en euskera para su conservación, mantenimiento, evolución y explotación al servicio de la transformación digital de las lenguas minoritarias, tanto por las empresas, como por la sociedad y la Administración a través de modelos de lenguaje entrenados específicamente en euskera, de los modelos de la familia ALIA y otros.

Séptimo.

Que el artículo 3 de la Ley 40/2015, de 1 de octubre, establece como uno de los principios generales de actuación de las Administraciones públicas, el principio de cooperación, colaboración y coordinación de estas entre sí. A tal efecto, en su capítulo VI del título preliminar, la ley regula el convenio como instrumento con efectos jurídicos para la formalización de acuerdos entre Administraciones para un fin común. Asimismo, el artículo 48 de la Ley 40/2015, de 1 de octubre, establece que las Administraciones públicas podrán suscribir convenios con sujetos de derecho público.

Octavo.

Que el presente convenio es de naturaleza interadministrativa, quedando sometido al régimen jurídico de convenios previsto en el capítulo VI del título preliminar de la Ley 40/2015, de 1 de octubre, de conformidad con el artículo 47, apartado 2, letra a), de la misma ley, donde ambas Administraciones públicas se comprometen a realizar una serie de actuaciones orientadas a la transformación digital y vinculadas al PERTE de la Nueva Economía de la Lengua.

Noveno.

Que en la tramitación del presente convenio se ha recabado el Informe de Abogacía del Estado, de conformidad con la letra a), apartado dos, del artículo 50 de la Ley 40/2015, de 1 de octubre.

Igualmente, se ha procedido a recabar informe de los ministerios de Hacienda y Política Territorial.

Además, este convenio, al estar suscrito con una Comunidad Autónoma, deberá ser remitido al Senado por el Ministerio de Política Territorial, cumpliendo así con lo preceptuado en la letra e), apartado 2, del artículo 50, de la Ley 40/2015, de 1 de octubre.

Décimo.

Que las actuaciones contenidas en el presente convenio se financiarán conjuntamente por las partes suscribientes. Así, para la ejecución de las actuaciones recogidas en el presente convenio, el Ministerio para la Transformación Digital y de la Función Pública aportará una cuantía de 5 millones de euros y la Administración General de la Comunidad Autónoma de Euskadi, las financiará con 5,5 millones de euros.

Por todo ello, las partes acuerdan suscribir el presente convenio, que se regirá por las siguientes,

CLÁUSULAS

Primera. *Objeto del convenio.*

El presente convenio tiene por objeto establecer el marco de colaboración entre la Administración General del Estado y la Administración General de la Comunidad Autónoma de Euskadi para impulsar la presencia digital del euskera como patrimonio cultural inmaterial.

Para ello, se prevé la realización de un conjunto de iniciativas y actuaciones que tienen como fin la promoción, generación y desarrollo activo y metódico del corpus digital del euskera, conectado al marco europeo de recursos lingüísticos, para su conservación, mantenimiento, evolución y explotación al servicio de la transformación digital de las lenguas minoritarias, tanto por las empresas, como por la sociedad y la Administración, en el marco de la iniciativa del gobierno vasco Euskorpus.

Segunda. *Actuaciones y obligaciones de las partes.*

En cuanto a las actuaciones de las partes:

Actuaciones por parte del Ministerio, a través de la Secretaría de Estado de Digitalización e Inteligencia Artificial (en adelante SEDIA):

1. Impulsar la presencia del euskera en iniciativas presentes y futuras, en relación con las tecnologías del lenguaje, que sean promovidas por parte del Ministerio, a través de la Secretaría de Estado de Digitalización e Inteligencia Artificial.
2. Proporcionar datos relevantes para la correcta ejecución, asesoramiento y apoyo de las actuaciones a realizar en el marco de este convenio.
3. Incorporar los resultados de este convenio a la familia de modelos ALIA.
4. Proporcionar el apoyo institucional adecuado para fomentar el uso del euskera en el ámbito digital.
5. Llevar a cabo las actuaciones de supervisión, difusión y comunicación del desarrollo del proyecto. Previo acuerdo con el Gobierno del País Vasco, la SEDIA se compromete a comunicar los avances, logros y resultados que se vayan alcanzando. Esto incluye la difusión en las plataformas en línea, así como en las redes sociales y

otros canales de difusión que permitan mostrar de forma abierta y transparente las actividades que se están llevando a cabo y conduzcan de forma sencilla a los interesados, a los recursos que ya puedan ser consultados y utilizados.

Actuaciones por parte del Departamento de Industria, Transición Energética y Sostenibilidad del Gobierno Vasco:

1. Impulsar la ejecución de las actividades y tareas necesarias para garantizar la presencia del euskera en iniciativas presentes y futuras, en relación con las tecnologías del lenguaje.

2. Impulsar las siguientes iniciativas para incrementar la presencia del euskera en el ámbito digital, con especial foco para su uso en Inteligencia Artificial:

- Definición de los sectores, aplicaciones y servicios estratégicos que se puedan beneficiar de la compilación de corpus lingüísticos y del desarrollo de modelos base de código abierto en euskera.

- Realización de las tareas necesarias para el uso de los corpus, como la anotación y el etiquetado.

- Análisis y establecimiento de los mecanismos legales necesarios para compilar corpus lingüísticos y desarrollar modelos base de código abierto en euskera, de forma que puedan estar accesibles para todas aquellas partes interesadas en desarrollar las tecnologías digitales del euskera.

- Compilación de corpus lingüísticos en euskera para el procesamiento del habla, la traducción automática y el modelado de lenguaje, entre otros.

- Impulso del desarrollo de modelos base de código abierto en euskera de tecnología de procesamiento del habla, traducción automática y procesamiento de lenguaje, entre otros, para asegurar la validez de los corpus realizados.

- Puesta en marcha de la infraestructura hardware necesaria para almacenar y securizar los corpus lingüísticos que se vayan compilando, así como para testear y validar los modelos base de código abierto que se vayan desarrollando.

- Impulso de la transferencia y explotación de los corpus lingüísticos y de los modelos base de código abierto a las empresas, a la sociedad y a otras plataformas de datos europeas.

3. Contar con un sistema de gobernanza para la ejecución, creación de sinergias entre las administraciones y los actores implicados, incluidas entidades sin ánimo de lucro, en el marco de la iniciativa Euskorpus.

4. Poner a disposición del Ministerio los recursos lingüísticos digitales en euskera que sean generados dentro del marco de colaboración establecido en el presente convenio.

5. Elaborar la documentación e información sobre el desarrollo de las iniciativas que se pongan en marcha dentro del presente marco de colaboración.

6. Realizar actividades y eventos que permitan el conocimiento, por parte de los sectores económicos y del conjunto de la sociedad civil, de los objetivos perseguidos y los beneficios esperados con el Proyecto e iniciativas.

Obligaciones del Ministerio a través de la SEDIA:

1. Apoyar al Gobierno Vasco en la ejecución de las actuaciones descritas en el anexo I del presente convenio.

2. Colaborar en la recogida de la información que sea necesaria para la adecuada ejecución del convenio. La SEDIA, por su especial acceso a datos e información que pueden permitir la ejecución óptima del convenio, se compromete a recopilar y proporcionar toda la información que se considere de utilidad para la ejecución de las cláusulas y actividades estipuladas en el mismo. Asimismo, informará de cualquier novedad administrativa que pueda afectar al avance del convenio.

3. Designar a los equipos internos para el seguimiento del cumplimiento de las actuaciones del presente convenio en colaboración y coordinación con los órganos competentes designados por el Gobierno Vasco.

4. Asesorar sobre las líneas de acción a realizar, dadas sus competencias, en el ámbito de la Inteligencia Artificial, la compartición de datos y el emprendimiento digital. La SEDIA se compromete a poner al servicio del desarrollo de las actuaciones la experiencia que posee, como Administración pública encargada del impulso y fomento de la Inteligencia Artificial y sus avances en España.

5. Poner los datos, desarrollos y resultados de las diferentes actuaciones realizadas en el marco de este convenio (datos, corpus, información, herramientas técnicas, informes finales...) a disposición de las Administraciones Públicas.

6. Llevar a cabo las actuaciones de difusión y comunicación de las actuaciones realizadas a través del presente convenio. La SEDIA se compromete a comunicar los avances, logros y resultados que se vayan alcanzando. Esto incluye su comunicación en su página web y las páginas web del Gobierno de España en las que proceda anunciarse, así como la difusión en redes sociales y otros canales y plataformas que permitan mostrar, de forma abierta y transparente, las actividades que se están llevando a cabo y que conduzcan de forma sencilla a los interesados, a los recursos que ya puedan ser consultados y utilizados.

7. Aportar financiación a la Comunidad Autónoma de Euskadi con los límites establecidos en la cláusula siguiente para que la Comunidad Autónoma pueda ejecutar las actuaciones del anexo I.

Obligaciones del Departamento de Industria, Transición Energética y Sostenibilidad del Gobierno Vasco:

1. Ejecutar las actuaciones que corresponden al Gobierno Vasco descritas anteriormente, implementando los instrumentos y controles necesarios para su correcta ejecución.

2. Realizar las actuaciones necesarias para poner los datos, desarrollos y resultados de las diferentes actuaciones realizadas en el marco de este convenio, que tengan como resultado datos, corpus digital (gran base de datos lingüísticos estructurados en euskera), información, herramientas técnicas, así como los informes finales, a disposición de las Administraciones públicas, la ciudadanía y las empresas que quieran hacer uso de ellos, empleando el modelo de licenciamiento o utilización que más se ajuste a su objeto y características.

3. Específicamente, facilitar al Ministerio para la Transformación Digital y de la Función Pública la documentación y los procedimientos que se han seguido para producir estos recursos lingüísticos digitales con la finalidad de apoyar a la Iniciativa ALIA y a otros proyectos de los que sea competente el mismo.

4. Financiar las actuaciones detalladas anteriormente para dar cumplimiento al presente convenio, conforme a lo establecido en el anexo II.

Tercera. *Financiación, pago y liquidación.*

1. Las actuaciones a realizar por las partes, descritas en el anexo I del presente convenio, se financiarán de manera conjunta por la Administración General del Estado y la Administración Autonómica. La Administración General de la Comunidad Autónoma de Euskadi financiará la ejecución de las actuaciones descritas en el presente convenio con una cuantía de 5,5M€. Por otro lado, la SEDIA aportará un total de 5M€ procedentes de los Presupuestos Generales del Estado entre los años 2025-2028.

2. El Departamento de Industria, Transición Energética y Sostenibilidad del Gobierno de Euskadi financiará con una cuantía de 5,5M€, el desarrollo de las actuaciones recogidas en el anexo I, relativo a las «actividades financiadas», con el desglose establecido en el anexo II, que deban ejecutarse con arreglo a su presupuesto. A continuación, se recogen las aplicaciones presupuestarias y las anualidades.

Aplicación presupuestaria	Cuantía - Euros
2025.1000611100.1.54130.75301.2025/000700.	2.000.000
2026.1000611100.1.54130.75301.2025/000700.	3.000.000
2025.1000222220.1.47112.45200.2025/000900.	500.000
Total.	5.500.000

3. La SEDIA financiará con una cuantía de 5M€ las actuaciones recogidas en el anexo I, relativo a las actividades financiadas, con el desglose establecido en el anexo II, con cargo a la partida presupuestaria 33.04.467G.751 de los presupuestos de la Secretaría de Estado de Digitalización e Inteligencia Artificial, de acuerdo con la distribución de anualidades que se expone en el apartado 4.2 de esta cláusula Tercera.

4. Los pagos de las cantidades acordadas se realizarán, mediante transferencia bancaria de la Administración General del Estado, a la cuenta designada por la Administración General de la Comunidad Autónoma de Euskadi de acuerdo con los siguientes plazos:

1. Tras la firma y entrada en vigor de este convenio, el 10 % del presupuesto del Ministerio, es decir, 500.000 euros, en el ejercicio 2025.

2. A partir del año 2026, las transferencias, previa certificación de las actuaciones realizadas, se harán de acuerdo con el siguiente calendario:

Año	Euros
2026	1.500.000
2027	1.500.000
2028	1.500.000

Las transferencias tendrán periodicidad semestral.

5. Para efectuar el pago por parte de la Administración General del Estado no se exigirá garantía alguna, dado el carácter de Administración Pública de la Comunidad Autónoma de Euskadi.

6. Acreditación de la realización de las actuaciones conveniadas. La Administración General de la Comunidad Autónoma de Euskadi, a través del Departamento de Industria, Transición Energética y Sostenibilidad, en cumplimiento de los compromisos adquiridos, realizará la correspondiente liquidación de los fondos recibidos.

Cuarta. *Puesta a disposición de los resultados del Proyecto.*

De conformidad con el Real Decreto Legislativo 1/1996, de 12 de abril, que aprueba el texto refundido de la Ley de Propiedad Intelectual, la propiedad intelectual y los derechos de explotación sobre la generación del corpus digital, desarrollos y resultados del Proyecto Euskorpus corresponderán a las entidades que participen en el desarrollo de este proyecto, sin perjuicio de los derechos morales que pudieran corresponder a los autores de los mismos así como de los derechos morales y de explotación de los autores de las obras puestas a disposición para la generación del corpus digital.

Dichos resultados consistirán, entre otros, en un corpus digital de datos y modelos de propósito general.

En este sentido, la Administración General de la Comunidad Autónoma de Euskadi realizará las actuaciones necesarias para que se conceda al Ministerio para la Transformación Digital y de la Función Pública una licencia de uso, gratuita, no exclusiva y no sublicenciable, excepto a otras Administraciones públicas, sobre los resultados que

permitan su uso en los modelos ALIA (reproducibilidad y reutilización) y que sea permisiva con la licencia con la que se distribuyan estos modelos. En cuanto al acceso a contenidos para la generación del corpus digital en euskera, este no conllevará ninguna obligación de pago a autores, ni empresas, entidades, organizaciones, instituciones o personas físicas que los aporten y así se establecerá en los contratos de cesión de contenidos. En consecuencia, no será objeto de gasto a imputar al presente convenio ningún importe por el acceso a contenidos.

Quinta. Protección de datos y seguridad de la información.

El tratamiento de los datos personales y las posibles cesiones o comunicaciones de datos necesarios para la ejecución y cumplimiento del presente convenio se llevarán a cabo conforme a la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, y al Reglamento (UE) 679/2016 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales, y la legislación nacional en la materia.

Las partes del presente convenio se comprometen al cumplimiento de los requerimientos sobre protección de datos establecidos por la normativa referida anteriormente, en particular en lo referente a la finalidad del tratamiento, a la base de legitimación del tratamiento, el periodo de conservación de los datos, a las cesiones o comunicaciones de datos y a los derechos de los interesados, así como a la adopción de las medidas de seguridad necesarias.

Las partes quedan informadas de que los datos de contacto que constan en el convenio serán tratados con la finalidad de permitir el desarrollo, cumplimiento y control del mismo, siendo la base del tratamiento el cumplimiento de la relación establecida y conservándose los datos durante todo el tiempo en que esta subsista y aún después, hasta que prescriban las eventuales responsabilidades derivadas de ella.

Los datos de las partes podrán ser comunicados a terceros únicamente en la medida en que fuera necesario para cumplir con sus respectivas obligaciones contractuales o legales de conformidad con la normativa vigente. Los datos tratados a efectos de suscribir el presente convenio son necesarios para la efectiva suscripción del mismo; la negativa a facilitarlos podría suponer la imposibilidad de su suscripción. Los afectados podrán solicitar el acceso a los datos personales, su rectificación, supresión, portabilidad y la limitación de su tratamiento, así como oponerse al mismo, y podrán interponer, en caso de vulneración de derechos, una reclamación ante la Agencia Española de Protección de Datos, sita en la calle Jorge Juan número 6, Madrid (www.aepd.es). Se informa de que no se toman respecto de los datos decisiones automatizadas, incluida la elaboración de perfiles.

Sexta. Comisión de Seguimiento, vigilancia y control del convenio.

De acuerdo con el artículo 49, letra f), de la Ley 40/2015, de 1 de octubre, el seguimiento, vigilancia y control del convenio, se efectuará por una Comisión de Seguimiento compuesta por ocho (8) miembros, cuatro por cada una de las partes firmantes.

Esta Comisión estará constituida por cuatro representantes designados por parte del Ministerio para la Transformación Digital y de la Función Pública y cuatro representantes designados por parte de la Administración General de la Comunidad Autónoma de Euskadi.

La sustitución, con carácter definitivo o temporal, de cualquiera de los miembros de la Comisión de Seguimiento, será anunciada en la convocatoria de la reunión correspondiente y quedará debidamente reflejada en el acta de la misma.

La Comisión de Seguimiento se reunirá como mínimo con carácter semestral, y siempre que lo solicite uno de sus miembros. En su reunión constitutiva se establecerán las normas de funcionamiento de la misma, así como el calendario de reuniones.

La presidencia de la Comisión de Seguimiento corresponderá a la persona designada por parte del titular del Ministerio para la Transformación Digital y de la Función Pública, y tendrá carácter rotatorio. La rotación se llevará a cabo de manera anual.

A su vez, la secretaría de la Comisión de Seguimiento corresponderá a la persona de menor edad entre las personas que forman parte de la Comisión de Seguimiento, exceptuando a la persona que ejerza la presidencia de la misma, que actuará con voz y con voto. Al secretario, o secretaria, le corresponderá:

- La convocatoria de las reuniones a instancia de la Presidencia, cuando así lo decida la Comisión, la preparación de los trabajos y levantar las correspondientes actas.
- La recopilación y elaboración de informes para facilitar la toma de decisiones.
- Levantar acta de las sesiones y traslado de las mismas a las partes para su aprobación. El acta de las reuniones se aprobará por la Comisión, bien al concluir la sesión, bien en la reunión siguiente.
- La tramitación y, en su caso, ejecución de aquellos acuerdos o decisiones de la Comisión que se le encomienden expresamente.

A la Comisión de Seguimiento podrán asistir invitados de cada una de las partes, que actuarán con voz, pero sin voto.

Los responsables estarán en estrecho contacto, comunicándose o reuniéndose siempre que lo consideren conveniente para la buena coordinación de las acciones objeto del presente convenio y el óptimo desarrollo de las mismas.

En concreto, la Comisión de Seguimiento se encargará de:

- a) Conocer la actividad y situación de las actuaciones a desarrollar en el marco de este convenio.
- b) La resolución de las cuestiones relativas a la interpretación y cumplimiento de los compromisos derivados del presente convenio, así como proponer a las partes firmantes cualquier modificación del mismo y decisiones relativas al uso de los datos y resultados alcanzados en el marco del proyecto.
- c) Velar por el debido cumplimiento de este convenio y resolver los problemas de interpretación que pudieran plantearse respecto al mismo, así como la supervisión de la correcta ejecución de las actuaciones integrantes del presente convenio.
- d) La recepción de la comunicación del requerimiento por incumplimiento de las obligaciones y compromisos por parte de alguno de los firmantes prevista en el artículo 51, apartado 2, letra c), de la Ley 40/2015, de 1 de octubre.
- e) La propuesta prevista en el artículo 52, apartado 3, de la Ley 40/2015, de 1 de octubre, sobre las actuaciones en curso de ejecución cuando concurra una causa de resolución del convenio.
- f) Cualesquiera otras que se deriven de este convenio.

Las reuniones de la Comisión de Seguimiento se realizarán, preferentemente, por medios electrónicos.

Las partes acuerdan expresamente que las dudas que puedan surgir en el funcionamiento de esta Comisión de Seguimiento y que no estén reguladas en el presente convenio, se resolverán aplicando los principios y criterios establecidos en la Ley 40/2015, de 1 de octubre, en lo que se refiere al funcionamiento de los órganos colegiados.

La Comisión citada se constituirá en el plazo de treinta días naturales desde la entrada en vigor de este convenio.

Séptima. *Naturaleza del presente convenio.*

El presente convenio tiene naturaleza administrativa, y se registrará por lo dispuesto en el capítulo VI del título preliminar de la Ley 40/2015, de 1 de octubre, no teniendo

carácter de contrato administrativo ni subvención, sino que busca satisfacer una finalidad pública de la específica competencia de cada una de las partes.

El objeto y naturaleza de este convenio se enmarcan en el ámbito de una cooperación entre distintos organismos del sector público para el logro de un fin público, mediante la coordinación y el ejercicio de competencias entre todas ellas. Así pues, las partes llevarán a cabo las actividades científico-técnicas en el marco de sus propias competencias, fines y estrategias.

Octava. Extinción del convenio y efectos.

Las causas de extinción del convenio son las recogidas en el artículo 51 de la Ley 40/2015, de 1 de octubre, ya sea por el cumplimiento de las actuaciones que constituyen su objeto o bien por incurrir en causa de resolución.

Son causas de resolución:

- a) El transcurso del plazo de vigencia del convenio sin haberse acordado la prórroga del mismo, o término del plazo de vigencia de la prórroga en su caso.
- b) El acuerdo unánime de las partes.
- c) En el caso de que una de las partes considere que se ha producido incumplimiento de las obligaciones respectivas establecidas en la cláusula segunda, requerirá fehacientemente su cumplimiento dentro del plazo de quince días desde el envío del requerimiento, con las obligaciones o compromisos que se consideran incumplidos. Este requerimiento será comunicado a la Comisión de Seguimiento prevista en la cláusula sexta del presente convenio. Si trascurrido dicho plazo persistiera el incumplimiento, la parte que lo dirigió notificará fehacientemente a la otra parte la concurrencia de la causa de resolución y se entenderá resuelto el convenio, sin perjuicio de la liquidación de las obligaciones económicas y técnicas pendientes que se dirimirán en la citada Comisión de Seguimiento.
- d) La decisión judicial declaratoria de la nulidad del convenio.
- e) Por cualquier otra causa distinta de las anteriores prevista en el convenio o en otras leyes.

Asimismo, las partes firmantes podrán resolver este convenio por denuncia de cualquiera de las mismas mediante preaviso comunicado de forma fehaciente a la otra parte con, al menos, tres meses de antelación a la fecha de resolución propuesta.

El cumplimiento o la resolución del convenio darán 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, de 1 de octubre.

No obstante, si cuando concurra cualquiera de las causas de resolución del convenio existen actuaciones en curso de ejecución, las partes, a propuesta de la Comisión de Seguimiento, podrán acordar la continuación y finalización de aquellas que consideren oportunas, estableciendo un plazo improrrogable para su finalización, transcurrido el cual deberá realizarse la liquidación de las mismas en los términos establecidos en el párrafo anterior.

Novena. Régimen de resolución de conflictos.

Las partes firmantes se comprometen a solventar por acuerdo mutuo, en el seno de la Comisión de Seguimiento, descrita en la cláusula sexta, cuantas discrepancias resulten de la interpretación y cumplimiento de este convenio.

Las partes se consultarán inmediatamente para la solución de las dificultades que pudieran surgir en la aplicación del presente convenio.

Las cuestiones litigiosas surgidas sobre la interpretación, modificación, resolución y efectos que pudieran derivarse de la aplicación del presente convenio que se establezcan al amparo del mismo, se someterán ante el orden jurisdiccional contencioso-administrativo, de acuerdo con la Ley 29/1998, de 13 de julio, reguladora de la Jurisdicción Contencioso-administrativa.

Décima. *Eficacia, vigencia, duración, modificación y prórroga del convenio.*

El convenio tiene una vigencia de tres (3) años a contar desde la fecha de la inscripción en el Registro Electrónico estatal de Órganos e Instrumentos de Cooperación (REOICO) y podrá prorrogarse por el mismo plazo de común acuerdo de las partes.

De acuerdo con lo establecido en el artículo 48, apartado 8 de la Ley 40/2015, de 1 de octubre, el presente convenio se perfecciona por la prestación del consentimiento de las partes, en este caso a través de su firma, resultando eficaz desde el momento en que se inscriba en el Registro Electrónico estatal de Órganos e Instrumentos de Cooperación (REOICO) del sector público estatal.

Posteriormente, será publicado en el «Boletín Oficial del Estado» (BOE) en cumplimiento del artículo 48, apartado 8 de la Ley 40/2015, de 1 de octubre, y también será puesto a disposición de los ciudadanos en el Portal de Transparencia en la forma exigida por el artículo 8.1 b) de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno.

El presente convenio podrá modificarse por mutuo acuerdo de las partes, mediante la tramitación de adendas, cuando resulte necesario para la mejor realización de su objeto siguiendo los mismos trámites establecidos para su suscripción. Y en prueba de conformidad, y para que así conste a los efectos oportunos, las partes firman el presente documento, en la fecha indicada en la última firma digital.

Madrid, 23 de octubre de 2025.—El Ministro para la Transformación Digital y de la Función Pública, Óscar López Águeda. El Consejero de Industria, Transición Energética y Sostenibilidad, Mikel Jauregi Letemendia.

ANEXO I

Actividades financiadas

La Comunidad Autónoma de Euskadi llevará a cabo las actividades necesarias para realizar una primera fase de la iniciativa Euskorpus cuyo objetivo final es avanzar en las tecnologías de la lengua para situar el euskera entre las lenguas europeas con elevado nivel tecnológico, y para ello es importante que esté integrado en el desarrollo de la Inteligencia Artificial, de forma que se puedan ofrecer servicios de calidad en euskera, tanto a las Administraciones públicas como al conjunto de la sociedad.

Este proyecto va a posibilitar el desarrollo de recursos lingüísticos (datos y modelos lingüísticos), instrumentos para el tratamiento de la lengua (analizadores, etiquetadores, anonimizadores) y herramientas, además de fomentar su uso. De esta manera, se irá construyendo la infraestructura lingüística necesaria para seguir integrando el euskera en las tecnologías del lenguaje.

El euskera debe ser una herramienta clave para el desarrollo en Euskadi, en un escenario digital y globalizado, por lo que hay que contribuir a la creación del corpus digital en euskera. El corpus, escrito y hablado, el léxico computacional, las bases de datos terminológicas, la recopilación de voz, etc. son recursos lingüísticos, entendidos como un conjunto de datos y descripciones de voz o lenguaje en formato legible por máquina, utilizados para crear, mejorar o evaluar sistemas o algoritmos de lenguaje natural y de voz.

La complejidad inherente del patrimonio lingüístico, que abarca desde marcos legales hasta recursos lingüísticos y consideraciones tecnológicas, demanda un enfoque personalizado que considere las particularidades de cada comunidad lingüística y sus necesidades específicas. Esta iniciativa representa un paso crucial hacia el desarrollo de una metodología integral adaptable a otras lenguas minoritarias que enfrentan desafíos similares. Al atender cuestiones como la digitalización de recursos, los derechos de propiedad intelectual y el desarrollo de tecnología lingüística y modelos base, se allana el camino para futuros esfuerzos destinados a salvaguardar y revitalizar las lenguas.

A continuación, se describen más en detalle las principales actuaciones a poner en marcha para la consecución del objeto de este convenio agrupadas en tres grandes bloques:

1. Actuaciones encaminadas a la identificación de los sectores, datos y modelos que pueden generar un mayor impacto en el ámbito del euskera en digital.

Dado el escaso nivel de aplicación de las tecnologías del habla y el lenguaje en euskera en todos los sectores económicos, la identificación de ámbitos estratégicos en los que focalizar los desarrollos tecnológicos del euskera es la base para la consecución del objeto del presente convenio.

Se analizarán las áreas estratégicas donde estas tecnologías pueden tener un impacto más significativo, como la industria, la salud, los medios de comunicación y el entretenimiento, la educación o la administración pública, entre otras. Apuntar estratégicamente a estos sectores permitirá adaptar las tecnologías del habla y el lenguaje para satisfacer las necesidades y desafíos específicos de la realidad vascoparlante, así como para enriquecer la familia de modelos ALIA con la aportación de corpus en euskera.

En el marco de este bloque de actuaciones se determinará la tipología de corpus y modelos a desarrollar y se recopilarán los datos necesarios de habla y de texto, bilingües y monolingües en todas las variantes del euskera. Asimismo, se priorizarán aquellos sectores, aplicaciones y servicios relacionados con las tecnologías del lenguaje en euskera que se consideren más estratégicos por su impacto económico o social.

Dentro de este bloque de actuaciones se identificarán las entidades públicas o privadas que dispongan de datos relevantes y cuya participación se considere de utilidad para el desarrollo del proyecto.

2. Preparación del corpus y operación del corpus en euskera.

La creación del corpus digital del euskera llena un vacío crítico en los recursos lingüísticos existentes. Los datos compilados servirán como recursos invaluable para facilitar a administraciones, investigadores, desarrolladores y empresas avanzar en la implementación y la evaluación de sistemas de procesamiento del habla y lenguaje diseñados específicamente para el euskera.

El objetivo es construir un corpus que permita la generación activa de recursos lingüísticos accesibles para el desarrollo tecnológico del euskera en esta nueva era digital, recursos lingüísticos que resultan estratégicos para el desarrollo de modelos y aplicaciones digitales avanzadas.

Para desarrollar este corpus en euskera se han de definir los estándares de calidad a cumplir y, tras identificar las fuentes de datos en euskera más relevantes para los sectores, aplicaciones y servicios de alto impacto, se evaluará su calidad de acuerdo con los estándares definidos.

La compilación del corpus digital del euskera incluirá miles de horas de audios y millones de segmentos de textos bilingües y monolingües anotados en euskera, con calidad contrastada, lo que representa una novedad importante, sobre todo teniendo en cuenta la escasez de corpus de habla y texto disponible para este idioma en la actualidad. De esta forma se enriquecerán los modelos desarrollados en castellano y lenguas cooficiales, en los que el euskera tiene hasta ahora una presencia muy limitada.

Una vez recopilados y definidos los conjuntos de datos, se elaborarán guías de anotación. La anotación y el etiquetado son esenciales para los algoritmos de aprendizaje supervisado, ya que proporcionan la información necesaria para que los modelos aprendan patrones y hagan predicciones con precisión.

Por otra parte, para generar corpus digitales para procesamiento de habla, los audios a compilar y generar serán de la mayor diversidad posible en términos de dialectos, formas de habla, variabilidad acústica, hablantes, etc., y serán anotados con diferentes niveles de profundidad para poder entrenar modelos para múltiples aplicaciones, como

reconocimiento y síntesis de habla, reconocimiento de emociones, sistemas biométricos, anti-fraude o modelos para la generación de biomarcadores para ser aplicados en sectores como la salud.

Sobre estos corpus se implementarán modelos base de código abierto para el procesamiento del habla y el lenguaje en euskera, lo que representa una novedad significativa, especialmente dada la escasez de dichos modelos disponibles actualmente. Además, la naturaleza de código abierto de los modelos fomentará la colaboración, la innovación y el intercambio de conocimientos dentro de la comunidad de tecnología lingüística, promoviendo el crecimiento y la sostenibilidad de las iniciativas tecnológicas del euskera a largo plazo y enriqueciendo los modelos de lenguaje en castellano y lenguas cooficiales.

Los modelos a desarrollar ofrecerán oportunidades para hacer avanzar la tecnología lingüística en euskera. Aprovechando estos modelos, investigadores y desarrolladores podrán acelerar el desarrollo del reconocimiento de voz, la comprensión del lenguaje natural o los sistemas de traducción automática para el euskera, permitiendo una comunicación más precisa, eficiente y accesible en dicho idioma.

Para llevar a cabo estas actuaciones es necesario implementar y poner en marcha la infraestructura que permita almacenar de forma segura los distintos corpus que se vayan compilando, y que permita probar y validar los modelos de código abierto para el euskera que se vayan desarrollando.

La estructuración y el procesamiento de todos los datos, incluyendo la anotación de corpus y la generación de modelos de lenguaje, tiene como objetivo garantizar su explotación, de manera que pueda ponerse al servicio del desarrollo de asistentes conversacionales, de traducción y otras aplicaciones digitales avanzadas vinculadas al procesamiento del lenguaje natural en euskera, y se consiga así la transferencia de dichos corpus y modelos a las Administraciones públicas, las empresas, la sociedad y otras plataformas de datos europeas.

3. Actuaciones de acompañamiento para garantizar la legalidad, así como la promoción y difusión del euskera en digital.

El uso de recursos o datos lingüísticos para el desarrollo de tecnologías de procesamiento del habla y lenguaje implica que dichos datos estén a disposición de desarrolladores en condiciones técnicas y jurídicas que legitimen y faciliten su utilización. Para ello, resulta esencial tener en cuenta la normativa aplicable y los aspectos legales relacionados con los datos lingüísticos de cara a determinar con qué finalidad y en qué condiciones es posible su utilización y compartición.

De esta forma, el uso de los recursos lingüísticos conecta de forma directa con cuestiones relativas a la propiedad intelectual y la protección de datos personales. Asimismo, deberá tenerse en cuenta la normativa europea específica que regula el uso, la compartición y puesta a disposición de los datos.

Por todo ello, en el marco del convenio habrá de definirse el abanico de licencias que permita llevar a cabo la explotación de los corpus y modelos resultantes de acuerdo con la estrategia fijada.

Fomentar la explotación de las tecnologías del habla y el lenguaje en euskera en diferentes sectores económicos representa un esfuerzo pionero, sobre todo teniendo en cuenta la escasez actual de este tipo de aplicaciones. La iniciativa de integrar estas tecnologías en varios sectores estratégicos presenta ventajas y oportunidades. Al aprovechar el poder del reconocimiento de voz, el procesamiento del lenguaje natural o los sistemas de traducción automática diseñados específicamente para el euskera, las organizaciones pueden mejorar su eficiencia operativa, la atención al cliente e impulsar la innovación.

Además, la integración de las tecnologías del habla y el lenguaje promueve la inclusión digital y la diversidad lingüística, permitiendo a los vascoparlantes acceder e interactuar con contenidos y servicios digitales en su lengua materna. Al fomentar la colaboración y el intercambio de conocimientos entre las partes interesadas, esta

iniciativa puede estimular el crecimiento económico, crear oportunidades de empleo y posicionar el euskera como un modelo de innovación tecnológica lingüística en el mercado global.

Para promocionar y difundir las actuaciones recogidas en este convenio, es necesario establecer un canal de comunicación estable, dinámico y efectivo con todas las instituciones y organizaciones implicadas en el impulso de las tecnologías del lenguaje en euskera, pero también articular actuaciones de carácter divulgativo orientadas a hacer partícipe al conjunto de la sociedad a través de actividades como la organización de talleres, sesiones de capacitación e iniciativas de colaboración para crear conciencia sobre la disponibilidad y utilidad de estos recursos lingüísticos, fomentar su adopción e integración en diversas aplicaciones y servicios, y facilitar la colaboración y el intercambio de conocimientos entre los sectores implicados.

ANEXO II

Desglose presupuestario de las actuaciones

Los costes asociados las iniciativas anteriormente contempladas, a ejecutar en 2025-2028, se especifican en la siguiente tabla según las actuaciones a poner en marcha en el marco de este convenio de colaboración.

	MTDFP	Consejería de Industria, Transición Energética y Sostenibilidad	Total
Identificación de los sectores, datos y modelos.	550.000	642.650	1.192.650
Preparación y operación del corpus.	3.950.000	4.269.850	8.219.850
Actuaciones en el ámbito legal, promoción y difusión.	500.000	587.500	1.087.500
Total.	5.000.000	5.500.000	10.500.000