

III. OTRAS DISPOSICIONES

MINISTERIO DE LA PRESIDENCIA Y PARA LAS ADMINISTRACIONES TERRITORIALES

- 538** *Resolución de 9 de enero de 2018, de la Subsecretaría, por la que se publica la Adenda al Convenio marco entre la Real Academia Española y la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital, para el desarrollo de la industria del lenguaje.*

El Director de la Real Academia Española y el Secretario de Estado para la Sociedad de la Información y la Agenda Digital del Ministerio de Energía, Turismo y Agenda Digital han suscrito, con fecha 20 de diciembre de 2017, una Adenda al Convenio Marco entre la Real Academia Española y la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital del Ministerio de Energía, Turismo y Agenda Digital para el desarrollo de la industria del lenguaje.

Para general conocimiento, y en cumplimiento de lo establecido en el artículo 48.8 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, dispongo la publicación en el «Boletín Oficial del Estado» del referido Convenio, como anexo a la presente Resolución.

Madrid, 9 de enero de 2018.—El Subsecretario de la Presidencia y para las Administraciones Territoriales, José María Jover Gómez-Ferrer.

ANEJO

Adenda al Convenio Marco entre la Real Academia Española y la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital del Ministerio de Energía, Turismo y Agenda Digital para el desarrollo de la industria del lenguaje

En Madrid, a 20 de diciembre de 2017.

REUNIDOS

De una parte, don Darío Villanueva Prieto, Director de la Real Academia Española desde el 8 de enero de 2015 por Acuerdo del Pleno de la Real Academia Española de 11 de diciembre de 2011, en el ejercicio de las competencias atribuidas por el Real Decreto 1109/1993, de 9 de julio, por el que se aprueban los Estatutos de la Real Academia Española.

De otra parte, don José María Lassalle Ruiz, Secretario de Estado para la Sociedad de la Información y la Agenda Digital del Ministerio de Energía, Turismo y Agenda Digital, en virtud del nombramiento efectuado mediante Real Decreto 516/2016, de 18 de noviembre, en el ejercicio de las competencias delegadas por el Ministro mediante Orden IET/556/2012, de 12 de marzo, «BOE» de 19 de marzo, conforme a lo establecido en la disposición adicional tercera del Real Decreto 424/2016, de 11 de noviembre, por el que se establece la estructura orgánica básica de los departamentos ministeriales.

Reconociéndose ambas partes la capacidad legal para suscribir la presente adenda y, en su virtud,

EXPONEN

Primero.

Que con fecha 19 de mayo de 2016 la Real Academia Española (en adelante RAE) y la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital (en adelante SESIAD) del Ministerio de Energía, Turismo y Agenda Digital suscribieron un Convenio Marco para el Desarrollo de la Industria del Lenguaje cuyo objetivo es fijar las bases de la colaboración entre la RAE y la actual SESIAD en la implementación del Plan Nacional de Impulso a las Tecnologías.

Segundo.

La Real Academia Española, según establece el artículo primero de sus actuales estatutos, aprobados mediante el Real Decreto 1109/1993, de 9 de julio, es una institución con personalidad jurídica propia que tiene como misión principal velar por que los cambios que experimente la lengua española en su constante adaptación a las necesidades de sus hablantes no quiebren la esencial unidad que mantiene en todo el ámbito hispánico. Debe cuidar igualmente de que esta evolución conserve el genio propio de la lengua, tal como este ha ido consolidándose con el correr de los siglos, así como de establecer y difundir los criterios de propiedad y corrección, y de contribuir a su esplendor.

Tercero.

El artículo 62 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, señala que corresponde a los Secretarios de Estado la responsabilidad de la ejecución de la acción del Gobierno en un sector de actividad específica de un Departamento actuando bajo la dirección de su titular, siendo el apartado 2.g) del mencionado artículo 62 el que le atribuye la competencia para suscribir convenios.

De acuerdo con lo dispuesto en el artículo 5 del Real Decreto 903/2017, de 13 de octubre, por el que se desarrolla la estructura orgánica básica del Ministerio de Energía, Turismo y Agenda Digital, le corresponde a la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital de dicho Ministerio, entre otras funciones, la promoción y desarrollo de las infraestructuras y servicios avanzados de telecomunicaciones y de la Sociedad de la Información, así como la elaboración, gestión y seguimiento de planes, proyectos tecnológicos y programas de actuaciones orientados al desarrollo de las infraestructuras, el fomento de la oferta de nuevas tecnologías, servicios, aplicaciones y contenidos en el ámbito de las telecomunicaciones y la Sociedad de la Información.

Cuarto.

El Consejo de Ministros, a propuesta del extinto Ministro de Industria, Energía y Turismo y del Ministro de Hacienda y Administraciones Públicas, aprobó, en su sesión de 15 de febrero de 2013, la Agenda Digital para España. En el marco de dicha Agenda, el 7 de octubre fue aprobado el Plan de Impulso de las Tecnologías del Lenguaje promovido por la extinta Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información. El Plan tiene un alcance temporal de cinco años y tiene por objeto fomentar el desarrollo del sector del procesamiento del lenguaje natural y la traducción automática en España.

Este Plan se diseña en apoyo a un sector, el de las tecnologías del lenguaje, cuyo diagnóstico de situación se podría resumir en las siguientes ideas clave:

– Alto potencial de crecimiento y desarrollo: las tecnologías del lenguaje representan un sector emergente transversal vinculado a la innovación con capacidad para promover crecimiento, competitividad y empleo de calidad.

– Oportunidad única: su desarrollo industrial, paralelo al desarrollo tecnológico, es imparable, y dada la importancia del idioma español en el mundo, la oportunidad es única para que ocupemos un espacio que sin duda alguien ocupará.

– Recursos disponibles pero dispersos: España dispone de los recursos necesarios para acometer ese desarrollo, pero para tener éxito, es imprescindible impulsar y coordinar las actuaciones pertinentes desde la Administración General del Estado, en coordinación con las Comunidades Autónomas y en colaboración con Iberoamérica y la Unión Europea, para aprovechar esta oportunidad.

El Plan de Impulso de las Tecnologías del Lenguaje se articula en torno a cuatro ejes: Eje I. Desarrollo de infraestructuras lingüísticas; Eje II. Impulso de la industria de las tecnologías del lenguaje; Eje III. Las Administraciones Públicas como impulsoras de la industria del lenguaje; y Eje IV. Proyectos faro.

Quinto.

La cláusula tercera del Convenio Marco para el Desarrollo de la Industria del Lenguaje establece, entre las obligaciones de la SESIAD, promover la participación de la Real Academia Española en el Plan de impulso de Tecnologías del Lenguaje. En especial, facilitar su cooperación en el desarrollo de infraestructuras lingüísticas previsto en el Eje I y su participación en proyectos faro del Eje IV del Plan.

Por otro lado, la cláusula cuarta establece, entre las obligaciones de la RAE, asesorar técnica y científicamente a la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital en la implementación del Plan de Impulso de las Tecnologías del Lenguaje.

Sexto.

Asimismo, la cláusula tercera del Convenio Marco para el Desarrollo de la Industria del Lenguaje establece, entre las obligaciones de la SESIAD, la siguiente:

«Colaborar en el desarrollo e impulso de estándares aplicados a recursos lingüísticos (anotaciones lingüísticas, estructura de recursos, metadatos, etc.). Promover el desarrollo de recursos lingüísticos estandarizados en dominios de información especializados, en aquellas áreas de trabajo propias de la Real Academia Española (ej. impulsar el uso de una planta común de diccionarios especializados).»

Por otro lado, la cláusula cuarta establece, entre las obligaciones de la RAE, «Colaborar en el diseño de la arquitectura o planta para el desarrollo de aquellas infraestructuras lingüísticas (recursos, procesadores y campañas de evaluación lingüísticos) de interés para el Plan de Impulso de las Tecnologías del Lenguaje», «Compartir herramientas abiertas y participar en el desarrollo de herramientas abiertas para la creación y mantenimiento de recursos lingüísticos (entornos de trabajo lexicográfico, etc.)», «Incrementar el número de datos abiertos sobre recursos lingüísticos de la Real Academia Española de interés industrial y académico». Por ello, conviene a ambas partes concretar la construcción de un corpus de español actual que sirva de base y referencia para la tecnología cuyo desarrollo pretende fomentar el Plan de Impulso de las Tecnologías del Lenguaje.

La tecnología no solo ha contribuido a la modernización de la metodología de trabajo de la RAE y a potenciar la difusión de sus resultados, sino que favorece el mantenimiento de la unidad en un idioma hablado por quinientos millones de personas en cuatro continentes. Mediante el Convenio Marco y la presente Adenda se pretende que represente el mismo rol en las tecnologías del lenguaje. A su vez, se pretende utilizar las relaciones de la RAE con las restantes Academias de la lengua con el fin de introducir estas tecnologías en el resto de países americanos.

Los corpus anotados lingüísticamente no solo son de utilidad para estudiar una lengua desde el punto de vista teórico, sino que son imprescindibles para el desarrollo de las tecnologías del lenguaje humano. Actualmente, un alto porcentaje de algoritmos para el procesamiento de lenguaje utiliza aprendizaje automático y obtiene los datos de

entrenamiento de corpus anotados lingüísticamente. Además de su uso para la creación de tecnología lingüística, los corpus anotados son imprescindibles para poder evaluar de una manera apropiada distintos sistemas de procesamiento automático. Para el español, existen algunos corpus pero tanto su número, su acceso, como su extensión y profundidad de análisis están lejos de los corpus que existen para otras lenguas, especialmente si se compara para el inglés. Con este corpus, se pretende proporcionar a la industria, datos actuales con los que poder desarrollar tecnologías para el español.

En base a lo expuesto anteriormente, las Partes acuerdan suscribir la presente adenda que se registrá por las siguientes

CLÁUSULAS

Primera. *Objeto de la Adenda.*

Son objetos de la presente adenda:

1. La construcción de un corpus textual de español actual que sirva de base y referencia para la tecnología cuyo desarrollo pretende fomentar el Plan de Impulso de las Tecnologías del Lenguaje. La construcción de este corpus contempla tanto su diseño, adquisición, y marcación, así como la gestión de derechos y licencias, y la documentación y preparación para su acceso y posterior distribución.

2. La anotación semiautomática con revisión manual de un subcorpus del corpus de referencia que permita la construcción y evaluación de procesadores básicos. Por anotación básica se entenderá la segmentación en oraciones y unidades de análisis (palabras, números, signos de puntuación, etc.); el reconocimiento y clasificación de entidades con nombre, fechas, etc.; la lematización y el análisis morfosintáctico; y la identificación de regiones en otras lenguas.

3. La incorporación al subcorpus anterior de anotaciones avanzadas (funciones sintácticas), que igualen los ya existentes para inglés que puedan ser de interés para el desarrollo de aplicaciones innovadoras en procesamiento de lenguaje natural.

4. De forma coordinada con la metodología de evaluación establecida en el Plan de Impulso de las Tecnologías del Lenguaje, la RAE participará en campañas de evaluación nacional e internacional, tanto suministrando datos específicos, como mediante su participación en tareas de evaluación.

Segunda. *Actuaciones.*

Para la construcción y anotación de un corpus de español actual, las actuaciones serán las siguientes:

1. Creación de un grupo de seguimiento.
2. Gestión de derechos.
3. Diseño del corpus.
4. Adquisición de corpus.
5. Preprocesamiento morfosintáctico automático del corpus.
6. Preprocesamiento sintáctico automático del corpus.
7. Creación de un entorno de revisión de la anotación del corpus.
8. Revisión manual del preprocesamiento automático.
9. Explotación.
10. Puesta en producción del acceso público.
11. Desarrollo de un plan de comunicación y promoción.

En el anexo I se incluye mayor detalle de las actuaciones junto con el cronograma correspondiente.

Tercera. *Obligaciones de las Partes.*

1. Obligaciones de la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital.

La SESIAD se obliga a:

A. Colaborar en la recogida de toda la información que sea necesaria para la adecuada ejecución de la Adenda.

B. Poner a disposición del proyecto los medios humanos y materiales necesarios para que se puedan desarrollar las actuaciones previstas en la cláusula segunda.

C. Aportar la financiación prevista en la cláusula sexta de la presente Adenda.

D. Participar en la formulación del proyecto junto con RAE, involucrando para ello si es necesario a expertos en lingüística de corpus.

E. Participar en el seguimiento de la construcción del Corpus de apoyo al Plan de Impulso a las Tecnologías del Lenguaje.

F. Participar en la elaboración de los Planes Anuales de Ejecución para su aprobación por la Comisión de Seguimiento.

2. Obligaciones de la Real Academia Española.

La RAE se obliga a:

A. Colaborar en la recogida de toda la información que sea necesaria para la adecuada ejecución de las actuaciones previstas en la cláusula segunda.

B. Poner a disposición del proyecto recursos que pudieran considerarse necesarios y de los que disponga la RAE para que se puedan desarrollar las actuaciones previstas en la cláusula segunda.

C. Incoar los procedimientos de contratación para la ejecución de aquellas actividades del proyecto que lo requieran.

D. Participar en la formulación del Proyecto junto con la SESIAD.

E. Participar en la elaboración de los Planes Anuales de Ejecución para su aprobación por la Comisión de Seguimiento.

F. Elaborar las Memorias Anuales de Actuaciones para su aprobación por la Comisión de Seguimiento.

Las Partes se comprometen a realizar cuantas acciones resulten precisas para la correcta y completa ejecución de las actuaciones previstas en esta adenda.

En caso de incumplimiento de dichas obligaciones y compromisos por parte de alguno de los firmantes, será de aplicación lo previsto al efecto en la cláusula undécima del Convenio Marco de fecha 19 de mayo de 2016.

Cuarta. *Adhesión de otras entidades.*

La adenda permite la adhesión de otras entidades que quieran participar en el desarrollo del Corpus de apoyo al Plan de Impulso a las Tecnologías del Lenguaje.

A tal fin el órgano de gobierno de dicha entidad interesada presentará su solicitud de adhesión a la adenda ante la Comisión de Seguimiento que se regula en la cláusula séptima. Estas adhesiones se aprobarán siguiendo el trámite establecido para la aprobación de convenios en cada una de las instituciones firmantes.

Las entidades se comprometen a adherirse expresamente a todas y cada una de las cláusulas del Convenio Marco y sus adendas, asumiendo las obligaciones derivadas del mismo y con sujeción a todas sus cláusulas.

Quinta. *Titularidad de los resultados obtenidos.*

La titularidad de los resultados obtenidos en ejecución de la presente Adenda corresponderá de forma conjunta y a partes iguales a la RAE y la SESIAD.

Todos los resultados obtenidos de la ejecución de la presente Adenda se pondrán a disposición de las distintas administraciones públicas para su uso y libre acceso.

Asimismo, los resultados obtenidos podrán –a decisión de SESIAD– ser puestos a disposición del público y la industria a los efectos de su reutilización, en los términos que se definan en el marco del Plan de Impulso de las Tecnologías del Lenguaje. A tales efectos, la RAE acepta que los resultados obtenidos se sujeten a cualesquiera condiciones y licencias de propiedad intelectual que permitan su reutilización (incluidas las licencias de tipo Creative Commons).

La RAE se asegurará de que los resultados del trabajo desempeñado por las empresas a las que se subcontrate permitan el libre acceso y la reutilización en los términos previstos en los dos párrafos anteriores.

Sexta. Obligación de disociación de datos personales.

De acuerdo con el RD 1495/2011, de 24 mayo, por el que se desarrolla la LRISP para el ámbito del sector público estatal, y en previsión de su posible sujeción a la reutilización por parte del público, la RAE se compromete a disociar (anonimizar) cualesquiera datos de carácter personal contenidos en los recursos que pudieran estar afectados por dicho real decreto, en ejecución de la presente Adenda.

Séptima. Aportación económica.

1. Aportación económica de la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital.

La aportación financiera de la SESIAD asciende a un importe máximo de 586.027,60 euros, conforme a la siguiente distribución:

	Porcentaje	Importe – Euros
Contratación de servicios a terceros (por la RAE)	89,93	527.027,60
Recursos materiales.	4,61	27.000,00
Otros gastos	5,46	32.000,00
Total.	100,00	586.027,60

La SESIAD anticipará el 10 % (58.602,76 €) de acuerdo con el artículo 21.3 de la Ley 47/2003, de 26 de noviembre, General Presupuestaria. Dicha aportación se hará efectiva a la firma de la presente Adenda.

2. Aportación económica de la Real Academia Española.

La RAE no llevará a cabo aportación financiera alguna, siendo su aportación exclusivamente en especie, justificándose porque se trata de un organismo sin parangón, tanto por el conocimiento del corpus del español como por el número de especialistas dedicados a dicha tarea.

En el anexo II.1 se incluye el desglose presupuestario y de ejecución por actuaciones tanto de la SESIAD como de la RAE, recogándose en el anexo II.2 el desglose presupuestario por anualidades.

Los pagos se realizarán de forma trimestral.

En el año 2017 se abonará el anticipo del 10% de acuerdo con el artículo 21.3 de la Ley 47/2003, de 26 de noviembre, General Presupuestaria.

En 2018 se realizará la mayor parte de los abonos, dado que el grueso del trabajo se efectuará en dicha anualidad. A su vez, los recursos materiales y la mitad de «otros gastos» se imputarán al año 2018.

En el año 2019 se abonarán los últimos pagos, correspondientes al primer trimestre de 2019, según la tabla de entregables. A partir de ese momento realizarán las actividades de difusión, de menor importe, hasta la finalización de la Adenda. A dichas actividades se imputará la otra mitad del concepto de «otros gastos».

A su vez, en el anexo III se recoge la aportación en especie de cada una de las partes firmantes.

Sumadas ambas –económicas y en especie–, se observa que la SESIAD aportará el 59,98 % del importe total, frente al 40,02 % de la RAE, tal y como se recoge en el anexo III.3.

Octava. *Entregables.*

Se han definido los entregables asociados a las actividades de los distintos objetivos. En el anexo II.1 se recoge el listado completo de dichos entregables junto a su valoración económica.

Novena. *Seguimiento de las actuaciones.*

El seguimiento de la presente adenda corresponde a la Comisión de Seguimiento del Convenio Marco entre la Real Academia Española y la Secretaría de Estado para la Sociedad de la Información y la Agenda Digital para el desarrollo de la Industria del Lenguaje.

Las funciones de la Comisión de Coordinación y Seguimiento en relación con esta adenda serán las siguientes:

- Evaluar el grado y eficacia de cumplimiento de la adenda y velar por el cumplimiento de las estipulaciones suscritas.
- Interpretar, en caso de duda, las normas y previsiones de la presente adenda. Resolver las controversias que pudieran plantearse sobre la interpretación y ejecución de la presente adenda, que deberán solventarse de mutuo acuerdo entre las partes.
- Cualesquiera otras que se encaminen al cumplimiento de esta adenda o a su desarrollo, conforme a los intereses de las partes.

Décima. *Naturaleza y legislación aplicable.*

La presente adenda se regirá por las cláusulas establecidas en la misma, en el Convenio Marco y, por la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, quedando excluido del ámbito de aplicación del Real Decreto Legislativo 3/2011, de 14 de noviembre, por el que se aprueba el Texto Refundido de la Ley de Contratos del Sector Público al amparo de lo dispuesto en su artículo 4.1.d). No obstante, los principios de dicho texto sí serán de aplicación para resolver las dudas y lagunas que puedan plantearse y no se resolvieran por la comisión de seguimiento y control a que se refiere la cláusula cuarta, como establece el artículo 4.2 del mismo texto legal.

Undécima. *Desarrollo de las actuaciones.*

Las actuaciones de la adenda se realizarán desde su firma hasta la finalización de la vigencia. La comprobación de la ejecución de los compromisos respectivos y su certificación por los firmantes se realizará antes de la finalización de la vigencia de adenda.

Duodécima. *Modificación de la Adenda.*

Para la modificación de la presente Adenda será de aplicación lo previsto en la cláusula novena del Convenio Marco de fecha 19 de mayo de 2016.

La Comisión de Seguimiento podrá precisar o modificar las actuaciones a realizar o reasignar las partidas establecidas en el presupuesto inicial, sin superar en ningún caso el 20 % del presupuesto total de la Adenda.

Décimotercera. *Resolución de la Adenda y consecuencias en caso de incumplimiento.*

Las causas de resolución de la presente Adenda, así como sus consecuencias en caso de incumplimiento, serán las previstas en el Convenio Marco de fecha 19 de mayo de 2016; es decir:

- a) El cumplimiento de las actuaciones que constituyen su objeto.
- b) El transcurso del plazo de vigencia del convenio sin haberse acordado la prórroga del mismo.
- c) Por mutuo acuerdo de las partes o por denuncia de cualquiera de las partes mediante preaviso a la otra parte con al menos un mes de antelación a la fecha de resolución propuesta.
- d) El incumplimiento de las obligaciones y compromisos asumidos por parte de alguno de los firmantes.

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 la Comisión Mixta de Coordinación y Seguimiento.

Si trascurrido 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.

- e) Por decisión judicial declaratoria de la nulidad del convenio.
- f) Por cualquier otra causa distinta de las anteriores prevista en la ley.

Decimocuarta. *Vigencia.*

La presente Adenda al convenio marco entrará en vigor y resultará eficaz una vez inscrito en el Registro estatal de Órganos e Instrumentos de Cooperación del sector Público Estatal y publicado en el Boletín Oficial del Estado, de conformidad con lo dispuesto en el apartado 8 del artículo 48 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público.

La presente Adenda tendrá una vigencia de dos años. En cualquier momento antes de su finalización, los firmantes podrán acordar unánimemente su prórroga por un periodo de hasta dos años adicionales.

Y en prueba de conformidad, las Partes firman la presente adenda en el lugar y fecha señalados en el encabezamiento.–El Secretario de Estado para la Sociedad de la Información y la Agenda Digital, José María Lassalle Ruiz.–El Director de la Real Academia Española, Darío Villanueva Prieto.

ANEXO I

Detalle de actuaciones y cronograma

Actuación	Actividades	Inicio	Finalización
1. Creación grupo de seguimiento.	1.1 Identificar el grupo de seguimiento y definir el acta de constitución del proyecto.	T1/2018	T1/2018
	1.2 Definir el alcance, planificación y presupuesto.	T1/2018	T1/2018
	1.3 Definir y desarrollar la información de gestión y control de proyecto.	T1/2018	T1/2018
	1.4 Gestión, seguimiento y coordinación del proyecto.	T1/2018	T4/2018
	1.5 Campañas de evaluación: Especificar los datos y asistir en la validación para crear los datos <i>gold standard</i> .	T1/2018	T4/2018

Actuación	Actividades	Inicio	Finalización
2. Gestión de derechos.	2.1 Identificación, selección y análisis de proveedores de contenidos.	T1/2018	T4/2018
	2.2 Generar documentación para las entidades de gestión de derechos de la propiedad intelectual.	T1/2018	T2/2018
	2.3 Negociación y seguimiento de la gestión de derechos.	T2/2018	T4/2018
	2.4 Diseño de la licencia para distribución del corpus.	T3/2018	T1/2019
3. Diseño del corpus.	3.1 Establecimiento y desarrollo de parámetros de equilibrio y representatividad.	T1/2018	T2/2018
	3.2 Diseño del esquema de codificación.	T1/2018	T2/2018
	3.3 Diseño del esquema de anotación morfosintáctica.	T1/2018	T2/2018
	3.4 Diseño del esquema de anotación sintáctica.	T1/2018	T2/2018
	3.5 Diseño de metodología y controles de calidad.	T3/2018	T1/2019
4. Adquisición de corpus.	4.1 Adquisición automatizada de textos y metadatos a través de gestores de descarga u otros medios. Control de duplicados.	T1/2018	T2/2018
	4.2 Conversión de formatos e identificación de material codificable.	T2/2018	T3/2018
	4.3 Aplicación de esquema de codificación.	T2/2018	T3/2018
	4.4 Revisión de los archivos codificados.	T2/2018	T3/2018
5. Preprocesamiento morfosintáctico automático del corpus.	5.1 Segmentación de oraciones.	T1/2018	T4/2018
	5.2 Segmentación en unidades de análisis.	T1/2018	T4/2018
	5.3 Análisis léxico.	T1/2018	T4/2018
	5.4 Desambiguación morfosintáctica.	T1/2018	T4/2018
	5.5 Reconocimiento de entidades nombradas.	T1/2018	T4/2018
	5.6 Identificación de regiones en otra lengua.	T1/2018	T4/2018
	5.7 Conversión a formatos estándar.	T2/2018	T1/2019
6. Preprocesamiento sintáctico automático del corpus.	6.1 Entrenamiento y adaptación de un analizador de dependencias.	T3/2018	T1/2019
	6.2 Análisis de dependencias funcionales.	T3/2018	T1/2019
7. Creación de un entorno de revisión de la anotación.	7.1 Instalación y configuración de la plataforma de revisión.	T3/2018	T1/2019
	7.2 Conversión de formatos de la plataforma.	T3/2018	T1/2019
	7.3 Gestión de usuarios de la plataforma.	T3/2018	T1/2019
8. Revisión manual del preprocesamiento automático.	8.1 Revisión según el esquema de anotación morfosintáctico.	T3/2018	T1/2019
	8.2 Revisión según el esquema de anotación sintáctico.	T3/2018	T1/2019
	8.3 Aplicaciones de los controles de calidad.	T2/2018	T1/2019
9. Explotación.	9.1 Adaptación de los indexadores.	T2/2018	T1/2019
	9.2 Adaptación de las herramientas de explotación.	T2/2018	T1/2019
	9.3 Implementación de la interfaz y API de consulta y descarga.	T2/2018	T1/2019
10. Puesta en producción del acceso público.	Puesta en producción del acceso público.	T2/2018	T1/2019
11. Desarrollo de un plan de difusión y comunicación.	Organización de jornadas de difusión.	T1/2018	T1/2019

ANEXO II

II.1 Desglose presupuestario y de ejecución por actuaciones

La valoración económica de las actividades contenidas en este convenio se ha realizado utilizando el conocimiento existente en la SESIAD y la RAE sobre la contratación de trabajos que requieren una dedicación de recursos similar, así como estimando las horas necesarias para llevar a cabo las distintas tareas, aplicando los costes de hora/persona de las consultoras que trabajan con la SESIAD y la RAE.

Actuación	Actividades	Entregables	Presupuesto – (Euros)	Financia	Ejecuta
1. Creación grupo de seguimiento.	1.1 Identificar el grupo de seguimiento y definir el acta de constitución del proyecto.	Documento identificativo de creación del grupo de trabajo y acta de constitución del proyecto.	0,00 ^(*)		RAE.
	1.2 Definir el alcance, planificación y presupuesto.	Reuniones y plan de ejecución.			
	1.3 Definir y desarrollar la información de gestión y control de proyecto.	Plan de gestión.			
	1.4 Gestión, seguimiento y coordinación del proyecto.	Documentos de gestión.			
	1.5 Campañas de evaluación: Especificar los datos y asistir en la validación para crear los datos gold standard.	La validación de las colecciones de datos gold standard.	2.481,60	SESIAD.	RAE.
2. Gestión de derechos.	2.1 Identificación, selección y análisis de proveedores de contenidos.	Relación de fuentes y proveedores.	0,00 ^(*)		RAE.
	2.2 Generar documentación para las entidades de gestión de derechos de la propiedad intelectual.	Documentación para la gestión de los derechos de la propiedad intelectual.			
	2.3 Negociación y seguimiento de la gestión de derechos.	Reuniones y documentos de gestión de derechos.			
	2.4 Diseño de la licencia para distribución del corpus.	Licencias firmadas.			
3. Diseño del corpus.	3.1 Establecimiento y desarrollo de parámetros de equilibrio y representatividad.	Parámetros y distribución.	0,00 ^(*)		RAE.
	3.2 Diseño del esquema de codificación.	Esquema de codificación.			
	3.3 Diseño del esquema de anotación morfosintáctica.	Esquema de anotación morfosintáctica.			
	3.4 Diseño del esquema de anotación sintáctica.	Esquema de anotación sintáctica.			
	3.5 Diseño de metodología y controles de calidad.	Metodología y plan de control de calidad.			
4. Adquisición de corpus.	4.1 Adquisición automatizada de textos y metadatos a través de gestores de descarga u otros medios. Control de duplicados.	Material adquirido.	89.738,00	SESIAD.	RAE.
	4.2 Conversión de formatos e identificación de material codificable.	Material convertido a formatos necesarios.			
	4.3 Aplicación de esquema de codificación.	Material codificado.			
	4.4 Revisión de los archivos codificados.	Archivos codificados y revisados.			

Actuación	Actividades	Entregables	Presupuesto – (Euros)	Financia	Ejecuta
5. Preprocesamiento morfosintáctico automático del corpus.	5.1 Segmentación de oraciones.	Oraciones segmentadas.	52.404,00	SESIAD.	RAE.
	5.2 Segmentación en unidades de análisis.	Texto segmentado.			
	5.3 Análisis léxico.	Texto analizado léxicamente.			
	5.4 Desambiguación morfosintáctica.	Anotación morfosintáctica desambiguada.			
	5.5 Reconocimiento de entidades nombradas.	Texto con entidades identificadas.			
	5.6 Identificación de regiones en otra lengua.	Fragmentos en otras lenguas marcados.			
	5.7 Conversión a formatos estándar.	Textos convertidos a formatos estándar.			
6. Preprocesamiento sintáctico automático del corpus.	6.1 Entrenamiento y adaptación de un analizador de dependencias.	Analizador adaptado y entrenado.	19.536,00	SESIAD.	RAE.
	6.2 Análisis de dependencias funcionales.	Dependencias funcionales adaptadas.			
7. Creación de un entorno de revisión de la anotación.	7.1 Instalación y configuración de la plataforma de revisión.	Plataforma instalada.	8.888,00	SESIAD.	RAE.
	7.2 Conversión de formatos de la plataforma.	Conversión de formatos realizada.			
	7.3 Gestión de usuarios de la plataforma.	Acceso de usuarios a la plataforma.			
8. Revisión manual del preprocesamiento automático.	8.1 Revisión según el esquema de anotación morfosintáctico.	Anotación morfosintáctica revisada.	324.148,00	SESIAD.	RAE.
	8.2 Revisión según el esquema de anotación sintáctico.	Anotación sintáctica revisada.			
	8.3 Aplicaciones de los controles de calidad.	Revisión de calidad.			
9. Explotación.	9.1 Adaptación de los indexadores.	Indexadores adaptados.	23.936,00	SESIAD.	RAE.
	9.2 Adaptación de las herramientas de explotación.	Herramientas de explotación.			
	9.3 Implementación de la interfaz y API de consulta y descarga.	API e interfaz.			
10. Puesta en producción del acceso público.	Puesta en producción del acceso público.	Acceso público.	5.896,00	SESIAD.	RAE.
11. Desarrollo de un plan de difusión y comunicación.	Organización de jornadas de difusión.	Tres jornadas de difusión.	0,00 ^(*)		RAE/SESIAD.
Subtotal			527.027,60		
Recursos materiales.	Equipamiento informático. Software.		27.000,00	SESIAD.	RAE.
Subtotal			27.000,00		
Otros gastos.	Viajes y desplazamientos (expertos, stakeholders, equipo de trabajo, etc.). Participación/asistencia workshops. Adquisición documentación. Difusión de resultados.		32.000	SESIAD.	RAE/SESIAD.
Subtotal			32.000,00		
Total			586.027,60 ^(**)		

(*) Las actuaciones con importe igual a 0,00 euros se financiarán íntegramente mediante aportaciones en especie.

(**) El importe será subcontratado en su totalidad.

II.2 Desglose presupuestario por anualidades

	2017			2018			2019			Total		
	RAE – (Euros)	SESIAD – (Euros)	TOTAL – (Euros)									
Contratación de servicios a terceros (por la RAE).	0,00	0,00	0,00	0,00	415.868,36	415.868,36	0,00	111.159,24	111.159,24	0,00	527.027,60	527.027,60
Recursos materiales.	0,00	0,00	0,00	0,00	27.000,00	27.000,00	0,00	0,00	0,00	0,00	27.000,00	27.000,00
Otros gastos.	0,00	0,00	0,00	0,00	16.000,00	16.000,00	0,00	16.000,00	16.000,00	0,00	32.000,00	32.000,00
Total.			0,00			458.868,36			127.159,24	0,00	586.027,60	586.027,60

No obstante, teniendo en cuenta el anticipo del 10%, las anualidades de la SESIAD quedarían:

	2017	2018	2019	Total
SESIAD	58.602,76 €	458.868,36 €	68.556,48 €	586.027,60 €

Sin embargo, en caso de que la fecha de la firma de la presente Adenda no permita la imputación del 10% al ejercicio 2017, se llevará a cabo con cargo a 2018.

ANEXO III

Aportación en especie

Las siguientes aportaciones de las partes no se incluyen en las aportaciones económicas previstas en la cláusula quinta.

III.1 Aportación en especie de la SESIAD

Actuación	Dedicación recursos propios (número de horas)			Importe por actuación – (Euros)
	Coordinador	Jefe de área	Consejero técnico	
1. Creación grupo de seguimiento	32,85	25,23	8,80	1.946,14
2. Gestión de derechos	39,89	35,79	25,23	2.902,32
3. Diseño del corpus	6,45	6,45	6,45	552,31
4. Adquisición de corpus	0,00	0,00	0,00	0,00
5. Preprocesamiento morfosintáctico automático del corpus ..	0,00	0,00	0,00	0,00
6. Preprocesamiento sintáctico automático del corpus	0,00	0,00	0,00	0,00
7. Creación de un entorno de revisión de la anotación	0,00	0,00	0,00	0,00
8. Revisión manual del preprocesamiento automático	0,00	0,00	0,00	0,00
9. Explotación	0,00	0,00	0,00	0,00
10. Puesta en producción del acceso público	0,00	0,00	0,00	0,00
11. Desarrollo de un plan de difusión y comunicación	5,87	5,87	2,93	422,18
Total horas	85,06	73,34	43,41	
Coste/hora (€)	30,89	27,37	27,37	
Total (€)	2.627,50	2.007,32	1.188,13	5.822,95

Por parte de la SESIAD, se ha calculado con la siguiente tabla de sueldos con una productividad media (sin tener en cuenta trienios ni retribuciones extras por incentivos):

Puestos tipo	Vocal Asesor Nivel 30 Grupo A1		Coordinador Nivel 29 Grupo A1		Jefe de Área nivel 28 Grupo A1		Jefe de servicio nivel 26 Grupo A2		Administrativo Nivel 22 Grupo C1		Puesto tipo Nivel 15 Secretaría Grupo C2	
	Mensual	Anual	Mensual	Anual	Mensual	Anual	Mensual	Anual	Mensual	Anual	Mensual	Anual
Sueldo (14 mensualidades y 2 pagas extra.	1.120,15	14.824,22	1.120,15	14.824,22	1.120,15	14.284,22	968,57	13.035,60	727,23	9.983,82	605,25	8.462,46
Complemento Destino (14 pagas iguales).	978,44	13.698,16	877,62	12.286,68	840,73	11.770,22	705,19	9.872,66	514,94	7.209,16	330,72	4.630,08
Complemento Específico anual (14 pagas iguales).	1.415,94	19.823,16	1.415,94	19.823,16	1.134,35	15.880,90	797,61	11.166,54	329,94	4.619,16	437,84	6.129,76
Productividad Media (12 pagas iguales).	700,00	8.400,00	620,00	7.440,00	520,00	6.240,00	400,00	4.800,00	350,00	4.200,00	320,00	3.840,00
Total mes.	4.214,53		4.033,71		3.615,23		2.871,37		1.922,11		1.693,81	
Total anual.		56.745,54		54.374,06		48.175,34		38.874,80		26.012,14		23.062,30

III.2 Aportación en especie de la RAE

Actuación	Dedicación recursos propios (número de horas)			Importe por actuación — (Euros)
	Director	Jefe de proyecto	Técnico	
1. Creación grupo de seguimiento.	288,93	187,73	45,47	31.248,52
2. Gestión de derechos.	643,87	352,00	230,27	70.960,62
3. Diseño del corpus.	566,13	720,13	142,27	80.757,32
4. Adquisición de corpus.	176,00	498,67	234,67	45.701,62
5. Preprocesamiento morfosintáctico automático del corpus.	157,66	157,66	78,83	21.758,08
6. Preprocesamiento sintáctico automático del corpus.	58,66	58,66	29,33	8.095,08
7. Creación de un entorno de revisión de la anotación.	25,66	40,33	66,00	6.188,70
8. Revisión manual del preprocesamiento automático.	557,33	586,66	1202,66	111.641,86
9. Explotación.	58,66	58,66	22,00	7.831,20
10. Puesta en producción del acceso público.	29,33	29,33	88,00	6.687,60
11. Desarrollo de un plan de difusión y comunicación.	29,33	29,33	14,66	4.047,36
Total horas	2591,56	2719,16	2154,16	
Coste/hora (€)	70,00	50,00	36,00	
Total (€)	181.409,20	135.958,00	77.549,76	394.916,96

III.3 Proporción entre ambas aportaciones

	SESIAD	RAE	Total
Aportación económica	586.027,60 €	—	
Aportación en especie	5.822,95 €	394.916,96 €	
Total	591.850,67 €	394.916,96 €	986.767,51 €
Proporción	59,98 %	40,02 %	100 %