



**D.ª ELENA MACULAN, SECRETARIA GENERAL DE LA
UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA,**

C E R T I F I C A: Que en la reunión del Consejo de Gobierno, celebrada el día siete de marzo de dos mil veintitrés, fue adoptado, entre otros, el siguiente acuerdo:

03. Estudio y aprobación, si procede, de las propuestas del Vicerrectorado de Investigación, Transferencia y Divulgación Científica.

03.02. El Consejo de Gobierno aprueba la creación del Centro de Investigación de Procesamiento del Lenguaje Natural y Recuperación de Información (LENAR), según anexo.

Y para que conste a los efectos oportunos, se extiende la presente certificación haciendo constar que se emite con anterioridad a la aprobación del Acta y sin perjuicio de su ulterior aprobación en Madrid, a ocho de marzo de dos mil veintitrés.

Centro UNED de investigación en Procesamiento del Lenguaje Natural y Recuperación de Información (LENAR)

Finalidad del centro

El centro UNED de investigación en Procesamiento del Lenguaje Natural (LENAR) enfocará su trabajo en el área de las tecnologías de la lengua (el Procesamiento del Lenguaje Natural y la Recuperación de Información). Sus objetivos prioritarios son la investigación, innovación, transferencia y formación en los aspectos científico-tecnológicos del acceso y tratamiento de la información en la nueva economía de la Lengua, con un énfasis especial en el desarrollo de tecnología, datos y metodologías relacionados con la Inteligencia Artificial (IA) en español.

Para llevar a cabo estos objetivos, el centro impulsará (i) la formación de expertos en el área, (ii) la organización de reuniones científicas, cursos, seminarios y conferencias que contribuyan a la difusión y el avance del conocimiento en los ámbitos de conocimiento propios del centro, (iii) la investigación en innovación en el área mediante la participación en proyectos y redes de ámbito autonómico, nacional e internacional, (iv) la colaboración con otros centros y universidades participando en proyectos y campañas conjuntas, (v) la divulgación de sus actividades y la transferencia de sus resultados de investigación.

En definitiva, el centro se propone ser un referente en el área de las tecnologías lingüísticas a nivel internacional, especialmente en los aspectos que conciernen a la Inteligencia Artificial en español.

Necesidad científica y relevancia social

Las tecnologías de Procesamiento del Lenguaje Natural (PLN) posibilitan el análisis y la generación automatizada de texto y voz para su explotación en todo tipo de sectores (ámbitos médico, legal, educativo, turístico, administración del estado, ocio, creación artística, humanidades digitales, comercio electrónico, etc.). Los buscadores web, los chatbots, los sistemas de traducción automática y los asistentes de voz son ejemplos de tecnologías del lenguaje de alcance universal en nuestra sociedad actual.

El mercado global en tecnologías del lenguaje superó la barrera de los 40,000 millones de dólares en 2016, y ha seguido creciendo por encima de un 5% anual (según la consultora Common Sense Advisory). Sólo en Europa hay más de 500 empresas relacionadas con el sector, que es clave en la consecución del objetivo de la Comisión Europea de alcanzar un mercado digital único en una sociedad plurilingüe como la europea.

En el ámbito del estado español, la Estrategia Nacional de Inteligencia Artificial (ENIA) lanzada por la secretaría de Estado de Digitalización e Inteligencia Artificial, señala las tecnologías del lenguaje en español (y resto de idiomas oficiales del estado) como una

de las líneas prioritarias para impulsar el liderazgo de España dentro del campo de la Inteligencia Artificial y para acercar a los ciudadanos el potencial de la IA en las mismas condiciones que las de los ciudadanos anglosajones. El PERTE (proyecto estratégico para la recuperación y transformación económica) “Nueva economía de la lengua” aprobado por el consejo de ministros el 1 de Marzo de 2022, sitúa la Inteligencia Artificial en español como uno de sus cinco ejes estratégicos, indicando que *“Se invertirán más de 330 millones de euros para poner en marcha proyectos como la creación de una Red de Excelencia en Inteligencia Artificial y los Programas de Lenguas accesibles, con el objetivo de proporcionar herramientas digitales para personas con discapacidad y diversidad funcional y actuaciones del Plan Nacional de Tecnología del Lenguaje y su industria”*.

El centro será una herramienta esencial para la participación de la UNED en las actividades marcadas por la ENIA y el PERTE Nueva economía de la lengua, junto con el Laboratorio de Investigación en Humanidades Digitales de la UNED (LINHD) y las herramientas adicionales que se impulsen, en este sentido, desde el rectorado.

Grupo promotor

El grupo de investigación en Procesamiento del Lenguaje Natural y Recuperación de Información de la UNED (nlp.uned.es), dirigido actualmente por el catedrático Julio Gonzalo, está formado por unos treinta investigadores, de los que dieciocho son personal docente en plantilla. Comenzó sus actividades en 1996, y desde entonces se ha convertido en uno de los grupos de referencia a nivel nacional en su ámbito. Desde su fundación por parte de la catedrática M. Felisa Verdejo, ha participado en más de cuarenta proyectos europeos, nacionales y autonómicos; ha publicado más de 650 artículos y monografías de investigación en el ámbito del PLN y la Recuperación de Información; ha colaborado con más de cien universidades y centros de investigación de Europa, Asia y América; ha recibido numerosas distinciones, entre ellas: *Premio Nacional de Informática José García Santesmases* para M. Felisa Verdejo en 2016; *Google Faculty Research Award* para Julio Gonzalo y Enrique Amigó en 2012; *Generation Google Scholarship for Women in Computer Science* y *Premio Archiletras de la Lengua de investigación* para Elena Álvarez Mellado en 2022; y *Premio Hispanic Digital Humanities Award* para Elena Álvarez Mellado en 2021. Desde la perspectiva docente, el grupo constituye la columna vertebral del máster oficial en Tecnologías de la Lengua de la UNED, y una de las tres líneas de investigación dentro del doctorado en Sistemas Inteligentes de la UNED; sus miembros han dirigido 50 tesis doctorales relacionadas con el Procesamiento del Lenguaje Natural, la mayoría con mención de doctorado internacional/europeo.

Líneas de investigación del centro

El grupo promotor tiene una trayectoria de excelencia en las siguientes líneas de investigación, en las que guarda características únicas en el ámbito español, y que continuarán su desarrollo en el ámbito del centro:

- *Metodologías, métricas y campañas de evaluación*: el trabajo del grupo en esta área ha sido reconocido con un Google Faculty Research Award, y es la base de

un convenio entre UNED y red.es (2022-2024) para crear un espacio de observación de la Inteligencia Artificial en español en que se monitoriza la distancia entre la IA en español y su idioma hegemónico, el inglés. El grupo ha sido co-fundador del CLEF, la iniciativa europea para la evaluación competitiva para sistemas de PLN e IR en idiomas de ámbito europeo; y del IberLEF, iniciativa homóloga para español.

- *PLN aplicado al dominio biomédico.* El grupo destaca por su experiencia en (i) el ámbito de los sistemas de ayuda a la toma de decisiones en el dominio médico. En particular, en el marco de varios proyectos competitivos de investigación, estamos aplicando, en colaboración con el Servicio de Psiquiatría del Hospital Clínico San Carlos, técnicas de PLN al ámbito de la salud mental identificando patrones que se relacionen con situaciones de riesgo, como suicidio, anorexia, etc. Nuestros sistemas han sido de los más eficientes en competiciones internacionales sobre estas temáticas, como CLEF eRisk; (ii) la caracterización de enfermedades raras: en colaboración con la Consejería de Sanidad de la Comunidad de Madrid trabajamos en la identificación automática de rasgos asociados a este grupo de enfermedades. Hemos organizado una tarea competitiva de evaluación sobre este problema en el marco de IberEval 2018; (iii) la clasificación de historia clínica no estructurada, a través de colaboraciones con el Hospital universitario Fundación de Alcorcón; y (iv) el enriquecimiento de bases de datos terminológicas médicas en español a través de técnicas translingües.
- *PLN aplicado a la lucha contra la desinformación y los contenidos tóxicos.* El grupo ha desarrollado metodologías para recopilar evidencias tanto al nivel de los mensajes como al nivel de redes sociales para inferir la intención y las narrativas de las campañas de desinformación, ha participado en varios proyectos coordinados de ámbito nacional sobre el tema, y organiza anualmente una campaña internacional de evaluación sobre la detección y caracterización de contenido sexista en redes sociales.
- *Búsqueda de Respuestas en documentos y bases de conocimiento:* el grupo ha trabajado en diversas líneas relativas a ofrecer respuestas concisas a preguntas formuladas en lenguaje natural. Las fuentes sobre las que se han trabajado son tanto de bases de conocimiento estructuradas, como por ejemplo DBpedia, como documentos digitales. Se ha trabajado tanto desde la perspectiva de la evaluación, como en el desarrollo de sistemas, participando en proyectos como VIGICOVID, financiado por el Fondo Supera COVID-19, centrado en la búsqueda de información concisa sobre COVID-19 en publicaciones científicas. Además, el grupo siempre ha participado en la coordinación de la tarea de búsqueda de respuestas del CLEF, siendo el organizador principal a partir de 2007.

Además, ha contribuido a muchos otros ámbitos de investigación dentro del Procesamiento del Lenguaje Natural y la Recuperación de Información, entre otros:

- Aplicación de técnicas de análisis de grafos para diversas tareas de extracción de información: desambiguación del sentido de las palabras, extracción de relaciones, etc.
- Desambiguación del sentido de las palabras en contextos multilingües y en el dominio médico con técnicas basadas en grafos.
- Detección de spam en distintos ámbitos como páginas web, Twitter o Wikipedia.
- Minería de textos y búsqueda de información en contextos monolingües y multilingües.

Estimación de gastos de funcionamiento

Se prevé que los ingresos que se obtengan a través de proyectos competitivos, contratos de investigación y convenios cubrirán los gastos previstos, por lo que no se precisará de ninguna ayuda directa por parte de la UNED.

Adscripción preliminar de personal docente e investigador

Profesorado:

M. Lourdes Araujo Serna (CU)
 Julio Gonzalo Arroyo (CU) – director del centro
 Raquel Martínez Unanue (CU) – subdirectora del centro
 Anselmo Peñas Padilla (CU)
 M. Felisa Verdejo Maillo (colaboradora honorífica)

Roberto Centeno Sánchez (TU)
 Víctor Fresno Fernández (TU)
 Ana García Serrano (TU)
 Juan Martínez Romo (TU)
 Laura Plaza Morales (TU)
 Álvaro Rodrigo Yuste (TU)

Enrique Amigó Cabrera (CD)
 Jorge Carrillo de Albornoz (CD)
 Juan M. Cigarrán Recuero (CD)
 Agustín D. Delgado Muñoz (CD)
 Andrés Duque Fernández (CD)
 Fernando López Ostenero (profesor colaborador)
 Alberto Pérez García-Plaza (Ayudante Doctor)

Personal contratado:

Elena Álvarez Mellado
 Adrián Ghajari Espinosa
 Alicia Lara Clares
 Guillermo Marco Remón

Constantino Martín Valmaseda
Pablo Moral Martín
Roser Morante
Jacobo Pedrosa Marín
Alejandro Sánchez de Castro
Eva Sánchez Salido
Víctor Sánchez Sánchez

Compromiso de igualdad y no discriminación

El centro reconoce como objetivo estratégico el desarrollo de relaciones laborales basadas en la igualdad de oportunidades, la no discriminación y el respeto a la diversidad, promoviendo un entorno favorable, facilitando medidas de conciliación de la vida personal y laboral, respetando la legislación vigente y procurando seguir las mejores prácticas.

Reglamento de funcionamiento interno

Reglamento de funcionamiento interno del centro de investigación en Procesamiento del Lenguaje Natural y Recuperación de Información de la UNED.

TITULO 1

DISPOSICIONES GENERALES (NATURALEZA Y FINES)

Artículo 1 – Denominación y régimen jurídico

1. El centro de investigación en Procesamiento del Lenguaje Natural y Recuperación de Información (en adelante, *el centro*) es un centro de investigación propio de la UNED.
2. El centro es una entidad sin ánimo de lucro.
3. El centro se regirá:
 - a. Por la legislación y procedimiento que le sea de aplicación.
 - b. Por las disposiciones de este reglamento de funcionamiento interno.

Artículo 2 – Objetivos y actividades

El objetivo del centro es ser un referente en el área de las tecnologías lingüísticas a nivel internacional, especialmente en los aspectos que conciernen a la Inteligencia Artificial en español. En consecuencia:

1. El ámbito funcional de actuación será el de la investigación básica y aplicada dentro de las áreas y/o líneas específicas de investigación que se aprueben por el comité científico, dentro del ámbito de las tecnologías de la lengua (el Procesamiento del Lenguaje Natural y la Recuperación de Información).

2. Sus objetivos prioritarios son la investigación, innovación, transferencia y formación en los aspectos científico-tecnológicos del acceso y tratamiento de la información en la nueva economía de la Lengua, con un énfasis especial en el desarrollo de tecnología, datos y metodologías relacionados con la Inteligencia Artificial (IA) en español.
3. Para llevar a cabo estos objetivos, el centro impulsará
 - (i) la formación de expertos en el área,
 - (ii) la organización de reuniones científicas, cursos, seminarios y conferencias que contribuyan a la difusión y el avance del conocimiento en los ámbitos de conocimiento propios del centro.
 - (iii) la investigación en innovación en el área mediante la participación en proyectos y redes de ámbito autonómico, nacional e internacional.
 - (iv) la colaboración con otros centros y universidades participando en proyectos y campañas conjuntas.
 - (v) la divulgación de sus actividades y la transferencia de sus resultados de investigación.

TITULO II – DEL PERSONAL DEL CENTRO

Artículo 3 – Miembros y normas de admisión

1. Los miembros fundacionales son el personal docente e investigador permanente adscrito al grupo de investigación en Procesamiento del Lenguaje Natural y Recuperación de Información de la UNED.
2. El acceso de nuevos miembros investigadores permanentes deberá ir precedido de una solicitud del interesado/a avalada por dos miembros ya pertenecientes al centro, y de la aceptación posterior realizada por el Comité Científico por mayoría simple en la primera reunión que se celebre.
3. El personal no permanente será miembro a petición del responsable de la investigación a realizar. Su aceptación, por el periodo requerido, será aprobada por la dirección del centro.
4. El centro podrá contar con personal de administración y servicios adscrito al mismo.

Artículo 4 – Pérdida de la condición de miembro del centro

1. Los miembros investigadores permanentes causarán baja por alguna de las causas siguientes:
 - a. Por voluntad propia, comunicada por escrito al secretario/a del centro.
 - b. Por traslado a otro centro en el caso de personal técnico.
 - c. Por exclusión por parte del comité científico del centro.
2. El personal no permanente causará baja automática tras la finalización del contrato o, con antelación, por solicitud razonada por parte del responsable de investigación que solicitó su admisión.

TITULO III DE LOS ORGANOS DE GOBIERNO DEL CENTRO

Artículo 5 – Órganos de gobierno

1. El gobierno del centro estará compuesto de órganos colegiados y unipersonales encargados de establecer las políticas de funcionamiento y gestionar el día a día del centro.
2. Son órganos colegiados de gobierno del centro el Consejo Asesor, el Comité Científico y la Comisión de Dirección, más las comisiones específicas ad hoc que en cada momento se constituyan a propuesta de alguno de los órganos colegiados anteriores para estudiar aspectos o temáticas concretas.

Artículo 6 – Comité Científico

1. El comité científico es el órgano de gobierno del centro, y estará compuesto por:
 - a. El director/a, que lo preside.
 - b. El secretario/a.
 - c. Los doctores miembros permanentes del centro.
 - d. Una representación de los doctores no permanentes y del personal investigador en formación del centro que podrá alcanzar, como máximo, un 30% del total del comité.
 - e. Una representación de miembros del PAS adscritos al centro que podrá alcanzar, como máximo, un 5% del total del comité.
2. Corresponde al comité científico del centro el control e impulso de sus funciones, entre otras:
 - a. Admisión de miembros.
 - b. Propuesta de nombramiento del director/a.
 - c. Aprobación de líneas de investigación.
 - d. Aprobación de la memoria de ejecución presupuestaria.
 - e. Cualquiera otra recogida en este reglamento o en las normativas aplicables en la UNED y en la legislación vigente.
3. Inclusión y mandato
 - a. Los miembros del comité cesarán por renuncia o por pérdida de la condición por la que formaban parte de este.
 - b. Los miembros del comité tendrán el derecho y el deber de asistir con voz y voto a las sesiones de comité, así como a las comisiones de que formen parte.
4. Presidencia.
 - a. El director/a del centro, y en su ausencia el subdirector/a, presidirá el comité científico y tendrá como función propia la de asegurar el cumplimiento de las leyes, moderando el curso de los debates.
5. Convocatoria
 - a. La convocatoria del comité científico corresponde al director/a, y será acordada y notificada con una antelación mínima de cinco días hábiles.
 - b. Se realizará al menos una reunión ordinaria del comité científico al año. El director/a convocará al comité de forma extraordinaria cuando lo estime oportuno, o cuando lo soliciten por escrito al menos 2/5 de los

- miembros de este, en un plazo no superior a cinco días naturales tras la solicitud.
- c. Una sesión del comité científico podrá ser aplazada cuando lo aprueben los miembros del comité por mayoría simple.
6. Orden del día
- a. A la convocatoria del comité se acompañará el orden del día, que será fijado por la comisión de dirección, teniendo en cuenta, en su caso, las peticiones de los demás miembros formuladas con suficiente antelación.
 - b. Para que sea exigible la inclusión de algún punto en el orden del día de las sesiones deberá ser suscrito al menos por una quinta parte de los miembros del consejo.
7. Comisiones
- a. El comité científico podrá constituir las comisiones que entienda convenientes, las cuales se disolverán al concluir su cometido. La única comisión permanente será la comisión de dirección.
8. Comisión de dirección
- a. La comisión de dirección es el órgano colegiado ordinario de gobierno del centro por delegación del comité científico.
 - b. La comisión de dirección estará formada por el director/a que la preside, el secretario/a, dos representantes del personal docente doctor y uno del personal investigador en formación de entre los adscritos al centro, elegidos por el comité científico.
 - c. Las competencias de la comisión de dirección son:
 - i. Elaboración del orden del día de las sesiones del comité.
 - ii. Entender y decidir sobre los asuntos diarios y de trámite.
 - iii. Proponer las tarifas de cursos de especialización y servicios.
 - iv. Participar en la selección de personal en el ámbito de sus competencias.
 - v. Gestionar los servicios económicos y administrativos del centro.
 - vi. Decidir sobre aquellas materias urgentes, competencia del comité científico, cuya premura no permita tratarlas por éste. En ese caso, deberán ser posteriormente refrendadas por el comité científico.
 - d. La comisión de dirección se reunirá siempre que sea convocada por el director/a del centro. Actuará como secretario/a el del centro, y sus acuerdos constarán en un acta firmada por el secretario/a y visada por el director/a.

Artículo 7. Consejo asesor

1. El consejo asesor de expertos estará constituido por cuatro miembros. Estará presidido por el Vicerrector/a competente en investigación, y actuará como secretario/a, con voz, pero sin voto, el director del centro. Los dos restantes miembros del comité serán nombrados por el rector/a de la UNED a propuesta del comité científico entre personas externas a la UNED de reconocido prestigio, siendo al menos una de ellas de un centro no español.
2. Las responsabilidades del consejo asesor son:

- a. Emitir informes de valoración sobre las actividades de investigación del centro.
- b. Asesorar sobre la planificación y organización científica del centro en relación con cambios en las actividades científicas, docentes o de servicios que puedan surgir.
- c. Sugerir y promover el establecimiento de convenios de colaboración con empresas o con otros centros de investigación.

Artículo 8. Dirección

1. El director/a del centro ostenta su representación y ejerce las funciones de dirección y gestión ordinaria.
2. El nombramiento de director/a corresponde al rector/a, a propuesta del comité científico.
3. La duración de su mandato será de un máximo de cuatro años.

Artículo 9. Subdirección

1. A propuesta del director/a, el Rector/a nombrará un subdirector/a del centro, de entre el personal con vinculación permanente a la UNED integrado en el centro.
2. Corresponde al subdirector/a dar fe de los acuerdos y resoluciones de los órganos de gobierno del centro, garantizar la difusión y publicidad de los acuerdos, resoluciones, convenios, reglamentos y demás normas generales de funcionamiento institucional entre los miembros del centro, llevar el registro y custodiar el archivo, expedir las certificaciones que le correspondan y desempeñar aquellas otras competencias que le sean delegadas por el director/a, sin perjuicio de las funciones que le asigne el presente reglamento de régimen interno.

DISPOSICIÓN FINAL

Sobre la modificación del presente reglamento

1. El desarrollo del centro podría hacer necesaria la modificación del presente reglamento, bien sea por la incorporación de un número importante de nuevos miembros y/o por la adquisición o el uso de determinados servicios e infraestructuras. La iniciativa para su modificación corresponderá al comité científico, mediante acuerdo del más del 50% de sus constituyentes.