



**D.<sup>a</sup> REBECA DE JUAN DÍAZ, 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 ocho de marzo de dos mil veintidós, fue adoptado, entre otros, el siguiente acuerdo:

**05. Estudio y aprobación, si procede, de las propuestas del Vicerrectorado de Grado y Posgrado.**

**05.05.** El Consejo de Gobierno aprueba la modificación de la memoria del Máster Universitario en Tecnologías del Lenguaje, 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 nueve de marzo de dos mil veintidós.

**Juan José Escribano Ródenas**  
**Secretario Académico**



**D. JUAN JOSÉ ESCRIBANO RÓDENAS, SECRETARIO ACADÉMICO DE LA ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA DE LA UNED**

**CERTIFICA:** Que en la Junta de la Escuela Técnica Superior de Ingeniería Informática, celebrada el día quince de diciembre de dos mil veintiuno fue adoptado, entre otros, el siguiente acuerdo:

**4.b.- Aprobación, si procede, de los acuerdos de la Comisión Coordinadora del Máster Universitario en Tecnologías del Lenguaje, necesarios para iniciar el procedimiento de modificación de su Memoria Verificada.**

Se aprueba.

Para que así conste, a los efectos oportunos, se extiende el presente certificado en Madrid a diecisiete de diciembre de dos mil veintiuno.

\*Se adjunta documentación

## ACTA DE LA REUNIÓN DE LA COMISIÓN DEL MÁSTER DE "TECNOLOGÍAS DEL LENGUAJE" DE LA E.T.S.I. DE INFORMÁTICA

### Asisten:

Rafael Pastor Vargas  
Director de la ETSI de Informática

Víctor Fresno Fernández  
Coordinador del Máster en Tecnologías del Lenguaje

Roberto Centeno Sánchez  
Secretario del Máster en Tecnologías del Lenguaje

Lidia Segovia Orellana  
Representante de Personal de Administración y Servicios

Olga C. Santos  
Coordinadora del Máster en Inteligencia Artificial

Aitor García Moreno  
Representante de Estudiantes

En Madrid, siendo las 12:00h del día 2 de diciembre de 2021, se reúne la Comisión del Máster en Tecnologías del Lenguaje con el siguiente orden del día:

- 1. Ratificación acuerdos de la Subcomisión de Ordenación Académica**
- 2. Ruegos y preguntas**

El coordinador da comienzo a la sesión.

### **1. Ratificación acuerdos de la Subcomisión de Ordenación Académica**

El coordinador del máster informa que el pasado día 24 de noviembre de 2021 se reunió la Subcomisión de Ordenación Académica, tomando los acuerdos presentes en el acta adjunta a este documento. Tras esta reunión se propone elevar a esta comisión de coordinación del máster la ratificación de dichos acuerdos.

Se aprueba por unanimidad.

### **2. Ruegos y preguntas**

No hay ruegos ni preguntas.

Y sin más asuntos que tratar, siendo las 12:30, finaliza la reunión de la Comisión.

Fdo: Roberto Centeno Sánchez  
Secretario Académico del Máster  
en Tecnologías del Lenguaje.

**ACTA DE LA REUNIÓN DE LA SUBCOMISIÓN DE ORDENACIÓN  
ACADÉMICA DEL MÁSTER EN TECNOLOGÍAS DEL LENGUAJE DEL DÍA 24  
DE NOVIEMBRE DE 2021**

**Asisten:**

Víctor Fresno Fernández, *Coordinador*  
Roberto Centeno, *Secretario Académico*  
Agustín D. Delgado  
Álvaro Rodrigo Yuste  
Anselmo Peñas Padilla  
Juan Martínez Romo  
Andrés Duque Fernández  
Fernando López Ostenero  
Felisa Verdejo  
Raquel Martínez Unanue  
Lourdes Araujo  
Laura Plaza  
Juan Manuel Cigarrán Recuero  
Jorge Amando Carrillo de Albornoz Cuadrado  
Enrique Amigó Cabrera  
Manuel Luque Gallego

En Madrid, siendo las 10.30 horas del día 24 de noviembre de 2021, se reúne la subcomisión de Ordenación Académica del Máster en Lenguajes y Sistemas Informáticos con el siguiente orden del día:

- 1. Informe del coordinador.**
- 2. Modificaciones del plan de estudios para el curso 2022/2023**
- 3. Propuesta de gasto del Presupuesto del curso 2020-2021**
- 4. Ruegos y preguntas**

El coordinador da comienzo a la sesión.

**1. Informe del coordinador**

El coordinador comienza la reunión recordando que dentro de la línea de trabajo de la coordinación del Máster en Tecnologías del Lenguaje centrada en la actualización y reorganización de los contenidos del máster, para el curso

presente (2021-22) se solicitaron a la ANECA las modificaciones siguientes, que fueron APROBADAS:

1. **Inclusión** de la asignatura **“REDES NEURONALES PARA EL PROCESAMIENTO DEL LENGUAJE NATURAL”**

o Equipo Docente: Álvaro Rodrigo y Andrés Duque.

2. **Inclusión** de la asignatura **“REPRESENTACION DE TEXTOS EN ESPACIOS VECTORIALES Y PROBABILÍSTICOS”**

o Equipo Docente: Enrique Amigó y Víctor Fresno.

A continuación, el coordinador informa de la intención de solicitar la aprobación de nuevas modificaciones para entrar en vigor en el curso 2022/2023, por lo que se deben aprobar en la Junta de Escuela convocada el 15 de diciembre de 2021 y luego deben ser aprobadas por la COA de la Universidad para, posteriormente, introducirlas en el programa MODIFICA para su evaluación por parte de la ANECA.

Del mismo modo, siguiendo con la línea de actualización del máster, el coordinador presenta un calendario tentativo de modificaciones futuras que se solicitarán para su aprobación en futuras comisiones, con el objetivo de entrar en vigor en el curso 2023/2024. Entre las modificaciones que se están considerando encontramos las siguientes:

1. Eliminación de asignatura “Minería de la web”
2. Inclusión de la asignatura “Minería de Textos” y eliminación de “Descubrimiento de Información en Textos”
3. Inclusión de nuevas asignaturas dentro de alguna o varias de las siguientes temáticas:
  1. Análisis de sentimiento
  2. Sistemas de diálogo
  3. Comprensión del Lenguaje Natural (NLU)
  4. Explicabilidad/interpretabilidad en PLN
  5. Simplificación de textos
  6. Traducción automática
  7. Asignatura con contenidos dinámicos de temáticas novedosas

Debido a que algunas de las asignaturas del máster son compartidas con el Máster en Investigación en Inteligencia Artificial, el coordinador informa que, si se aprueba en esta comisión la inclusión de una nueva asignatura propuesta

“Fundamentos del Procesamiento Lingüístico”, ésta pasaría a sustituir a la actual “Procesamiento del Lenguaje Natural” en dicho máster, a partir del curso 2022/2023, previa aprobación por todas las comisiones y organismos pertinentes.

Del mismo modo, la asignatura actual de “Descubrimiento de Información en Textos”, impartida en ambos másteres, se seguirá manteniendo en el máster en Investigación en Inteligencia Artificial, aunque en un futuro calendario de modificaciones se plantee sustituir en el máster en Tecnologías del Lenguaje por la asignatura “Minería de Textos”.

Finalmente, el coordinador informa que las dos asignaturas nuevas incluidas en el máster en este curso 2020/2021 (“Redes neuronales para PLN” y “Representación de textos en espacios vectoriales y probabilistas”) fueron ofrecidas al máster en Investigación en Inteligencia Artificial para ser incluidas en su plan de estudios, pero no fue considerado por la coordinación del citado máster.

## 2. Modificaciones del plan de estudios para el curso 2022/2023

El coordinador comienza este segundo punto informando de las modificaciones a considerar, y aprobar (si procede):

1. **Reducción** de la carga docente de la asignatura de TFM “**Trabajo: Procesamiento Del Lenguaje Natural Y Recuperación De Información**” de 30 a 18 ECTS.
  - Los estudiantes que deseen continuar con la investigación y hacer el doctorado deberían cursar la asignatura de “Iniciación a la investigación en Procesamiento del Lenguaje Natural”.
  - Esta reducción en el número de créditos del TFM implica que el número de asignaturas obligatorias a cursar pase de 5 a 7 asignaturas.
2. **Inclusión** de una asignatura con título (tentativo) “**Iniciación a la investigación en Procesamiento del Lenguaje Natural**” (6 ECTS)  
El temario (tentativo) sería el siguiente, hasta que sea definido por el equipo docente:
  - Estructura de un trabajo de investigación
  - Metodología de investigación
  - Elaboración del estado del arte
  - Métricas de evaluación
  - Colecciones de entrenamiento y evaluación
  - Estándares de anotación
  - Experimentación

De este modo mantenemos el carácter de investigación del máster para aquellos estudiantes que quieran después hacer el doctorado con nosotros.

Profesor responsable: Julio Gonzalo

3. **Eliminación** de la asignatura “**Procesamiento del Lenguaje Natural**” (6 ECTS) y solicitud de **inclusión** de una asignatura con título (tentativo) “**Fundamentos del procesamiento lingüístico**” en su lugar.
- Carga de 6 ECTS y carácter optativo,

El temario (tentativo) sería el siguiente, hasta que sea definido por el equipo docente:

### **Fundamentos del procesamiento lingüístico (6 ECTS)**

1. Introducción al Procesamiento del Lenguaje Natural
2. Análisis **morfosintáctico**
3. Análisis **sintáctico** (constituyentes, dependencias)
4. Análisis semántico (semántica léxica y semántica proposicional - gramáticas extendidas-)
5. **Herramientas de procesamiento lingüístico** (librerías, etc.)

Profesor responsable: Anselmo Peñas

4. Modificación del temario en la asignatura “**Semántica y pragmática en la web**” (6 ECTS) para eliminar los temas correspondientes a “Semántica Computacional (semántica léxica)”, que pasaría a impartirse en “Fundamentos del procesamiento lingüístico”, y “Diseño inicial de proyecto de investigación”, que pasaría a impartirse en “Iniciación a la investigación en Procesamiento del Lenguaje Natural”.

Además, se solicita el cambio en el nombre de un tema de “La web de los datos” a “Modelado de Información y conocimiento con ontologías”, quedando el temario como sigue:

- TEMA 1. INTRODUCCION
- TEMA 2. MODELADO DE INFORMACIÓN Y CONOCIMIENTO CON ONTOLOGÍAS
  - ONTOLOGÍAS Y SU POBLACIÓN
  - LENGUAJES SEMÁNTICOS
  - ACCESO A LA INFORMACIÓN
  - ENRIQUECIMIENTO DE ONTOLOGÍAS
- TEMA 3. LA WEB DE LOS DATOS
  - LINKED OPEN DATA (LOD)

- LINGUISTIC LINKED OPEN DATA (LLOD)
- RETOS Y EJEMPLOS DE CASOS DE ÉXITO
- TEMA 4. PRAGMATICA Y PRAXIS
  - LENGUAJE NATURAL
  - HUMANIDADES DIGITALES
  - BIOMEDICINA
  - OTRAS ÁREAS (FINANZAS ...)

Se informa que esta propuesta de modificación ha sido acordada previamente con el equipo docente de la asignatura.

5. Considerar las nuevas asignaturas “Fundamentos del Procesamiento Lingüístico” y “Representación de textos en espacios vectoriales y probabilistas” como asignaturas obligatorias de facto durante el proceso de admisión para aquellos estudiantes que accedan SIN haber cursado una asignatura básica de PLN.

Se incluiría en la carta de admisión la obligatoriedad o recomendación de realizar las dos asignaturas consideradas como fundamentales y si el objetivo es una carrera investigadora, realizar la nueva asignatura "Introducción a la investigación en Procesamiento del Lenguaje Natural".

Se acuerda por unanimidad la aprobación de todas las modificaciones propuestas para entrar en vigor en el curso 2022/2023, nombrando además como ponentes de las dos nuevas asignaturas a incluir a los profesores:

- D. Anselmo Peñas (*Fundamentos del procesamiento lingüístico*)
- D. Julio Gonzalo (*Iniciación a la investigación en Procesamiento del Lenguaje Natural*),

Ambos serán los encargados de preparar la ficha de las asignaturas para su posterior aprobación en las diferentes comisiones necesarias.

### **3. Propuesta de gasto del Presupuesto del curso 2020-2021.**

Se informa que el presupuesto de este año asciende a 1.658,59€, correspondiéndose 1.523,53€ al máster en Tecnologías del Lenguaje y 135,06€ al máster en LSI.

Tras propuesta por parte de la coordinación a todos los miembros de la comisión del máster sobre la posibilidad de financiar la impartición de seminarios o charlas por expertos en el área, con cargo al presupuesto del máster, se solicitó el pago del siguiente seminario:

- Vincenzo della Mea, que tendrá lugar el día 16 de diciembre de 2021

Aprobado por unanimidad.

#### **4. Ruegos y preguntas**

No hay ruegos ni preguntas.

Y sin más asuntos que tratar, siendo las 12:30, finaliza la reunión.

Fdo: Roberto Centeno Sánchez  
*Secretario Académico del Máster en  
Tecnologías del Lenguaje.*

## PROPUESTA DE MODIFICACIÓN DE LOS CONTENIDOS Y ESTRUCTURA DEL MÁSTER EN TECNOLOGÍAS DEL LENGUAJE

Dentro de la línea de trabajo de la coordinación del *Máster en Tecnologías del Lenguaje* centrada en la actualización y reorganización de los contenidos del máster, para el curso presente (2021-22) se solicitaron a la ANECA las modificaciones siguientes, que fueron APROBADAS:

1. **Inclusión de la asignatura “REDES NEURONALES PARA EL PROCESAMIENTO DEL LENGUAJE NATURAL”**
  - Equipo Docente: Álvaro Rodrigo y Andrés Duque.
  
2. **Inclusión de la asignatura “REPRESENTACIÓN DE TEXTOS EN ESPACIOS VECTORIALES Y PROBABILÍSTICOS”**
  - Equipo Docente: Enrique Amigó y Víctor Fresno.

El objetivo es la actualización del máster de forma paulatina. La propuesta de modificación para el próximo curso es la siguiente:

1. **Reducción de la carga docente** de la asignatura de TFM “**Trabajo: Procesamiento Del Lenguaje Natural Y Recuperación De Información**” de 30 a 18 ECTS.

La justificación de esta reducción es que, de este modo, permitimos que se puedan realizar TFM de carácter más aplicado, no necesariamente trabajos de inicio a la investigación en NLP, y de menos carga. Este máster se implantó como uno de los primeros después de la entrada de la LOU en un momento en el que los másteres oficiales podían tener un carácter investigador o profesional. Posteriores modificaciones de la normativa nacional eliminaron esta consideración.

La experiencia de estos años nos anima a plantear TFM de menor carga docente pensando en el caso de estudiantes que no tengan la intención de seguir con sus estudios dentro un programa de doctorado. Los estudiantes que deseen continuar con la investigación y hacer el doctorado deberían cursar la asignatura de “Iniciación a la investigación en Procesamiento del Lenguaje Natural” (que se plantea como propuesta de modificación en el siguiente punto), y así se les indicará durante el proceso de admisión al título.

2. **Inclusión de una asignatura con título (tentativo) “Iniciación a la investigación en Procesamiento del Lenguaje Natural” (6 ECTS)**

La carga sería de 6 ECTS y de carácter optativo.  
El temario sería el siguiente:

- Tema 1: Estructura de un trabajo de investigación
- Tema 2: Metodología de investigación
- Tema 3: Elaboración del estado del arte
- Tema 4: Métricas de evaluación
- Tema 5: Colecciones de entrenamiento y evaluación
- Tema 6: Estándares de anotación
- Tema 7: Experimentación

De este modo mantenemos el carácter de investigación del máster para aquellos (pocos) estudiantes que quieran después hacer el doctorado, ofreciéndoles unos contenidos fundamentales para adquirir un perfil de investigación. Aquellos estudiante que no deseen continuar con la investigación pueden cursarla también, adquiriendo con ellas las bases de la metodología científica.

3. **Eliminación** de la asignatura **“Procesamiento del Lenguaje Natural”** (6 ECTS) y solicitud de **inclusión** de una asignatura con título **“Fundamentos del procesamiento lingüístico”** (6ECTS) en su lugar.

En este caso el objetivo es la actualización del temario de esta asignatura de contenidos básicos dentro de la línea de investigación en Tecnologías del Lenguaje. El área del Procesamiento del Lenguaje Natural ha experimentado enormes cambios en los últimos años, presentándose nuevas arquitecturas de computación que han hecho evolucionar los paradigmas existentes dentro del Procesamiento Lingüístico. Esta nueva realidad quedará plasmada en la estructura de la asignatura que proponemos, y que se complementa con la asignatura ya implantada en el máster desde el curso 2021-2022 de **“Representación de textos en espacios vectoriales y probabilísticos”**.

La carga de la nueva asignatura sería de 6 ECTS y su carácter sería optativo, aunque se recomendará su matrícula durante el proceso de admisión para aquellos estudiantes que accedan sin haber cursado una asignatura básica de PLN.

El temario sería el siguiente:

**Fundamentos Procesamiento lingüístico (6 ECTS)**

- Tema 1: Metodología de investigación en Procesamiento del Lenguaje Natural: introducción.
- Tema 2: Planteamiento del problema: elección del tema, tareas y escenarios, cuestiones de investigación (“research questions”), objetivos y determinación del estado del arte.
- Tema 3: Abstracción del problema y métricas de evaluación.
- Tema 4: Diseño experimental: A/B testing, sistemas de referencia y sistemas contrastivos, experimentos de ablación, etc. Leaderboards.
- Tema 5: Recolección y anotación de datasets. Crowdsourcing. Identificación y prevención de sesgos. Tratamiento de desacuerdos entre anotadores.
- Tema 6: Análisis de resultados: análisis cuantitativo, uso de múltiples métricas, tests de relevancia estadística,

análisis de errores, análisis cualitativo, identificación de alcance y limitaciones.

- Tema 7: Comunicación científica: elaboración de artículos de investigación en Procesamiento del Lenguaje Natural. Peer review. Revistas académicas y congresos científicos en el área, y medidas de impacto científico. Publicación de código, datasets y otros recursos. Reproducibilidad. Repositorios de investigación. Comunicación oral.

4. **Modificación del temario** en asignatura “**Semántica y pragmática en la web**” para eliminar los temas correspondientes a “Semántica Computacional (semántica léxica)”, que pasaría a impartirse en la asignatura de “Fundamentos del procesamiento lingüístico”, y el tema de “Diseño inicial de proyecto de investigación”, que pasaría a impartirse en la asignatura “Iniciación a la investigación en Procesamiento del Lenguaje Natural”.

Además, se solicita el cambio en el nombre de un tema de “La web de los datos” a “Modelado de Información y conocimiento con ontologías”, quedando el temario como sigue:

- TEMA 1. INTRODUCCION
- TEMA 2. MODELADO DE INFORMACIÓN Y CONOCIMIENTO CON ONTOLOGÍAS
  - ONTOLOGÍAS Y SU POBLACIÓN
  - LENGUAJES SEMÁNTICOS
  - ACCESO A LA INFORMACIÓN
  - ENRIQUECIMIENTO DE ONTOLOGÍAS
- TEMA 3. LA WEB DE LOS DATOS
  - LINKED OPEN DATA (LOD)
  - LINGUISTIC LINKED OPEN DATA (LLOD)
  - RETOS Y EJEMPLOS DE CASOS DE ÉXITO
- TEMA 4. PRAGMATICA Y PRAXIS
  - LENGUAJE NATURAL
  - HUMANIDADES DIGITALES
  - BIOMEDICINA
  - OTRAS ÁREAS (FINANZAS ...)

5. **Modificación del temario** en asignatura “**Minería de Datos**”

Revisión del temario de la asignatura para actualizar contenidos en un área de conocimiento que ha sufrido un desarrollo espectacular en los últimos años, debido al éxito de nuevos modelos de aprendizaje profundo, y corregir así inconsistencias con los contenidos que se indican en la guía de la asignatura, quedando el temario como sigue:

- TEMA 1. INTRODUCCION
- TEMA 2. APRENDIZAJE SUPERVISADO
- TEMA 3. SELECCIÓN DE CARACTERÍSTICAS
- TEMA 4. EVALUACIÓN Y SELECCIÓN DE RASGOS

- TEMA 5. REDES NEURONALES ARTIFICIALES (EL PERCEPTRÓN MULTICAPA)
- TEMA 6. APRENDIZAJE PROFUNDO (CNN)
- TEMA 7. BOSQUES ALEATORIOS
- TEMA 8. CONSECUENCIAS ÉTICAS Y SOCIALES

6. **Inclusión** del profesor **Andrés Duque Fernández** como profesor del máster.

La coordinación del máster en tecnologías del Lenguaje considera que la incorporación del profesor doctor del departamento de Lenguajes y Sistemas Informáticos Andrés Duque Fernández puede ayudar a mejorar la calidad del máster, desarrollando nuevos contenidos y ayudando a dar mejor servicio al alumno. Se plantea que forme parte del Equipo Docente de la asignatura REDES NEURONALES PARA EL PROCESAMIENTO DEL LENGUAJE NATURAL.

Fdo: Víctor Fresno Fernández  
Coordinador del Máster en Tecnologías de la UNED

## FICHA DE ASIGNATURA DE MÁSTER

<b>NOMBRE DE LA ASIGNATURA</b>	<b>Fundamentos del Procesamiento Lingüístico</b>
<b>CARÁCTER</b>	Opcional
<b>ECTS</b>	6 ECTS (150 horas): 70 horas de estudio teórico (HT) + 80 horas de trabajo práctico (HP)
<b>SEMESTRE</b>	Anual
<b>IDIOMA</b>	Español
<b>ESPECIALIDAD EN QUE SE IMPARTE</b>	
<b>RESULTADOS DE APRENDIZAJE</b>	<p>El objetivo principal de la asignatura consiste en conocer y aplicar las técnicas de procesamiento lingüístico automático de los textos como paso previo a su conversión en representaciones simbólicas manipulables por los ordenadores.</p> <p>De acuerdo con esto, los resultados del aprendizaje son:</p> <ol style="list-style-type: none"> <li>1. Conocer y aplicar las técnicas de análisis morfosintáctico de las palabras en contexto.</li> <li>2. Conocer y aplicar las técnicas de análisis sintáctico de las oraciones, tanto el análisis de constituyentes como de dependencias.</li> <li>3. Conocer y aplicar las técnicas para obtener representaciones semánticas de las oraciones.</li> <li>4. Saber utilizar algunas herramientas existentes de procesamiento lingüístico de los textos.</li> </ol>
<b>CONTENIDO</b>	<p>Tema 1: Introducción al procesamiento lingüístico de los textos</p> <p>Tema 2: Análisis morfosintáctico de las palabras. Formas base, etiquetas gramaticales y reconocimiento de entidades.</p> <p>Tema 3: Análisis sintáctico de las oraciones. Análisis de constituyentes y análisis de dependencias.</p> <p>Tema 4: Análisis semántico. Desambiguación léxica, resolución correferencias, enlazado de entidades, etiquetado de roles semánticos, extracción de relaciones y representación lógica.</p>
<b>OBSERVACIONES</b>	<p>Conocimientos previos necesarios:</p> <ul style="list-style-type: none"> <li>● Lectura fluida en inglés</li> <li>● Conocimientos de programación</li> </ul>

	<p>Conocimientos previos recomendables:</p> <ul style="list-style-type: none"> <li>● Conocimientos básicos de álgebra, cálculo y lógica</li> <li>● Aprendizaje automático</li> <li>● RDF y SPARQL</li> </ul> <p>Esta asignatura puede ser cursada aisladamente.</p>
<p><b>Nº DE LAS COMPET. GENERALES</b></p>	<p>Con la superación del curso se espera que el alumno complete las siguientes competencias generales especificadas en la memoria del máster:</p> <p>CPG1 - Adquirir capacidad de abstracción, análisis, síntesis y relación de ideas.</p> <p>CPG2 - Adquirir capacidad crítica y de decisión</p> <p>CPG3 - Adquirir capacidad de estudio y autoaprendizaje</p> <p>CPG4 - Adquirir capacidad creativa y de investigación</p>
<p><b>Nº DE LAS COMPET. ESPECÍFICAS</b></p>	<p>Con la superación del curso se espera que el alumno complete las siguientes competencias específicas indicadas en la memoria del máster:</p> <p>CE1 - Adquirir capacidad de comprender y manejar de forma básica los aspectos más importantes relacionados con los lenguajes y sistemas informáticos en general y, de manera especial, en los siguientes ámbitos: Tecnologías del lenguaje y de acceso a la información en web.</p> <p>CE3 - Adquirir capacidad de estudio de los sistemas y aproximaciones existentes para distinguir las aproximaciones más efectivas.</p> <p>CE4 - Adquirir capacidad para detectar carencias en el estado actual de la ciencia y la tecnología.</p> <p>CE5 - Adquirir capacidad para proponer nuevas aproximaciones que den solución a las carencias detectadas.</p> <p>CE6 - Adquirir capacidad de especificar, diseñar, implementar y evaluar tanto cualitativa como cuantitativamente los modelos y sistemas propuestos.</p> <p>CE7 - Adquirir capacidad para proponer y llevar a cabo experimentos con la metodología adecuada como para poder extraer conclusiones y determinar nuevas líneas de actuación e investigación.</p>

<p><b>ACTIVIDADES FORMATIVAS CON Nº DE HORAS POR CADA UNA</b></p>	<p>Tema 1: Introducción al procesamiento lingüístico de los textos  T1.1. Lectura de material sugerido (5 HT)  T1.2. Demostradores e instalación de herramientas (10 HP)  T1.3. Descarga y preparación de colecciones de texto (5 HP)</p> <p>Tema 2: Análisis morfosintáctico de las palabras  T2.1. Lectura de material sugerido (15 HT)  T2.2 – Ejercicios prácticos obligatorios (15 HP)</p> <p>Tema 3: Análisis sintáctico de las oraciones  T3.1. Lectura de material sugerido (20 HT)  T3.2 – Ejercicios prácticos obligatorios (20 HP)</p> <p>Tema 4: Análisis semántico  T4.1. Lectura de material sugerido (30 HT)  T4.2 – Ejercicios prácticos obligatorios (30 HP)</p>
<p><b>METODOLOGÍA DOCENTE</b></p>	<p>La asignatura no tiene clases presenciales. Los contenidos teóricos se imparten a distancia y de acuerdo con las normas y estructuras de soporte telemático de la enseñanza en la UNED.</p> <p>El material docente incluye los contenidos de cada tema y distintos tipos de actividades relacionadas con la consulta bibliográfica, la realización de ejercicios prácticos y la utilización de herramientas asociadas a las tecnologías y técnicas presentadas en el curso. Junto a las actividades y enlaces con fuentes de información externas, existe material didáctico propio preparado por el equipo docente con ejercicios que organizan las actividades prácticas del alumno.</p> <p>Las actividades de aprendizaje se estructuran alrededor de la adquisición de conceptos en cada uno de los temas del curso y a su aplicación práctica mediante ejercicios que el alumno deberá realizar en cada tema. Se utilizarán los recursos de la UNED para las actividades complementarias que se propondrán a lo largo del curso como, por ejemplo, la asistencia virtual o presencial a los seminarios de temas avanzados impartidos por investigadores de reconocido prestigio.</p> <p>El plan de trabajo marcará unas pautas adecuadas para que el alumno alcance los objetivos al final del curso. Incluye una descripción ordenada y secuenciada de los contenidos de cada tema y orientaciones sobre las actividades a realizar, tanto de estudio teórico como práctico. Los ejercicios prácticos correspondientes a cada tema sirven de apoyo y refuerzo del aprendizaje, y su realización es obligatoria para superar la asignatura.</p> <p>Las actividades pueden ser de los siguientes tipos:</p> <ol style="list-style-type: none"> <li>1. <b>L:</b> Lectura y estudio del material correspondiente</li> <li>2. <b>P:</b> Ejercicios prácticos de entrega obligatoria (exige una entrega en el aula virtual)</li> </ol>

<b>SISTEMAS DE EVALUACIÓN CON PONDERACIÓN</b>	Evaluación continua a través de la elaboración de prácticas obligatorias por tema. Cada práctica será evaluada por separado y la calificación final de la asignatura corresponderá a la suma ponderada de las calificaciones de dichas prácticas.
---	---

## FICHA DE ASIGNATURA DE MÁSTER

<b>NOMBRE DE LA ASIGNATURA</b>	<b>Introducción a la investigación en Procesamiento del Lenguaje Natural</b>
<b>CARÁCTER</b>	Opcional
<b>ECTS</b>	6 ECTS (150 horas): 80 horas de estudio teórico (HT) + 70 horas de trabajo práctico (HP)
<b>SEMESTRE</b>	Anual
<b>IDIOMA</b>	Español
<b>ESPECIALIDAD EN QUE SE IMPARTE</b>	
<b>RESULTADOS DE APRENDIZAJE</b>	<p>El objetivo principal de la asignatura es que el estudiante se familiarice con la metodología de investigación en Procesamiento del Lenguaje Natural, de forma que pueda plantear, planificar, desarrollar y comunicar un trabajo de investigación de forma autónoma y efectiva. Los resultados de aprendizaje esperados son:</p> <ol style="list-style-type: none"> <li>1. Capacidad para elegir y plantear un tema de investigación adecuado en función de su interés práctico y científico, de su viabilidad y del estado del arte.</li> <li>2. Capacidad para focalizar, abstraer y formalizar el tema elegido en términos de funcionalidad, problemas abstractos y métricas de evaluación.</li> <li>3. Capacidad para elaborar un diseño experimental adecuado en función de las cuestiones de investigación planteadas y el escenario de aplicación.</li> <li>4. Capacidad para establecer resultados cuantitativos y cualitativos de la experimentación e identificar el alcance de los mismos.</li> <li>5. Capacidad para compartir y comunicar con la comunidad científica los resultados de investigación de forma efectiva.</li> </ol>
<b>CONTENIDO</b>	<p>Tema 1: Metodología de investigación en Procesamiento del Lenguaje Natural: introducción.</p> <p>Tema 2: Planteamiento del problema: elección del tema, tareas y escenarios, cuestiones de investigación (“research questions”), objetivos y determinación del estado del arte.</p> <p>Tema 3: Abstracción del problema y métricas de evaluación.</p> <p>Tema 4: Diseño experimental: A/B testing, sistemas de referencia y sistemas contrastivos, experimentos de ablación, etc. Leaderboards.</p>

	<p>Tema 5: Recolección y anotación de datasets. Crowdsourcing. Identificación y prevención de sesgos. Tratamiento de desacuerdos entre anotadores.</p> <p>Tema 6: Análisis de resultados: análisis cuantitativo, uso de múltiples métricas, tests de relevancia estadística, análisis de errores, análisis cualitativo, identificación de alcance y limitaciones.</p> <p>Tema 7: Comunicación científica: elaboración de artículos de investigación en Procesamiento del Lenguaje Natural. Peer review. Revistas académicas y congresos científicos en el área, y medidas de impacto científico. Publicación de código, datasets y otros recursos. Reproducibilidad. Repositorios de investigación. Comunicación oral.</p>
<p><b>OBSERVACIONES</b></p>	<p>Conocimientos previos recomendables:</p> <ul style="list-style-type: none"> <li>● Conocimientos básicos de matemáticas, estadística y programación.</li> <li>● Conocimientos básicos de aprendizaje automático</li> <li>● Lectura fluida en inglés</li> </ul> <p>Esta asignatura puede ser cursada aisladamente.</p>
<p><b>Nº DE LAS COMPET. GENERALES</b></p>	<p>Con la superación del curso se espera que el alumno complete las siguientes competencias generales especificadas en la memoria del máster:</p> <p>CPG1 - Adquirir capacidad de abstracción, análisis, síntesis y relación de ideas.</p> <p>CPG2 - Adquirir capacidad crítica y de decisión</p> <p>CPG3 - Adquirir capacidad de estudio y autoaprendizaje</p> <p>CPG4 - Adquirir capacidad creativa y de investigación</p>
<p><b>Nº DE LAS COMPET. ESPECÍFICAS</b></p>	<p>Con la superación del curso se espera que el alumno complete las siguientes competencias específicas indicadas en la memoria del máster:</p> <p>CE1 - Adquirir capacidad de comprender y manejar de forma básica los aspectos más importantes relacionados con los lenguajes y sistemas informáticos en general y, de manera especial, en los siguientes ámbitos: Tecnologías</p>

	<p>del lenguaje y de acceso a la información en web.</p> <p>CE3 - Adquirir capacidad de estudio de los sistemas y aproximaciones existentes para distinguir las aproximaciones más efectivas.</p> <p>CE4 - Adquirir capacidad para detectar carencias en el estado actual de la ciencia y la tecnología.</p> <p>CE5 - Adquirir capacidad para proponer nuevas aproximaciones que den solución a las carencias detectadas.</p> <p>CE6 - Adquirir capacidad de especificar, diseñar, implementar y evaluar tanto cualitativa como cuantitativamente los modelos y sistemas propuestos.</p> <p>CE7 - Adquirir capacidad para proponer y llevar a cabo experimentos con la metodología adecuada como para poder extraer conclusiones y determinar nuevas líneas de actuación e investigación.</p>
<p><b>ACTIVIDADES FORMATIVAS CON Nº DE HORAS POR CADA UNA</b></p>	<p>Tema 1: Metodología de investigación en Procesamiento del Lenguaje Natural: introducción. Actividad 1.1: [5 HT] Lectura de material sugerido</p> <p>Tema 2: Planteamiento del problema: elección del tema, cuestiones de investigación (“research questions”) y determinación del estado del arte. Actividad 2.1: [13 HT] Lectura de material sugerido Actividad 2.2: [11 HP] Ejercicios prácticos obligatorios</p> <p>Tema 3: Abstracción del problema y métricas de evaluación. Actividad 3.1: [13 HT] Lectura de material sugerido Actividad 3.2: [11 HP] Ejercicios prácticos obligatorios</p> <p>Tema 4: Diseño experimental Actividad 4.1: [13 HT] Lectura de material sugerido Actividad 4.2: [11 HP] Ejercicios prácticos obligatorios</p> <p>Tema 5: Recolección y anotación de data sets Actividad 5.1: [13 HT] Lectura de material sugerido Actividad 5.2: [11 HP] Ejercicios prácticos obligatorios</p> <p>Tema 6: Análisis de resultados Actividad 6.1: [13 HT] Lectura de material sugerido Actividad 6.2: [11 HP] Ejercicios prácticos obligatorios</p> <p>Tema 7: Comunicación científica Actividad 7.1: [10 HT] Lectura de material sugerido Actividad 7.2: [15 HP] Ejercicios prácticos obligatorios</p>

<b>METODOLOGÍA DOCENTE</b>	<p>La asignatura no tiene clases presenciales. Los contenidos teóricos se imparten a distancia y de acuerdo con las normas y estructuras de soporte telemático de la enseñanza en la UNED.</p> <p>El material docente incluye un resumen de los contenidos de cada tema y distintos tipos de actividades relacionadas con la consulta bibliográfica, la realización de ejercicios prácticos y, en ocasiones, la utilización de herramientas asociadas a las tecnologías y técnicas presentadas en el curso. Junto a las actividades y enlaces con fuentes de información externas, existe material didáctico propio preparado por el equipo docente con ejercicios que organizan las actividades prácticas del alumno.</p> <p>Tratándose de un máster orientado a la investigación, las actividades de aprendizaje se estructuran alrededor del estado del arte en cada uno de los temas del curso y a la parte práctica que el alumno deberá realizar en cada sección. Se utilizarán los recursos de la UNED para las actividades complementarias que se propondrán a lo largo del curso, actividades de asistencia o visualización (on-line, off-line) de seminarios de temas avanzados, impartidos por investigadores de reconocido prestigio.</p> <p>El plan de trabajo marcará unas pautas adecuadas para que el alumno alcance los objetivos al final del curso. Incluye una descripción ordenada y secuenciada de los contenidos de cada tema y orientaciones sobre las actividades a realizar, tanto de estudio teórico como práctico. Los ejercicios prácticos correspondientes a cada tema, así como la práctica final, son para apoyo y refuerzo del aprendizaje, y su realización es obligatoria para superar la asignatura.</p> <p>Las actividades pueden ser de los siguientes tipos:</p> <ol style="list-style-type: none"><li>1. <b>L:</b> Lectura y estudio del material correspondiente</li><li>2. <b>P:</b> Ejercicios prácticos de entrega obligatoria (exige una entrega en el aula virtual)</li><li>3. <b>PF:</b> Tarea relacionada con la práctica final obligatoria (exige al menos una entrega en el aula virtual)</li></ol>
<b>SISTEMAS DE EVALUACIÓN CON PONDERACIÓN</b>	<p>Evaluación continua a través de la elaboración de las tareas encomendadas a lo largo del curso. La ponderación de cada una de las tareas en la calificación final se comunicará a los estudiantes al comienzo del curso. En la evaluación se tendrá en cuenta, además de la calidad de los trabajos presentados, la interacción con el equipo docente y con el resto de estudiantes.</p>