

EL INGLÉS Y EL ESPAÑOL EN LA LINGÜÍSTICA COMPUTACIONAL (2017-18)

DATOS GENERALES

Código 38610

Créditos ECTS 6

Departamentos y áreas

Departamento	Área	Dpt. Resp.	Dpt. Acta
FILOL. ESPAÑOLA, LINGÜÍSTICA GENERAL Y TEORIA DE LA LITERATURA	LENGUA ESPAÑOLA		
LENGUAJES Y SISTEMAS INFORMÁTICOS	LENGUAJES Y SISTEMAS INFORMATICOS	SÍ	SÍ

Estudios en que se imparte

MÁSTER UNIVERSITARIO EN INGLÉS Y ESPAÑOL PARA FINES ESPECÍFICOS

Contexto de la asignatura

La lingüística computacional es un área de investigación de la inteligencia artificial cuyo objetivo es modelar automáticamente la capacidad de comunicación lingüística del ser humano, es decir la lengua mediante la que nos comunicamos. Se trata por tanto, de un campo interdisciplinar, en el que participan lingüistas, informáticos, y matemáticos, entre otros, pero también son necesarios expertos en un dominio concreto, como por ejemplo, médicos, psicólogos, juristas, etc. para poder modelar de manera precisa las características específicas de un contexto especializado.

Esta asignatura se plantea como una introducción general a los aspectos básicos de la Lingüística Computacional. En ella, se estudiarán sus fundamentos, así como conceptos relacionados con la lingüística de corpus; recursos y herramientas para el modelado computacional de una lengua y realización de forma automática de diferentes tipos de análisis lingüístico sobre un texto (léxico, sintáctico y semántico). Además, se explicarán aplicaciones prácticas. En concreto la traducción automática, con una gran relevancia en la sociedad actual, ya no sólo debido a la gran cantidad de información existente, sino a que esta está presente en un gran número de idiomas.

La demanda de lingüistas computacionales ha aumentado en los últimos años. Con el aumento y disponibilidad de información no estructurada y en distintos idiomas, las empresas están buscando formas de procesar y comprender mejor dicha información. Empresas de gran prestigio internacional como Facebook, Yahoo, Google, Microsoft o IBM, por nombrar algunas de ellas, tienen departamentos especializados en Lingüística Computacional, donde expertos en esta materia con distintos perfiles crean y desarrollan nuevos recursos con el objetivo de mejorar la capacidad de comprensión y generación de un idioma de forma automática.



OBJETIVOS

Objetivos específicos aportados por el profesorado (2017-18)

El objetivo de esta asignatura es que los estudiantes adquieran unos conocimientos mínimos y básicos sobre la Lingüística Computacional, desde una perspectiva fuertemente computacional centrada en análisis automático de corpus, modelado computacional del lenguaje natural, herramientas existentes y aplicaciones actuales, como la traducción automática, haciendo especial hincapié para los idiomas inglés y español.

CONTENIDOS

Contenidos teóricos y prácticos (2017-18)

Los contenidos abordados en esta asignatura están estructurados de tal forma que, después de cursarla, los alumnos del máster serán capaces de recopilar, analizar y transformar textos mediante el uso de herramientas que les permiten conocer mejor un idioma y procesarlo de manera automática o semi-automática. De este modo pueden colaborar en tareas de análisis y generación de una lengua de forma automática, con el objetivo de extraer conocimiento adaptado a la finalidad que se persiga, sin importar el dominio o género textual al que pertenezca dicha información o idioma en la que esté escrita. Además también serán capaces de determinar la idoneidad o no de introducir la traducción automática para reducir coste de traducción sin que se vea comprometida la calidad de las traducciones.

En concreto, los temas impartidos en la asignatura son los que a continuación se detallan:

1. Fundamentos de la Lingüística Computacional

- 1.1. Introducción a la Lingüística Computacional y el Procesamiento del Lenguaje Natural
- 1.2. Módulos de un sistema de Lingüística Computacional y problemas derivados del tratamiento automático del lenguaje

2. Lingüística de corpus

- 2.1. Diseño y compilación de un corpus
- 2.2. Análisis automático de un corpus
- 2.3. Extracción automática de terminología de un corpus
- 2.4. Diseño y evaluación de expresiones regulares
- 2.5. Anotación de un corpus

3. Análisis computacional del lenguaje natural

- 3.1. Análisis léxico desde la perspectiva computacional
- 3.2. Análisis sintáctico desde la perspectiva computacional
- 3.3. Análisis semántico desde la perspectiva computacional

4. Aplicación práctica: Traducción Automática

- 4.1. Definición, usos y aplicaciones de la traducción automática
- 4.2. Evaluación de la traducción automática
- 4.3. Técnicas de traducción automática

5. Investigación en Lingüística Computacional: problemas abiertos

- 5.1. Áreas de investigación y disciplinas afines. Retos a resolver.

La asignatura tiene un carácter eminentemente práctico, ya que está estructurada en seminarios teórico-prácticos y clases con ordenador, con un mayor número de sesiones de ordenador para permitir que los alumnos pongan en práctica todos los conocimientos adquiridos.

EVALUACIÓN

Instrumentos y criterios de Evaluación 2017-18

Todos los contenidos de la asignatura están fuertemente interrelacionados y conectados, de tal manera que lo que se aprenda en un tema servirá para aplicarlo en temas posteriores. Por tanto, la evaluación de la asignatura en la convocatoria ordinaria consistirá en la realización y entrega de un portafolios global con tres partes diferenciadas pero relacionadas entre sí, donde cada una de ellas tendrá un peso de un 30% de la nota global:

• Parte I. Lingüística de corpus (30%). En esta parte del trabajo, el alumno tendrá que recopilar y anotar un corpus y realizar análisis superficial (frecuencias, n-gramas, expresiones regulares, extracción de terminología, etc.)

• Parte II. Fundamentos de la Lingüística Computacional y análisis computacional del lenguaje (30%). En esta parte, el alumno utilizará el corpus desarrollado en la parte I y lo procesará con herramientas de lingüística computacional, valorando los resultados obtenidos y analizando en detalle los errores producidos.

• Parte III. Traducción Automática (30%). En esta parte, se evaluará un traductor automático a elección del alumno y se determinará su idoneidad para la traducción del corpus de la parte I. Además se analizarán los patrones de errores más habituales cometidos por el traductor automático elegido.

El portafolios supondrá el 90% de la nota final. **El 10% restante se obtendrá mediante la entrega de un resumen de los contenidos de la asignatura en relación con el perfil de cada estudiante.** La nota final de la asignatura se calculará de 0 a 10, siendo necesario al menos un 5 para aprobar.

En caso de suspenso o no presentado en la convocatoria ordinaria, la nota final de la asignatura para la convocatoria extraordinaria, calculada de 0 a 10, corresponderá únicamente a la evaluación del portafolios que los alumnos deberán entregar en la fecha que se les indique y que deberá incluir, además de los tres apartados anteriormente descritos, un apartado extra (Parte IV) con el resumen de los contenidos de la asignatura y una discusión relacionada con los problemas abiertos y retos existentes en la Lingüística computacional.

Para ambas convocatorias, los trabajos teórico/prácticos realizados han de ser originales. La detección de copia o plagio supondrá la calificación de "0" en la prueba correspondiente. Se informará la dirección de Departamento y de la facultad sobre esta incidencia. La reiteración en la conducta en esta u otra asignatura conllevará la notificación al vicerrectorado correspondiente de las faltas cometidas para que estudien el caso y sancionen según la legislación (Reglamento de disciplina académica de los Centros oficiales de Enseñanza Superior y de Enseñanza Técnica dependientes del Ministerio de Educación Nacional BOE 12/10/1954).

Tipo	Criterio	Descripción	Ponderación
ACTIVIDADES DE EVALUACIÓN DURANTE EL SEMESTRE	<p>Se evaluará la entrega de un resumen de los contenidos de la asignatura en relación con el perfil de cada estudiante.</p> <p>Actividad recuperable. Para la convocatoria extraordinaria, esta actividad se incorporará al portafolios como un apartado extra (apartado IV) que además de contener el resumen de la asignatura, deberá incluir una discusión relacionada con los problemas abiertos y retos existentes en la Lingüística computacional.</p>	Entrega de un resumen sobre Lingüística Computacional	10



<p>ACTIVIDADES DE EVALUACIÓN DURANTE EL SEMESTRE</p>	<p>Se evaluará la entrega de un portafolios global con tres partes diferenciadas pero relacionadas entre sí,:</p> <ul style="list-style-type: none">• Parte I. Lingüística de corpus (30%). En esta parte del trabajo, el alumno tendrá que recopilar y anotar un corpus y realizar análisis superficial (frecuencias, n-gramas, expresiones regulares, extracción de terminología, etc.)• Parte II. Fundamentos de la Lingüística Computacional y análisis computacional del lenguaje (30%). En esta parte, el alumno utilizará el corpus desarrollado en la parte I y lo procesará con herramientas de lingüística computacional, valorando los resultados obtenidos y analizando en detalle los errores producidos.• Parte III. Traducción Automática (30%). En esta parte, se evaluará un traductor automático a elección del alumno y se determinará su idoneidad para la traducción del corpus de la parte I. Además se analizarán los patrones de errores más habituales cometidos por el traductor automático elegido. <p>Actividad recuperable.</p>	<p>Portafolios Lingüística Computacional</p>	<p>90</p>
--	--	--	-----------