

RESPUESTA A LAS ALEGACIONES RECIBIDAS DE FECHA 03-06-2013

CRITERIO I. DESCRIPCIÓN DEL TÍTULO

CRITERIO 5. PLANIFICACIÓN DE LAS ENSEÑANZAS

Se incluye el sistema de evaluación en la nueva asignatura: Técnicas basadas en grafos aplicados al procesamiento del lenguaje de la especialidad de Tecnologías del lenguaje en la web.

MODIFICACIONES EN EL MÁSTER DE LENGUAJES Y SISTEMAS INFORMÁTICOS PARA SU APROBACIÓN POR LA ANECA:

Justificación para la creación de nueva asignatura

NOMBRE DE LA ASIGNATURA: Técnicas basadas en grafos aplicados al procesamiento del lenguaje

CARÁCTER Optativo

ECTS 6

CUATRIMESTRE Anual

IDIOMA Español

ESPECIALIDAD EN QUE SE IMPARTE: Tecnologías del Lenguaje en la web

Justificación:

En la especialidad de Tecnologías del lenguaje en la web del Máster universitario en lenguajes y sistemas informáticos, en la que se enmarca la asignatura que se propone, se estudian distintos aspectos de las áreas de Procesamiento del Lenguaje Natural y Recuperación de Información en las distintas asignaturas.

En muchas aplicaciones del Procesamiento del Lenguaje Natural las entidades pueden representarse de forma natural como nodos de un grafo y las relaciones entre ellas como enlace entre dichos nodos.

Se ha desarrollado una amplia investigación que ha demostrado que las representaciones basadas en grafos de las unidades lingüísticas, como palabras, oraciones y documentos dan lugar a nuevas y eficientes soluciones en numerosas tareas de procesamiento del lenguaje. Estas aplicaciones van desde el etiquetado léxico, el análisis sintáctico y la desambiguación del sentido de las palabras, a la extracción de información, o la generación de resúmenes.

Así mismo, la teoría de redes, tiene una gran relevancia en los sistemas de recuperación información, que ha permitido grandes avances en el desarrollo de aplicaciones en la web, que van desde el desarrollo de buscadores hasta el análisis de redes sociales.

Por ello consideramos fundamental contar una asignatura que permita a los estudiantes conocer los fundamentos de la teoría de grafos, algunas herramientas para su manejo y algunas de sus principales aplicaciones al Procesamiento del Lenguaje, que no se estudian en otras asignaturas de la especialidad.

2. Justificación del programa

2.1 Referentes académicos

2.1.1. Objetivos generales y su relación con las competencias y perfiles.

Los objetivos generales del master en Lenguajes y Sistemas Informáticos son enlazar los conocimientos básicos asociados a las áreas de Ciencias de la Computación e Inteligencia Artificial, y Lenguajes y Sistemas Informáticos, dentro de los estudios de grado, con las fronteras actuales de la investigación en ambas áreas con dos propósitos básicos:

1. Aprovechar las competencias de los profesores de ambos departamentos para planificar e impartir un programa modular y estructurado con óptima utilización de recursos,

2. Establecer sinergias en contenidos y temas de investigación de forma que nuestros alumnos tengan una visión clara de dónde están los problemas de la computación y cuáles son los métodos disponibles para resolverlos, más allá de las fronteras siempre difusas entre departamentos y entre áreas de conocimiento.

De esta forma se harán más evidentes los potenciales solapes en los contenidos de las asignaturas y serán más eficientes los esfuerzos de profesores y alumnos. Adicionalmente, es previsible un beneficio mutuo en tareas de investigación y una mejor integración laboral de nuestros estudiantes.

Finalmente, la competencia docente e investigadora queda garantizada al considerar que los profesores de este programa de posgrado han sido y son los responsables de la docencia en IA y LSI en los actuales planes de estudio y han participado y/o participan en proyectos nacionales e internacionales en las áreas que define la línea de especialización curricular que se oferta.

2.1.2. Adecuación a los objetivos estratégicos de la UNED.

El programa del master en Lenguajes y Sistemas Informáticos es de iniciación a la investigación. La línea curricular sobre procesamiento de lenguaje natural y recuperación de información muestra la característica distintiva de sinergia y reutilización del conocimiento porque toda la licitación del conocimiento en los sistemas de IA está basada en descripciones declarativas en lenguaje natural de los procedimientos que usan los humanos para resolver problemas. Por consiguiente, el alumno potencial de nuestro programa conjunto encontrará enlaces, coincidencia y posibilidades de reutilización del conocimiento.

2.1.3. Refinamiento de los referentes académicos.

El resto de los argumentos que nos ayudan a profundizar en los referentes académicos que justifican este master se describen en los siguientes puntos:

2.1.A. Referentes Académicos Máster “Lenguajes y Sistemas Informáticos”

2.1.1.A Objetivos generales y su relación con las competencias y perfiles

El master en Lenguajes y Sistemas Informáticos, con perfil investigador, se plantea con un

doble objetivo:

Dar continuidad, dentro del nuevo marco de estudios europeos superiores, al programa de doctorado del departamento de Lenguajes y Sistemas Informáticos, y con él a las actividades formativas de los grupos de investigación asociados (ver "líneas de investigación asociadas" más abajo).

Servir como piedra angular para, a partir de la experiencia obtenida con este master, proponer dentro de este programa de posgrado otros títulos de master profesionales que respondan de forma directa a las demandas sociales y del mercado de trabajo en informática. La relación entre los estudios con perfil investigador que proponemos y los futuros estudios profesionales se establece de forma natural, puesto que las dos líneas de investigación del departamento son de carácter marcadamente aplicado.

Para dar continuidad a los programas de doctorado, y teniendo en cuenta los conocimientos que se adquirirán en los futuros estudios de grado, este master pretende dar a los alumnos una formación general en temas avanzados de Lenguajes y Sistemas Informáticos (que complemente la formación básica recibida en los estudios de grado), junto con una formación específica en los temas propios de investigación, y además con las técnicas y corpus de conocimiento generales (ya sean propios de Lenguajes y Sistemas Informáticos como de Inteligencia Artificial) que son de aplicación directa en estas líneas de investigación. Finalmente, en la etapa final del master los alumnos se iniciarán a la metodología de investigación de forma práctica, de forma que su ingreso en el programa de doctorado se produzca de forma idónea.

El programa de doctorado en Sistemas Inteligentes de la UNED se fundamenta en éste Master y supone la cantera principal de personal para la actividad investigadora de las líneas de investigación del departamento.

2.1.2.A Adecuación de los Objetivos Estratégicos de la Universidad

La combinación de master y doctorado de Lenguajes y Sistemas Informáticos garantiza la continuidad –y posibilita la mejora- de dos equipos de investigación consolidados y con amplia proyección internacional (ver abajo), que representan un porcentaje significativo de las actividades investigadoras en la UNED (la Escuela Técnica Superior de Informática es, en la actualidad, la escuela/facultad que aporta mayores ingresos a la UNED en el capítulo de investigación). Además sienta las bases para futuros master profesionales en Lenguajes y Sistemas Informáticos, un área de máximo interés desde el punto de vista del mercado laboral y la naciente sociedad de la información.

2.1.3.A Interés y Relevancia Académica-Científica-Profesional

El área de Lenguajes y Sistemas Informáticos es uno de los pilares de los estudios de informática; su interés científico y profesional no necesita ser puesto de relieve. Por otro lado, la línea de especialización cubre algunos de los aspectos tecnológicos más relevantes de la sociedad "en red": el acceso, exploración y análisis de grandes volúmenes de información textual en la WWW. El impacto científico de estos temas ha aumentado de forma creciente a lo

largo de los últimos años hasta acaparar la atención de una gran comunidad de investigadores.

2.1.4.A Equivalencia en el Contexto Internacional

El programa está elaborado a partir de la experiencia de colaboración estrecha (en el marco, sobre todo, de proyectos europeos) con más de una decena de universidades europeas y norteamericanas. Desde la perspectiva del perfil investigador del master/doctorado que proponemos, estas relaciones (detalladas en el punto 2.1.7 más abajo) garantizan la vigencia e interés del programa propuesto.

2.1.5.A Adecuación del Título al Nivel Formativo del Postgrado (descriptores de Dublín)

La base del programa la forman asignaturas avanzadas de Lenguajes y Sistemas Informáticos que en estos momentos se dan en quinto curso – y que posiblemente desaparezcan en los estudios de grado -, junto con las asignaturas del actual programa de doctorado del departamento y una serie de asignaturas del master en Inteligencia Artificial de este mismo posgrado cuyos contenidos son relevantes para la línea de especialización de LSI. Esta combinación es la más natural para unos estudios de posgrado orientados a la investigación y, al mismo tiempo, aprovecha la flexibilidad de los nuevos estudios de grado para establecer sinergias con áreas afines (Inteligencia Artificial, en nuestro caso).

El plan de estudios del master en LSI se ajusta a los descriptores de Dublín para segundo ciclo:

- *"knowledge and understanding that is founded upon and extends and/or enhances that typically associated with Bachelor's level, and provides a basis or opportunity for originality in developing and/or applying ideas, often within a research context."* Los conocimientos de este master comienzan con los temas más avanzados de la actual licenciatura, que son tanto de diseño software en general como de introducción a las líneas de investigación; esto asegura la continuidad y profundización respecto a los futuros estudios de grado. El segundo año está reservado para temas de investigación – búsqueda de información, entornos colaborativos de enseñanza y aprendizaje- que experimentan cambios a gran velocidad, y en los que los alumnos deben reflexionar sobre multitud de problemas sin resolver, contando con el bagaje de técnicas generales que tienen posibilidades creativas de aplicación.
- *"can apply their knowledge and understanding, and problem solving abilities in new or unfamiliar environments within border (or multidisciplinary) contexts related to their field of study".* El carácter multidisciplinar de las líneas de especialización del master – la búsqueda de información textual está muy relacionada con la lingüística computacional, y los entornos de enseñanza y aprendizaje con la pedagogía- garantizan la práctica de los alumnos en aplicar sus conocimientos a disciplinas con las que no están previamente familiarizados.
- *"have the ability to integrate knowledge and handle complexity, and formulate judgements with incomplete or limited information, but that include reflecting on social and ethical responsibilities linked to the application of their knowledge and judgements".* Las dos líneas de especialización del master son inseparables de los aspectos sociales y éticos de la investigación tecnológica; aunque no se han reservado créditos específicos para este tipo de

consideraciones, varias asignaturas tocan aspectos éticos y de responsabilidad social. Por ejemplo, en la asignatura sobre “motores de búsqueda” se reflexiona sobre el impacto económico de los algoritmos de posicionamiento en resultados de búsqueda y sobre cómo se intentan manipular sus resultados de forma fraudulenta. La línea de investigación en enseñanza y aprendizaje, por otro lado, es una reflexión continua sobre el uso de la tecnología para fines de impacto social.

- *“can communicate their conclusions, and the knowledge and rationale underpinning these, to specialist and non-specialist audiences clearly and unambiguously”*. Mientras el primer curso del master se evalúa de forma relativamente tradicional – a través de exámenes presenciales -, el segundo curso ofrece un abanico más amplio y flexible de formas de evaluación, en las que el trabajo de análisis, síntesis y reflexión juega un papel fundamental.
- *“have the learning skills to allow them to continue to study in a manner that may be largely self-directed or autonomous”*. Los trabajos de síntesis y reflexión de las asignaturas de especialización, junto con el trabajo de iniciación a la investigación – que puede significar el germen de una futura tesis doctoral- están orientados a garantizar la madurez de los alumnos para comenzar su trabajo de investigación aportando iniciativa y autonomía.

2.1.6.A Líneas de Investigación Asociadas (Grupos, Proyectos, Tesis y Publicaciones)

El master/doctorado en Lenguajes y Sistemas Informáticos se sustenta sobre dos grupos de investigación que agrupan a casi todos los miembros del departamento homónimo de la UNED:

- **Grupo de Procesamiento de Lenguaje Natural y Recuperación de Información.**

Comenzó su actividad investigadora en 1995, y desde entonces ha estado involucrado en proyectos de investigación básica y aplicada en torno a el diseño y evaluación de sistemas inteligentes de recuperación de información (mono y multilingües), la creación y aplicación de bases de datos léxico-semánticas a gran escala y las interfaces de lenguaje natural.

La actividad de este grupo en los últimos cinco años, en cifras, ha sido la siguiente:

- Participación en **cinco proyectos de la Comisión Europea** (V y VI programas marco), dos proyectos CICYT coordinados, un proyecto CyTED y un proyecto financiado por la UNED. El grupo se ha integrado, además, en tres redes de excelencia europeas, una red iberoamericana y una red española.
- Más de **90 artículos científicos**, la mayoría en publicaciones de ámbito internacional, incluyendo las revistas del ramo de más prestigio (Computational Linguistics, Information Processing and Management, Information Retrieval) y los congresos mundiales de referencia (SIGIR, ACL, CLEF). Algunos de estos trabajos están elaborados con la participación de autores europeos y norteamericanos.
- Estos artículos han sido **referenciados en más de 200 ocasiones** (excluyendo autorreferencias) según scholar.google.com.
- Además de artículos científicos, el grupo ha **generado recursos para la investigación en el área**, como la base de datos léxica EuroWordNet (en colaboración con universidades

europeas y españolas) y las colecciones CLEF de evaluación para sistemas de recuperación de información multilingüe y sistemas de búsqueda de respuestas; estas colecciones son actualmente el estándar internacional para evaluar este tipo de sistemas, y han sido utilizadas por más de 90 grupos de investigación de todo el mundo.

- En los últimos cinco años se han defendido **cuatro tesis doctorales** (dos de ellas premio extraordinario de doctorado de la UNED, y una premio a la mejor monografía de la Sociedad Española para el Procesamiento del Lenguaje Natural), y se terminarán cuatro más en el curso 2005-2006.
- Los miembros del grupo han **participado en más de cuarenta comités** que incluyen los congresos mundiales y asociaciones internacionales de mayor prestigio, presidiendo algunos de ellos, como el comité de programa y el comité de organización de la 10th European Conference on Digital Libraries o el comité de programa del ACL Workshop on Recent Advances in Natural Language Processing and Information Retrieval, y co-editando seis volúmenes de actas en Springer-Verlag LNCS.
- Sus estudiantes han realizado **estancias de investigación** en centros como el NIST (National Institute for Standard Technologies, EEUU), la U. Maryland, la U. Southern California, la New York University, la Universidad de Brighton o la Universidad de Chile.
- El grupo ha **organizado el congreso mundial más relevante de Procesamiento del Lenguaje Natural** (congreso de la ACL, International Association for Computational Linguistics) en 1997, y **organiza** (junto con la U. Alicante y la Biblioteca Virtual Miguel de Cervantes) el **congreso europeo de Bibliotecas Digitales (ECDL)** en 2006. Desde 2001 participa activamente en la organización de la **competición científica CLEF (Cross-Language Evaluation Forum)**, referencia para la investigación en Recuperación de Información Multilingüe.

2.1.7.A Situación de la I+D+I del Sector Profesional

El impacto en la industria de las líneas de investigación de este master ha sido enorme: baste recordar que Google es, en estos momentos, la empresa de medios con mayor valor en bolsa, por encima de gigantes como Microsoft. En Europa, de momento, la distancia entre investigación y mercado es todavía mucho mayor que en Estados Unidos, y por tanto existe todavía un enorme mercado potencial por desarrollar. Para ello es necesario contar con profesionales de la más alta cualificación universitaria, puesto que la industria y la innovación deben ir necesariamente de la mano (Google es la empresa estadounidense con mayor ratio de doctorados).