ASIGNATURA DE MÁSTER:

MODELOS COMPUTACIONALES DEL DISCURSO: SEMÁNTICA Y PRÁGMÁTICA

Curso 2012/2013

(Código:31101292)

1.PRESENTACIÓN

Esta asignatura presenta los contenidos relacionados con la teoría y la práctica de

1. la semántica computacional, que incluye además una presentación sobre la web semántica y
2. los modelos computacionales de discurso, que por un lado permiten resolver problemas lingüísticos como la anáfora y por otro, son de gran relevancia para aplicaciones como los sistemas pregunta y respuesta o los sistemas de diálogo.

La planificación de esta asignatura OPTATIVA y ANUAL, esto es la dedicación esperada del estudiante (en horas), dado que es una asignatura de 6 créditos (150h de dedicación), son:

Horas de estudio teórico: 70
Horas de trabajo práctico: 60
Horas de actividades complementarias: 20

2.CONTEXTUALIZACIÓN

Esta asignatura se encuadra en el módulo “Tecnologías del lenguaje en la web” dentro de la especialidad “LSI-1 Tecnologías del lenguaje en la red” de la titulación de posgrado “Master en Lenguajes y Sistemas Informáticos” en la ETSI de Informática de la UNED (http://posgrados.informatica.uned.es/master-IA-SI/).

3.REQUISITOS PREVIOS RECOMENDABLES

No hay ningún prerrequisito diferente de los generales de acceso a este programa de posgrado orientado a la investigación, aunque es necesario un cierto dominio del inglés para el estudio de los temas y el uso de algunas herramientas informáticas.

Esta asignatura puede ser cursada aisladamente, aunque el estudiante se beneficiaría si hubiera cursado previamente una asignatura de *Procesamiento del Lenguaje Natural* (como la ofertada en este programa).

4.RESULTADOS DE APRENDIZAJE

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
01550FE673FA5AA7442933480AEC8928

El objetivo general de la materia es proporcionar al alumno una visión global de las técnicas y tecnologías involucradas en el estudio de la semántica y pragmática computacional y sus aplicaciones. Los objetivos concretos o destrezas y competencias son:

OB1: Saber los conceptos básicos sobre la semántica y la pragmática de un lenguaje.

OB2: Saber distinguir los diferentes niveles de información que se pueden encontrar en los textos.

OB3: Conocer los modelos semánticos y pragmáticos (contexto y discurso) del estado del arte.

OB4: Conocer las principales técnicas y tecnologías implicadas.

OB5: Saber representar la semántica y pragmática tanto de textos como diálogos y tener criterios para seleccionar las formas de gestión más adecuada para el uso.

OB6: Conocer algoritmos para etiquetado análisis semántico y pragmático y tareas asociadas (desambiguación, anáfora, identificación de los hilos de discurso etc), siendo capaz de criticarlos de acuerdo a la solución necesaria para problemas concretos.

OB7: Identificar una metodología básica de evaluación de los resultados a partir de los requisitos iniciales.

5.CONTENIDOS DE LA ASIGNATURA

Tema 1.- Introducción.
Tema 2.- Semántica Computacional
- 2.1 Representación del significado
- 2.1 Semántica léxica
- 2.2 Desambiguación del sentido de las palabras

Tema 3.- La web semántica (Web 3.0)
- 3.1 Stándares
- 3.2 Ontologías y su población
- 3.3 Modelos para recuperación de información

Tema 4.- Análisis del discurso (I)
- 4.1 Estructura del discurso
- 4.2 Análisis de referencias
- 4.3 Sistemas Pregunta y Respuesta

Tema 5.- Análisis del discurso (II)
- 5.1 Modelo DRT
- 5.2 Generación de presentaciones multimedia

Tema 6.- Interacción persona-ordenador.
- 6.1 Conceptos básicos
- 6.2 Sistemas de diálogo

Tema 7.- Evaluación
- 7.1 Evaluación de tareas semánticas
- 7.2 Componentes de evaluación de diálogos
- 7.3 Evaluación en la web semántica

Dentro de las actividades de aprendizaje se especifican las lecturas más adecuadas para cada uno de los temas. Se presentarán y se facilitará, en su caso, el acceso a diferentes ejemplos y herramientas accesibles desde el aula virtual.

Respecto a los objetivos generales de cada tema:

o El objetivo global del Tema 1 es presentar al alumno aquellos conceptos y conocimientos preliminares sin los que no podría ubicar los contenidos de la asignatura. Se pretende, además, justificar el interés de la asignatura, motivar al alumno en su estudio y presentar las posibles aplicaciones de los contenidos.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
01550FE673FA5AA7442933480AEC8928

o En el Tema 2 se presentan diferentes modelos semánticos, y diversos criterios comúnmente utilizados en la bibliografía para estudiar la cobertura y adecuación de los mismos.

o El Tema 3 incluye los conceptos básicos y algunos de los estándares de la web semántica. Además se estudiarán recursos existentes y aspectos candentes como la población de ontologías. Finalmente se estudiarán algunos modelos de recuperación de información basada en ontologías.

o Los Temas 4 y 5 se dedican a presentar el concepto de discurso y las teorías relacionadas, que además tienen una proyección aplicada, como en el caso de las presentaciones o interfaces que gestionan y combinan adecuadamente datos multimedia en aplicaciones o los sistemas de pregunta y respuesta.

o En el Temas 6 se dará a conocer al alumno la problemática relacionada con la interacción persona y ordenador con un planteamiento de ingeniería del software. La segunda parte se dedica a un tema de elevado interés, la gestión de diálogos para aplicaciones.

o Finalmente, el tema 7, se dedicará a la presentación de algunas de las técnicas que se utilizan para la evaluación de los aspectos relacionados con algunos de los temas presentados en esta asignatura.

El plan de trabajo del estudiante, de acuerdo con el contenido de la asignatura se organiza como sigue:

Tema	Horas de teoría	Horas de práctica
Tema 1	4	0
Tema 2	10	6
Tema 3	10	6
Tema 4	10	6
Tema 5	10	6
Tema 6	10	6
Tema 7	10	6

Práctica 50h

6.EQUIPO DOCENTE

- ANA Mª GARCIA SERRANO
- JUAN MANUEL CIGARRAN RECUERO
- ALEJANDRO RODRIGUEZ ASCASO

7.METODOLOGÍA

La asignatura no tiene clases presenciales. Los contenidos teóricos se impartirán a distancia, de acuerdo con las normas y estructuras de soporte telemático de la enseñanza en la UNED.
El material docente incluye distintos tipos de actividades relacionadas con la consulta bibliográfica y, en ocasiones, la utilización de herramientas asociadas a las tecnologías y técnicas presentadas en el curso. Además, será necesario el estudio del libro de base propuesto por el equipo docente.
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.
Tratándose de un master orientado a la investigación, las actividades de aprendizaje se estructuran en torno al estado del arte en cada uno de los temas del curso y a los aspectos en los que se va a focalizar la 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 o presencial.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
01550FE673FA5AA7442933480AEC8928

8.BIBLIOGRAFÍA BÁSICA

Comentarios y anexos:

Material de estudio

Capítulos seleccionados de los libros:

1. Daniel Jurafsky and James H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing,* Computational Linguistics and Speech Recognition, (2nd Ed.), Pearson International Edition, 2009, ISBN: 013504196-1
   (http://www.cs.colorado.edu/~martin/slp2.html).
2. Verdejo, F. y García Serrano A. (Edrs), 2009, UNED (libro electrónico)

En concreto:

Tema 1.- Introducción. (Jurafski&Martin 2009, cap 1)
Tema 2.- Semántica Computacional (Jurafski&Martin 2009, cap 17, 19 y 20)
Tema 3.- La semántica de la web. Web 3.0 (Verdejo, García-Serrano, 2009)
Tema 4.- Análisis del discurso (I) (Jurafski&Martin 2009, secciones 21.1, 21.2 y cap 23)
Tema 5.- Análisis del discurso (II) (Jurafski&Martin 2009, cap 21 y Dale et al 2000, cap 12)
Tema 6.- Interacción persona-ordenador. (Jurafski&Martin 2009 cap 24 y Verdejo, García-Serrano, 2009)
Tema 7.- Evaluación (Verdejo, García-Serrano, 2009)

9.BIBLIOGRAFÍA COMPLEMENTARIA

Comentarios y anexos:

Bibliografía de consulta

Allen, J.F. *Natural Language Understanding*, Benjamin Cummings, 1987, Second Edition, 1994, ISBN: 0-8053-0335-9

Cole et al, 97 *Survey of the State of the Art in Human Language Technology*. Cambridge University Press (http://www.dfki.de/~hansu/HLT-Survey.pdf)

Dale, R. ; Moisi H.; Somers H. "*Handbook of natural language processing*" 2000, ISBN: 0-8247-9000-6

Curso universitario de lingüística general. *Tomo II : Semántica, pragmática, morfología y fonología.* Autor: Moreno Cabrera, Juan Carlos. Editorial: Síntesis. Año de edición: 2000. Páginas: 679. ISBN: 9788477382508

10.RECURSOS DE APOYO AL ESTUDIO

La plataforma de e-Learning *Alf* proporcionará el aula virtual y el adecuado interfaz de interacción entre el alumno y sus profesores. *Alf* permite impartir y recibir formación, gestionar y compartir documentos, crear y participar en comunidades temáticas, así como realizar proyectos online. Se usarán las herramientas necesarias para que, tanto el equipo docente como el alumnado, encuentren la manera de compaginar el trabajo individual con el aprendizaje cooperativo.

Para las actividades complementarias se utilizarán recursos audiovisuales de la UNED (televisión, visualización etc), de acuerdo a la planificación que el equipo docente publicará anualmente.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
01550FE673FA5AA74429334804EC8928

11.TUTORIZACIÓN Y SEGUIMIENTO

La tutorización de los alumnos se llevará a cabo por la profesora de la asignatura, a través de la plataforma de e-Learning *Alf*, por teléfono y por correo electrónico (agarcia@lsi.uned.es, 91 398 7993 o juanci@lsi.uned.es, 91 3987620).

Breve reseña: Ana García Serrano, ha participado y coordinado asignaturas de forma continua en diferentes programas de doctorado. Ha dirigido 5 tesis doctorales y mas de 10 trabajos para el DEA o tesis de master. Su ámbito de investigación se relaciona con la interacción persona-ordenador, la representación del conocimiento y los sistemas multiagente.

Es profesora titular en la ETSI de Informática de la UNED y miembro del Grupo de investigación NLP&IR dirigido por la profesora F. Verdejo. En los últimos años su investigación se orienta al procesamiento del lenguaje natural, la recuperación de información y la gestión de la interacción (diálogos) y en sus aplicaciones.

Breve reseña: Juan Manuel Cigarran Recuero es profesor ayudante doctor en la ETSI de Informática de la UNED y miembro del Grupo de investigación NLP&IR dirigido por la profesora F. Verdejo. Sus intereses científicos se centran en el procesamiento del lenguaje natural, la recuperación de información multilingue y multimedia, así como en la minería y extracción de información aplicada principalmente a la web. Actualmente, compagina la labor docente e investigadora con la participación en proyectos tanto nacionales como europeos.

Mas información en http://nlp.uned.es/.

12.EVALUACIÓN DE LOS APRENDIZAJES

La evaluación se realizará a partir de los ejercicios propuestos a lo largo del curso, así como de un trabajo práctico final individual que organizará y ampliará el trabajo realizado en los ejercicios, según el calendario que se publicará en el aula virtual al comienzo del periodo lectivo de la asignatura.
Estas actividades prácticas se irán generando de forma dinámica en el curso virtual, y a modo de ejemplo, se refieren a:

- Representación de información. Estándares de anotaciones semánticas.
- Representación de diálogos. Gestión de metainformación.
- Metodologías de recogida de corpus (de diálogos) y de evaluación.

La calificación consistirá en un 60% según el seguimiento del trabajo teórico-práctico y de autoevaluación realizado y en un 40% el proyecto final del estudiante.

13.COLABORADORES DOCENTES

Véase equipo docente.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
01550FE673FA5AA7442933480AEC8928