26-27

GUÍA DE ESTUDIO PÚBLICA

WEB SEMÁNTICA Y ENLAZADO DE DATOS

CÓDIGO 31108018

UNED

26-27

WEB SEMÁNTICA Y ENLAZADO DE DATOS

CÓDIGO 31108018

ÍNDICE

PRESENTACIÓN Y CONTEXTUALIZACIÓN
REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA
EQUIPO DOCENTE
HORARIO DE ATENCIÓN AL ESTUDIANTE
COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE
RESULTADOS DE APRENDIZAJE
CONTENIDOS
METODOLOGÍA
SISTEMA DE EVALUACIÓN
BIBLIOGRAFÍA BÁSICA
BIBLIOGRAFÍA COMPLEMENTARIA
RECURSOS DE APOYO Y WEBGRAFÍA
IGUALDAD DE GÉNERO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

Representación del Conocimiento e Inferencia. La asignatura se inscribe en esta tarea genérica de los sistemas inteligentes: la construcción de una representación interna, significativa y operativa que hace el agente de su entorno exterior. Representación que, en este área, se produce sobre *algún lenguaje simbólico*, con su semántica concreta y sus *cálculos de inferencia*. En sentido amplio, estos lenguajes se vienen proponiendo como refinamiento de diversas familias de formalismos, cada una con sus asunciones ontológicas y epistemiológicas: desde las derivadas de las 'lógicas clásicas' (proposicional, lógicas de predicados, descriptivas, ...), las lógicas modales y temporales o las aproximaciones probabilistas hasta las lógicas difusas.

La relación de la asignatura con el programa del Máster. Prácticamente todas las asignaturas describen procesos que permiten a un agente inteligente variar su estado interno para responder más adecuadamente a variaciones exteriores. En algunas de ellas ese estado interno y sus procesos de adaptación son más legibles directamente por un observador humano ya que se expresan en alguno de esos lenguajes simbólicos. Desde esta perspectiva, la asignatura formalmente más relacionada con ésta es *Métodos Simbólicos*, con la que además comparte los lenguajes de representación (proposicional, predicados, descriptivas). En el mismo marco formal, pero con otra opción de lenguaje, se encuentra *Métodos Probabilistas*. Y estos formalismos, junto a los no simbólicos de otras asignaturas (como Aprendizaje o Minería de Datos), se ofrecen para la resolución de procesos en asignaturas que se pueden considerar más finalistas: Visión, Robótica, Descubrimiento de Información en Textos o Fundamentos del Procesamiento Lingüístico.

El contexto específico de la asignatura. La disponibilidad de grandes conjuntos de datos está favoreciendo el diseño y el entrenamiento de agentes inteligentes, especialmente desde la perspectiva renovada de los sistemas conexionistas en IA. La Inteligencia Artificial Simbólica, como línea complementaria (e hibridable), también puede verse beneficiada por la disponibilidad masiva de datos curados, estructurados y compartidos.

La Web de Datos Públicos Enlazados articula colectivamente los procesos de publicación y uso de fuentes de datos diversas. Es una arquitectura común sobre la que además se superpone el diseño compartido de conceptualizaciones (vocabularios, taxonomías, ontologías) que estructuran estos datos para determinados objetivos.

Estas bases de conocimiento, publicas e interconectadas, facilitan tareas de recopilación de información y de inferencia. Siempre dentro de los estándares, con sus restricciones, en que se conforman estas bases en la Web Semántica. Desde una perspectiva más general, la investigación sobre Grafos de Conocimiento (Knowledge Graphs) estudia la creación, refinamiento y explotación de bases de conocimiento similares, no necesariamente ubicadas en la Web Semántica.

Objetivos de la asignatura. La aproximación a la Web Semántica se hace desde la perspectiva del consumidor de datos y no tanto como ingeniero responsable de la publicación de datos de una institución. Se revisan los estándares y funcionalidades de la Web de Datos así como las conceptualizaciones superpuestas (como vocabularios u ontologías) a estas bases de conocimiento públicas. Este objetivo, de la primera parte de la

asignatura, sirve de fundamento para la consecución del segundo objetivo: la construcción, refinamiento y explotación de un Grafo de Conocimiento propio.
Contribución al perfil profesional. Esta asignatura pretende contribuir a la formación del estudiante como ingeniero de conocimiento. Desde una aproximación actualizada, donde la construcción de bases de conocimiento ya no se realiza exclusivamente mediante entrevistas a expertos del dominio sino mediante la recopilación masiva de datos del dominio y la conformación semiautomatizada de ontologías a partir de esos datos.

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

Más allá de los requisitos fijados para cursar el máster, no hay requisitos adicionales para el estudio de esta asignatura.
Sobre la línea básica de la asignatura hay dos posibilidades de ampliación o de intensificación. La primera consiste en ahondar más en el estudio funcional y conceptual del formalismo, que lleva a una intensificación en sistemas lógicos, inferencia y gestión de ontologías. La segunda sigue una línea más operativa, donde se pueden aprovechar competencias previas en programación para ensamblar el uso de herramientas ya disponibles en procesos de construcción, refinamiento y explotación de estas bases de conocimiento.

EQUIPO DOCENTE

Nombre y Apellidos	JOSE LUIS FERNANDEZ VINDEL (Coordinador/a de asignatura)
Correo Electrónico	jlvindel@dia.uned.es
Teléfono	91398-7181
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	INTELIGENCIA ARTIFICIAL

Nombre y Apellidos	RAFAEL MARTINEZ TOMAS
Correo Electrónico	rmtomas@dia.uned.es
Teléfono	91398-7242
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	INTELIGENCIA ARTIFICIAL

Nombre y Apellidos	MIGUEL ANGEL PORTAZ COLLADO
Correo Electrónico	mportaz@dia.uned.es
Teléfono	
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	INTELIGENCIA ARTIFICIAL

HORARIO DE ATENCIÓN AL ESTUDIANTE

Para la atención telefónica o presencial se pueden utilizar las siguientes direcciones:
- Horario de Guardias: Martes (lectivos), de 15.00-19.00 hrs.

José Luis Fernández Vindel:
•Correo-e: jlvindel@dia.uned.es
•Tfno: +34 91 3987181
•Dpto. de Inteligencia Artificial (Despacho: 3.11)
Rafael Martínez Tomás:
•Correo-e: rmtomas@dia.uned.es
•Tfno: +34 91 3987242
•Dpto. de Inteligencia Artificial (Despacho: 3.15)
E.T.S.I. Informática (UNED)
C/ Juan del Rosal, n. 16
28040 Madrid (España)

- Horario de Atención al Estudiante: Martes y Miercoles de 10 a 14 h

Las consultas por correo electrónico se pueden remitir a la dirección jlvindel@dia.uned.es, aunque recomendamos que se utilice preferentemente el foro del grupo virtual de la asignatura para la remisión de dudas. De esta forma se comparten cuestiones y respuestas entre todos los participantes del curso.

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

COMPETENCIAS BÁSICAS

CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades. sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.

CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

COMPETENCIAS GENERALES

CG1 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CG2 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.
CG3 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
CG4 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

COMPETENCIAS ESPECÍFICAS

CE1 - Conocer los fundamentos de la Inteligencia Artificial y las fronteras actuales en investigación.
CE2 - Conocer un conjunto de métodos y técnicas tanto simbólicas como conexionistas y probabilistas, para resolver problemas propios de la Inteligencia Artificial.
CE3 - Conocer los procedimientos específicos de aplicación de estos métodos a un conjunto relevante de dominio (educación, medicina, ingeniería, sistemas de seguridd y vigilancia, etc.), que representan las áreas más activas de investigación en IA.

RESULTADOS DE APRENDIZAJE

RA1: Ser capaz de modelar un dominio mediante la utilización de las facilidades expresivas de los lenguajes de la Web Semántica
RA2: Hacer uso de los razonadores para la Web Semántica tanto en la extracción inferida de información implícita como en la ayuda a la depuración de ontologías y el desarrollo de aplicaciones.
RA3: Ser capaz de enlazar, consultar y de integrar los datos de fuentes diversas en la nube de datos enlazados.
RA4: Conocer los casos de éxito, los potenciales escenarios de interés y la tecnología implicada.

CONTENIDOS

Tema 1. Motivación e introducción

Motivación. El papel de los agentes inteligentes como asistentes personales o mediadores en nuestras tareas. La Web Semántica como espacio de interacción cognitiva de estos sistemas. La nube de Datos Públicos Enlazados como infraestructura de recopilación, curado, enriquecimiento y enlazado de datos, dentro de los objetivos de la Web Semántica.
Introducción. Perspectiva funcional de los agentes y procesos que confluyen en la Web de Datos Enlazados.

Tema 2. RDF. Modelo de datos y diseño de vocabularios

El modelo de datos RDF. Recursos identificados unívocamente con IRIs. Expresión de enunciados como tripletas <sujeto> <predicado> <objeto>. Nodos en blanco. Grafo de un conjunto de enunciados. Identificación de grafos distintos mediante IRIs.
RDF Schema. Un lenguaje para para la declaración de taxonomías: asignación de entidades a clases, jerarquización de clases via subclases, propiedades que relacionan elementos o a éstos con datos tipados, dominio y rango de una propiedad, jerarquización de propiedades vía subpropiedades. Facilidades adicionales del lenguaje para el etiquetado y la anotación.
Vocabularios e inferencia. Diseño de conceptualizaciones sencillas. Estudio de su semántica (interpretaciones, niveles de implicación) y de sus servicios de inferencia.
Serializaciones del modelo de datos RDF. Especialmente Turtle, dentro de su familia de serializaciones (N-Triples, Turtle, Trig y N-Quads), por su facilidad de lectura. Y JSON-LD por su impacto en la gestión de datos enlazados a través de la Web.

Tema 3. Publicación y consulta en la Web de Datos

Introducción a Linked Data. Una arquitectura de información granulada, conceptualizada y compartida. Navegación por indexadores de datasets públicos: la nube de datos abiertos enlazados.
Vocabularios más reutilizados entre datasets. Navegación por indexadores de vocabularios. Estudio de los vocabularios más comunes: FOAF, DC, SKOS, Schema, ORG, Geo ...
Opciones de consulta de un dataset. El lenguaje SPARQL de consulta. Ejecución de consultas a través de interfaces Web y a través de APIs diversas. Consultas federadas.
Estudio, consulta y explotación de ciertos datasets como caso de estudio. Inicialmente, análisis y consulta a Wikidata, con seguimiento a otros datasets temáticamente especializados. Enlazado de datos.

Tema 4. Flujos de información en la Web de Datos

Flujos de información hacia o desde datasets de datos enlazados. Poblamiento de un dataset a partir de fuentes estructuradas. Poblamiento de un dataset a partir de fuentes no estructuradas. Recuperación selectiva de datos desde uno o varios dataset. Reutilización y refinamiento de vocabularios, como apoyo a la gestión estructurada de datos.

Tema 5. Ontologías. Diseño e inferencia

OWL: un lenguaje de declaración de ontologías. Comparativa de expresividad respecto a las conceptualizaciones básicas construibles con RDFS. Constructores OWL (básicos y avanzados) aplicables a individuos, clases y propiedades. Semántica de estos constructores (referida a la semántica de los constructores de lógicas descriptivas e indirectamente a la semántica de la lógica de primer orden).
Inferencia sobre ontologías OWL. OWL DL y OWL Full. Perfiles OWL (EL, QL, RL) y su complejidad computacional. Razonadores y servicios de inferencia. Reglas (SWRL, RIF).

Tema 6. Aplicaciones de interés. Grafos de conocimiento

Estudio de casos de éxito. Discusión de potenciales aplicaciones. Escenarios de integración.

METODOLOGÍA

En el grupo virtual de la asignatura se presenta la secuencia de actividades propuestas, con los contenidos, instrucciones y sugerencia de herramientas necesarios para su realización. Sobre esta propuesta básica se pueden acordar algunas modificaciones personalizadas.
La lectura de los materiales y la búsqueda de información complementaria ocupan aproximadamente el 50% de las horas de dedicación. Para el desarrollo de las actividades prácticas (incluida la evaluación) se prevé el restante 50%.
Aunque se sugiere una planificación temporal para las entregas de prácticas, se facilitará la flexibilidad en los plazos de remisión de actividades, dentro del periodo lectivo del curso.

SISTEMA DE EVALUACIÓN

TIPO DE PRIMERA PRUEBA PRESENCIAL

Tipo de examen: No hay prueba presencial

TIPO DE SEGUNDA PRUEBA PRESENCIAL

Tipo de examen2: No hay prueba presencial

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad: No

Descripción

Aunque las pruebas de evaluación continua se publican atendiendo a un calendario prefijado, los plazos para su entrega se flexibilizan mucho atendiendo a las restricciones habituales de los estudiantes a tiempo parcial.

Criterios de evaluación

Ponderación de la prueba presencial y/o los trabajos en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

PRUEBAS DE EVALUACIÓN CONTINUA (PEC)

¿Hay PEC? Si,PEC no presencial

Descripción

La evaluación se llevará a cabo por trabajos (actividades de evaluación) cuyo enunciado se publica en el curso virtual.

Estas actividades permitirán confirmar la adquisición de los resultados de aprendizaje.

Criterios de evaluación

Se evaluará la precisión y claridad de los resultados así como la calidad metodológica y grado de cumplimiento de lo solicitado en el enunciado. También se valorará las iniciativas de ampliación del trabajo.

Ponderación de la PEC en la nota final La nota final se obtendrá de la nota media de las actividades evaluables ponderada por la estimación de duración y dificultad de cada una de ellas.

Fecha aproximada de entrega

Comentarios y observaciones

OTRAS ACTIVIDADES EVALUABLES

¿Hay otra/s actividad/es evaluable/s? No

Descripción

Criterios de evaluación

Ponderación en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

¿CÓMO SE OBTIENE LA NOTA FINAL?

Se valorarán todas las actividades de evaluación y se calculará como nota final la media ponderada con los pesos asignados a estas actividades en función de su complejidad.

BIBLIOGRAFÍA BÁSICA

En el curso virtual se incluirán enlaces a toda la documentación básica de la asignatura, en su mayoría formada por materiales propios con enlaces a otros documentos públicos.

BIBLIOGRAFÍA COMPLEMENTARIA

ISBN(13): 9780262045094
Título: KNOWLEDGE GRAPHS: FUNDAMENTALS, TECHNIQUES, AND APPLICATIONS
Autor/es: Mayank Kejriwal; Pedro Szekely; Craig A. Knoblock
Editorial: THE MIT PRESS

ISBN(13): 9781450376143
Título: SEMANTIC WEB FOR THE WORKING ONTOLOGIST: EFFECTIVE MODELING FOR LINKED DATA, RDFS, AND OWL Third edition edición
Autor/es: James Hendler; Dean Allemang; Fabien Gandon
Editorial: ¿ACM Books

ISBN(13): 9781484265512
Título: ONTOLOGIES WITH PYTHON: PROGRAMMING OWL 2.0 ONTOLOGIES WITH PYTHON AND OWLREADY2
Autor/es: Lamy Jean-Baptiste
Editorial: APRESS

ISBN(13): 9781636392356
Título: KNOWLEDGE GRAPHS
Autor/es: Aidan Hogan; Antoine Zimmermann; Michael Cochez; Claudia D'Amato; Gerard De Melo; Claudio Gutierrez; Sabrina Kirrane; Jose Emilio Labra Gayo; Roberto Navigli; Sebastian Neumaier; Axel-Cyrille Ngonga Ngomo; Axel Polleres; Sabbir M. Rashid; Anisa Rula; Juan Sequeda; Lukas Schmelzeisen; Steffen Staab; Eva Blomqvist
Editorial: Morgan & Claypool Publishers

ISBN(13): 9783030515829
Título: THE WEB OF DATA 1st edition edición
Autor/es: Aidan Hogan
Editorial: SPRINGER-VERLAG

RECURSOS DE APOYO Y WEBGRAFÍA

Para la exposición de los contenidos más conceptuales se dispone de material propio o de enlaces a textos consultables electrónicamente a través de la Biblioteca UNED.
Este campo de trabajo está soportado por un marco común de estándares y de lenguajes de declaración de conceptualizaciones, con su fundamentación semántica e inferencial. Entre las referencias técnicas básicas de esta infraestructura se encuentran:
[RDF 1.1 Primer] https://www.w3.org/TR/rdf11-primer/
[Linked Open Data Cloud] https://lod-cloud.net
[Linked Open Vocabularies] https://lov.linkeddata.es/dataset/lov/
[SPARQL Query Language] https://www.w3.org/TR/2013/REC-sparql11-query-20130321/
[A Description Logic Primer] https://arxiv.org/abs/1201.4089

[SKOS] https://www.w3.org/TR/skos-primer/
[OWL 2 Primer] https://www.w3.org/TR/owl2-primer/

IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no se hayan sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.