

19-20

MÁSTER UNIVERSITARIO EN LAS
TECNOLOGÍAS DE LA INFORMACIÓN Y
LA COMUNICACIÓN EN LA ENSEÑANZA
Y EL TRATAMIENTO DE LENGUAS

GUÍA DE ESTUDIO PÚBLICA



ONTOLOGÍAS, DICCIONARIOS Y LEXICONES RESTRINGIDOS COMPUTACIONALES

CÓDIGO 24409193

UNED

19-20

ONTOLOGÍAS, DICCIONARIOS Y
LEXICONES RESTRINGIDOS
COMPUTACIONALES
CÓDIGO 24409193

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
ADENDA AL SISTEMA DE EVALUACIÓN CON MOTIVO DE LA
PANDEMIA COVID 19

Nombre de la asignatura	ONTOLOGÍAS, DICCIONARIOS Y LEXICONES RESTRINGIDOS COMPUTACIONALES
Código	24409193
Curso académico	2019/2020
Título en que se imparte	MÁSTER UNIVERSITARIO EN LAS TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN EN LA ENSEÑANZA Y EL TRATAMIENTO DE LENGUAS
Tipo	CONTENIDOS
Nº ETCS	5
Horas	125.0
Periodo	SEMESTRE 2
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

La asignatura **Ontologías, diccionarios y lexicones restringidos computacionales** ofrece una visión introductoria de la codificación computacional del significado, donde se estudian las características de estos tres formatos que tienen que ver con el tratamiento del léxico. Asimismo se analizan herramientas como los corpora y se inicia al alumnado en el tratamiento informático de los datos textuales.

La lexicología ha experimentado una rápida evolución, desde la simple concepción como nomenclatura hasta el tratamiento actual a través del análisis informático del corpus. Por ello, el uso de las herramientas informáticas abre nuevas perspectivas al análisis lingüístico, ya que es posible disponer de una masa considerable de textos en formato electrónico, lo que permite la extracción de datos de manera fiable.

Asimismo, la codificación del conocimiento a través del desarrollo de ontologías abre amplias perspectivas para el manejo de facilidades computacionales en entornos como internet, que cada vez en mayor medida son el principal medio en el que se desenvuelve el conocimiento. Los estudiantes deberán familiarizarse con el uso del software existente de extracción del léxico, la noción de frecuencia, las co-ocurrencias, el análisis estadístico de las extracciones, etc., con el objetivo de iniciarse en un tema complejo como es la creación y manipulación de un corpus. Asimismo, deberán familiarizarse con los diferentes tipos de ontologías de orientación y/o soporte lingüístico, con su configuración, su desarrollo y sus aplicaciones.

Esta asignatura, que se imparte en el segundo semestre del **Máster de las TIC en la Enseñanza y el Tratamiento de las Lenguas**, consta de 5 créditos ECTS y está encuadrada en el módulo temático de contenidos formativos propios con asignaturas fundamentales para la consecución de los objetivos del máster.

Hay otras asignaturas de estas mismas enseñanzas de posgrado que también emplean el análisis de corpus para sus diferentes objetivos. Asimismo, en otras asignaturas se usan las ontologías como bloques fundamentales del desarrollo de la llamada red semántica. No obstante, en este curso se trata de ofrecer a los alumnos simplemente una perspectiva general de los distintos tipos de herramientas de análisis conceptual y textual y los fundamentos del desarrollo de recursos léxicos computacionales y ontológicos, independientemente del uso que se les vaya a dar.

Mientras que el análisis de corpus es un elemento común a muchas de las asignaturas que

componen este máster, el tratamiento y uso de los corpora (como se verá a lo largo del curso) difiere en cada una de ellas. Por otra parte, el estudio de la codificación conceptual, aunque se pueda tratar en otras asignaturas, en ésta adquiere una relevancia determinante.

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

En esta asignatura se aplican los requisitos generales de admisión al máster. Además, es necesario estar en posesión de un nivel B1/B2 de inglés y se recomienda el conocimiento de una segunda lengua extranjera (francés / alemán) para poder leer bibliografía especializada y artículos científicos y técnicos.

El alumno deberá poseer conocimientos de informática a nivel de usuario y disponer de los medios técnicos y conocimientos para poder acceder al curso virtual, dado que esta asignatura sigue la metodología “a distancia” propia de la UNED.

EQUIPO DOCENTE

HORARIO DE ATENCIÓN AL ESTUDIANTE

Los profesores de la Sede Central (equipo docente) están a disposición del alumno a través del curso virtual, y también por teléfono o personalmente (para esto último deberán concertar una cita) en los siguientes horarios, direcciones y teléfonos:

Dra. Margarita Goded Rambaud
Tel. 91398 8181
Despacho 622, planta 6
e-mail: margarita.goded@flog.uned.es
Horario de atención al alumno:

- Martes: de 11:00 a 14:00 horas.
- Miércoles: de 10:00a 14:00 horas (Atención telefónica)
- Jueves: de 10:00a 14:00 horas.

Dr. Antonio Pareja Lora
Tif.: 91394 7578 [DESACONSEJADO: mejor correo electrónico]
e -mail: aplora@ucm.es
Horario de atención al alumno:
•Viernes de 17:30 a 19:30.

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:

CG04 - Adquirir competencias cognitivas superiores en análisis y síntesis

CG05 - Ser capaz de aplicar los conocimientos a la práctica

CG06 - Resolver problemas en entornos nuevos o poco conocidos

CG12 - Adquirir competencias en el uso de las TIC

CG14 - Desarrollar competencias en la gestión y organización de la información y la documentación

CG15 - Desarrollar competencias en la recolección y análisis de datos, el manejo de bases de datos y su presentación

CG20 - Adquirir compromiso ético (por ejemplo, en la realización de trabajos sin plagios, etc.)

CG21 - Adquirir ética profesional (abarca también la ética como investigador)

Competencias Específicas:

CE07 - Evaluar el trabajo cooperativo entre lingüistas e informáticos dentro de la lingüística computacional.

CE11 - Diseñar e implementar un programa y un informe de investigación para resolver una tarea de extracción de información y generación de resúmenes.

CE12 - Diseñar e implementar un trabajo de investigación relativo a la red semántica que esté basado en la utilización de las TIC.

CE13 - Desarrollar las destrezas analíticas necesarias para la construcción de un corpus en formato digital.

CE14 - Adquirir las competencias necesarias para manejar de forma eficaz las herramientas computacionales para el análisis léxico.

CE15 - Reflexionar sobre los modelos segmentales y prosódicos adecuados para los

módulos lingüísticos en los proyectos de tecnologías del habla.

CE16 - Identificar sus propias carencias formativas y sus necesidades de especialización individual en metodologías fonéticas como la transcripción y notación de corpus.

CE17 - Familiarizarse con las herramientas disponibles para el traductor profesional de forma aislada e integrada en la mesa del traductor.

CE18 - Reconocer la importancia del uso de las TIC en el estudio de la diversidad lingüística y la multiculturalidad.

CE19 - Apreciar, manejar y combinar las diferentes técnicas de investigación según el tipo de corpus con el que se trabaje.

CE20 - Diseñar e implementar un trabajo de investigación dentro de la lingüística diacrónica que esté basado en la utilización de las TIC.

RESULTADOS DE APRENDIZAJE

Al finalizar el curso el estudiante, tras un aprovechamiento adecuado del curso, habrá obtenido los siguientes resultados:

1. Resultados relacionados con el aprendizaje:

1. Conocerá las principales aplicaciones de la lingüística de corpus a la semántica computacional, como el desarrollo y uso de diccionarios electrónicos, lexicones computacionales y ontologías.

2. Habrá desarrollado las destrezas analíticas necesarias para la gestión, extracción e interpretación de datos semánticos de un corpus en formato digital y/o para la creación de ontologías.

3. Habrá adquirido los conocimientos lingüísticos necesarios para manejar de forma eficaz las herramientas computacionales para el análisis léxico y/o semántico.

4. Podrá desarrollar proyectos de investigación lingüística e interdisciplinar relacionados con el análisis léxico.

2. Resultados relacionados con las actitudes:

1. Habrá desarrollado una actitud positiva hacia el valor del análisis de los datos empíricos en el avance del conocimiento.

CONTENIDOS

Tema 1. Introducción al tratamiento computacional del léxico.

Tema 2. El almacenamiento computacional del léxico

Tema 3. Revisión sobre los córpora

Tema 4. La codificación del significado en diccionarios y tesauros

Tema 5. Bases de datos

Tema 6. Dicionarios

Tema 7. Ontologías

METODOLOGÍA

La asignatura tiene una carga de 5 créditos ECTS, que equivalen a 125 horas de trabajo (1 crédito ECTS = 25 horas de trabajo), de las cuales, el 40 % se dedicará al trabajo de interacción con el equipo docente y el 60% restante al trabajo autónomo.

Dentro de estas dos modalidades, las actividades formativas se distribuyen de la siguiente manera:

INTERACCIÓN CON EL EQUIPO DOCENTE (40% = 50 horas = 2 créditos)

- Lectura de las unidades didácticas y/o materiales impresos (recomendados y complementarios).
- Lectura de las orientaciones de la asignatura para la realización de las actividades prácticas.
- Trabajo con las herramientas computacionales (software).
- Resolución de dudas en línea, a través del curso virtual o a través del teléfono.
- Revisión de actividades y trabajos propuestos.

TRABAJO AUTÓNOMO (60% = 75 horas = 3 créditos)

- Estudio de los temas de la asignatura.
- Participación en grupos de estudio.
- Interacción con los compañeros en el foro.
- Búsqueda de información y datos en bibliotecas, Internet u otras fuentes documentales.

Ambas modalidades irán destinadas a la adquisición de los conocimientos necesarios para realizar **ACTIVIDADES PRÁCTICAS y DESARROLLO DE DESTREZAS TRANSVERSALES**, tales como las siguientes:

- Resolución de problemas y ejercicios.
- Análisis y comprensión de textos.

- Elaboración de informes, resúmenes, redacciones o trabajos diversos.
- Trabajo en grupo.
- Discusión de temas y solución de problemas en grupo dentro del curso virtual en el foro de alumnos.

SISTEMA DE EVALUACIÓN

TIPO DE PRUEBA PRESENCIAL

Tipo de examen No hay prueba presencial

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad No

Descripción

No hay prueba presencial final. La evaluación continua se realiza mediante trabajos evaluables, cuya suma ponderada constituye la nota final de la asignatura.

Criterios de evaluación

Ponderación de la prueba presencial y/o los trabajos en la nota final No existe prueba presencial final. La nota final se obtiene de la suma ponderada de todos los trabajos de evaluación continua o PECs

Fecha aproximada de entrega 20/06/2019

Comentarios y observaciones

PRUEBAS DE EVALUACIÓN CONTINUA (PEC)

¿Hay PEC? Si,PEC no presencial

Descripción

Los trabajos evaluables son uno por cada uno de los temas del curso.

Criterios de evaluación

El principal criterio de evaluación es la capacidad de trasladar la información sobre distintas teorías y aplicaciones a los ejercicios propuestos en cada tema.

Ponderación de la PEC en la nota final

Fecha aproximada de entrega La fecha de comienzo de cada tema coincide con la fecha inicial de entrega del ejercicio. La fecha final improrrogable es el 20 de junio para todos los ejercicios

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?

La nota final es la suma ponderada de todos los trabajos de evaluación continua.

BIBLIOGRAFÍA BÁSICA

ISBN(13):9788448607722

Título:THE CODIFICATION OF MEANING IN ENGLISH (1ª)

Autor/es:Margarita Goded Rambaud ;

Editorial:MCGRAWHILL

Aguado de Cea, G., Álvarez de Mon Rego, I., Pareja-Lora, A. 2004. "OntoTag's Linguistic Ontologies: Enhancing Higher Level and Semantic Web Annotations". En *Proceedings of the IV International Conference on Language Resources and Evaluation (LREC 2004)*, Vol. VI, pp. 1905 - 1908. European Language Resources Association (ELRA). Lisbon, Portugal.

Aguado de Cea, G., Álvarez de Mon Rego, I., Pareja-Lora, A. 2005. "Conocimiento lingüístico para la Web Semántica: los nombres de persona en el dominio del cine". *Perspectivas Interdisciplinarias de la Lingüística Aplicada, Tomo II*, pp. 369 - 380. Asociación Española de Lingüística Aplicada (AESLA), 05/2005.

Aguado de Cea, G., Álvarez de Mon Rego, I., Pareja-Lora, A. 2009. "Una visión interdisciplinar de la anotación semántica". *Terminología y Sociedad del Conocimiento*, pp. 219 - 254. Peter Lang Verlag, A.G.

De Miguel, E. 2009. *Panorama de la lexicología*. Barcelona: Ariel.

Fellbaum, Christiane (ed.). 1998. *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

García, A., Pareja-Lora, A., Pradana, D. (2008). "Reutilización de tesauros: el documentalista frente al reto de la web semántica". *El Profesional de la Información*, 17(1):8-21

Goded, M., A. Ibañez, and V. Hoste. 2014. *On the Architecture of Words: applications of meaning studies*. Madrid: UNED.

Huang, Chu-Ren, N. Calzolari, A.Gangemi, A.Oltramari, L. Prevot. (editors). 2010. *Ontology and the Lexicon*. Cambridge University Press.

Lipka, L. 1990. *An Outline of English Lexicology*, Tubinga, Niemeyer.

Mitkov, R. 2003. *Oxford Handbooks in Linguistics*. Oxford University Press

Ooi, Vincent B. Y.1998. *Computer Corpus lexicography*. Edinburgh University Press. Edinburgh.

SE RECOMIENDA ENCARECIDAMENTE la visualización de los vídeos del VII seminario de investigación TIC-ETL, titulado **El concepto de ontologías y sus aplicaciones para el**

tratamiento y aprendizaje de las lenguas, disponibles en CANAL UNED:

<https://canal.uned.es/serial/index/id/1235>.

BIBLIOGRAFÍA COMPLEMENTARIA

Aguado de Cea, G., Álvarez de Mon Rego, I., Pareja-Lora, A. 2003. "Primeras aproximaciones a la anotación lingüístico-ontológica de documentos de la Web Semántica: OntoTag". *Inteligencia Artificial*, Vol. 2003, n. 18, pp. 37 - 50.

NeOn Toolkit (2014) http://neon-toolkit.org/wiki/Main_Page.

Niremburg, S & V. Raskin. 2005. *Ontological Semantics*. MIT Press.

Pareja-Lora, A. 2012. *Providing Linked Linguistic and Semantic Web Annotations - The OntoTag Hybrid Annotation Model*. Saarbrücken: LAP Lambert Academic Publishing.

[[https://www.lap-publishing.com/catalog/details/store/gb/book/978-3-659-25526-7/providing-linked-linguistic-and-semantic-web-annotations?search=linked annotations](https://www.lap-publishing.com/catalog/details/store/gb/book/978-3-659-25526-7/providing-linked-linguistic-and-semantic-web-annotations?search=linked%20annotations)]

Pareja-Lora, A. 2013. "An Ontology-Driven Methodology to Reuse, Link and Merge Terminological and Language Resources". En *Proceedings of the 10th International Conference on Terminology and Artificial Intelligence – Terminology for a networked society: Recent advances in multilingual knowledge-based resources*, pp. 189-196. París (Villetaneuse), Francia. 28-30 de octubre de 2013 [<https://lipn.univ-paris13.fr/tia2013/Proceedings/actesTIA2013.pdf#page=187>].

Pareja-Lora, A. 2014. "The Pragmatic Level of OntoLingAnnot's Ontologies and Their Use in Pragmatic Annotation for Language Teaching". En E. Bárcena et al. (editores), *Languages for Specific Purposes in the Digital Era. Series: Educational Linguistics, Vol. 19*, pp.323-344.

Suiza: Springer International Publishing Switzerland. DOI: 10.1007/978-3-319-02222-2_15 [http://link.springer.com/chapter/10.1007/978-3-319-02222-2_15].

Pareja-Lora, A., Aguado de Cea, G. 2010. "Modelling Discourse-related Terminology in OntoLingAnnot's Ontologies". En *Proceedings of TKE 2010: Presenting terminology and knowledge engineering resources online: models and challenges*, pp. 547-574. Ass. for Terminology and Knowledge Transfer (Gesellschaft für Terminologie und Wissenstransfer, GTW). Dublín, Irlanda. 11/08/2010 [<http://oa.upm.es/6249/>].

Pareja-Lora, A., Blume, M., Lust, B. 2013. "Transforming the Data Transcription and Analysis Tool Metadata and Labels into a Linguistic Linked Open Data Cloud Resource". En *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, pp. 34–43. ACL Anthology – Association for Computational Linguistics (ACL) [<http://aclweb.org/anthology/W/W13/W13-5506.pdf>].

Pareja-Lora, A., María Liakata, Stefanie Dipper (editores). 2013. En *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse (LAW VII & ID)*.

Association for Computational Linguistics, pp. 1-256. Sofía, Bulgaria. 8-9 de agosto de 2013.

ISBN 978-1-937284-58-9 [<http://www.aclweb.org/anthology/W13-23>].

WebProtégé (2014) <http://webprotege.stanford.edu/>.

Schalley, A. C., D. Zaefferer eds, 2007. *Ontologistics*. Berlin. New York: Mouton De Gruyter

OTROS RECURSOS:

Asimismo, como complemento de su aprendizaje, el alumno podrá consultar los enlaces de internet relacionados con la asignatura, que serán facilitados por el equipo docente.

RECURSOS DE APOYO Y WEBGRAFÍA

Como recursos de apoyo, la asignatura contará con un curso virtual, en el que el alumno se comunicará con el profesor y completará los contenidos de la asignatura y resolverá sus dudas.

La plataforma virtual permite también la comunicación entre los propios alumnos que estudian a distancia, lo que fomenta el aprendizaje colaborativo y la consecución de objetivos comunes y compartidos por diversos alumnos.

Es importante que el alumno utilice el curso virtual, en el que encontrará los materiales para trabajar en cada unidad, enlaces web recomendados, además de otros recursos de apoyo como ejercicios interactivos multimedia, tutoriales y diverso material que sirve de ayuda para el estudio de la asignatura.

Por otra parte, es aconsejable la participación del alumno en los distintos foros de comunicación. En dichos foros se aclaran numerosas dudas sobre el contenido y el funcionamiento de la asignatura, además de servir también como instrumento de aprendizaje.

La UNED cuenta con una biblioteca extensa y rica en bibliografía, a la cual se puede acceder también de forma virtual. Asimismo, en la UNED existe la posibilidad de la comunicación a través de la videoconferencia o la conferencia en línea, herramienta que podrá usarse para mantener un contacto directo con el grupo de alumnos.

ADENDA AL SISTEMA DE EVALUACIÓN CON MOTIVO DE LA PANDEMIA COVID 19

<https://app.uned.es/evacaldos/asignatura/adendasig/24409193>

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.