



Universidad de Puerto Rico
Recinto Universitario de Mayagüez
Colegio de Artes y Ciencias
Departamento de Estudios Hispánicos
Programa de Bachillerato en Ciencias en Biolingüística



PRONTUARIO OFICIAL

**COMPUTADORAS Y LENGUAJE
LING 4080/CCOG4080**

Horas crédito: 3	Horas contacto: 2 horas de conferencia y 2 horas de laboratorio
Requisitos previos: LING 4010, INGL 3225 o permiso del director	Requisitos concurrentes: N/A
Descripción del curso (español): Introducción a la aplicación e implementación de programas para codificar lenguas naturales en computadoras con un lenguaje de programación. Estudio y práctica de cómo construir representaciones explícitas de lenguas naturales en una computadora mediante modelación lingüística usando estructuras de datos y de cómo esta puede obtener conocimiento lingüístico mediante acercamientos basados en datos o en teoría. Manejo básico de tipos de datos y variables, estructuras de control, textos y archivos, módulos, funciones y programación orientada a objetos. Discusión sobre cómo realizan las computadoras tareas relacionadas con lenguas naturales, así como las posibilidades y limitaciones de las tecnologías de lenguaje humano.	
Descripción del curso (inglés): Introduction to the application and implementation of programs to encode natural language in computers with a programming language. Study and practice of how to build explicit natural language representations in a computer by means of linguistic modeling using data structures, and how the computer can obtain knowledge about language by means of both data-driven and theory-driven approaches. Basic handling of data types and variables, control structures, texts and files, modules, functions, and object-oriented programming. Discussion about how computers support language-related tasks, as well as the possibilities and limitations of human language technologies.	
Objetivos: Al finalizar el curso, el estudiantado estará capacitado para: <ol style="list-style-type: none">1. interpretar y diseñar algoritmos sencillos para el procesamiento de lenguas naturales.2. aplicar conocimientos básicos de programación y computación en el diseño e implementación de códigos.3. escribir programas con una interfaz gráfica de usuario.4. implementar funciones iterativas y recursivas.5. usar expresiones regulares para el reconocimiento de patrones.6. construir implementaciones simples para aspectos específicos del procesamiento de lenguas naturales.7. reconocer y evaluar el alcance y las limitaciones computacionales de las tecnologías de lenguaje humano y el impacto de estas en la vida contemporánea.	

Bosquejo de contenido:

<i>Temas a cubrir</i>	<i>Horas contacto</i>
Introducción: instalación, ambientes, interacción, edición, compilación	2 horas
Tipos de datos y variables	2 horas
Estructuras de control	4 horas
Manipulación de archivos	2 horas
Subrutinas, módulos y funciones	4 horas
Expresiones regulares	2 horas
Manipulación de textos	2 horas
Datos de internet	2 horas
Objetos: Clases e instancias	5 horas
Interfaces gráficas de usuario	2 horas
Programación funcional	3 horas
Laboratorio	30 horas
Total de horas: (deben ser equivalentes a las horas contacto del curso)	60 horas

Estrategias instruccionales:

conferencia discusión cómputos laboratorio investigación problemas especiales

seminario con presentación formal seminario sin presentación formal taller

taller de arte práctica viaje tesis tutoría otros, especifique:

Recursos mínimos disponibles:

computadora y sistema de proyección digital

Estrategias de evaluación y su peso relativo:

	Por ciento
<input checked="" type="checkbox"/> pruebas escritas	40
<input checked="" type="checkbox"/> proyectos	30
<input checked="" type="checkbox"/> asignaciones o pruebas cortas	20
<input checked="" type="checkbox"/> laboratorio	10
TOTAL: 100%	100

Sistema de calificación:

cuantificable (de letra) no cuantificable

Curva estándar

100-90 A; 89-80 B; 79-70 C; 69-60 D; 59-50 F

Bibliografía:

TEXTO: Hammond (2020) *Python for Linguists*.

Bird, Steven; Klein, Ewan & Loper, Edward (2009). *Natural Language Processing with Python*. Sebastopol, CA: O'Reilly Media.

Clark, Alexander; Fox, Chris & Lappin, Shalom, eds. (2013). *The Handbook of Computational Linguistics and Natural Language Processing*. Oxford: Blackwell.

Dickinson, Markus; Brew, Chris & Meurers, Detmar (2013). *Language and Computers*. Oxford: Wiley-Blackwell.

- Eisenstein, Jacob (2019). *Introduction to Natural Language Processing*. Cambridge, MA: The MIT Press.
- Goodrich, Michael; Tamassia, Roberto & Goldwasser, Michael (2013) *Data Structures and Algorithms in Python*. Hoboken, NJ: Wiley.
- Grus, Joel (2015) *Data Science from Scratch*. Sebastopol, CA: O'Reilly Media.
- Gutttag, John V. (2016) *Introduction to Computation and Programming Using Python*. 2da. edición. Cambridge, MA: MIT Press.
- Hammond, Michael (2020). *Python for Linguists*. Cambridge: CUP.
- Jurafsky, Daniel & Martin, James. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. 2da. edición. New Jersey: Prentice Hall.
- Lambert, Kenneth A. (2019) *Fundamentals of Python: Data Structures*. Boston: Cengage.
- Lutz, Mark (2013) *Learning Python*. 5ta. edición. Sebastopol, CA: O'Reilly.
- Ojeda, Almerindo. (2013) *A Computational Introduction to Linguistics*. Stanford: CSLI Publications.
- Ruslan, Mitkov (2005). *The Oxford Handbook of Computational Linguistics*. Oxford: Oxford University Press.
- Sportiche, Dominique; Koopman, Hilda & Stabler, Edward (2014) *An Introduction to Syntactic Analysis and Theory*. Oxford: Wiley-Blackwell

RECURSOS EN INTERNET

- Anaconda Navigator** <https://www.anaconda.com/products/individual>
Interfaz gráfica de Python para computación científica.
- Association for Computational Linguistics** <https://www.aclweb.org/portal/>
Portal de la Asociación de Lingüística Computacional y enlace a sus revistas académicas
- Contemporary Linguistics**
<https://www.macmillanlearning.com/college/us/product/Contemporary-Linguistics/p/1319039774>
Portal electrónico del libro homónimo de O'Grady et. al (2018).
- FrameNet** <https://framenet.icsi.berkeley.edu/fndrupal/>
Base de datos léxica legible por máquinas y humanos.
- International Phonetic Association** <https://www.internationalphoneticassociation.org/>
Plataforma de recursos sobre el Alfabeto Fonético Internacional.
- An Introduction to Language and Linguistics** <http://www.cambridge.org/fasold>
Portal electrónico del libro homónimo de Fasold y Connor-Linton (2008).
- The LINGUIST List** <http://www.linguistlist.org/>
Colección de recursos lingüísticos y profesionales en la internet.
- Linguistic Data Consortium** <https://www ldc.upenn.edu/>

Repositorio y punto de distribución de recursos lingüísticos.

Natural Language Toolkit

<http://www.nltk.org/>

Plataforma para construir programas de procesamiento de lenguas naturales en Python

Penn TreeBank Project

<https://catalog.ldc.upenn.edu/LDC99T42>

Banco de árboles sintácticos a partir de un corpus natural anotado.

Peter Jurgec's PhonoApps

<http://www.phonology.us/>

Página electrónica del herramientas computacionales de análisis fonológico.

Python

<http://www.python.org/>

Página electrónica del herramientas computacionales de análisis fonológico.

Speech and Language Processing

<https://web.stanford.edu/~jurafsky/slp3/>

Portal del texto de Jurafsky & Martin con recursos para el instructor.

The Stanford Parser

<http://nlp.stanford.edu/software/lex-parser.shtml>

Analizador sintáctico estadístico desarrollado por el Stanford NLP Group.

WordNet

<http://wordnet.princeton.edu/>

Base de datos léxica del inglés.

Acomodo Razonable:

El estudiantado puede orientarse y solicitar los servicios del Programa de Acomodo Razonable, que está ubicado en el Departamento de Consejería y Servicios Psicológicos (Decanato de Estudiantes, oficina DE21, en los bajos de Asistencia Económica). Para aclarar cualquier duda u obtener más información, puede comunicarse al teléfono directo [787-265-3864](tel:787-265-3864) o a través del cuadro [787-832-4040](tel:787-832-4040), extensiones 2040, 3372 o 3864.

Integridad Académica:

La Universidad de Puerto Rico promueve los más altos estándares de integridad académica y científica. El Artículo 6.2 del Reglamento General de Estudiantes de la UPR (Certificación Núm. 13, 2009-2010, de la Junta de Síndicos) establece que “la deshonestidad académica incluye, pero no se limita a: acciones fraudulentas, la obtención de notas o grados académicos valiéndose de falsas o fraudulentas simulaciones, copiar total o parcialmente la labor académica de otra persona, plagiar total o parcialmente el trabajo de otra persona, copiar total o parcialmente las respuestas de otra persona a las preguntas de un examen, haciendo o consiguiendo que otro tome en su nombre cualquier prueba o examen oral o escrito, así como la ayuda o facilitación para que otra persona incurra en la referida conducta”. Cualquiera de estas acciones estará sujeta a sanciones disciplinarias en conformidad con el procedimiento disciplinario establecido en el Reglamento General de Estudiantes de la UPR vigente.

Hostigamiento Sexual: La certificación 130-2014-2015, indica:

El hostigamiento sexual en el empleo y en el ambiente de estudio es una práctica ilegal y discriminatoria, ajena a los mejores intereses de la Universidad de Puerto Rico. Toda persona que entienda ha sido objeto de actuaciones constitutiva de hostigamiento sexual en la Universidad de Puerto Rico podrá quejarse para que se investigue, de ser necesario, y se tome la correspondiente acción por parte de las autoridades universitarias. Si quien reclama fuera estudiante, deberá referir su queja a la Oficina de la Procuradora Estudiantil o al Decanato de Estudiantes.

La certificación 06-43 del Senado Académico indica:

“Las guías académicas para el ofrecimiento de cursos en línea”, define: Cursos presenciales son aquellos que tienen menos de un 25% de las horas contacto regular del curso a través de la Internet. Así, un curso de 3 créditos, será considerado “presencial” si, de las 45 horas de contacto regular, 11 o menos son a través de la Internet.

Incluye anejos:

Si

No