



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales

Planilla a completar para presentación de Cursos de Posgrado

1.- DEPARTAMENTO de COMPUTACIÓN

2.- NOMBRE DEL CURSO: **Introducción al Procesamiento del Lenguaje Natural**

3.- DOCENTES:

RESPONSABLE/S: **José M. Castaño**
COLABORADORES:
AUXILIARES: **Viviana Cotik**

4.- CARRERA de DOCTORADO

5.- AÑO: **2011**

CUATRIMESTRE/S: **PRIMERO**

6.- PUNTAJE PROPUESTO PARA CARRERA DE DOCTORADO: **4 puntos**

7.- DURACIÓN (anual, cuatrimestral, bimestral u otra): **Cuatrimstral**

8.- CARGA HORARIA SEMANAL:

Teóricas:
Problemas: **3 horas**
Laboratorio:
Seminarios:
Teórico - Práctico: **3 horas**
Salida a Campo:

9.- CARGA HORARIA TOTAL: **96 horas**

10.- FORMA DE EVALUACIÓN: **Trabajo Práctico Final**

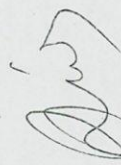
11.- PROGRAMA ANALÍTICO:

Introducción al PLN. Enfoques racionalista y empirista del lenguaje. Breve reseña histórica Conceptos lingüísticos básicos. El lenguaje como fenómeno probabilístico. Algunas aplicaciones básicas de PLN.

- **Principales conceptos de lingüística.** Distintos niveles de análisis; Léxico, categorías gramaticales y morfología. Sintaxis. Semántica y Pragmática. Los corpora. Construcción y anotación de un corpus. Tokenization (segmentación, morfología y oraciones). Maración de documentos y etiquetado.
- **El nivel de la Palabra.** Expresiones Regulares y Automatas de estados finitos. Análisis morfológico y Transductores.
- **N-gramas.** Modelos Markovianos visibles y ocultos. Estimación de parámetros.
- **Etiquetado (tagging) de categorías sintácticas.** Principales técnicas de etiquetamiento. Aplicaciones del etiquetamiento. Evaluación: Precision, Recall y medida-F.
- **El nivel de las frases.** Chunking: parsing con automatas finitos, o parsing parcial.
- **Gramáticas libres de Contexto, Compiladores, Gramáticas probabilísticas, Gramáticas Lexicalizadas.** Evaluación.
- **Nivel Semántico.** Representación del significado. Desambiguación léxica. Clases de ambigüedad léxica. Adquisición léxica. Papel y aplicaciones de la adquisición léxica. Adquisición de distintas propiedades de las palabras: subcategorización, preferencias de selección, rasgos semánticos

12.- BIBLIOGRAFÍA:

- **Speech and Language Processing (SLP).** Jurafsky, Martin, 2da edición, 2009. Prentice Hall. La primera edición también es una opción.
- **Foundations of Statistical Natural Language Processing (FSNLP).** Manning, Schütze. MIT Press, 1999
- **Finite State Morphology (FSM).** Bescley, Karttunen.
- **Introduction to Automata Theory, Languages, and Computation(ATLC).** Hopcroft, Ullman. Addison Wesley


Dr. Alejandro N. Ríos
Departamento de Computación
FCEN UBA

UNIVERSIDAD DE BUENOS AIRES FACULTAD DE CIENCIAS EXACTAS Y NATURALES	
ENTRO	SALIO
2011	

REGISTRADO



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales

Referencia Expte. N° 487.712/06

Buenos Aires, 14 NOV 2011

VISTO:

la nota presentada por el Dr. Alejandro Ríos, miembro titular de la Comisión de Doctorado de esta Facultad por el Departamento de Computación, mediante la cual eleva la información y el programa del curso de posgrado **INTRODUCCION AL PROCESAMIENTO DEL LENGUAJE NATURAL**, dictado durante el primer cuatrimestre de 2011 por el Dr. José M. Castaño como docente responsable y Viviana Cotik como docente auxiliar,

CONSIDERANDO:

lo actuado por la Comisión de Doctorado de esta Facultad el 25/10/2011,
lo actuado por la Comisión de Enseñanza, Programas, Planes de Estudio y Posgrado,
lo actuado por este Cuerpo en la sesión realizada en el día de la fecha,
en uso de las atribuciones que le confiere el Artículo 113° del Estatuto Universitario,

**EL CONSEJO DIRECTIVO DE LA FACULTAD DE
CIENCIAS EXACTAS Y NATURALES
RESUELVE:**

Artículo 1°: Dar validez al dictado del curso de posgrado **INTRODUCCION AL PROCESAMIENTO DEL LENGUAJE NATURAL** de 96 hs. de duración.

Artículo 2°: Aprobar el programa del curso de posgrado **INTRODUCCION AL PROCESAMIENTO DEL LENGUAJE NATURAL**, obrante a fs 26 del expediente de referencia.

Artículo 3°: Aprobar un puntaje de cuatro (4) puntos para la Carrera del Doctorado.

Artículo 4°: Aprobar un arancel de 20 módulos. Disponer que los montos recaudados serán utilizados conforme a lo dispuesto por Resolución CD N° 072/03.

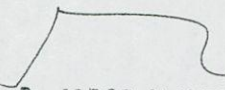
Artículo 5°: Comuníquese a la Dirección del Departamento de Computación, a la Biblioteca de la FCEyN y a la Subsecretaría de Postgrado (con fotocopia del Programa incluido) y a la Dirección de Alumnos y Graduados sin fotocopia del Programa. Cumplido Archívese.

Resolución CD N°

SP/ga 28/10/11

2850 ==


Dr. JAVIER LÓPEZ DE CASENAVE
SECRETARIO ACADEMICO


Dr. JORGE ALIAGA
DECANO