



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE

FACULTAD DE LETRAS

Perspectivas de desarrollo para el área de procesamiento de lenguaje natural en la PUC

César Antonio Aguilar
Facultad de Lenguas y Letras
15/12/2011

caguilara@uc.cl

Introducción



La Pontificia Universidad Católica de Chile (UC) actualmente tiene un fuerte interés por desarrollar líneas de investigación interdisciplinarias, las cuales tengan impacto directo en la formación de recursos humanos de alta calidad, así como también establecer vínculos con el sector empresarial.



En este sentido, y tomando en cuenta la evolución que ha tenido el procesamiento de lenguaje natural (o PLN) en distintos medios académicos, éste puede convertirse en una línea de investigación y desarrollo importante para la UC.

Casos de éxito en Chile: Yahoo Research



New User? Register | Sign In | Help

Trending: Powerball

Y Yahoo! Mail My

YAHOO! LABS **Web Search**

Site Search

- Labs Home
- About Yahoo! Labs
- Areas
- News
- Events
- Job Opportunities
- People
- Publications
- Locations
- Contact Us



Locations: Santiago, Chile



En el año 2006, bajo la coordinación de Ricardo Baeza-Yates, Yahoo Research se establece en Santiago, apoyando el desarrollo del Centro de Investigación de la WEB (CWR), el cual se orienta sobre todo en los siguientes rubros:

1. Minería de la WEB
2. Recuperación y extracción de información
3. Modelado matemático de la WEB
4. Sistemas distribuidos y paralelismos

El potencial de la UC en PLN



En estos momentos, se están empezando a establecer lazos de colaboración entre la Facultad de Letras y el Departamento de Ciencias de la Computación de la UC.



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

FACULTAD DE LETRAS



DEPARTAMENTO DE CIENCIA DE LA COMPUTACIÓN
ESCUELA DE INGENIERIA

Estos lazos se ubican de momento en dos rubros: docencia e investigación. En un futuro, esperamos abordar también la gestión de proyectos patrocinados por instituciones gubernamentales (CONCYT), así como de dependencias privadas.

Lo que ahora hay (1)



Desarrollo de hipermedios: Pablo Chiuminatto (FL) y Renato Verdugo (DCC) han impartido una serie de cursos para estudiantes de pregrado en Letras e Ingeniería sobre la creación y uso de hipermedios.

Un producto que se generará de esta colaboración es el establecimiento de un **Laboratorio de Hipermedios**, el cual tendrá como fin de alfabetizar y fomentar el empleo de recursos electrónicos para la publicación y análisis de páginas WEB, así como otros medios electrónicos de comunicación.



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE LETRAS
SALA DE LECTURA

<http://www.letraspuc.cl/>

Lo que ahora hay (2)



Curso de PLN para Magíster y doctorado: a partir de este semestre se impartirá de forma continua un seminario sobre PLN, con miras a establecerlo como una línea terminal de investigación dentro del programa de doctorado.

Posgrado
DOCTORADO EN LINGÜÍSTICA 

http://www.uc.cl/letras/html/4_d_linguistica2.html

Del mismo modo, existe el interés por involucrar a los alumnos de pregrado a esta línea de investigación. Si bien estos esfuerzos se realizan en la FL, esperamos interés por parte de los estudiantes de Ingeniería en Computación.

Análisis y captura de conceptos para la ontología FunGramKB: Carlos González coordina el grupo *Los Conspiradores*, conformado por estudiantes y profesores de la FL, quienes colaboran de manera activa con el proyecto **FunGramKB**, una ontología multilingüe, la cual ha logrado avances significativos en la solución de un formalismo fiable y económico para la representación de conceptos.



<http://www.fungramkb.com/>

Lo que ahora hay (4)



Magíster en Procesamiento y Gestión de Información (MPGI):

quizá el proyecto más ambicioso que la UC se plantea en cuestiones de PLN es la implementación de un magíster orientado hacia la enseñanza de modelos, métodos y técnicas de análisis de grandes fuentes de información textual.

Este magíster será coordinado por la FL, el DCC y el Sistema de Bibliotecas (SIBUC).

<http://mpgi.ing.puc.cl/>

**MAGÍSTER EN
PROCESAMIENTO
Y GESTIÓN DE LA
INFORMACIÓN**

MPGI
INGENIERÍA - LETRAS - SIBUC



¿Por qué resulta estratégico este programa de magíster?:

1. Ofrece una auténtica visión interdisciplinaria sobre cómo abordar el tratamiento de información bibliográfica.
2. Está orientado a estudiantes con distintos perfiles, los cuales tengan interés por alguna de las líneas de investigación considerada en el magíster.
3. Se concibe a dicho magíster como un semillero de proyectos de I + D en tratamiento de información.
4. Tomando en cuenta estos intereses, este magíster es una ventana para la colaboración con otros grupos de PLN relacionados con los temas que aquí se tratan.



El potencial que podemos tener como colaboradores en PLN se orienta en las siguientes líneas:

1. Recuperación y extracción de información
2. Minería de textos
3. Clasificación de documentos
4. Anotación morfosintáctica y semántica de corpus
5. Desarrollo de ontologías



Proyectos viables a corto plazo:

1. Colaboración en la planeación y enseñanza de cursos, talleres y seminarios sobre PLN
2. Co-dirección de tesis
3. Explotación de bases de datos textuales
4. Intercambio de estudiantes de grado y posgrado
5. Organización de foros de PLN en Chile, Latinoamérica y España



Gracias por su atención

Para mayores detalles:

<http://cesaraguilar.weebly.com/>