

PROGRAMA JUAN DE LA CIERVA
HISTORIAL CIENTÍFICO-TÉCNICO DE LOS ÚLTIMOS CINCO AÑOS DEL EQUIPO DE
INVESTIGACIÓN
(SCIENTIFIC/TECHNICAL RECORD DURING THE LAST FIVE YEARS OF THE RESEARCH TEAM)

(se deberá adjuntar un historial científico-técnico y el Currículum Vitae del investigador principal para cada uno de los investigadores candidatos incluidos en el apartado B (Relación de investigadores candidatos))

Indicar la labor de investigación que se desarrolla, publicaciones, proyectos financiados, contratos con empresas, patentes, tesis doctorales dirigidas, y cualquier otro aspecto que se considere de interés. De acuerdo con lo indicado en el apartado vigésimo sexto.1. de la Resolución de convocatoria, el equipo de investigación deberá estar formado por los mismos investigadores que recibieron la ayuda para el proyecto de I+D que se presenta.

CUMPLIMENTAR PREFERIBLEMENTE EN INGLÉS – FILL IN BETTER IN ENGLISH

Natural Language Processing Group
TALP Research Center
Software Department, UPC

The Natural Language Processing Group (<http://www.lsi.upc.edu/~nlp>), created in 1986, has among its members 6 associate professors, 4 researchers, 10 assistant professors and 11 PhD students. The professors teach at several Computer Science courses of the UPC and at the AI PhD program of the Software Department, and 9 theses supervised by members of the group have been defended in the last 5 years.

Ever since its creation the NLPG has devoted to technologies and applications of automatic natural language processing. The languages dealt with by the group are Spanish, Catalan, and English. The group's main areas of activity are the collection and management of multilingual lexical resources, information extraction from documents and question answering applications, statistical machine translation, machine learning for NLP, design of natural language interfaces, acquisition and exploitation of semantic information, and the development of basic techniques for language processing (morphosyntactic and semantic disambiguation, syntactic analysis, named entity recognition, semantic parsing, etc.). The group has focused on the constitution of a solid infrastructure of resources and linguistic processors that constitutes the basis of the different research and development lines that are being followed at present.

The group's members are involved in a large number of publications (they can be checked in the URL mentioned above) and have regularly participated in the organisation and programme committees of several national and international conferences in the area. Since 1998, the group's members have published at least 15 journal papers and more than 100 papers in conferences, and have participated in the preparation of 3 research books.

The NLPG has participated and is currently participating in several national and international projects with both state and private funding. The external funding obtained for the projects within the 2002-2007 period (see list below) amounts to over 3.5 million euro. Moreover, the group has the status of a Consolidated Research Group (2005 SGR00130) by DURSI, the Research Department of the Catalan Government (*Direcció General de Recerca de la Generalitat de Catalunya*).

Main projects covering the 2002-2007 period:

HERMES	TIC2000-0335-C03-02	75.913 €	2000-2003
PETRA	TIC2000-1735-C02	66,265 €	2001-2003
FAME	EU-IST-2001-28323	405.648 €	2001-2004
MEANING	EU-IST-2001-34460	350.805 €	2002-2005
MEANING (AC)	TIC2002-10675-E	54.000 €	2002-2005
LC-STAR	EU-IST-2001-32216	170.000 €	2002-2005
ALIADO	TIC2002-04447	98.260 €	2002-2005
3LB	FIT-150500-2002-244	10.800 €	2002-2004
CHIL	EU-IST-2004-506909	560.000 €	2004-2007
HOPS	EU-IST-2004-507967	380.937 €	2004-2007
HOPS (AC)	TIN2004-0169-E	20.500 €	2004-2007
OpenTRAD	FIT-340101-2004-3	101.880 €	2004-2005
CESS-ECE	HUM2004-21127-E	18.000 €	2005-2007
Arabic WordNet DOI	REFLEX-NBCHC050031	939.268 \$US	2005-2007
KNOW	TIN2006-15049-C03-03	100.000 €	2006-2008
OpenMT	TIN2006-15307-C03-02	72.000 €	2006-2009
Text-Mess(SAMiT)	TIN2006-15265-C06-05	110.000 €	2006-2009

PhD theses during the 2002-2007 period:

TITLE: An Information Extraction System Portable to New Domains

PhD STUDENT: Jordi Turmo

UNIVERSITY: Universitat Politècnica de Catalunya

CENTER: Departament de Llenguatges i Sistemes Informàtics

YEAR: 2002

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dr. Horacio Rodriguez Hontoria

TITLE : An Island-Driven Parsing System

PhD STUDENT: Alicia Ageno

UNIVERSITY: Universitat Politècnica de Catalunya

CENTER: Departament de Llenguatges i Sistemes Informàtics

YEAR: 2003

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dr. Horacio Rodriguez Hontoria

TITLE: Acquiring Information Extraction Patterns from Unannotated Corpora

PhD STUDENT: Neus Catala Roig

UNIVERSITY: Universitat Politècnica de Catalunya

CENTER: Departament de Llenguatges i Sistemes Informàtics

YEAR: 2003

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dra. Nuria Castell Arino

TITLE :“Adquisicio d’informacio lexica i morfosintactica a partir de corpus sense anotar: aplicacio al rus i al croat”

PhD STUDENT : Antoni Oliver Gonzalez

UNIVERSITY: Universitat de Barcelona

CENTER: Departament de Linguistica General

YEAR: 2004

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dr. Lluis Marquez Villodre
(co-dirigida con la Dra. Irene Castellon Masalles-UB)

TITLE: Representing Discourse for Automatic Text Summarization via Shallow NLP

PhD STUDENT: Laura Alonso i Alemany

UNIVERSITY: Universitat de Barcelona

CENTER: Departament de Llingüística General

YEAR: 2005

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dr. Lluís Padró Cirera (co-supervised with Irene Castellón)

TITLE: Un Mètode per a la Comprovació de Restriccions d'Integritat en Bases de Dades Deductives Bitemporals

PhD STUDENT: Carme Martín Escofet

UNIVERSITY: Universitat Politècnica de Catalunya

CENTER: Departament de Llenguatges i Sistemes Informàtics

YEAR: 2005

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dra. Nuria Castell Ariño

TITLE: Enlace de Jerarquías Usando el Etiquetado por Relajación

PhD STUDENT: Jordi Daudé

UNIVERSITY: Universitat Politècnica de Catalunya

CENTER: Departament de Llenguatges i Sistemes Informàtics

YEAR: 2005

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dr. Lluís Padró Cirera (co-supervised with German Rigau)

TITLE: Automatic Construction of Wide-Coverage Domain-Independent Lexico-Conceptual Ontologies

PhD STUDENT: Javier Farreres

UNIVERSITY: Universitat Politècnica de Catalunya

CENTER: Departament de Llenguatges i Sistemes Informàtics

YEAR: 2005

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dr. Horacio Rodriguez Hontoria

TITLE: Learning and Inference in Phrase Recognition: A Filtering-Ranking Architecture using Perceptron

PhD STUDENT: Xavier Carreras Pérez

UNIVERSITY: Universitat Politècnica de Catalunya

CENTER: Departament de Llenguatges i Sistemes Informàtics

YEAR: 2005

RESEARCHER FROM THE GROUP WHO HAS SUPERVISED IT: Dr. Lluís Márquez Villodre