

# Introduction to Natural Language Processing

## Overview

Natural Language processing (NLP) is a field of computer science and linguistics concerned with the interactions between computers and human (natural) languages. In theory, natural-language processing is a very attractive method of human-computer interaction. Natural-language understanding is sometimes referred to as an artificial intelligence-complete problem, because natural-language recognition seems to require extensive knowledge about the outside world and the ability to manipulate it. NLP has significant overlap with the field of computational linguistics, and is often considered a sub-field of artificial intelligence. The course will also cover a number of standard algorithms that are used throughout language processing. The course will also put light on the Information Retrieval and Machine Translation. Application areas within Language Information Processing include automatic (machine) translation between languages; dialogue systems, which allow a human to interact with a machine using natural language; and information extraction, where the goal is to transform unstructured text into structured (database) representations that can be searched and browsed in flexible ways.

Course participants will learn these topics through lectures and hands-on experiments. Also case studies and assignments will be shared to stimulate research motivation of participants.

<b>Modules</b>	<b>Course Duration:</b> : 2-7 May, 2016 <b>This course will cover following topics:</b>  1:Overview of Natural Language Processing 2:Formal Language Theory 3:Markov Model and N-gram Model 4:Part-of-Speech Tagging 5:Word Sense Disambiguation 6:Probabilistic Parsing 7:Machine Learning Approaches in NLP 8:Information Retrieval 9:Dialogue System 10:Machine Translation 11:Knowledge Acquisition for NLP
<b>You Should Attend If...</b>	<ul style="list-style-type: none"><li>▪ You are an executive/ engineer/ researcher from manufacturing, service and government organizations including R&amp;D laboratories.</li><li>▪ You are a student at all levels (BTech/MSc/MTech/PhD) or Faculty from reputed academic institutions and technical institutions.</li></ul>
<b>Fees</b>	The participation fees for taking the course is as follows: <b>Not-for-profit R&amp;D and Educational Institutions: Rs. 10000/-</b> <b>Industry/ Research Organizations: Rs. 30000/-</b> <b>Academic Institutions (Full time students): Rs. 1000/-</b> <b>Note: The course fee will be made half for SC/ST students.</b>  The above fee include all instructional materials, computer use for tutorials and assignments, laboratory equipment usage charges, 24 hr free internet facility. The participants will be provided with accommodation and food on payment basis.

## The Faculty



Prof. Sadao Kurohashi is currently a professor of the Graduate School of Informatics at Kyoto University, Japan. He has presented scientific papers in many national and international conferences and published in various journals and books in the field of Natural Language Processing.

His current interests include machine translation, information retrieval, and knowledge engineering. He is and has been a regular reviewer for several journals and conferences, and has served in program committees of many conferences. He has been doing pioneer work in the domain of Machine Translation and Information Extraction. He has received 2009 IBM Faculty Award and 2010 NTT DOCOMO Mobile Science Award.



**Prof. Pushpak Bhattacharyya** is a Professor of Computer Science and Engineering at IIT Bombay. He is currently the Professor and Director of IIT Patna. He was a Visiting research fellow in Massachusetts Institute of Technology in 1990. He is well known for his contributions to natural

language processing and has several distinctions in that field. He has also currently been appointed as Vijay and Sita Vashee Chair Professor. Prof. Bhattacharyya specializes in Natural Language Processing (NLP), Machine Learning (ML), Machine Translation (MT), Cross Lingual IR (CLIR) and Information Extraction (IE). A highly recognized researcher, Prof. Bhattacharyya is a member of National Knowledge Commission. For his outstanding contribution to computer science, he has received many accolades including IBM Innovation award (2007), Yahoo Faculty Award (2011), P. K. Patwardhan Award for Technology Development (2008), and VNMM Award of IIT Roorkee (2014).



**Dr. Asif Ekbal** is currently a faculty member of the department of Computer Science and Engineering, Indian Institute of Technology Patna, India. His current research interests include Natural Language Processing, Data Mining and Machine Learning Applications, Information Extraction, Bio-

text Mining and Machine Learning in Social Networks. His h-index is 20 and total citation count of his papers is 1100 (according to Google scholar). He is the recipient of the *Best Innovative Project Award* from the INAE, best paper award in ICACCI 2012 etc.



**Dr. Sriparna Saha** is currently a Faculty Member of the Department of Computer Science and Engineering, Indian Institute of Technology Patna, India. She is the author of a book published by Springer-Verlag. Her current research interests include pattern recognition, multiobjective

optimization and biomedical information extraction. Her h-index is 15 and total citation count of her papers is 1336 (according to Google scholar). She is the recipient of the Lt Rashi Roy Memorial Gold Medal from ISI Kolkata, Google India Women in Engineering Award, 2008, Junior Humboldt Research Fellowship etc.

## Course co-ordinators

**Professor Pushpak Bhattacharya**  
Director, Indian Institute of Technology Patna  
Vijay and Sita Vashee Chair Professor  
Department of Computer Science and Engineering  
Indian Institute of Technology Bombay  
Ph: 91-22-25764729 (o), 25721955(h)  
Email: pb@cse.iitb.ac.in  
<http://www.cse.iitb.ac.in/~pb>

**Dr. Asif Ekbal**  
Assistant Professor  
Department of Computer Science and Engineering  
Indian Institute of Technology Patna  
Bihta, Kanpa Road, District – Patna Bihar, India – 801118  
Ph: +91-8809559190(m)  
Email: asif@iitp.ac.in  
<http://www.iitp.ac.in/~asif/>

**Dr. Sriparna Saha**  
Assistant Professor  
Department of Computer Science and Engineering  
Indian Institute of Technology Patna  
Bihta, Kanpa Road, District – Patna Bihar, India – 801118  
Ph: +91-8809559190(m)  
Email: sriparna@iitp.ac.in  
<http://www.iitp.ac.in/~sriparna/>