

Diptesh Kanojia

MALE, 26 YEARS

Email: dipteshkanojia@gmail.com



Contact: +91 8879058668, 9969054037

EDUCATIONAL QUALIFICATION

B. Tech. (CSE)	2013	Gautam Buddha Technical University, Lucknow	72.9%
Higher Secondary Examination	2005	Greenfields Public School, Dilshad Garden, Delhi (CBSE)	61.4%
Secondary Examination	2003	St. Josephs Academy, Savita Vihar, Delhi (CBSE)	75%

▪ Ranked among the **top three** students of college, in all academic years during graduation.

PUBLICATIONS

1. Arindam Chatterjee, Salil Joshi, Diptesh Kanojia, Akhlesh Meena and Pushpak Bhattacharyya, **A Study of the Sense Annotation Process: Man v/s Machine**, Global WordNet Conference (GWC 2011), Matsue, Japan, January, 2012.
2. Diptesh Kanojia, Arindam Chatterjee, Salil Joshi, and Pushpak Bhattacharyya, **Discrimination Net for Hindi**, International Conference for Computational Linguistics (COLING 2012), Mumbai, India, December 2012.
3. Salil Joshi, Diptesh Kanojia, and Pushpak Bhattacharyya, **More than meets the eye: Study of Human Cognition in Sense Annotation**, North American Chapter of Association of Computational Linguistics (NAACL 2013), Atlanta, U.S.A., June 2013.
4. Diptesh Kanojia, Pushpak Bhattacharyya, Raj Dabre, Siddhartha Gunti & Manish Shrivastava, **Do not do processing, when you can look up: Towards a Discrimination Net for WSD**, Global Wordnet Conference (GWC 2014), Tartu, Estonia, January 2014.
5. Diptesh Kanojia, Manish Shrivastava, Raj Dabre, & Pushpak Bhattacharyya, **PaCMan: Parallel Corpora Management Workbench**, International Conference on Natural Language Processing (ICON 2014), Goa, India, December 2014.
6. Neha Prabhugaonkar, Apurv Nagvenkar, Diptesh Kanojia, Jyoti Pawar, Pushpak Bhattacharyya, & Manish Shrivastava, **PanchBhoota: Hierarchical Phrase Based Machine Translation Systems for Five Indian Languages**, International Conference on Natural Language Processing (ICON 2014), Goa, India, December 2014.
7. Hanumant Redkar, Sudha Bhingardive, Diptesh Kanojia & Pushpak Bhattacharyya, **World WordNet Database Structure: An Efficient Schema for Storing Information of WordNets of the World**, Association for the Advancement of Artificial Intelligence (AAAI 2015), Texas, USA, January 2015.
8. Diptesh Kanojia, Aaditya Joshi, Pushpak Bhattacharyya, and Mark J. Carman, **Using multilingual topic models for improved alignment in English-Hindi MT**, International Conference on Natural Language Processing (ICON 2015), Goa, India, December 2015.
9. Diptesh Kanojia, Naman Gupta, Abhijit Mishra, Shehzaad Dhuliawala and Pushpak Bhattacharyya, **TransChat: Cross-Lingual Instant Messaging for Indian Languages**, International Conference on Natural Language Processing (ICON 2015), Goa, India, December 2015.
10. Meghna Singh, Rajita Shukla, Jaya Jha, Laxmi Kashyap, Diptesh Kanojia, and Pushpak Bhattacharyya, **Mapping it differently: A solution to the linking challenges**, Global WordNet Conference (GWC 2016), Bucharest, Romania, January 2016.
11. Diptesh Kanojia, Raj Dabre, and Pushpak Bhattacharyya, **Sophisticated Lexical Databases - Simplified Usage: Mobile Applications and Browser Plugins for Wordnets**, Global WordNet Conference (GWC 2016), Bucharest, Romania, January 2016.
12. **Diptesh Kanojia**, Shehzaad Dhuliawala, and Pushpak Bhattacharyya, **A picture is worth a thousand words: Using OpenClipArt library for enriching IndoWordNet**, Global WordNet Conference (GWC 2016), Bucharest, Romania, January 2016.
13. Abhijit Mishra, Diptesh Kanojia, and Pushpak Bhattacharyya, **Predicting Readers' Sarcasm Understandability by Modeling Gaze Behavior**, Thirtieth Association for the Advancement of Artificial Intelligence Conference (AAAI 2016), Phoenix, Arizona, U.S.A., February 2016.

WORK EXPERIENCE

Research Engineer (2013 – Present)	<ul style="list-style-type: none">▪ Currently employed as a Research Engineer at Centre for Indian Languages Technology, IIT Bombay.▪ Working on maintenance Hindi, Marathi, Sanskrit WordNets, and IndoWordNet.▪ Parallel Corpora Management Workbench (PaCMan) – a machine translation aid allowing translation / validation / alignment for parallel corpora.▪ PanchaBhoota - Hierarchical Phrase based Statistical Machine Translation system for five language pairs.▪ Cognitive studies for Sentiment annotation, Sarcasm Understandability, and Sarcasm Detection.▪ Mobile App and Browser Plugin development for WordNets on Android, iOS, and Windows Phones.▪ Core Systems Administrator for Centre for Indian Language Technology handling / troubleshooting / maintaining both server systems, and desktop computers.
---	---

Diptesh Kanojia

MALE, 26 YEARS

Email: dipteshkanojia@gmail.com



Contact: +91 8879058668, 9969054037

RESEARCH WORK

"WordNet and its Applications"

Summer Internship
CSE Department, IIT Bombay

Guide: Prof. Pushpak Bhattacharyya
July 2011 – Aug 2011

WordNet	<ul style="list-style-type: none">WordNet is a rich repository of knowledge of words and plays a key role in many Information Retrieval and Language Processing applications.Extensive study of WordNet (English and Hindi) and various state of the art algorithms.
IWSD Algorithm	<ul style="list-style-type: none">Work based on Error Analysis of the IWSD algorithm, divided in various parts, viz. POS based, Ontology based, Ablation test on IWSD parameters.Aim being the importance of contextual part of the scoring function of IWSD algorithm.

"Man vs. Machine Annotation for WSD"

Research Project
CSE Department, IIT Bombay

Guide: Prof. Pushpak Bhattacharyya
Jan 2012 – Mar 2012

Study of Man / Machine Annotation Process	<ul style="list-style-type: none">Our work aimed at understanding the meaning and significance of "contextual evidence", with respect to the human sense disambiguation task.The experiments probed into the disambiguation time across different POS categories and ontological categories, they also investigated into the set of clues in the neighbouring words of a pre-defined target word.
--	---

"Mechanisms for Sense Discrimination"

Minor Research Project
CSE Department, IIT Bombay

Guide: Prof. Pushpak Bhattacharyya
Jun 2012 – Aug 2012

Development of "Sense Discrimination Tool"	<ul style="list-style-type: none">Emphasized on importance of context and clue words around a target word synset word.Developed "Sense Discrimination Tool" which facilitates the addition of clue words by lexicographers manually into the WordNet database.Laid groundwork for automation of the tool in future using the clues present in the database, using various heuristics.
---	---

"Mechanisms for Sense Discrimination"

Major Research Project
CSE Department, IIT Bombay

Guide: Prof. Pushpak Bhattacharyya
Jan 2013 – Mar 2013

Improvements in "Sense Discrimination Tool" and its Future Implications	<ul style="list-style-type: none">Improved the context word searching scenario by providing the lexicographers with a rich corpus.Facilitated the clue addition to database by including Google Hindi Transliteration API within the tool, and by including Concordancer Corpus search.Automated the clue searching by providing with an automatically generated possible clue set.Work on automatic clue addition in future using PMI and G² indexing of the clues.Laid groundwork for a possible future rule based framework to be called, "Discrimination Net".
--	---

"College Website in HTML5 and JS"

Double Major Project
CSE Department, KSJIT

Guide: Mrs Aishwarya Jain
Jan 2013 – Mar 2013

Developed a College Website in HTML5 & CSS3	<ul style="list-style-type: none">Completed College Website development in a month.Development of website in HTML5 and CSS3.Improved the interface and did most of the designing (logos etc.) myself.Made use of JavaScript and jQuery plugins for some features.
--	---

PROGRAMMING AND DIGITAL ART

C / C++	<ul style="list-style-type: none">Advanced knowledge of C / C++ programming language, successful implementation of "Data Structures" and basic algorithms done during academic courses.
JAVA and Python	<ul style="list-style-type: none">Knowledge of JAVA and Python, Implementation of various algorithms for NLP and data manipulation during research work at CFILT.
HTML, PHP, CSS, MySQL, JS	<ul style="list-style-type: none">Proficient website designer, using the aforementioned languages and methods.Designed the complete College Website, in a short span of time.Tweaking of web based interfaces in CFILT Lab, IIT Bombay.

EXTRA CURRICULAR ACTIVITIES

Sports	<ul style="list-style-type: none">Passionate about sports like Basketball, Table Tennis and Volleyball.Enthusiastic runner, have participated in cross country events.
Reading	<ul style="list-style-type: none">Avid admirer of English Fiction, Read many novels and short stories, both in classical English Literature and in Modern day Literature.
Photography	<ul style="list-style-type: none">Zealous about digital photography, keen to learn more about it.Have and Anthology of photography done by self, using both Semi-DSLR and DSLR cameras.
Singing	<ul style="list-style-type: none">Winner of first prize in a school level singing competition.
Debating & Quizzes	<ul style="list-style-type: none">Participant in various school level debating competitions, and head of the college debating club.Winner of Intra College quiz competition, and many other quizzes at school level.