

# Diptesh Kanojia

MALE, 28 YEARS

Email: dipteshkanojia@gmail.com



Contact: +91 8879058668

## EDUCATIONAL QUALIFICATION

Ph. D.	Pursuing	IITB – Monash Research Academy	--
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 the academic years during undergraduate studies.

## 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, Dec. 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), Austin, Texas, U.S.A., 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, Shehzaad Dhuliawala, Abhijit Mishra, Naman Gupta 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.
14. Shehzaad Dhuliawala, Diptesh Kanojia, and Pushpak Bhattacharyya, **SlangNet: A WordNet like resource for English Slang**, Language Resource and Evaluation Conference (LREC 2016), Portorož, Slovenia, May 2016.
15. Diptesh Kanojia, Aaditya Joshi, Pushpak Bhattacharyya, and Mark J. Carman, **That'll do fine!: A coarse lexical resource for English-Hindi Machine Translation, using polylingual topic models**, Language Resource and Evaluation Conference (LREC 2016), Portorož, Slovenia, May 2016.

## CURRENT COURSEWORK

Ph. D. (2016 – Present)	<ul style="list-style-type: none"><li>▪ Course on Linear Optimization (CS750)</li><li>▪ Course on Kernel Programming (CS746)</li><li>▪ Course on Information and Communication Technology for Socio-Economic Development (CS671)</li><li>▪ Course on Communication Skills (CS792 + HS791)</li></ul>
----------------------------	---

# Diptesh Kanojia

MALE, 28 YEARS

Email: dipteshkanojia@gmail.com



Contact: +91 8879058668

## WORK EXPERIENCE

<b>Research Engineer, CFILT, IITB (2013 – 2015)</b>	<ul style="list-style-type: none"><li>Previously employed as a Research Engineer at Centre for Indian Languages Technology, IIT Bombay.</li><li>Cognitive studies for Sentiment annotation, Sarcasm Understandability, and Sarcasm Detection, through Gaze Tracking.</li><li>Multilingual Topic Modelling to improve Indian Language Machine Translation.</li><li>Working on maintenance Hindi, Marathi, Sanskrit WordNets, and IndoWordNet.</li><li>Parallel Corpora Management Workbench (PaCMan) – a machine translation aid allowing translation / validation / alignment for parallel corpora.</li><li>PanchaBhoota - Hierarchical Phrase based Statistical Machine Translation system for five language pairs.</li><li>Mobile App and Browser Plugin development for WordNets on Android, iOS, and Windows Phones.</li><li>Core Systems Administrator for Centre for Indian Language Technology handling / troubleshooting / maintaining both server systems, and desktop computers.</li></ul>
---	--

## PREVIOUS RESEARCH WORK & PROJECTS

<b>"WordNet and its Applications"</b>		<b>Summer Internship CSE Department, IIT Bombay</b>	<b>Guide: Prof. Pushpak Bhattacharyya July 2011 – Aug 2011</b>
<b>WordNet</b>	<ul style="list-style-type: none"><li>WordNet is a <b>rich repository of knowledge of words</b> and plays a <b>key role in many Information Retrieval and Language Processing applications</b>.</li><li>Extensive study of WordNet (English and Hindi) and various <b>state of the art algorithms</b>.</li></ul>		
<b>IWSD Algorithm</b>	<ul style="list-style-type: none"><li>Work based on <b>Error Analysis of the IWSD algorithm</b>, divided in various parts, viz. POS based, Ontology based, Ablation test on IWSD parameters.</li><li>Aim being the <b>importance of contextual part of the scoring function of IWSD algorithm</b>.</li></ul>		
<b>"Man vs. Machine Annotation for WSD"</b>		<b>Research Project CSE Department, IIT Bombay</b>	<b>Guide: Prof. Pushpak Bhattacharyya Jan 2012 – Mar 2012</b>
<b>Study of Man / Machine Annotation Process</b>	<ul style="list-style-type: none"><li>Our work aimed at <b>understanding the meaning and significance of "contextual evidence"</b>, with respect to the human sense disambiguation task.</li><li>The experiments probed into the <b>disambiguation time across different POS categories and ontological categories</b>, they also investigated into the <b>set of clues in the neighbouring words</b> of a pre-defined target word.</li></ul>		
<b>"Mechanisms for Sense Discrimination"</b>		<b>Minor Research Project CSE Department, IIT Bombay</b>	<b>Guide: Prof. Pushpak Bhattacharyya Jun 2012 – Aug 2012</b>
<b>Development of "Sense Discrimination Tool"</b>	<ul style="list-style-type: none"><li>Emphasized on importance of context and clue words around a target word synset word.</li><li><b>Developed "Sense Discrimination Tool" which facilitates the addition of clue words by lexicographers manually into the WordNet database.</b></li><li>Laid groundwork for <b>automation of the tool</b> using clues present in the database, using various heuristics.</li></ul>		
<b>"Mechanisms for Sense Discrimination"</b>		<b>Major Research Project CSE Department, IIT Bombay</b>	<b>Guide: Prof. Pushpak Bhattacharyya Jan 2013 – Mar 2013</b>
<b>Improvements in "Sense Discrimination Tool" and its Future Implications</b>	<ul style="list-style-type: none"><li><b>Improved the context word searching scenario</b> by providing the lexicographers with a rich corpus.</li><li><b>Facilitated the clue addition to database</b> by including <b>Google Hindi Transliteration API</b> within the tool, and by including Concordancer Corpus search.</li><li><b>Automated the clue searching by providing with an automatically generated possible clue set.</b></li><li><b>Work on automatic clue addition in future using PMI and G<sup>2</sup> indexing of the clues.</b></li><li>Laid groundwork for a <b>possible future rule based framework to be called, "Discrimination Net"</b>.</li></ul>		
<b>"College Website in HTML5 and JS"</b>		<b>Double Major Project CSE Department, KSJIT</b>	<b>Guide: Mrs Aishwarya Jain Jan 2013 – Mar 2013</b>
<b>Developed a College Website in HTML5 &amp; CSS3</b>	<ul style="list-style-type: none"><li>Completed <b>College Website development</b>.</li><li>Development of website in <b>HTML5 and CSS3</b>.</li><li>Made use of <b>JavaScript</b> and <b>jQuery</b> plugins for many features.</li></ul>		

## PROGRAMMING AND DIGITAL ART

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

Email: dipteshkanojia@gmail.com

Mob: +91 8879058668

# Diptesh Kanojia

MALE, 28 YEARS

Email: dipteshkanojia@gmail.com



Contact: +91 8879058668

## EXTRA CURRICULAR ACTIVITIES

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