

# Diptesh Kanojia

MALE, 30 YEARS

Email: dipteshkanojia@gmail.com, diptesh@cse.iitb.ac.in, diptesh.kanojia@monash.edu

Contact: +91 8879058668



IITB-Monash  
Research Academy



## EDUCATIONAL QUALIFICATION

Ph. D.	Pursuing	IITB – Monash Research Academy	8.01
B. Tech. (CSE)	2013	Gautam Buddha Technical University, Lucknow	72.90%
Higher Secondary Examination	2005	Greenfields Public School, Dilshad Garden, Delhi (CBSE)	61.40%
Secondary Examination	2003	St. Josephs Academy, Savita Vihar, Delhi (CBSE)	75.00%

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

## AWARDS

- **Awarded "TA of the Semester"** for the course Software Systems Lab (CS251) working with Prof. Kavi Arya.
- **Awarded the CISCO Fellowship** among all applicants from the CSE Department in 2016.
- **Best Poster (Early Researcher Category)**, IITB-Monash Research Academy Annual Awards 2016.
- **Winner of ICON SMT Competition (2014)**, our work on Hierarchical Phrase based Statistical Machine Translation outperformed all other systems, thus making us winners of the competition.

## MAIN CONFERENCE PUBLICATIONS

1. Diptesh Kanojia, Preethi Jyothi, Pushpak Bhattacharyya, **Synthesizing Audio for Hindi Wordnet**, Global WordNet Conference (GWC 2018), January 2018.
2. Diptesh Kanojia, Kevin Patel, Pushpak Bhattacharyya, **Indian Language Wordnets and their Linkages with Princeton WordNet**, Language Resources and Evaluation Conference (LREC 2018), May 2018.
3. Jayashree Gajjam, Diptesh Kanojia, Malhar Kulkarni, **New Vistas to study Bhartrhari: Cognitive NLP**, World Sanskrit Conference (WSC 2018), July 2018.
4. Ritesh Panjwani, Diptesh Kanojia, and Pushpak Bhattacharyya, **pyiwn: A Python-based API to access Indian Language WordNets**, Global WordNet Conference (GWC 2018), January 2018.
5. Hanumant Redkar, Rajita Shukla, Sandhya Singh, Jaya Saraswati, Lakshmi Kashyap, Diptesh Kanojia, Preethi Jyothi, Malhar Kulkarni and Pushpak Bhattacharyya, **Hindi Wordnet for Language Teaching: Experiences and Lessons Learnt**, Global WordNet Conference (GWC 2018), January 2018.
6. Diptesh Kanojia, Nikhil Wani, Pushpak Bhattacharyya, **Is your Statement Purposeless? Predicting Computer Science Graduation Admission Acceptance based on Statement of Purpose**, 14th International Conference on Natural Language Processing (ICON 2017), December 2017.
7. Abhijit Mishra, Diptesh Kanojia, Kuntal Dey, Seema Nagar and Pushpak Bhattacharyya, **Scanpath Complexity: Modeling Reading Effort using Gaze Information**, Thirty-first Association for the Advancement of Artificial Intelligence Conference (AAAI 2017), San Francisco, California, USA, February 2017.
8. Abhijit Mishra, Diptesh Kanojia, Seema Nagar, Kuntal Dey, and Pushpak Bhattacharyya, **Leveraging Cognitive Features for Sentiment Analysis**, The SIGNLL Conference on Computational Natural Language Learning (CoNLL 2016), Berlin, Germany, August 2016.
9. Abhijit Mishra, Diptesh Kanojia, Seema Nagar, Kuntal Dey, and Pushpak Bhattacharyya, **Harnessing Cognitive Features for Sarcasm Detection**, Association for Computational Linguistics (ACL 2016), Berlin, Germany, August 2016.
10. Diptesh Kanojia, Aditya 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.
11. 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.
12. Abhijit Mishra, Diptesh Kanojia, and Pushpak Bhattacharyya, **Predicting Readers' Sarcasm Understandability by Modelling Gaze Behaviour**, Thirtieth Association for the Advancement of Artificial Intelligence Conference (AAAI 2016), Phoenix, Arizona, U.S.A., February 2016.



13. 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.
14. 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.
15. 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.
16. 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.
17. Diptesh Kanojia, Aditya 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.
18. 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.
19. 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.
20. 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.
21. 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.

## CONFERENCE DEMONSTRATION TRACK PUBLICATIONS

22. Aditya Joshi, Diptesh Kanojia, Pushpak Bhattacharyya, Mark Carman, ***Sarcasm Suite: A browser-based engine for sarcasm detection and generation***, Thirty-first Association for the Advancement of Artificial Intelligence Conference: Demonstrations Track (AAAI 2017), San Francisco, California, USA, February 2017.
23. 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.
24. Diptesh Kanojia, Arindam Chatterjee, Salil Joshi, and Pushpak Bhattacharyya, ***Discrimination Net for Hindi***, International Conference for Computational Linguistics (COLING 2012), Mumbai, India, December 2012.

## CONFERENCE WOKRSHOP & COMPETITION PUBLICATIONS

25. Kevin Patel, Diptesh Kanojia, and Pushpak Bhattacharyya, ***Semi-automatic WordNet Linking using Word Embeddings***, Workshop on Word Embeddings at Global WordNet Conference (GWC 2018), January 2018.
26. Diptesh Kanojia, Vishwajeet Kumar, Krithi Ramamritham, ***Civique: Using Social Media to Detect Urban Emergencies***, Workshop on Social Data Analytics and Management (SoDAM 2016), at VLDB 2016, New Delhi, India, September 2016.
27. 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.

# Diptesh Kanojia

MALE, 30 YEARS

Email: dipteshkanojia@gmail.com, diptesh@cse.iitb.ac.in, diptesh.kanojia@monash.edu



Contact: +91 8879058668

## COURSEWORK

<b>Ph. D. Courses (2016 – Present)</b>	<ul style="list-style-type: none"><li>▪ Linear Optimization (CS750)</li><li>▪ Kernel Programming (CS746)</li><li>▪ Information and Communication Technology for Socio-Economic Development (CS671)</li><li>▪ Communication Skills (CS792 + HS791)</li><li>▪ Artificial Intelligence (CS621)</li><li>▪ Software Lab (CS699)</li><li>▪ Embedded Systems (CS684)</li><li>▪ Automatic Speech Recognition (CS753)</li><li>▪ Software Architecture (CS718)</li><li>▪ Seminar on “Computational Phylogenetics” (CS801)</li></ul>
--	---

## 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 detection, Sarcasm Understandability, and 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>▪ 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>
---	--

## POSITIONS OF RESPONSIBILITY

Web Nominee & Design Lead	PG Academic Council, IIT Bombay	Mar 2017 – Mar 2018
<b>Design Team lead &amp; Doctoral Council</b>	<ul style="list-style-type: none"><li>▪ <b>Worked with the Institute Doctoral Council as an internal group lead and conceptualized the PhD Qualifier Document, also lead the design team for various activities.</b></li><li>▪ <b>Discussed the issue of tagging for PhD Students and proposed an alternative scheme for PG Students where Best X out of Y courses can be used to calculate CPI for IITB-Monash Research Academy Students.</b></li><li>▪ Mentored the website work for PG Academic Council.</li><li>▪ <b>Created an online interface for student grievances to be later clubbed with a central portal.</b></li><li>▪ Proposed an exit degree for IITB-Monash Research Academy students.</li><li>▪ Proposed more Faculty Advisors for PhD Students (X faculty on every Y students).</li><li>▪ <b>Direct Transfer of CISCO Fellowship to student accounts instead of the release of a cheque.</b></li></ul>	
Research Scholars Representative	CSE Department, IIT Bombay	July 2016 – July 2017
<b>Research Scholars Representative</b>	<ul style="list-style-type: none"><li>▪ Representing research scholars in the Departmental Post-Graduate Committee (DPGC) meetings.</li><li>▪ Organizing various meetings between potential employers, and interested research scholars.</li><li>▪ Helping organize intra-departmental meets between faculty-RS, and among RS, if need be.</li><li>▪ Organizing monthly ‘Chai-meet’ for RS, where they interact and share ideas across their research areas.</li><li>▪ Organizing RISC as a core-group member.</li></ul>	

## PREVIOUS RESEARCH WORK & PROJECTS

“WordNet and its Applications”	Summer Internship CSE Department, IIT Bombay	Guide: Prof. Pushpak Bhattacharyya July 2011 – Aug 2011
<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>	
“Man vs. Machine Annotation for WSD”	Research Project CSE Department, IIT Bombay	Guide: Prof. Pushpak Bhattacharyya Jan 2012 – Mar 2012
<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>	

# Diptesh Kanojia

MALE, 30 YEARS

Email: dipteshkanojia@gmail.com, diptesh@cse.iitb.ac.in, diptesh.kanojia@monash.edu



IITB-Monash  
Research Academy



Contact: +91 8879058668

<b>"Mechanisms for Sense Discrimination"</b>		<b>Minor Research Project</b> <b>CSE Department, IIT Bombay</b>	<b>Guide: Prof. Pushpak Bhattacharyya</b> <b>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</b> <b>CSE Department, IIT Bombay</b>	<b>Guide: Prof. Pushpak Bhattacharyya</b> <b>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 possible future rule based framework to be called, <b>"Discrimination Net"</b>.</li></ul>
--	---

<b>"College Website in HTML5 and JS"</b>		<b>Double Major Project</b> <b>CSE Department, KSJNET</b>	<b>Guide: Mrs. Aishwarya Jain</b> <b>Jan 2013 – Mar 2013</b>
--	--	--	---

<b>Developed the 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

<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>
----------------------------------	---

## EXTRA CURRICULAR ACTIVITIES

<b>Tech-Geek</b>	<ul style="list-style-type: none"><li>▪ <b>Technology enthusiast.</b> Like being up-to-date with the latest technology and gadgets around. Helps me make conscious decision on what technology / gadget is worth investing into.</li></ul>
<b>Sports</b>	<ul style="list-style-type: none"><li>▪ Played sports like <b>Basketball, Table Tennis</b> at school, and college level.</li></ul>
<b>Reading</b>	<ul style="list-style-type: none"><li>▪ English / Hindi Fiction &amp; Non-fiction, <b>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 Advanced Point-and-Shoot &amp; 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>