



EDUCATIONAL QUALIFICATION

Ph. D.	Pursuing	IITB – Monash Research Academy	7.67
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 undergraduate college, in all the academic years during undergraduate studies.

AWARDS

- **Best Poster Presentation (Early Researcher Category)**, IITB-Monash Research Academy Annual Awards.
- **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. 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.
2. 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.
3. 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.
4. 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.
5. 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.
6. 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.
7. **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.
8. 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.
9. 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.
10. 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.
11. 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.



12. 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.
13. 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.
14. 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.
15. 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

16. 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.
17. 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.
18. 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 WORKSHOP & COMPETITION PUBLICATIONS

19. 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.
20. 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.

COURSEWORK

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

WORK EXPERIENCE

Research Engineer, CFILT, IITB (2013 – 2015)	<ul style="list-style-type: none">▪ Previously employed as a Research Engineer at Centre for Indian Languages Technology, IIT Bombay.▪ Cognitive studies for Sentiment annotation, Sarcasm Understandability, and Sarcasm Detection, through Gaze Tracking.▪ Multilingual Topic Modelling to improve Indian Language Machine Translation.▪ 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.▪ 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.
---	--



POSITIONS OF RESPONSIBILITY

Research Scholars Representative

CSE Department, IIT Bombay

July 2016 – Present

Research Scholars Representative	<ul style="list-style-type: none"> Representing research scholars in the Departmental Post-Graduate Committee (DPGC) meetings. Organizing various meetings between potential employers, and interested research scholars. Helping organize intra-departmental meets between faculty-RS, and among RS, if need be. Organizing monthly 'Chai-meet' for RS, where they interact and share ideas across their research areas. Organizing RISC as a core-group member.
---	--

PREVIOUS RESEARCH WORK & PROJECTS

"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 using 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 the College Website in HTML5 & CSS3	<ul style="list-style-type: none"> Completed College Website development. Development of website in HTML5 and CSS3. Made use of JavaScript and jQuery plugins for many features.
--	---

PROGRAMMING

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.
PHP, HTML, 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 and Maintenance of web based interfaces in CFILT Lab, IIT Bombay. Database Normalization for WordNets, CFILT, IIT Bombay.

Diptesh Kanojia

MALE, 29 YEARS

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

Contact: +91 8879058668



EXTRA CURRICULAR ACTIVITIES

Tech-Geek	▪ Technology enthusiast. 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.
Sports	▪ Played sports like Basketball, Table Tennis at school, and college level.
Reading	▪ English / Hindi Fiction & Non-fiction, novels and short stories , both in classical English Literature and in Modern day Literature .
Photography	▪ Zealous about digital photography , maintain an anthology of photography done by myself, using both Advanced Point-and-Shoot & DSLR cameras.
Singing	▪ Winner of first prize in a school level singing competition .
Debating & Quizzes	▪ 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.