Diptesh Kanojia

MALE, 27 YEARS

Email: dipteshkanojia@gmail.com   Contact: +91 8879058668, 9969054037

EDUCATIONAL QUALIFICATION

B. Tech. (CSE)	2013	Gautam Buddha Technical University, Lucknow	72.96 %
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 **top three** students of college, in all academic years during graduation.

PUBLICATIONS

1. Arindam Chatterjee, Salil Joshi, Diptesh Kanojia, Akhlesh Meena & Pushpak Bhattacharyya, **A Study of the Sense Annotation Process: Man v/s Machine**, International Conference on Global WordNet (GWC 2011), Matsue, Japan, Jan, 2012.
2. Diptesh Kanojia, Arindam Chatterjee, Salil Joshi, & 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.

WORK EXPERIENCE

Research Engineer (2013 – Present)	▪ Currently employed as a Research Engineer at Centre for Indian Languages Technology, IIT Bombay. ▪ Working simultaneously on Hindi, Marathi, Sanskrit WordNets, IndoWordNet and Development of DiscrimiNET. ▪ DiscrimiNET – a resource based disambiguation engine with the ability to disambiguate b/w word senses based on clue words. ▪ 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. ▪ Systems Administrator for Centre for Indian Language Technology handling / troubleshooting / maintaining both server systems, and desktop computers.

RESEARCH WORK

"WordNet and its Applications"	Summer Internship CSE Department, IIT Bombay	Guide: Prof. Pushpak Bhattacharyya July 2011 – Aug 2011
WordNet	▪ 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	▪ 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	▪ 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"	▪ 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	▪ **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 G^2 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, KSJIET	Guide: Mrs Aishwarya Jain Jan 2013 – Mar 2013
Developed College Website in HTML5 & CSS3	▪ Completed **College Website development** in less than a month. ▪ Development and hosting of website in **HTML5 and CSS3**. ▪ Complete website interface from scratch and **did most of the designing (logos etc.) myself**. ▪ Deployed **JavaScript** and **jQuery** plugins for many features.	

POSITIONS OF RESPONSIBILITY

Overall Coordinator	▪ Headed **Abhinithi - 2010**, College cultural festival and fresher's festival. ▪ **Conceptualized and executed** a one day festival, in our college campus. ▪ **Lead a Handpicked team of 25 core-group members** to **handle over 400 guests**. ▪ **Mentored and motivated** the fresher's to take part in the festival and perform various acts on stage.
Overall Coordinator	▪ **Headed Adios - 2012**, college farewell fest. ▪ **Lead a Handpicked team of 15 team members** to **handle over 400 guests.** ▪ Organizational help varying from dance choreography to stage and catering issues.
Editor	▪ Served the **editorial board for our 1st college newsletter, Lakshya**, also conceptualising its **design and layout.** ▪ Served the **editorial board** for "**National Conference on Microwaves and its Applications '10**" **Journal**, held at KSJIET.
Anchor	▪ **Anchored** in the "National Conference on Microwaves and its applications '10" successfully. ▪ **Hosted Abhinithi - 2010**, College cultural and fresher's festival, as a stage anchor. ▪ **Anchored in "Bonvenon '13"** successfully, Fresher's festival in college**.**

PROGRAMMING AND DIGITAL ART

C / C++	▪ **Advanced knowledge of C / C++** programming language, successful implementation of "Data Structures" and basic algorithms done during academic courses.
HTML, PHP, CSS, MySQL, JS	▪ **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.**
Digital Art	▪ **Skilled designer** with knowledge of art designing and DTP software like **Adobe Photoshop & CorelDraw.**

EXTRA CURRICULAR ACTIVITES

Sports	▪ Passionate about sports like **Basketball, Table Tennis and Volleyball**. ▪ Enthusiastic runner, have participated in cross country events. ▪ Participated in intra-college Table tennis competition, leading our batch.
Reading	▪ Avid admirer of English Fiction, **Read many novels and short stories**, both in **classical English Literature** and in **Modern day Literature**.
Photography	▪ 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	▪ 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.