

Naman Kumar Gupta

CONTACT INFORMATION

5412 Gates Hillman Center
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

Voice: (412) 620-3153
Email: nkgupta@cs.cmu.edu
Web: <http://www.cs.cmu.edu/~nkgupta>

EDUCATION

Masters in Language Technologies, Carnegie Mellon University, Pittsburgh
Expected Graduation: August 2010. Current QPA: 3.58

Bachelor of Technology in Information Technology,
Indian Institute of Information Technology (IIIT-A), Allahabad
2002–2006 GPA: 3.4/10

RESEARCH INTERESTS

Personalization, Information Retrieval, User Modeling, Natural Language Processing, Machine Learning

PUBLICATIONS

Naman K. Gupta and Carolyn P. Rose “**Mining Search Behavior Characteristics for non-native English speakers**” *Submitted to the 16th annual ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2010.*

Sourish Chaudhuri, Naman K. Gupta, Noah A. Smith and Carolyn P. Rose. “**Leveraging Structural Relations for Fluent Compressions at Multiple Compression Rates**”. *Proceedings of the Joint Conference of the Annual Meeting of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing, ACL-IJCNLP, 2009.*

Naman K. Gupta, Sourish Chaudhuri and Carolyn P. Rose. “**Evaluating the Syntactic Transformations in Gold Standard Corpora for Statistical Sentence Compression**”. *Proceedings of Human Language Technologies- North American Association for Computational Linguistics conference, HLT-NAACL, 2009.*

Saurabh Garg, Naman K. Gupta, Ratna Sanyal “**Pronoun Resolution and Summary Extraction from English Documents**”. *Proceedings of UDL workshop on Optical Character Recognition with Workflow and Document Summarization, OCR-DS, 2005.*

Naman K. Gupta, Saurabh Garg, Ratna Sanyal. “**Using Anaphoric Resolution and Sentence Splitting for English Document Summarization**”. *Proceedings of International Conference of Speech and Language Technology, ICSLT, 2004.*

RESEARCH EXPERIENCE

Language Technologies Institute (LTI), Carnegie Mellon University **August 2008–Present**
Graduate Research Assistant

Personalization of Search Technologies for non-native English speakers

- Current goal is to identify the impact of low English proficiency and minimal Web experience on Search Capabilities of users from rural areas in developing countries.
- Presently conducting a user study with 12000 users in India.
- Final objective is to design a Personalized search algorithm for such users, catering to the above identified specific needs and limitations, to ensure an efficient and productive search experience.

Sentence Compression

- Worked on enhancing algorithms in Sentence Compression Paradigm.
- Incorporated discourse features to generate shorter grammatical sentences retaining the relevant information in the given context.
- Identified the limitations of the current algorithms by doing an error analysis on the commonly used Sentence Compression corpora.

RESEARCH
EXPERIENCE

UDL (Universal Digital Library), IIIT-A
Research Assistant

June 2004–May 2006

UDL is an initiative taken by IIIT-A and group of institutions in India together with CMU to capture millions of digital books and develop speech, summarization and Optical Recognition technologies.

Document Summarizer

Its goal is to produce a precise coherent summary of an English document which can be absorbed in a single glance and used for rapid relevance assessment. Devised a novel intra-sentential segmentation algorithm pronoun resolution. And customized the Sidners Centering Algorithm to generate short and coherent summaries.

Speech to Text Conversion for Hindi

Its aim is to translate speech in Hindi Language, spoken by illiterate people, to text in order to help them convey their message across the Internet. Developed a speech recognition system using MFCC for feature extraction and k-means LBG algorithm for vector Quantization. We used Hidden Markov Models (HMM) to recognize the features and decoding the speech excerpts.

PROFESSIONAL
EXPERIENCE

IBM, New Delhi, India
Application Programmer

June 2006- June 2008

Responsibilities included working as developer, designer, Team Lead, SCA (Software Configuration Administrator). Worked upon the following Technologies: Java, JSP, Servlets, Struts, Hibernate, SOA (Web Technologies), Oracle 9i, DB2 (DBMS).

AOL (Activation Online)

Web-based application for automated activation of Airtel Postpaid Connections across India. The project aims at automating the activation process which includes customer address credit verification and document validation. My role in the project was to interact with the clients to discuss and analyze the feasibility of their requirements. Development of Exception Tracking Framework and the number search module was part of my role as a Developer.

NMS (Number Management System)

Web-based application used for management of Airtel Postpaid Numbers across India and UK. Used for adding new Postpaid Numbers to the repository for Booking and Activation. My role in the project involved design (HLD, LLD), development and Configuration Management during the duration of the project.

RELEVANT
COURSEWORK

Information Retrieval (Jaime Callan); Experimental Information Retrieval (Jaime Callan)
Machine Learning (Roni Rosenfield); Machine Learning PhD (Carlos Guestrin);
Optimization (Carlos Guestrin); Probability and Statistics (Chad Schafer)
Grammar and Lexicon (Lori Levin); Grammar Formalisms (Lori Levin)

GRANTS AND
AWARDS

- Supported by Research Assistantship at Carnegie Mellon University (2008–2010).
- Honored, twice, with Bravo Award for outstanding performance at IBM.
- Qualified the Joint Entrance Examination (IIT-JEE) in 2002 among the top 1% of 30K students.
- Sun Certified Java Programmer (SCJP 5.0)
- Certification in Modern Chinese Language from Indian Institute of Foreign Trade, New Delhi (IIFT).
- Gold Medalist and Scholar for 6 years at DPS Noida.

POSITIONS OF
RESPONSIBILITY

- Graduate Student Assembly Representative for Language Technologies Institute (LTI) 2009-10.
- Membership Chair, Indian Graduate Student Assembly, 2009-10.
- Member and Treasurer IIIT-A Sports Club, 2004-05.
- Member IIIT-A Soccer and captain IIIT-A Basketball Team, 2004-06.
- Member of Organizing Committee IIIT-A Cultural Fest, 2004 & 2005.
- Volunteering activities as part of the IBM On-Demand Community (IBM ODC) in 2007-08.
- Mentoring as part of the GrassRoots Community (GRC) in IBM.

SKILLS

- Programming Languages: Java, C, C++, Matlab, Perl.
- Web Technologies: JSP, Servlets, J2EE, Struts Framework, SOA, AJAX.
- Tools: Weka, Rational Rose, WSAD, Clear Case, Clear Quest.
- Databases: DB2, Hibernate, Oracle, MySQL.