

Niket Tandon

EDUCATION

- 2012 – *2016 **Ph.D., Computer Science**
Max Planck Institute, Germany
Advisor: Prof. Dr. Gerhard Weikum
Thesis: Commonsense Knowledge Acquisition from the Web
IMPRS Scholarship
- 2010 – 2012 **M.Sc., Computer Science**
Max Planck Institute & Saarland Univ., Germany
GPA 1.2/1.0, Honors Deg.
IMPRS Scholarship
- 2003 – 2007 **B.Tech, Computer Science**
VIT Vellore, India
GPA 9.2/10, Rank: 3/139
Undergrad Scholarship

EMPLOYMENT

- OCT-DEC 2015 **Microsoft Research, Seattle**
Role: Research Intern
Topic: Domain independent Knowledge extraction
- 2013 - *2015 **PQRS Research**
Role: Founder
Topic: Matching Mentee-Project-Mentor for students in non top-tier institutions in developing countries, starting as a pilot project
Mentored award-winning thesis
- AUG-OCT 2011 **Microsoft Research, Seattle**
Role: Research Intern
Topic: Spelling Correction using Language Models with linear and generalized interpolation
Prototype shipped to Bing
- JUN- DEC 2009 **LTRC Lab, IIIT Hyderabad**
Role: Research Engineer
Topic: Cross Lingual IR: Language identification and constructing bilingual dictionaries using Wikipedia
Catalyzed a stagnant project
- 2007 – 2009 **IBM Software Lab, Gurgaon**
Role: Software Engineer
Topic: Agile based complete SDLC for Clearcase Source Code Management
Best internal (PBC) rating: 1.0
- JAN-JUN 2007 **Yahoo R&D, Bangalore**
Role: Undergrad Intern
Topic: Wrapper Induction on HTML pages for Information Extraction
Best undergrad thesis score

✉ Max Planck Institute for Informatics
66123 – Saarbruecken, Germany

☎ +49 176 3533 2047

✉ nntandon@mpi-inf.mpg.de

ℹ bit.ly/nntandon

ACHIEVEMENTS

- GRAD STUDIES IMPRS scholarship awarded to less than 2% applicants
- INDUSTRY Top achiever rating at IBM
- UNDERGRAD Merit scholarship for every year during undergrad
- ENTRANCE EXAMS Top 0.2% at State Level'03, Top 1% at National Level

RESEARCH INTERESTS

Machine Intelligence, Automated Knowledge Acquisition, Data Mining, Information Retrieval, Machine Learning, NLP and its applications, Accessibility.

SELECTED PUBLICATIONS

- [1] [Tandon](#), Hariman, Urbani, Rohrbach, Rohrbach, Weikum: Mining Part-Whole Relations from the Web and Image Tags: AAAI 2016
Introduces fine grained part-whole commonsense
- [2] [Tandon](#), de Melo, De, Weikum: Mining Activity Commonsense from Hollywood Narratives: CIKM 2015
Introduces activity commonsense from scripts
- [3] Chen, [Tandon](#), de Melo: Word Representations from Large-Scale Commonsense Knowledge: WI 2015.
word2vec guided by commonsense
- [4] Shutova, [Tandon](#), de Melo: Perceptually Grounded Selectional Preferences: ACL 2015
Introduces visual tags for selectional preferences
- [5] Rohrbach, Rohrbach, [Tandon](#), Schiele: A Dataset for Movie Description: CVPR 2015
Introduces audio descriptions for rich visual semantics
- [6] [Tandon](#), de Melo, De, Weikum: Knowledge Extraction from Movie Scripts: WWW 2015, short paper
Introduces activity commonsense as semantic frames from Movie scripts
- [7] [Tandon](#), de Melo, Weikum: Acquiring Comparative Commonsense from the Web: AAAI 2014
Open IE and semantic organization over 1B. documents
- [8] [Tandon](#), de Melo, Suchanek, Weikum: WebChild: Harvesting and Organizing Commonsense Knowledge from the Web: WSDM 2014
Semi-supervised modeling of a novel problem
- [9] [Tandon](#), de Melo, Weikum: Deriving a Web Scale Commonsense Fact Database: AAAI 2011
Novel statistical techniques for pattern scoring

SOFTWARE SKILLS

LANGUAGES	Java (proficient), C#, Matlab, V.B., C, C++, Javascript (past experience)
SCRIPTING	Python, Bash
TOOLS	Weka, Solr, Eclipse
SCALE	Hadoop, BigData

SELECTED PROJECTS

MARCH 2014 – JAN 2015

PhD project: Mining Activity Commonsense from Scripts
Max Planck Institute, Germany

To automatically acquire commonsense of activities, we first apply information extracting techniques over nearly 2 million scenes from movies and TV scripts. We then feed the output into a novel technique for semantic parsing, based on identifying clauses, mapping words and phrases to WordNet and VerbNet, and using integer linear programming (ILP) for the final disambiguation and construction of candidate activity frames. We use the output data of the first stage to construct a preliminary activity knowledge graph, with noise and false positives. We then use Probabilistic Similarity Logic (PSL) for efficient inference to construct a cleaner graph as consistent, high-quality output. Demo: bit.ly/tinylnp, bit.ly/knowlywood

APRIL 2012 – FEB 2014

PhD project: Fine-grained Commonsense Mining
Max Planck Institute, Germany

To automatically acquire fine-grained commonsense of activities and comparative knowledge, we combine pattern-based candidate gathering from Web corpora with semi-supervised Label Propagation over judiciously constructed weighted graphs. The edge-weights are derived from sense relatedness, pattern statistics, and co-occurrence statistics. Demo: bit.ly/webchild

JULY 2010 – JULY 2011

Masters thesis
Max Planck Institute, Germany

To automatically acquire several commonsense relations, we rely on a Web-scale n-gram dataset, which gives us a synopsis of a significant fraction of all text found on the Web. Unlike standard bootstrapping approaches, we rely on novel scoring functions to very carefully determine which patterns are likely to lead to good extractions. Unlike previous unsupervised outputs, we rely on a semi-supervised approach for scoring the output facts. Our system performs an efficient large scale information extraction.

DECEMBER 2008 – DECEMBER 2009

Independent project
IBM Labs, IIT Hyderabad

AI in Folksonomy: The major contribution of this work is: understanding the weight models and similarity measures useful for searching social bookmarking systems. The contribution of this effort was a new weighted similarity measure; this was employed in query expansion, improving precision and recall.

SOFT SKILLS

LANGUAGES	Hindi: Native, English: Proficient, TOEFL: 111/120, German: Beginner, Tamil: Beginner
LEADERSHIP	Led projects during undergrad, graduate and as founder of PQRS Research.
MULTICULTURAL	Exciting work experience with colleagues from different cultures.

TEACHING AND SEMINARS

GRAD THESIS	Supervised Masters thesis - Saarland Univ., 2014-15
UNDERGRAD	Supervised Several undergrad thesis with PQRS, 2013-14
TEACHING	TA for Grad course: IR & Data Mining - Saarland Univ., 2011
INVITED TALK	Commonsense Mining. NUS, Singapore - 2012
SEMINAR	AI and the Web: organized 3 days seminar at VIT, India - Oct 2011

PROFESSIONAL ACTIVITIES

PC MEMBER	Scholarly Big Data Workshop, AAAI 2016
PC MEMBER	Scholarly Big Data Workshop, AAAI 2015
PC MEMBER	Novel Computational Approaches to Keyphrase Extraction Workshop, ACL 2015

REFERENCES

Dr. Gerhard Weikum,
Director, Max Planck Institute for Informatics,
Saarbruecken, Germany.

Dr. Gerard De Melo,
Asst. Professor, Tsinghua University,
Beijing, China.

Dr. Martin Theobald,
Professor, Ulm University,
Ulm, Germany.

Please contact me for email ids.