# NIKITA BHUTANI

nbhutani@umich.edu
(734) 548-2891
2260 Hayward Street, Ann Arbor, MI, 48105

## RESEARCH INTERESTS

Knowledge base construction and modeling, knowledge-based question answering and search, data management, information extraction, natural language processing

## RESEARCH EXPERIENCE

**University of Michigan,** Ann Arbor, MI
Graduate Student with Prof. H. V. Jagadish | August 2014-Present
Open-domain information extraction, knowledge-based question answering, natural language interfaces for databases

**Megagon Labs, Inc.,** CA
Research Intern with Wang-Chiew Tan | June 2018 - August 2018
Knowledge base construction from conversational corpus

**IBM Research,** Almaden, CA
Research Intern with Yunyao Li | May 2017 - August 2017
On-demand curation and integration of natural language text

Research Intern with Yunyao Li | June 2016 - Sep 2016
Structured representation of named entities for entity resolution

**Technical University of Liberec,** Czech Republic
Summer Intern | May 2008 - July 2008
Electrospinning from free liquid surfaces

## INDUSTRIAL EXPERIENCE

**Megagon Labs, Inc.,** CA
Research Intern | June 2018 - August 2018
Developed a deep learning method for constructing knowledge bases from conversational question-answer datasets

**IBM Research,** Almaden, CA
Research Intern | May 2017 - August 2017, June 2016 - Sep 2016
Developed a hybrid system for querying structured and textual data, that curates and integrates textual data at runtime.
Developed an active-learning based framework for learning structured representations of named entities

**Ubiquiti Consultants Pvt. Ltd.,** Delhi, India
Software Engineer/ Analyst | July 2010 - June 2014
Worked on UX/UI of automative software suite for analytics and search. Curated ontologies for processing automative data.

**Arvind Limited,** Ahmedabad, India
Summer Intern | Quality Insurance | May 2009 - July 2009

## EDUCATION

**University of Michigan,** Ann Arbor, MI
Ph.D. Candidate | Expected April 2019
Computer Science and Engineering
- Advisor: H. V. Jagadish
- Committee: Michael Cafarella, Rada Mihalcea, Walter Lasecki, Yunyao Li, Qiaozhu Mei
- Thesis: Answering Complex Questions with Heterogeneous Structured Knowledge Sources derived from Text

M.S.E | May 2016
Computer Science and Engineering
- GPA: 4.0/4.0

**Indian Institute of Technology,** Delhi
Bachelor of Technology | June 2010
Textile Technology
- GPA: 8.73/10, Rank: 2

## COURSEWORK

Advanced Database Systems (EECS584) • Advanced Artificial Intelligence (EECS692) • Machine Learning (EECS545) • Natural Language Processing (EECS595) • Advanced Compilers (EECS583) • Information Retrieval and Web Search (EECS498)

## AWARDS AND ACCOLADES

- Nominated by UM-CSE for Rackham Barbour Scholarship, 2018
- IBM PhD Fellowship, 2017
- GHC Travel Scholarship, 2016
- UMich PhD Fellowship, 2014
- Merit Award (5 semesters), IIT Delhi
- Best B.Tech. Thesis, IIT Delhi
- Merit Certificate in Math, AISSCE

## SELECTED RESEARCH PROJECTS

Hybrid KB-QA over open and curated knowledge bases
Developing a KB-QA system that decouples the querying methods for curated and open KBs for inference using effective query decomposition and planning.

Multi-constraint QA with heterogeneous open knowledge bases
Developed a KB-QA system for questions with multiple relations/constraints. It uses an alignment-based algorithm to infer answers from heterogeneous representations in open KBs.

Nested propositions in open information extraction
Developed an open-domain extractor that uses bootstrapping to extract multiple complex assertions as nest-tuples from textual data with no pre-specified relations or training data.

Canonicalization of open knowledge bases (in collaboration with *IBM Cognitive Horizons Network*)
Clustering entity and relation phrases to canonicalize redundant and ambiguous facts in open KBs.

Template-based NLI for relational databases
Developed an NLIDB system that models natural language queries as a set of weighted SQL templates describing the likely query logics and their likelihood to be queried.

Representing news articles as RDF triples
Developed a rule-based system to extract nominal, temporal, spatial, event-based relations between entities in news articles

Optimizing loop unroll factors using machine learning
Developed a supervised learning approach to identify profitable loop candidates and optimal unroll factors

Melt electrospinning of nano-fibres (*B.Tech. Thesis*)
Designed and developed the first in-house prototype for melt electrospinning of nano-fibres, as part of the largest and highest funded research group at IIT Delhi.

## ACADEMIC SERVICE

- External Reviewer for: VLDB 2018, VLDB 2019

## PUBLICATIONS

- "Exploiting Structure in Representation of Named Entities with Active Learning", Nikita Bhutani, Yunyao Li, H V Jagadish, Kun Qian, Mauricio Hernandez, Mitesh Vasa. COLING 2018
- "LUSTRE: An Interactive System for Entity Structuring and Variant Generation", Kun Qian, Nikita Bhutani, Yunyao Li, H V Jagadish, Kun Qian, Mauricio Hernandez. ICDE Demo 2018
- "Nested Propositions in Open Information Extraction", Nikita Bhutani, H V Jagadish, Dragomir Radev, EMNLP 2016
- "Multi-constraint Question Answering with Open Knowledge Bases", Nikita Bhutani, H V Jagadish, SIGMOD 2019 (In submission)
- "Electrohydrodynamics of free liquid surface in a circular cleft: An application to Electrospinning". Bhutani N, Lukas D, Fiber Society Technical Conference, 2008

## PATENTS

- "Entity Structured Representation and Variant Generation", Nikita Bhutani, Yunyao Li, Mauricio A. Hernández, Kun Qian, Min Li (Patent Pending)

## SKILLS

Programming
Java • JavaScript • Python • C++ • Scala • Matlab • HTML • CSS

Web Development / Graphic Design
GWT • jQuery • Spark • Pixelmator • Inkscape • GIMP • D3

Databases / Frameworks
ElasticSearch • Lucene • MySQL • SQLite • MsSQL • Jena