

NIKITA BHUTANI

✉ nbhutani@umich.edu
☎ +1-734-548-2891
🏠 [nikibhutani.github.io](https://github.com/nikibhutani)

RESEARCH INTERESTS

Knowledge base construction, knowledge-based question answering, data management, information extraction

RESEARCH EXPERIENCE

University of Michigan, Ann Arbor, MI

Graduate Student with H V Jagadish | 08/2014 - Present
Open information extraction, knowledge-based question answering, natural language interfaces for databases

Megagon Labs, Inc. (formerly Recruit Institute of Technology), CA

Research Intern with Wang-Chiew Tan | 06/2018 - 08/2018
Open information extraction from conversational corpus

IBM Research, Almaden, CA

Research Intern with Yunyao Li | 05/2017 - 08/2017
On-demand curation of text and integration with structured KB
Research Intern with Yunyao Li | 05/2016 - 09/2016
Learning structured representations of named entities

Technical University of Liberec, Czech Republic

Summer Intern | 05/2008 - 07/2008
Electrospinning from free liquid surfaces

INDUSTRIAL EXPERIENCE

Megagon Labs, Inc. (formerly Recruit Institute of Technology), CA

Research Intern | 06/2018 - 08/2018
Developed an end-to-end system for extracting information from conversational question-answer pairs

IBM Research, Almaden, CA

Research Intern | 05/2017 - 08/2017, 05/2016 - 09/2016
Developed a hybrid system for querying structured and text data, which curates and integrates text data online
Developed an active-learning based framework for learning structured representations of named entities

Ubiquiti Consultants Pvt. Ltd., Delhi, India

Software Engineer/ Analyst | 07/2010 - 06/2014
Developed UX/UI of software suite for analytics and search of automotive data. Curated ontologies for extracting information from semi-structured and unstructured data.

Arvind Limited, Ahmedabad, India

Summer Intern | Quality Insurance | 05/2009 - 07/2009

EDUCATION

University of Michigan, Ann Arbor, MI

Ph.D. Candidate | Expected June 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 | April 2016

Computer Science and Engineering
• GPA: 4.0/4.0

Indian Institute of Technology, Delhi

Bachelor of Technology | August 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
- Rackham Conf. Travel Grant, 2018
- IBM PhD Fellowship, 2017
- Rackham Conf. Travel Grant, 2017
- GHC Travel Scholarship, 2016
- Rackham Conf. Travel Grant, 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 combines both automatically extracted and carefully curated (but incomplete) information to answer complex questions.

Online schemaless querying of heterogeneous open knowledge bases

Developed a querying method for open KBs that is agnostic about query specification and finds answers from facts having diverse representations.

Open Information Extraction from Question-Answer Pairs

Developed a multi-encoder, constrained-decoder framework that extracts tuples from multiple sentences in a conversational question-answer pair.

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

- "Open Information Extraction from Question-Answer Pairs", Bhutani N, Suhara Y, Tan W, Halevy A, Jagadish H V, NAACL 2019
- "Exploiting Structure in Representation of Named Entities with Active Learning", Bhutani N, Li Y, Jagadish H V, Qian K, Hernandez M, Vasa M. COLING 2018
- "LUSTRE: An Interactive System for Entity Structuring and Variant Generation", Qian K, Bhutani N, Li Y, Jagadish H V, Hernandez M. ICDE Demo 2018
- "Nested Propositions in Open Information Extraction", Bhutani N, Jagadish H V, Radev D, EMNLP 2016
- "Online Schemaless Querying of Heterogeneous Open Knowledge Bases", Bhutani N, Jagadish H V, Under Review
- "Learning to Answer Complex Questions over Knowledge Bases with Query Composition", Bhutani N, Zhang X, Jagadish H V, Under Review
- "Answering Complex Questions over Hybrid Knowledge Bases with Query Composition", Bhutani N, Zhang X, Jagadish H V, 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

SKILLS

Programming

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

Databases / Frameworks

Virtuoso • ElasticSearch • Lucene • MySQL • SQLite • MsSQL • Jena