

# Arindam Nandi

205 Springville Avenue, Buffalo, NY 14226

arindamnandidbl@gmail.com • +1 (716) 907-2000

Website: legacy25.github.io • LinkedIn: arindamnandi • GitHub: legacy25

## EDUCATION

### University at Buffalo, The State University of New York, Buffalo, New York, USA

- Master of Science (M.S.) in Computer Science and Engineering Aug 2014 – May 2016
  - Databases and Programming Languages, Distributed Systems, Machine Learning, Operating Systems
  - Thesis (in progress) : Mimir, bringing CTables into practice
  - Cumulative GPA: 4.0 / 4.0

### West Bengal University of Technology, Kolkata, West Bengal, India

- Bachelor of Technology (B.Tech.) in Computer Science and Engineering Aug 2010 – Jul 2014
  - Data Structures and Algorithms, Networking and Cryptography, Compilers, Computer Architecture
  - Cumulative GPA: 8.17 / 10.00

## RESEARCH

## EXPERIENCE

### University at Buffalo, The State University of New York, Buffalo, New York, USA

- Research Assistant Jun 2015 ongoing
  - Project: Mimir, automated ETL data cleaning with provenance tracking to streamline data exploration
  - Adviser: Professor Oliver Kennedy
  - Research areas: Probabilistic databases, data provenance, scalability.
  - Accomplishments: Introduced novel strategies for better join query performance with provenance tracking
  - Technologies: Scala, Java, Play Framework

## PUBLICATIONS

A. Nandi, Y. Yang, O. Kennedy, B. Glavic, R. Fehling, Z.H. Liu, D. Gawlick, "Mimir: Bringing CTables into Practice" under submission, SIGMOD 2016

## SKILLS

C, Java, Scala, SQL, Python, C++, JavaScript, CSS3, HTML5, Git, \*nix

## PROJECTS

### ValkyrieDB Sep 2015 – Dec 2015

- ValkyrieDB is a SQL query processing engine that generates efficient query plans and compiles them to LLVM assembly code before execution, running 50% faster than equivalent Java implementations
  - Technologies: Java, C++, LLVM, Python

### SimpleDynamo Feb 2015 – May 2015

- Implemented a distributed key-value storage similar to Amazon Dynamo as an Android app
  - Chain data replication for data consistency, load sharing based on consistent hashing and partition tolerance for random node failures and rejoins
  - Technologies: Android SDK

### OS/161 Feb 2015 – May 2015

- Augmented a barebones kernel to a fully-functional multiprocessing operating system
  - Added synchronization primitives, processes, file-system system calls and a virtual memory subsystem to the OS/161 operating system kernel
  - Technologies: C, GDB

### Machine Learning Projects Sep 2014 – Dec 2014

- Regression - Query document relevance
  - Used a Microsoft Asia research dataset to train a regression model to predict query to document relevance ratings
- Classification - Handwritten Digit Recognizer
  - Built a handwritten digit recognizer using logistic regression, neural networks and support vector machine classification with training data of 15,000 images, achieving 98.4% accuracy
  - Technologies: MATLAB