

# Sami Hadouaj

shadouaj@umich.edu | linkedin.com/in/sami-hadouaj/ | Michigan, USA

## SUMMARY

---

PhD candidate at the University of Michigan focused on integrating Bayesian machine learning into database cores to natively handle uncertain and probabilistic data. Proficient in C++, with deep expertise in approximate inference, probabilistic programming, and high-performance systems, passionate about solving complex problems at the intersection of databases and ML.

## EDUCATION

---

### University of Michigan (Rackham Graduate School)

Ph.D in Computer and Information Science

Dearborn, Michigan

Jan 2022 – Dec 2026

### Tunisia Polytechnic School

Master of Science in Computer Science

Tunis, Tunisia

Sep 2020 – Jun 2021

### National Institute of Applied Science and Technology

Bachelor in Software Engineering

Tunis, Tunisia

Sep 2016 – Jun 2021

## SKILLS

---

**Languages:** C++, Python, Java, SQL, Bash

**Systems & Tooling:** Linux, Git, Docker, CMake, Ninja, gdb/Valgrind, perf, OpenMP, LLVM/Clang (ClangJIT), Apache Arrow

**Data/ML Libraries:** NumPy, Pandas, scikit-learn, Matplotlib, Eigen

**Specializations:** Databases, Apache Arrow, Probabilistic Programming, Approximate Bayesian Inference, Machine Learning, Parallel Computing, Optimization, System Level Programming, Probabilistic Circuits.

## RESEARCH EXPERIENCE

---

### Graduate Research Assistant

University of Michigan-Dearborn

Jan 2022 – Present

Michigan, USA

- Designed a novel integration of Stochastic Variational Inference for Datalog probabilistic programming+. The system automatically compiles a custom inference algorithm tailored to each probabilistic program.
- Accelerated inference performance by applying knowledge compilation techniques to minimize the number of variational parameters, enhancing efficiency while maintaining model fidelity.
- Engineered a high-performance general purpose probabilistic programming system in C++. The system integrates Apache Arrow for efficient in-memory data handling, LLVM/ClangJIT for just-in-time compilation of inference operations, and OpenMP for parallelized execution. It achieved performance competitive with specialized, model-specific algorithms in benchmark evaluations.

## PUBLICATIONS

---

Ouael Ben Amara \*, **Sami Hadouaj** \*, Niccolo Meneghetti. (*SIGMOD 2024*) “*StarfishDB: A Query Execution Engine For Relational Probabilistic Programming*.”

**Sami Hadouaj** \*, Ouael Ben Amara, Niccolo Meneghetti. (Under Review) “*Stochastic Variational Inference for Datalog Probabilistic Programming*.”

## SERVICE

---

**BIGDATA** (2022, 2023, 2024, 2025): Reviewer

**ICDM** (2025): Reviewer

**SIGMOD ARI** (2024): Reviewer

## HONORS & AWARDS

---

**Rackham Doctoral Intern Fellowship Program:** Awarded the Rackham Doctoral Intern Fellowship Program to conduct research during summer 2024.

## TALKS & PRESENTATIONS

---

**SIGMOD 2024 Conference:** Talk and presentation of StarfishDB in the main research track.  
(Santiago, Chile June 2024)

**Northeast Database Day:** Talk and presentation about query-driven inference in probabilistic databases.  
(Boston, MA March 2023)