Mihail Stoian

Email: mihail.stoian@utn.de
URL: stoianmihail.github.io

Work Experience

Research Assistant — Nov. 2023–present
University of Technology Nuremberg — *Nuremberg, Germany*
Data systems lab (Andreas Kipf)

Research Intern — July 2025–Sept. 2025
Gray Systems Lab, Microsoft — *Barcelona, Spain*
Robust memory estimation (with Tiemo Bang)

Applied Scientist Intern — July 2023–Oct. 2023
Amazon Redshift — *Munich, Germany*
Learned systems group (Tim Kraska)

Student Research Assistant — Mar. 2019–Sept. 2023
TUM, Chair for Database Systems — *Munich, Germany*
`Umbra: A Flash-Based Database System with In-Memory Performance`
Implementing, improving, and testing the functionality

Student Research Assistant — Jan. 2023–Sept. 2023
TUM, Chair for Data Analytics and Machine Learning — *Munich, Germany*
Graph learning with differential privacy

Research Assistant Intern — Aug. 2022–Oct. 2022
Oracle Labs — *Zurich, Switzerland*
`Graph-in-DB` team (Vlad Haprian)

Quantum Software Engineer Intern — Mar. 2021–May 2021
Infineon Technologies — *Munich, Germany*
Quantum Algorithms group

Education

University of Technology Nuremberg — Nov. 2023-present
PhD, Database Systems — *Nuremberg, Germany*
Advisor: Andreas Kipf
Topic: Robust Query Processing

Ludwig-Maximilians-Universität — Oct. 2023-present
Diploma, Orthodox Theology — *Munich, Germany*
Topics: Early Church History, Patristics, Liturgics

Technical University of Munich — Oct. 2021-Aug. 2023
M.Sc., Elite Software Engineering — *Munich, Germany*
Passed with Honors (1.5/1.0)
Thesis: *Optimizing Linearized Dynamic Programming*
Supervisor: Thomas Neumann

Technical University of Munich — Oct. 2021-May 2023
M.Sc., Informatics — *Munich, Germany*
Passed with High Distinction (1.2/1.0)
Thesis: *On the Optimal Linear Contraction Order of Tree Tensor Networks, and Beyond*
Supervisor: Christian Mendl

Technical University of Munich Oct. 2018-July 2021
B.Sc., Informatics *Munich, Germany*
Passed with High Distinction (1.2/1.0)
Thesis: *An Efficient Implementation of Polynomial-Time Join Ordering*
Supervisor: Thomas Neumann

Awards

SIGMOD Honorable Mention (Best Paper Runner-Up) 2025
DPconv: Super-Polynomially Faster Join Ordering

EDBT Best Demonstration Award 2025
Virtual: Compresing Data Lake Files

BTW Best Paper Award 2025
Optimizing Linearized Join Enumeration by Adapting to the Query Structure

SIGMOD Student Travel Award 2023
Proposal: "Bridging the Gap Between Computational Fields"

Bronze Medal 2014
National Mathematics Olympiad, Romania

Scholarships

Deutschlandstipendium Apr. 2022-Mar. 2023
Allianz SE *Munich, Germany*
Scholarship awarded by the Ludwig Maximilian University of Munich

Research Projects

NVIDIA Research July 2022-Sept. 2023
Student Research Project *Remote*
Einsum optimization on GPU
Advisors: Jean Kossaifi, Anima Anandkumar

TUM, Visual Computing & Artificial Intelligence Lab Apr. 2022-Aug. 2022
Practical Course *Munich, Germany*
Outcome: Twofold improvement over DCP, the deep learning approach for iterative closest point (ICP)
Advisor: Matthias Niessner

Interdisciplinary Projects

INSIGHT Mar. 2022-Oct. 2022
Chair of Functional Materials (Prof. Peter Müller-Buschbaum) *Munich, Germany*
Improved the performance of INSIGHT, the package used by the chair for X-ray measurements
Published in Journal of Applied Crystallography.

PushQuantum Apr. 2021-Aug. 2021
IQM Quantum Computers *Munich, Germany*
Organiq-Q: Quantum simulations for OLED properties (pitch)

Programming Competitions

SIGMOD Programming Contest Feb. 2022-Apr. 2022
ACM SIGMOD *Munich, Germany*
We implemented a blocking system for Entity Resolution
Ranking: 6th place, Team: HyTUM

Technical Skills

Languages: {
  "expert" : { C/C++, Python, SQL, Assembler }
  "advanced" : { Java, Isabelle, HTML/CSS/JS }
}
Frameworks: PyTorch, Spark

Language Skills

Romanian: Native
English, German: C2
French: C1
Spanish, Ancient Greek: B1

Invited Talks

Parachute: Single-Pass Bi-Directional Information Passing. TUMuchData @TUM, January 2026

Instance-Optimized String Fingerprints. Joint Research Seminar, October 2025 (remote)

DPconv: Super-Polynomially Faster Join Ordering. Gray Systems Lab, Microsoft, January 2025 (remote)

Virtual: Compressing World's Parquet Files. TUMuchData @TUM, January 2025

What Selinger Forgot to Tell You About Query Optimization. Systems Group, TU Darmstadt, June 2024 (remote)

What do databases and tensor networks have in common? University of Jena, August 2023

Preprints

xBound: Join Size Lower Bounds
Mihail Stoian, Tiemo Bang, Hangdong Zhao, Jesús Camacho-Rodríguez, Yuanyuan Tian, Andreas Kipf
tba

Redbench: Workload Synthesis From Cloud Traces
Johannes Wehrstein, Roman Heinrich, **Mihail Stoian**, Skander Krid, Martin Stemmer, Andreas Kipf, Carsten Binnig, Muhammad El-Hindi
tba

Publications

Mind the Gap. Doubling Constant Parametrization of Weighted Problems: TSP, Max-Cut, and More
Mihail Stoian
STACS 2026

Waiting to Decompress: The Economics of LLM-Based Compression
Andreas Kipf, Tobias Schmidt, Ping-Lin Kuo, Skander Krid, Moritz Rengert, Luca Heller, Andreas Zimmerer, **Mihail Stoian**, Varun Pandey, Alexander van Renen
CIDR 2026

Parachute: Single-Pass Bi-Directional Information Passing
Mihail Stoian, Andreas Zimmerer, Skander Krid, Amadou Latyr Ngom, Jialin Ding, Tim Kraska, Andreas Kipf
VLDB 2025

Instance-Optimized String Fingerprints
Mihail Stoian*, Johannes Thürauf*, Andreas Zimmerer, Alexander van Renen, Andreas Kipf
AIDB @VLDB 2025

Redbench: A Benchmark Reflecting Real Workloads
Skander Krid, **Mihail Stoian**, Andreas Kipf
aiDM @SIGMOD 2025

DPconv: Super-Polynomially Faster Join Ordering
Mihail Stoian, Andreas Kipf
SIGMOD 2025

Virtual: Compressing Data Lake Files
Mihail Stoian, Alexander van Renen, Jan Kobiolka, Ping-Lin Kuo, Andreas Zimmerer, Josif Grabocka, Andreas Kipf
EDBT 2025

Optimizing Linearized Join Enumeration by Adapting to the Query Structure
Altan Birler, **Mihail Stoian**, Thomas Neumann
BTW 2025

Lightweight Correlation-Aware Table Compression
Mihail Stoian, Alexander van Renen, Jan Kobiolka, Ping-Lin Kuo, Josif Grabocka, Andreas Kipf
3rd Table Representation Learning Workshop (TRL@NeurIPS), 2024

Unified Mechanism-Specific Amplification by Subsampling and Group Privacy Amplification
Jan Schuchardt, **Mihail Stoian***, Arthur Kosmala*, Stephan Günnemann
37th Conference on Neural Information Processing Systems (NeurIPS), 2024

On the Optimal Contraction Order of Tree Tensor Networks, and Beyond
Mihail Stoian, Richard Milbradt, Christian B. Mendl
SIAM Journal on Scientific Computing, 2024

Approximate Min-Sum Subset Convolution
Mihail Stoian
22nd International Workshop on Approximation and Online Algorithms (WAOA), 2024

DataLoom: Simplifying Data Loading with LLMs
Alexander van Renen, **Mihail Stoian**, Andreas Kipf
Proceedings of the VLDB Endowment, Vol. 17, 2024

Corra: Correlation-Aware Column Compression
Hanwen Liu, **Mihail Stoian**, Alexander van Renen, Andreas Kipf
2nd Workshop on Cloud Databases (CloudDB @VLDB), 2024

Fast Joint Shapley Values
Mihail Stoian
Student Research Competition, Companion of the International Conference on Management of Data, 2023

Faster FFT-based Wildcard Pattern Matching
Mihail Stoian
Student Research Competition, Companion of the International Conference on Management of Data, 2023

Concurrent Link-Cut Trees
Mihail Stoian
Student Research Competition, Proceedings of the International Conference on Management of Data, 2022

PLEX: Towards Practical Learned Indexing
Mihail Stoian, Andreas Kipf, Ryan Marcus, Tim Kraska
3rd International Workshop on Applied AI for Database Systems and Applications (AIDB), 2021

Benchmarking Learned Indexes
Andreas Kipf, Ryan Marcus, Alexander van Renen, **Mihail Stoian**, Sanchit Misra, Alfons Kemper, Thomas Neumann, Tim Kraska
Proceedings of the VLDB Endowment, Volume 14, 2021

RadixSpline: A Single-Pass Learned Index
Andreas Kipf, Ryan Marcus, Alexander van Renen, **Mihail Stoian**, Alfons Kemper, Tim Kraska, Thomas Neumann
3rd International Workshop on Exploiting AI Techniques for Data Management (aiDM), 2020

SOSD: A Benchmark for Learned Indexes
Andreas Kipf, Ryan Marcus, Alexander van Renen, **Mihail Stoian**, Alfons Kemper, Tim Kraska, Thomas Neumann
NeurIPS Workshop on Machine Learning for Systems, 2019