

# Gaurav Tarlok Kakkar

3<sup>rd</sup> Year Ph.D. Student  
College Of Computing  
Georgia Institute of Technology  
[gkakkar7@gatech.edu](mailto:gkakkar7@gatech.edu)

## RESEARCH INTERESTS

---

My research interest lies in the intersection of databases and machine learning. Specifically, I focus on improving resource efficiency and enhancing usability and query capabilities for video analytics and LLM workloads. My contributions include novel query optimization and execution algorithms, along with improved storage layouts to accelerate these workloads. I am actively involved in leading the development of [EVADB](#), a novel database system tailored for accelerating emerging AI applications. It has garnered recognition with approximately 2.4K GitHub stars and acknowledgment on forums such as [HackerNews \(Discussion 1\)](#), [HackerNews \(Discussion 2\)](#), and [Decibel](#).

## EDUCATION

---

- **Georgia Institute of Technology** Atlanta, GA  
*Ph.D. in Computer Science* Aug. 2021 – ongoing
- **Georgia Institute of Technology** Atlanta, GA  
*Master's in Computer Science; GPA: 4.0/4.0* Aug. 2019 – May. 2021
- **Indian Institute of Technology** Kanpur, India  
*Bachelor of Engineering in Computer Science; GPA: 9.5/10.0* Aug. 2013 – July. 2017

## PUBLICATIONS

---

### A. Published Conference or Journal Articles

- **Gaurav Tarlok Kakkar\***, Jaeho Bang\*, Pramod Chunduri, Subrata Mitra, and Joy Arulraj.  
[Seiden: Revisiting Query Processing in Video Database Systems](#),  
VLDB23: 48th Intl Conf on Very Large Data Bases, Vancouver, Canada, 2023.  
Proceedings of the VLDB Endowment, Vol. 16, No. 11.
- **Gaurav Tarlok Kakkar\***, Zhuangdi Xu\*, Joy Arulraj, and Umakishore Ramachandran.  
[EVA: A Symbolic Approach to Accelerating Exploratory Video Analytics with Materialized Views](#),  
SIGMOD22: 49th ACM SIGMOD Intl Conf. on the Management of Data, Philadelphia, PA, 2022.

### B. Workshop Publications

- **Gaurav Tarlok Kakkar** et al.  
[EVA: An End-to-End Exploratory Video Analytics System](#)

DEEM @ SIGMOD23: 7th Workshop on Data Management for End-to-End Machine Learning, Seattle, WA, 2023.

## C. Demonstrations

- **Gaurav Tarlok Kakkar**, Aryan Rajoria, Myna Prasanna Kalluraya, Ashmita Raju, Jiashen Cao, Kexin Rong, and Joy Arulraj  
[Interactive Demonstration of EVA](#)  
Proceedings of the VLDB Endowment, 2023

## PATENTS

---

- **GT Kakkar**, M Singh Text Wrap Detection - US Patent 11,151,370 ([🔗 link](#))
- M. Rastogi, P. Mehrotra, S. Sinha, **G.Kakkar**. Mapping annotations to ranges of text across documents - US Patent 11,151,307 ([🔗 link](#))
- M. Rastogi, P. Mehrotra, S. Sinha, **G.Kakkar**. Digital Annotation And Digital Content Linking Techniques - US Patent 11,048,864 ([🔗 link](#))

## EXPERIENCE

---

### A. Research Experience

- **Graduate Research Assistant** GaTech  
*Advised by Prof. Arulraj Joy* *Aug 2021 - ongoing*  
**EVA** - A new database management system (DBMS) tailored to efficiently and accurately enable exploratory video analysis at scale. This project focuses on the following problem:
  - **Materialized Views**: EVA automatically materializes and reuses the results of expensive UDFs to facilitate faster exploratory data analysis. It differs from the state-of-the-art reuse algorithms in traditional DBMSs in three ways. First, it focuses on reusing the results of UDFs as opposed to those of sub-plans. Second, it takes a symbolic approach to analyze predicates and identify the degree of overlap between queries. Third, it factors reuse into UDF evaluation cost and uses the updated cost function in critical query optimization decisions like predicate reordering and model selection. Our empirical analysis of EVA demonstrates that it accelerates exploratory video analytics workloads by 4× with a negligible storage overhead (1.001×). The research was presented in SIGMOD 22.
  - **Cascades Optimizer**: EVA has a Cascades-style extensible optimizer implemented from scratch. Since visual queries are often UDF-dominant, EVA extends the Cascades framework to support new rules tailored for optimizing UDF execution. The optimizer aims to reduce query processing time while meeting user-provided accuracy requirements. Specific UDF-specific rules include: 1) materializing and reusing results of the UDFs, 2) adding derived models to reduce computation costs, 3) UDF reordering, and 4) UDF deduplication.
  - **Distributed Execution**: EVA's execution engine leverages heterogeneous computational units (CPUs, GPUs). To support distributed query execution, EVA leverages Ray, a distributed

framework. The modular and extensible nature of EVA enables users to write custom UDFs using deep learning frameworks like PyTorch, Tensorflow, etc.

- **Revisiting Query Processing:** We observe a diminishing gap in inference time between oracle and proxy models in VDBMSs, thereby challenging the assumptions made in SoTA VDBMSs. In this work, we leverage the oracle model and temporal video continuity to surpass SoTA query processing approaches. It incorporates a novel multi-arm bandit-based sampling algorithm for accelerated retrieval and aggregate queries, balancing exploration of unexplored regions while exploiting high-rewarding video segments. Empirical evaluations showcase a  $6.6\times$  speed-up.
- **Model-Aware Storage Management:** Visual queries are often bottlenecked by preprocessing, emphasizing the necessity to cache preprocessing outputs to accelerate queries. In this study, rather than treating preprocessing as a monolithic step, we dynamically optimize various video knobs, such as resolution and keyframe intervals. This approach better navigates the trade-off between query execution time and storage budget, as demonstrated by empirical analysis showing up to a  $2.2\times$  speed-up on queries bottlenecked by preprocessing.
- **Adaptive Query Processing:** Traditional static query optimization does not work well for VDBMSs, as UDFs often dominate the queries, and it is non-trivial to collect accurate estimates of UDF statistics, such as selectivity and cost. To address this limitation, in EVA, we propose an extensive and generalizable adaptive query processing framework to adjust the query plan at the execution stage dynamically. Our early empirical results show up to a  $2\times$  speed-up.

## B. Industrial Experience

- **Snowflake** Sunnyvale, USA  
*Software Development Intern in SQL Optimization Team* *Summer'21*
  - Led the research project to drive new query optimizations by analyzing production workloads based on query inter arrival patterns, query types, and execution statistics.
  - Designed architecture to collect and report back the query runtime statistics to the optimization engine to optimize future queries.
- **Google** Sunnyvale, USA  
*Software Development Intern in Cloud SQL* *Summer'20*
  - Led the project to accelerate OLAP (Online Analytical Processing) queries by automatically building columnar cache indexes
  - Improved the query statistics collection engine and building ML driven columnar cache index advisor.
  - 5x improvement in query execution time with no manual cost overhead and worked towards US patent.
- **Adobe Systems** Noida, India  
*Member of Technical Staff* *Jul 2017 - Aug 2019*
  - **Regenerate Layout from PDFs:** Research project

- \* Built a multitude of ML algorithms processing together along with Software level heuristics to provide a single click layout generator from an inspiration pdf
- \* Implemented a deep learning model, modified Faster RCNN to detect shape agnostic text wrap in a given pdf
- \* Tackled challenges viz. detecting white space cover, creating master pages, organizing raw text runs into well defined text frames and intelligently figuring out object styles.

**Key Achievements:** Filed patent in US on shape agnostic text wrap detection, implemented document analysis techniques and researched on core text properties.

- o **Import PDF Comments:** Keynote feature shipped with InDesign Max 2019
  - \* Implemented a novel approach to import and easily track the feedback made on a pdf version of document, solving the most in demand feature request of our million designers
  - \* Mastered the existing PDF library, to tackle the challenges of associating text or graphics with the annotation.

**Key Achievements:** Filed patent in US and mastered PDF Library APIs

- **Adobe Systems** Bangalore, India  
*Research Intern in Big Data Experience Lab* *May 2016 - July 2016*

- o Generating personalized bundles of products for customers of e-Commerce website that are needs-driven.
- o Formulated a novel approach of incorporated common sense knowledge, Concept net along with data driven insights.
- o Formed candidate set using hierarchical and minimum spanning tree-based clustering algorithm over customer-centric data enriched by semantic analysis. ( [↗ slides](#) )

## RECOGNITION AND AWARDS

---

- 2017** Dr. Elizabeth & Varkey Cherian Award(Best UG project with an impact on campus community)
- 2017** Academic Excellence Award, IIT Kanpur (awarded to top 7% students in the institute)
- 2014** Academic Excellence Award, IIT Kanpur (awarded to top 7% students in the institute)
- 2013** All India Rank 236, IIT-JEE Advanced (among 150,000 candidates).

## RELEVANT COURSES

---

- **Systems:** High Performance Parallel Computing, Advanced Operating Systems, Database Technologies, Computer Architecture, Compiler Design, Computer Networks, Computer Security, Advanced Data Structure and Algorithms
- **ML/Data Science:** Data analysis using Deep Learning, Recent Advances in Computer Vision, Machine Learning Tools and Techniques, Natural Language Processing, Data Visualization and Analysis

## OTHER PROJECTS

---

- **[ML Systems] Fast Array of Wimpy GPUs (FAWG)** GaTech  
*Research Project with Prof. Alexey Tumanov* *Sep. 2020 – Jan. 2021*
  - Serve memory-hungry models using model parallelism on cheap wimpy GPUs while meeting the latency SLOs
  - Proactive planner regresses over the batching parameters, model partitions, operator replicas, hardware types, and operator placement to search for a cost-effective model serving plan
  - The reactive planner behaves as a high-frequency tuner to auto-scale to meet tail latency goals in response to changes in the query arrival process.
  
- **[Databases] Cafeteria Automation System** IITK, India  
*Under-Graduate Project with Prof. Sumit Ganguly* *Jan. 2016 – Dec 2016*
  - Designed a desktop app in C# incorporating mess menu creation, consumption Records, items BOM management, worker Attendance and salary management.
  - Won **Dr. Elizabeth & Varkey Cherian Award - Best UG project** with an impact on campus community.
  - As of May 2017, managed over 2,00,000 transactions of worth greater than INR 3.4 million.  
( [↗ slides](#) ) ( [↗ code](#) )

## TECHNICAL SKILLS

---

**Languages** Python, C++, C, SQL, OpenMP, MPI, Js, Go, Bash, Assembly, HTML, CSS  
**ML** Tensorflow, Keras, scikit-learn, OpenCV, PyTorch