

# ANISH KUMAR KHARWAR

Senior Software Engineer | Scalable Microservices · Distributed Data Pipelines · Cloud (GCP/Azure)

[anishk7475@gmail.com](mailto:anishk7475@gmail.com) | +91 7905395419

LinkedIn: [anishk74](#) | GitHub: [anishk74](#) | Portfolio: <https://anishk74.github.io>

---

## Experience

**Senior Software Engineer | Nference Labs**

Python, Bash, Google Cloud Platform (GCP), Microsoft Azure, Django, Tesseract

August 2021 – Present

- Designed and deployed 4 microservices within a publisher-subscriber architecture to process and transform **480TB+** of high-volume, multi-format data. Implementations included pipelines for large-scale image parsing (**810GB**), structured data processing (**9TB**), high-throughput sequence data handling (**460TB**), and binary signal data processing (**21TB**), delivering reliable, scalable ingestion and transformation across diverse datasets.
- Extended image parsing microservices to support PDF parsing, implementing logic to handle overlapping characters and unordered text segments, enabling accurate text extraction from a **700GB** dataset.
- Enhanced structured data microservices to support **80+** additional data categories by analyzing raw datasets and extending processing workflows, enabling broader data coverage and improved system capabilities.
- Developed a pipeline to aggregate and analyze data loss during processing across **45TB** structured datasets, optimizing processing logic and reducing data loss from **23%** to **9%**.
- Integrated and optimized a new image model into image processing microservices, enabling processing of **12M+** files.
- Collaborated with client engineering teams to troubleshoot deployment issues, resolve incorrect data formats and address processing errors in deployed services, ensuring smooth onboarding for multiple deployments.
- Developed and maintained 60+ backend APIs for authentication and role-based access control, implementing secure management of users, groups, organizations, and applications to enforce granular permission models.

**Computer Vision and Machine Learning Engineer | Awiros**

Python, C++, MXNet

June 2021 – July 2021

- Designed data preprocessing pipelines for large-scale image datasets (**494K+** and **43K+** files), implementing logic for distribution analysis, aspect ratio handling, and quality checks to ensure robust downstream processing.
- Implemented backend automation for the ML lifecycle, including model creation, training, deployment, and inference, integrating with the UI to deliver streamlined, production-ready workflows.

**Backend Intern | Vaultedge Software**

Python, Flask, MongoDB, Grafana

March 2021 – May 2021

- Optimized job scheduling logic by migrating from FIFO to a priority queue system, improving task throughput and reducing latency for high-priority operations.
- Designed and implemented **6** REST APIs in Flask to serve performance data from MongoDB to Grafana dashboards, enabling low-latency and real-time visualization.
- Created **20+** system performance metrics and equations for monitoring and display, improving operational visibility.

## Skills

**Languages:** Python, Bash

**Backend:** Django, Flask, REST APIs

**Architecture:** Microservices, Distributed Systems, Pub/Sub

**Cloud & DevOps:** GCP, Azure, Docker, CI/CD, Git

**Databases & Queues:** Redis, MySQL, MongoDB, RabbitMQ

**Observability:** Grafana, Prometheus

## Education

2017 – 2021

Bachelor of Technology, Computer Science and Engineering

Indian Institute of Information Technology Kalyani

Cumulative GPA: **8.29**

## Interests

Distributed Computing, Cloud-Native Scalability, High-Throughput & Low-Latency Systems