

# Shreyas Kapale

+91-7892628145 | [kapale.shreyas@gmail.com](mailto:kapale.shreyas@gmail.com) | [github.com/shreyaskapale](https://github.com/shreyaskapale) | [linkedin.com/in/shreyas-kapale-82011513a](https://www.linkedin.com/in/shreyas-kapale-82011513a) | <https://shreyaskapale.com>

## Introduction

I like to build things and solve problems, I am a full stack developer with primary focus towards backend and scalable cloud infrastructure. I write clean code based on SOLID principles. I am hardworking and people friendly.

P. S. - Solutions = grep problems | Shreyas

Indie creator of Cheap GPT – served 200K + users from April 2022 – present

## TECHNICAL SKILLS


- **Programming Languages:** Python, Typescript, Go, Java, C++
- **Skills:** Micro-services, queuing/ ticketing systems, authentication, integrations, distributed systems, front-end development, databases, GRPC/Restful/GraphQL API's, Cloud Infra, ORM, design patterns.
- **Backend development technologies:** Python - (FastAPI, Flask, Django), JS/Typescript – (NestJS, PM2, Express), Go - (Gorilla, Gin)
- **Containerization, Orchestration, Caching and Micro-services:** Celery, **Serverless** (cloud formation), Redis, RabbitMQ, SQS, Docker, Apache Kafka, NATS & Kubernetes.
- **Cloud** – AWS & GCP
- **Mobile app development technologies:** Flutter ( Android & IOS ) , Ionic & Cordova
- **Front End web development frameworks** – React, NextJS, TailWind, VueJS, HTML5, CSS3, Material5
- **Database :** MySQL, MongoDB, InfluxDB, PostgreSQL, Firebase
- **Logging** – CloudWatch, GCloud Monitoring, ELK (Elastic, log stash and Kibana ), DataDog.
- **CD/CI** – Github Actions, Jenkins, CodeMagic & Travis CI, Mono-Repository.
- **Miscellaneous:** Open-Telemetry, Prometheus, Mixpanel, Pandas, OpenCV,Dask - distributed, Scikit-Learn, OpenWRT, Wireshark, TCPdump, Etherape, Bash, Linux, Rust, Routers, Git, TurnKey etc ...

## PROFESSIONAL EXPERIENCE

- NeuralGo - Lyzr**

Full Stack Developer Engineer


August 2023 - Current

 Remote, Newyork

- Built and orchestrated compliant **data integrations pipeline** leveraging Nango for Oauth and live sync capabilities for **GenAI** platform. Orchestrated seamless data flow and GDPR compliance by using non-sensitive Oauth scopes for handling data, deployment via combination of AWS EC2 instances and Python-based Lambda functions (Micro-services) served over AWS serverless IAAS.
  - Built API **Gateway with Authorizer for serverless** - cloud formation component fortified with a sophisticated **user membership system** (using memberstack for checking plan based resource restrictions) and bearer token authentication. Built Python-powered Lambda function development, ensuring a secure and performant API layer for critical data interactions. While making sure the latency is minimal.
  - I played pivotal role building field mapping AI algorithm for datasets by implementing **fuzzy matching algorithm using levenshtein distance** to calculate differences between sequences and deprecated GPT3.5-Turbo in favour of self-hosted LLAMA2 which resulted in response time **reduction of 30-40 %**.
  - Orchestrated the successful deployment of LLAMA2 with 65 billion params on an AWS EC2 instance with G5.4xlarge specifications.
  - Optimized the AI prompt recommender feature. Employed **REDIS caching** to achieve an **unprecedented 95% reduction in response time**
  - Built frontend in React ( Next.js ) integrating **SSO functionality for login and data integrations**
  - Led front and backend development for user API key feature and built whole flow for community cloud version.
  - **Technologies used** - Python, **FastAPI, AWS Serverless** – cloud formation ( IAAS ), **GPT –LLMs like GPT4, LLAMA ,** Dynamodb, Route53, CodeBuild, Memberstack, Docker, Redis.
- Yellow**

Software Development Engineer

March 2022 – July 2023

 Bangalore, India

- Being a former employee, led and built product features for Yellow mobile and web app.
  - Built and orchestrated a comprehensive communications layer for direct/scheduled template-based notification system for WhatsApp, Mobile Push Notifications, and Email.
  - Spearheaded the development and infrastructure of Yellow Plus backend, orchestrating a seamless amalgamation of the cart system with RazorPay for streamlined payment processing with webhooks.
  - Demonstrated advanced proficiency in backend development by significantly reducing GraphQL call times from **800ms to an impressive 200-250ms** by integrating Redis and utilizing NestJS context as cache which gained in performance.
  - I engineered an automated image compression pipeline that resulted in an **impressive 90% reduction in document generation time**. This system, which autonomously processes bucket uploads, utilizes the **Sharp library within a cloud function**.
  - Built entire backend **for CAS parser for parsing NSDL/ CDSL documents** into known asset classes on our infra.
  - Built entire end to end feature for demise trigger with user account locking feature and a dashboard page for the customer support/ ops team to manage the demise of a testator.
  - Built entire analytics layer on Mixpanel for the product on flutter and backend.
  - **Technologies used** – NestJS, Flutter, React, GraphQL, Firebase, GCP, Cloud functions, Mixpanel

## Synth ( YC W21 )

Product Engineer (backend focused)

October 2021 – March 2022

📍 Bangalore, India (Remote)

- Being a former employee, led and built product features for Synth electron app
- Written **ticket queue based micro-services** for async transcription and async diarization, with optimized transcript stitching function. **Nvidia Riva SDK deployment on g4dn.xLarge**, and written load testing automation scripts for Nvidia RIVA GRPC server for Speech to text (ASR).
- Built entire async file upload pipeline queued audio to text transcription which utilized AZURE SST.
- Built end to end user authentication using Oauth PKCE for Electron using AWS Amplify (cognito) with SSO login providers like Google sign in and Apple sign in.
- Feature to get current opened window details using WINDOWS native APIS and Linux XPROP and sending details over to electron app.
- Done AWS Route53 DNS delegation for different services on the same root domain and also built a API versioning system
- Built Azure batch transcription web-hook integration with auto re-verify mechanism.
- Built end to end Google Calendar and tasks API integration with Oauth integration on frontend.
- Built login components, speaker timestamp diarization user stitching UI components in VueJS and Electron.
- **Technologies used – Python , Electron, SST – ASR Azure , Nvidia RIVA, Cognito, Amplify**

## Finclus

Founding Engineer (Full stack)

Jan 2021 – October 2021

📍 Bangalore, India (Remote)

- Being a lead and built most of the vital infrastructure for building the product features.
- Built AWS lambda functions - OCR for Aadhaar and Pan Card using FastAPI, Amazon Rekognition, BOT03, Dynamo DB and S3. Whole pipeline from extraction and storage procedures.
- Built end to end user authentication using Oauth PKCE for Electron using AWS Amplify (cognito) with SSO login providers like Google sign in and Apple sign in.
- Built Typescript SDK for QR based authentication system for FinclD as part of SSI (self-sovereign identity) infrastructure.
- Setup whole internal network infrastructure (VPC) for internal services to talk to each other on subnet.
- Built a simple event ticketing bus for async services using lambda functions and dynamoDB
- Wrote smart contracts for issuer, holder and verifier in R3 Corda - Kotlin as part of the SSI infrastructure.
- Made flutter components and screens following the **MVC pattern**.

## - Informant Networks

Software Developer Internship (Router Development)

August 2018 – October 2018

📍 Belgaum, India

- OS/ Networking level ( **end to end** )
  - Router development, working on OpenWRT Linux.
  - **Multi-threaded ICMP link monitor** with uptime statistics (service for cloud router) which helps in knowing if a link/ Iface is down in Cpython.
  - Firmware update system with SHA based Integrity check
  - Writing make files for openwrt to run on different embedded boards
  - Parsing router usage and monitoring from proc data and piping it via **OpenWRT UCI**.
- Frontend -
  - Built PHP5 and AJAX based frontend for ( WIFI configuration and LAN configuration )
  - Built monitoring graph UI Components from scratch **for packet loss and CPU usage**.
- Notable Achievement in terms of **optimization**
  - Decreased the CPU usage by reducing AJAX calls made memory efficient code on backend which reduced **CPU usage by around 40-50%**.

## FREELANCE PROJECTS

### - Infigon Futures

Oct 2020-Jan 2021

- Managed the entire development team for building the application.
- Built microservice's with FastAPI and containerized on EC2 with docker and load balanced using AWS ELB
- Built a CD/CI pipeline integrated with AWS ECR. Built and deployed machine learning API's for personality testing (used for career guidance).
- Wrote machine learning algorithms using scikit-learn with a random forest classification for personality tests.
- Flutter App development for mobile and Web.
- Wrote automated python web scraping scripts.
- Technology stack: FastAPI, Python, Scikit-learn, AWS, Flutter

### - Rising Youth Foundation (OFY App)

Jan 2021- June 2021

- Built Complete Dashboard: Developed the entire dashboard for the OFY app using React and Material UI. This involved creating a responsive and user-friendly interface that allows users to access and manage various features and data within the app. The React framework and Material UI components were leveraged to ensure a modern and visually appealing design.
- Implemented Group and Direct Chat Features: Integrated group chat and direct chat functionalities within the OFY app using Flutter, Express, and Firebase Realtime Database. Users can now engage in real-time conversations with other individuals or within specific groups. This involved setting up server-side handling of messages using Express, utilizing Flutter for the mobile app interface, and Firebase Realtime Database for seamless and synchronized messaging.
- Technology stack: React, NodeJS, Firebase, Flutter.

### - PractWorks

Dec 2020- Jan 2021

- Designed and developed application features for PractWorks, an education app. specifically focused on building features for content display and a fine-grained video streaming system. Utilized Flutter, Node.js, and Express.js to create a seamless and immersive educational experience.
- Implemented a robust content display system, allowing users to access educational materials such as videos, documents, and interactive modules. Leveraged Flutter's UI toolkit to create an intuitive interface for browsing and accessing the content.
- Built a video streaming section with play/pause/seek features mapped with content.
- Technology stack: Flutter, NodeJS

## Dec 2020- Jan 2021

- ### - Sudina – A Grocery delivery app

May 2020 - Aug 2020

- **Informant networks – Person Tracker App**

*August 2020 - Oct 2020*

- **LaundroKart** – Emailing system

July 2019 - august 2019

- ## PERSONAL PROJECTS

*April 2023*

 [Demo](#)


- Technology stack: React, FastAPI, GPT3, Oracle Cloud, BS4, Requests.

## Jan 2020 – June 2020

 GitHub

- ## Jio-Fiber API ( Python Package )

June 2021 – July 2021

 Package

- **BFS based - Shortest Path Visualizer**

July 2023

 [Demo](#)

- RangeGPT

May 2023

 [Demo](#)


- QR Contact sharing app ( PanAm)

Oct 2020

 [Github](#)

- **Project ALPHA Robot**

Aug 2019

 [Youtube](#)

- **Dynamic Shoes** - charge mobile phones by walking

Oct 2016



- Won Gold in INSEF science and engineering fair, and got recognized by science society of India.

Talks

- **PyCON INDIA 2021:** Talk on **distributed computing** using Dask.
- **PyFLAMES 2021:** Lightning talk on converting normal routers into cloud routers.
- **Avalanche 2019:** Talk on "Heart Disease detection using KNN (K-Nearest Neighbors)."
- **Avalanche 2018:** Presentation on R-CNN (Region-based Convolutional Neural Networks).
- **Science Hack Day 2018:** Talk on "How to start programming as a teen."
- **Science Hack Day 2017:** Lightning talk on pen-testing wireless networks.
- **Science Hack Day 2022:** Talk on android development with flutter

EDUCATION

Degree/Certificate	Institute/Board	CGPA/Percentage	Year
B.E.	KLS Gogte Institute of Technology	8.8	2018-2022

KEY COURSES TAKEN

**Courses** - OOPS in Java and Python, DSA in C, DBMS and Compiler Design.  
**Electives** - Advanced Algorithms, Information Security & Networking, Cloud Computing.  
**Machine Learning:** Supervised vs Unsupervised machine learning, classification, regression, dimension reduction, and clustering.