

Sonu Kumar

Mail Id - sonuk.kumar@yahoo.com

Location - **Bengaluru** (Karnataka, India)

LinkedIn - <https://www.linkedin.com/in/sonu-kumar-55a218138/>

Mob- +918093155605

Career Overview

A Python developer with over five years of comprehensive experience spanning the entire project life cycle, from planning and evaluation to requirements gathering, design, development, testing, and deployment. Skilled in product and infrastructure improvements and well-versed in agile methodologies.

Summary

- Leveraging advanced NLP libraries like SpaCy, Recognizers Text, and transformer-based models such as GPT to extract and summarize critical data tailored for various business domains.
- Skilled in analysing large data sets and building systems to track Timescale data using Prometheus and Grafana.
- Skilled in development, unit testing, debugging, root cause analysis, and bug resolution, adhering to TDD and BDD methodologies.

Education

- 2018 – Bachelors of Technology (Computer Science and Engineering) - KIIT University - 8.01/10
- 2014 - 12th (Sr. Secondary) - 67.5%
- 2011 - 10th (Secondary) - 9.6/10

Technical Skills

- Programming Languages: Python, C++
- AI, LLM and ML : GPT, SpaCy, BERT, Recognizers Text, SWE-bench, Terminal-bench, Claude Code, SpecStory, Agent Trajectory Evaluation
- Python Libraries: Flask, Sanic, FastAPI, Langchain, Transformers, scikit-learn, pandas, numpy, pytest, paramiko, prometheus client, smtplib, matplotlib, Tkinter, Cookiecutter
- IDEs & Tools: PyCharm, Jupyter Notebook, Visual Studio
- Development & Collaboration Tools: Jira, Atlassian Suite, Bitbucket, Git, Docker
- Monitoring & Analysis: Prometheus, Prometheus Pushgateway, Grafana, Kibana
- Databases: Oracle, SQL, MySQL, PostgreSQL
- Operating Systems: Windows, Linux, MacOS
- Web Technologies: JSON, HTML5, CSS3, JavaScript
- Cloud Services: AWS MQ, EC2, CloudFormation Template, S3 Bucket, Azure Queue, Azure Container Registry
- BFSI: Fixed Income, Equities, Hedging, private investment

Professional Experience

AI Benchmarking Contributor at Mercor (AI Benchmarking Engineer)

Oct 2025 – Present

Worked on evaluating and improving agentic coding systems across SWE-bench and Terminal-bench environments.

- Generated structured trajectories (trajectory.md, trajectory.jsonl) and patches by solving real repository-level issues using Claude Code and Cursor with SpecStory tracking.

- Performed OpenCode baselining by reviewing agent reasoning, validating fixes through Docker-based oracle verification, and refining evaluation rubrics.
- Contributed to autonomous repo-review workflows enforcing structured iteration, correctness validation, and benchmark compliance.
- Strengthened reliability of LLM-assisted software engineering pipelines through reproducible verification workflows and artifact-driven submissions.

Manager at Morgan Stanley (Python Developer)

Apr 2024 – Oct 2025

Part of wealth management and investment management technology department. Working on deployment of microservices of Fixed Income and securities like Interest rate hedging and currency repatriation. Helping in code refactoring of code for getting and updating portfolio data.

- Developed an application to poll near real-time positions from Aladdin and save in DB2 database.
- Contributed to the migration of flask application to fastAPI and Kerberos authentication to OAuth.
- Developed apps to ease out the order submission for Muni/Corp Bonds, provided filters to sort the bonds based on different reference data.

Senior Product Software Engineer at Wolters Kluwer (Python Developer)

Dec 2022 – Apr 2024

Enhanced and implemented Annotator template leveraging Flask, NLP and LLMs to conduct semantic analysis, content summarization, and extraction from textual datasets, tailored for diverse business domains.

- Working on an annotator template utilizing Flask, langchain and cookiecutter, to provide a generic solution to create annotators for cognitive insight services with any LLM models. Developed prompts for multiple skills like summarization (stuff, refine ad map-reduce), QA, Translation.
- Developed an annotator utilizing GPT 3.5 Turbo, GPT-4, and GPT-4 32K models, enabling context-aware summarization of textual content and documents.
- Enhanced an existing annotator service by integrating the 'Recognizers Text' library, enabling it to accurately detect and extract date and temporal entities from textual content.
- Contributed to the migration of an annotator service from the Sanic framework to the Flask framework, optimizing performance and scalability.
- Implemented a specialized annotator service, utilizing the capabilities of SpaCy, to precisely identify and extract named entities from diverse textual sources. Trained SpaCy NLP model with custom training data-points.
- Fine-tuned state-of-the-art BERT model to enhance parts of speech tagging, optimizing named entity recognition with tailored datasets and advanced training strategies.

Worked on bug fixes for RDF based Semantic Content Enrichment application.

- Fixed API response for incorrect method and endpoint.
- Remove bibliographic resource info for non-entity RDF data.
- Script to call annotator workflows to compare extracted dates from RDF data.

Software Engineer at SymphonyAI (Python Backend Developer)

May 2022 – Dec 2022

Crafted monitoring utility modules with the Prometheus client to oversee the performance of industrial assets. This facilitates anomaly and error detection in sensors via counter and gauge metrics, leading to reduced maintenance costs for equipment and machinery.

- Engineered and sustained the Asset Template microservice, streamlining the provisioning of machinery and sensor data.
- Enhanced the Measurement microservice, enabling the assimilation of sensor readings via diverse APIs.

- Instituted a monitoring mechanism to oversee status codes across various APIs and modules, reducing the likelihood of API call failures.
- Constructed an HTTP-REST API using Flask for the intake of Timeseries data.

Senior Associate at Macquarie (Python Backend Developer)

Sep 2020 - May 2022

Designed and maintained products for key bank-wide initiatives, optimizing Transaction and Static Reference Data Caches. This central service benefits Middle and Back Office systems like Risk, Finance, and Operations. It resolves data discrepancies across trade platforms by offering a standardized, authoritative set of thoroughly vetted transaction and reference data.

- Designed and sustained the QRM (Quantitative Risk Management) application, producing diverse reports that illustrate the bank's and its clients' risk exposure.
- Retrieving deals and rates, generating reports in .csv format, and overseeing the deals using Prometheus pushgateway, Prometheus, and Grafana.
- Automating and scheduling macros to send an email with attached reports on a particular date and time.
- Management of Regulatory Reporting applications for daily trades to different regulatory bodies across all operational regions.

Software Developer at ALTRAN (C++/Python developer)

Oct 2018 - Aug 2020

Developed SIBs (Service Independent Building Block API) for the product Real-Time Charging control (RTCC), an SCF, which provides an interface to provide an online charging system employing IP Multimedia Systems core and the existing CAMEL/INAP services along with creating and analysing call data records for its preciseness and accuracy.

Roles and Responsibilities:

- Created and analysed CDR for accuracy and prepared records for release to clients using different python modules such as Pandas and NumPy.
- Developed and tested software applications to administer and support Call logic and automate different scenarios to test the m using Python.
- Designed and developed graphic user interfaces for subscriber data provisioning. Mass data provisioning using a text-oriented client.
- Planning and implementing the capacity expansion in order to ensure that the application's databases are scalable.
- Maintaining the standard of Protocols like INAP, MAP1, MAP2, Diameter in the designed application.
- Troubleshooting, bug reproducing and bug fixing.

Intern at ALTRAN (C++/Python Developer)

Jan 2018 - Jun 2018

Involved in development, maintenance and fixing bugs of the Alcatel Lucent's (Nokia) product Open Services Platform which allows rapid and flexible development and deployment of services coming under the purview of Intelligent Networks definition.

Roles and Responsibilities:

- Design, build, and maintain efficient and reliable python and C++ code.
- Worked on extensions and improvements of existing products.
- Plan, design, develop, manage, document, test, deploy, and support new and existing modules.
- Basic Database systems - indexes, filters.