Tanuj Soni

tanujsoni027@gmail.com | +91 94628 36377 | New Delhi, India | linkedin.com/in/tanuj-soni-372865126

Education

Indian Institute of Technology, Delhi *New Delhi, India*

B.Tech. in Engineering Physics *July 2014 - June 2018*

Technical Skills

Programming Languages: Python, Java, Scala, C++, JavaScript, TypeScript
Frameworks: Apache Spark, Spring Boot, Spring Cloud, Flask, Django
Cloud Services: AWS : All services, GCP : Compute Engine, Data Catalog, BigQuery, DataProc
Cloud Formation: Terraform, AWS : CDK, Code Commit, Serverless, GCP : Cloudbuild, Cloud Deploy
Databases: MongoDB, Oracle RDBMS, MySql, Postgres, AWS RedShift
Architectures: Micro-services Architecture, ETL: Event Driven Data Pipelines
Architectures: ETL: Data Driven Data Pipelines, Self Scaling and Configuring Glue Jobs

Professional Experience

Franconnect | MSA–Company API | FCSKY : FSC Scorecard *Noida, UP, India*

Lead Software Engineer | Technical Architect *January 2023 - Present*

- Led the architectural design of Data Pipelines for the FSC Scorecard Profit and Loss Dashboard.
- Spearheaded the design and implementation of a robust P&L pipeline utilizing Spark and AWS Glue, driving a 40% decrease in data processing time and enabling seamless generation of actionable insights on the analytics dashboard.
- Created an optimized Data Model for an Analytics dashboard, enabling the generation of end-user results within seconds. Reduced ETL time from 1 hour to 10 minutes.
- Deployed innovative design patterns and led the development of robust public APIs for the FCSKY application, resulting in a 25% reduction in API response time and improved system stability.
- Employed Spring Cloud and Micro-services Architecture to consolidate 50+ public APIs for both internal and external usage.

RxLogix Corporation | PV DataHub Serverless Pipeline *Noida, UP, India*

Senior Software Engineer *January 2022 - January 2023*

- Engineered an end-to-end ETL pipeline on AWS to handle high-volume Pharmacovigilance data, reducing data processing time by 60% and enhancing data accessibility for analysis and reporting purposes.
- Achieved near real-time processing of 1M daily load in 15 minutes using AWS Glue and custom S3 Streaming.
- Implemented data encryption with AWS KMS and Cognito to mask PII information during processing and data loading.
- Engineered a comprehensive technical architecture through stakeholder collaboration, optimizing scalability, reducing processing time by 75%, and enhancing data security within the data platform.
- Introduced an innovative Spark logic to convert complex nested JSON documents to multi relational Data-frames optimizing data processing speeds up by 70%.
- Streamlined a process to control and configure transformations via CSV and Excel files which reduced business initiation effort by 50%.
- Engineered an intelligent and interactive progress bar utilizing machine learning techniques, delivering instant updates on ETL completion and estimated time to finish; optimized user experience and streamlined workflow efficiency by 40%.
- Led the development and deployment of a developer dashboard, providing vital insights into failures, job executions, execution times, and areas for improvement. Boosted developer performance reduction in time spent on debugging and troubleshooting.

RxLogix Corporation | PV Signal–FAERS *Noida, UP, India*

Software Engineer *September 2021 - January 2022*

- Coded Drug Data Cleaning ML and NLP based Algorithm to clean FDA's public Drug Safety data with 95% accuracy.
- Hosted the ML server and exposed Python Django based Rest APIs to seamless integration of cleanup results within other PV Applications.
- Introduced Spark MLlib for distributed computation of cleanup logic and gained reduction in processing time by 90%.
- Reduced time from legacy cleanup taking 5 days to 15 minutes in new implementation.
- Implemented a Decision Support System within the PV Signal application, leveraging Python and C++. Utilized Bayesian Networks to provide end users with quick and accurate insights, enabling informed decision-making and optimizing operational efficiency.
- Re-engineered logics for EBGM and RR Scores calculation reducing calculation time from 20 minutes to 2-5 seconds using PL/SQL and SQL ETL pipelines.

RxLogix Corporation | PV Clinical *Noida, UP, India*

Software Engineer *September 2020 - September 2021*

- Automated an advanced data loading process for CDISC format Clinical Study Data into MongoDB, streamlining integration of CSV, JSON, and XML files; reduced manual data handling by 70%.
- Conceptualized a highly efficient Django-based Python Database micro-service to handle all data transactions and power machine learning dashboards, resulting in fast real-time analytics capabilities for the application.
- Optimized fault tolerance by leveraging a Celery-based load-balancer for the Django application, successfully conducting load testing with 5000+ active users without experiencing any issues or failures.

RxLogix Corporation | PV Signal *Noida, UP, India*

Associate Software Engineer *July 2018 - September 2020*

- Collaborated with Product team to enhance and maintain Oracle SQL and PL/SQL databases for PV-Signal and PV-Reports products, improving data integrity and reliability.
- Developed Java application controllers and DB querying functionalities to integrate PV-Database with application. Gained 20% performance boost by offloading complex calculations.
- Streamlined the integration of Oracle Argus Safety, ArisG, and FDA FAERS open data, along with other safety source data, into PV-Applications using ETL logic; enhanced reporting and signaling processes, leading to improved safety monitoring.
- Increased the query execution performance by 80% by implementing refined and scalable calculation logic for statistical calculations.
- Leveraged advanced statistical techniques, including EBGM, PRR, ROR, and CHI-Square, to revolutionize the calculation of PV safety statistical scores; enabled accurate identification of potential adverse events, leading to precise patient safety and regulatory compliance.
- Steered ad-hoc analysis of production data where necessary to develop solutions to reported incidents.
- Led cross-functional teams in resolving critical incidents, ensuring minimal impact on customer experience and achieving a high incident resolution rate within SLA.

Internship and Projects

FORENSIC ANALYSIS OF BLUE BALLPOINT PEN INKS : KMITL, BANGKOK THAILAND *May 2017 - July 2017*

Received a Letter of Recommendation from Prof. Bhanupol Klongratog, Physics Department, KMITL, for utilizing Image and Signal processing techniques to discriminate blue pen inks and detect forged documents using Thin Layer Chromatography (TLC). Developed a MATLAB program and GUI for easy identification of document similarity with an accuracy of 85%. *MATLAB*

SOLAR CELL DEVICE FABRICATION | IIT-D (B.Tech Major Project) (Guide : Dean Prof. B.R. Mehta) *Nanotechnology*

Achieved PLD growth and characterization of Copper Tin Sulphide (CTS) thin film photo-absorbers, providing a cost-effective alternative to Silicon-based solar cells. Fabricated and characterized CTS thin films with semiconductor-like band gaps, enhancing power conversion efficiency. Contributed to the production and synthesis of economically favorable thin films, advancing solar cell applications and device manufacturing.

IMPACT OF PACKAGING OF PRODUCTS IN INDIAN FMCG MARKET | IIT-D (B.TECH. DESIGN PROJECT) *Python, Pandas*

Studied 20+ elemental aspects of packaging; designed survey based on hypothesis obtained from studies across globe. Analyzed data obtained from 250+ responses; concluded importance of quality contents, nullifying hypothesis.

NUCLEAR REACTOR SIMULATION | IIT-D (B.TECH. COURSE PROJECT) *Python, MATLAB*

MATLAB and Python based program which simulates the shielding effect of slab intended to absorb or reflect most of the radiation from nucleus source. Monte Carlo simulations applied to probabilistic systems

Awards and PORs

RxLogix	**Received Spot Award 5 times.**, Recognized for outstanding contributions to development and RnD teams during the tenure.
Karakoram House	**Honoured as Mess Secratory POR	Karakoram Hostel	IIT-D**, Streamlined menu planning and introduced new dishes based on resident preferences, leading to a 25% increase in daily meal consumption and 40% reduced wastage.