

MANOJ BHATT

Senior Software Engineer | *Data Engineering & Intelligence Platforms*

Gurgaon, Haryana | +91 8859859556 | mbhatt2018@gmail.com | github.com/kingmanoj1 | linkedin.com/in/manoj-bhatt-504479153

Professional Summary

Results-driven Python Senior Developer with 6.5+ years of experience building high-throughput data pipelines, web scraping systems, and scalable backend infrastructure. Skilled in Python, Django, PostgreSQL, Scrapy, and Playwright, with expertise in ETL automation, large-scale data processing, API development, and performance optimization. Experienced with AWS and production deployments, with a strong focus on reliability, efficiency, and clean architecture. Passionate about solving complex engineering challenges, mentoring developers, and delivering high-impact solutions.

Experience

Contify

Software Engineer III (Oct 2025 - Present)

Software Engineer II (Aug 2023 - Oct 2025)

Software Developer (Dec 2019 - Aug 2023)

Stackmindz

Software Developer Intern Jun 2019 - Aug 2019

Key Projects

High-Volume Data Pipeline & PostgreSQL Optimization

- Architected a time-based partitioned PostgreSQL schema for 500GB+ of data, enabling efficient query execution at 500 QPS through advanced SQL and resource tuning.
- Reduced average query response time by over 60% post-partitioning by aligning partition keys with the most frequent access patterns and eliminating full-table scans.
- Implemented database indexing strategies and connection pooling to sustain throughput under concurrent load from multiple ingestion and read services.

Enterprise Web Scraping & Data Extraction Platform

- Built a multi-source data extraction system using Scrapy and Airflow, implementing proxy rotation and resilience patterns to handle complex, protected data sources at scale.
- Designed a modular spider architecture allowing rapid onboarding of new data sources with minimal code changes, significantly increasing team throughput for coverage expansion.

AI-Powered PDF Summarization Microservice

- Developed an asynchronous microservice using Django REST Framework and the OpenAI API, automating intelligent document extraction and summarization.
- Implemented async task queuing to decouple PDF ingestion from summarization, enabling the service to handle burst traffic without blocking upstream workflows.
- Designed a structured output schema for summaries, enabling downstream systems to index, search, and surface document insights through the intelligence platform.

AWS Cloud Infrastructure & Automation

- Collaborated with DevOps engineers to manage cloud infrastructure using AWS Console.
- Created AMIs, launched and configured EC2 instances, and optimized system performance by vertically scaling disk sizes.
- Monitored instance performance and health metrics via AWS Console, while deploying serverless Lambda functions to automate resource tracking, ensuring operational stability and driving cost optimization through the identification of underutilized infrastructure.

LLM-Driven Data Tagging & Post-Processing

- Designed an automated pipeline for post-processing extracted data, leveraging LLMs to assign relevant categorization tags and significantly improve search relevance and data discoverability.
- Optimized prompt engineering and implemented caching strategies to minimize API latency and operational costs while maintaining high classification precision.
- Built an evaluation framework to measure tagging accuracy against a labelled validation set, enabling continuous prompt iteration and quality assurance at scale.
- Integrated the tagging pipeline with Elasticsearch to surface categorized content through faceted search, directly improving end-user discoverability of intelligence data.

Technical Skills

- **Languages:** Python, SQL
- **Backend & Frameworks:** Django, Django REST Framework (DRF), Asynchronous Programming (AsyncIO)
- **Databases & Search:** PostgreSQL (Query Optimization, Partitioning), Elasticsearch
- **Infrastructure & Cloud:** AWS (EC2, S3, CloudWatch), Infrastructure Automation
- **Data Engineering:** ETL Pipelines, Apache Airflow, Web Scraping (Scrapy, Playwright)
- **Tools & Methodologies:** REST APIs, Microservices, Prompt Engineering (LLMs), Git

Education

- MCA, Amrapali Institute of Technology and Sciences, Haldwani (2017-2019).
- BCA, VCMT College of Management and Technology Haldwani (2014-2017).
- Intermediate, Govt Inter College Lohali 2014.
- High School, Govt Inter College Lohali 2012.