

# GOVARDHAN KHADAKKAR

(573) 639-5628 | khadakkargovardhan@gmail.com | [linkedin.com/in/govardhan-khadakkar](https://www.linkedin.com/in/govardhan-khadakkar)

## PROFESSIONAL SUMMARY

Experienced Software Engineer with 5+ years of experience building scalable data systems and applications. Expertise in Python, Java, SQL, and big data tools like Apache Spark and Kafka. Proven track record in optimizing ETL pipelines and cloud platforms (AWS, Azure), delivering 30% faster data processing and 25% better query performance. Collaborative problem-solver driving data-driven solutions for business impact.

## SKILLS

**Programming Languages:** Python, Java, Scala, SQL

**Data Engineering:** ETL Pipelines, Data Warehousing (Snowflake), Big Data Processing (Apache Spark, PySpark), Data Modeling, Real-time Data Processing (Kafka, RabbitMQ)

**Frameworks:** Spring Boot, Fast API, Flask, Event-driven Architecture, RESTful Services, SOAP

**Cloud & Infrastructure:** AWS (EC2, S3, RDS), Azure (Blob Storage), Databricks

**Databases:** MySQL, PostgreSQL, MongoDB, Snowflake, Hive, Oracle

**Web Development:** HTML, JavaScript, CSS, XML, JSON, PHP

**DevOps & CI/CD:** Git, GitHub, GitLab, Jenkins, Maven, Gradle, CI/CD Pipelines, Docker, Kubernetes, Linux

**Monitoring & Analytics:** Grafana, Prometheus, Qlik Cloud, Power BI

**Libraries & Tools:** JUnit, Mockito, PyTest, Selenium, Pandas, NumPy, Lombok

**Methodologies:** Agile, Scrum, Waterfall, JIRA, Confluence

## WORK EXPERIENCE

**Software Engineer** | University of Missouri-Columbia, USA

Jan 2024 – Present

- Engineered a cloud-based data platform utilizing microservices with Docker containerization and AWS infrastructure for plant-protein-phosphorylation research data analysis.
- Optimized MySQL and MongoDB, improving query performance by 20% through indexing and normalization.
- Developed and documented API gateways with over 25 RESTful endpoints using Python FastAPI and Flask, for seamless interaction between the frontend and backend components.
- Utilized GitHub for version control, code maintenance, and collaborative code reviews, ensuring code quality.
- Configured infrastructure on EC2 instances, including domain configuration, SSL certification, and server setup, ensuring secure data access with 100% compliance to security standards.
- Implemented automated deployment workflows with Docker and AWS, reducing release cycle time by 40%.

**Data Science Intern (Business Analytics)** | Siemens Healthineers, USA

May 2024 – Aug 2024

- Enhanced workforce planning models using Python, SQL, and statistics, improving prediction accuracy by 25%.
- Implemented data warehousing solutions with Snowflake, optimizing query performance and data accessibility.
- Designed over 20 interactive dashboards in Qlik Cloud, eliminating 11 redundant variables in regression model, and delivering insights for strategic decisions.
- Developed automation scripts with Python to streamline data preprocessing and transformation workflows.
- Leveraged Databricks to streamline data workflows and enhance collaboration among data teams.
- Presented technical solutions to stakeholders, influencing data-driven decision-making.

**Senior Software Engineer** | Tech Mahindra, India

Feb 2023 – Jul 2023

- Architected and deployed scalable microservices for Java and Spring Boot based applications, increasing system scalability and ensuring high availability across distributed architecture.
- Integrated Kafka message brokers for async communication between services, ensuring system resilience.
- Integrated AWS services, including EC2, S3, and CloudWatch, to enable scalability and real-time monitoring.
- Increased deployment efficiency by identifying and resolving workflow bottlenecks, using CI/CD pipeline optimizations to streamline release processes.
- Optimized PostgreSQL and SQL Server databases, improving data access performance by 25%.
- Led a team of 4 engineers, mentoring them in backend development and cloud deployment.

**Data Engineer** | LTIMINDTREE, India

May 2020 – Feb 2023

- Developed data-driven applications for distributed systems using Java, Apache Spark, Spring Boot, RESTful web services, and microservice architecture.
- Built and maintained ETL pipelines using Apache Spark and PySpark for processing various data formats at scale.
- Optimized ETL pipelines, through efficient data partitioning and caching strategies, maximizing data processing speed within distributed systems, while increasing throughput and data processing speed by 30%.

- Achieved a 28% optimization of the Hive database operations and implemented Kerberos-based authentication, enhancing product security and performance.
- Enhanced system stability by reducing memory consumption by 15% in high-traffic applications through advanced multi-threading and optimized memory allocation strategies.
- Maintained RabbitMQ for asynchronous communication between services, ensuring system resilience.
- Implemented comprehensive CI/CD pipeline with Jenkins for automated build, test, and deployment processes
- Established real-time monitoring with Prometheus and Grafana dashboards, enabling proactive system maintenance and optimization
- Achieved 89% code coverage by developing extensive JUnit test cases, enhancing code quality and reliability across application modules.

**Assistant System Engineer** | Tata Consultancy Services, India

Jun 2019 – Mar 2020

- Configured and managed comprehensive security platforms (Firewall, VPN, Proxy) to protect 3 client's networks, enhancing system security and reducing potential vulnerabilities by proactively identifying threats.
- Coordinated daily administrative tasks and progress reports, improving cross-team communication efficiency.
- Provided technical support and troubleshooting for 7+ security incidents daily to mitigate risks and minimize disruptions, with ServiceNow for tracking security incidents and resolution progress.
- Created and maintained regular reports on security incidents and platforms to analyze risk, identify trends, and inform stakeholders.

## PROJECTS

### Forecasting Twitter Engagement: Insights and Prediction

- Analyzed 80,000+ tweets with PySpark to predict engagement ( $R^2 = 0.96$ , MAE = 121.53)
- conducted sentiment analysis, and mapped geographic trends and sentiment (e.g., vaccine perceptions), providing insights for public health communication.

### Price Scraping and Tracking System

- Developed a Python-based web scraping tool using Selenium to track prices of multiple products across different states and stores.
- Designed automated data collection pipelines, storing results in a MySQL database for trend analysis.
- Implemented error handling and monitoring mechanisms, ensuring consistent and reliable data extraction.

## EDUCATION

**Master of Science in Computer Science**

Expected May 2025

University of Missouri, Columbia, USA

**Bachelor of Engineering in Information Technology**

June 2019

Sant Gadge Baba Amravati University, India

## ACTIVITIES & LEADERSHIP

**Treasurer**, University Fighting World Hunger (UFWH)

Aug 2024 – Present

**Volunteer**, Electrical Engineering and Computer Science Graduate Student Association

Oct 2023 – Present