

# Omkar Yadav

San Jose, CA • +1 (669) 288-4824 • [omkar.sy10@gmail.com](mailto:omkar.sy10@gmail.com) • [yadomkar.com](http://yadomkar.com) • [github.com/yadomkar](https://github.com/yadomkar)

## EDUCATION

---

### M.S. in Computer Science - San Jose State University, CA

Aug. 2023 – May 2025

Relevant Coursework: Advanced Algorithms, Computer Networks

### B.E. in Computer Engineering - University of Mumbai

Aug. 2016 – Oct. 2020

Relevant Coursework: Operating Systems, Database Systems, Data Structures

## EXPERIENCE

---

### CallHub

Jan. 2023 – Jun. 2023

#### Software Engineer – Python, Django, DRF, Celery, Redis, PostgreSQL.

- Led a 3-person team in modernizing a **400k LOC** Django project, transitioning from Python 2 to Python 3.
- Upgraded and refactored **200+** dependencies, including Django, Celery, RedisPy, and Django REST framework.
- Investigated and resolved **30%** of backlogged bugs, showcasing ability to modify large-scale software without disruption.

### Larsen & Toubro Infotech

Aug. 2021 – May 2022

#### Senior Data Engineer – Scala, Spark, Hadoop, Hive, SQL, Impala, Autosys, HBase.

- Served as a contractor at **Citigroup, Inc.**, contributing as a core developer to ICG's next-generation Data Fabric.
- Spearheaded a solution to tackle recurrent data loss across clusters, reconciling multi-table discrepancies of **100M+** records. Achieved an **80%** reduction in turnaround time.
- Automated cross-cluster data validation and reporting processes, saving **5 hours** each week.
- Crafted diverse Scala scripts, streamlining multiple internal processes.

### Larsen & Toubro Infotech

Aug. 2020 – Jul. 2021

#### GET – Data Engineer – Python, PySpark, Scala, Hadoop, Sqoop.

- Steered UAT processes, achieving a **50%** reduction in client onboarding time.
- Identified and resolved instances of critical data loss across **10+ tables**, safeguarding business data integrity.
- Acted as a liaison for QA on code matters, significantly improving resolution speed.
- Managed Autosys jobs and provided **24/7 support** for key releases, enhancing system reliability.

## TECHNICAL SKILLS

---

**Programming Languages:** Python, C++, Scala, JavaScript, HTML, CSS.

**Technologies and Frameworks:** Django, Celery, Redis, Pandas, React, NodeJS.

**Data and Cloud:** Apache Spark, Hadoop, Hive, HBase, PostgreSQL, Snowflake, AWS: S3, RDS, EC2.

**Developer Tools:** Git, Autosys, UrbanCode Deploy, TeamCity, Jira, GitHub, Postman, PyCharm, VS Code, Eclipse.

**Specialized Skills:** Linux, Distributed Systems, REST APIs, Full-Stack, Data Engineering, Object-Oriented Programming.

## PROJECTS

---

### Application of Ontology-Driven NLP for the Analysis of Text – Python, NetworkX, nltk

Aug. 2019 – Oct. 2020

- Devised a method to assess exam paper fairness, aiding teachers in question formulation.

### Cornerstone Cupcakes – Django, Python, JavaScript, Heroku, AWS RDS, AWS S3

Jun. 2020 – Jul. 2020

- Developed an e-commerce site for a cupcake bake sale, raising **INR 15,000+** for animal welfare during COVID-19.

## ACTIVITIES AND LEADERSHIP HIGHLIGHTS

---

- Active Open Source Contributor with contributions to projects including **NetworkX**, **Pandas**, and others.
- Active Competitive Programmer with CodeChef Ranks: October Long (Global Rank: **231**), Starters 59 (Global Rank: **429**), March Long (Global Rank: **502**); also ranked in the **top 2%** of 5M+ LeetCode users.
- Mentored an intern at **CallHub** and two junior developers at **Citigroup, Inc.**, fostering valuable contributions and maintaining code quality through rigorous reviews.