

# PAVEL YAKOVLEV

yvpavel@gmail.com  
github.com/paulyakovlev

## EXPERIENCE

---

### General Motors

February 2021 - Present

*Big Data Platform Software Engineer*

*Austin, TX*

- Contributed to migration from Cloudera to GM's custom Hadoop platform (GMDP), generating \$1.2M annual cost savings.
- Led migration from Cloudera Data Analytics Studio to Tez UI, eliminating closed-source dependencies and reducing licensing costs.
- Architected and implemented Oozie ETL pipeline for Ranger audit logs to Parquet format, enabling PowerBI analytics across production clusters.
- Contributed to Apache Ambari by implementing Livy as service in the Bigtop stack, enabling proper integration with Ambari deployment, configuration management, and monitoring. (PR #3790)
- Enhanced Docker-based Hadoop development environment with local RPM repos and auto-build + deployment of Ambari, significantly improving setup time for developers.

### Lawrence Livermore National Lab

June 2019 - September 2019

*Software Engineering Intern*

*Livermore, CA*

- Improved automated testing and software integration framework
- Contributed to JUnit test suite and provided technical documentation
- Prototyped new test automation processes

### remote.it

May 2018 - June 2019

*Software Engineering Intern*

*Palo Alto, CA*

- Prototyped user-end web portals for device monitoring and management
- Wrote a library of user-end Python and Shell scripts to manage IoT edge devices
- Provided technical documentation for internal and end-user software tools

## EDUCATION

---

### University of California, Santa Cruz

Class of 2020

*Bachelor's Degree in Computer Science*

## SKILLS

---

**Languages:** C++, Java, Python, Javascript, HTML, CSS

**Frameworks and Tools:** ReactJS, NodeJS, Flask, SQL, git, Docker, REST, Agile