PAVEL YAKOVLEV

yvpavel@gmail.com
github.com/paulyakovlev

EXPERIENCE

General Motors   February 2021 - Present
Big Data Platform Software Engineer   *Austin, TX*

· Contributed to development and deployment of GMDP (General Motors Data Platform), GM's custom open-source Hadoop distribution, successfully migrating from Cloudera Hadoop.

· Led migration from Cloudera Data Analytics Studio to Tez UI, eliminating dependencies on closed-source software while ensuring GMDP 2.0 compatibility and preparing open-source contributions.

· Architected and implemented Oozie-orchestrated ETL pipeline for Ranger audit logs, automating conversion to Parquet format and enabling critical PowerBI analytics across multiple production clusters.

· Contributed to Apache Ambari by implementing Livy as an individual service in the Bigtop Stack, enabling proper integration with deployment,configuration management, and monitoring (PR #3790)

· Enhanced Docker containerized Hadoop development environment by implementing local RPM repositories and enabled Ambari deployment workflows, significantly reducing setup time for developers.

· Migrated Greenplum SQL macros to Databricks with cross-functional teams, preserving functionality while improving data processing workflow efficiency

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 processsess

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

PROJECTS

Santa Cruz Stranding Map   github.com/lmlstrandingnetwork/lml-stranding-map
ReactJS, ExpressJS, Firebase

· Geospatial data analysis tool for the UCSC Long Marine Lab.

EDUCATION

University of California, Santa Cruz   Graduated June 2020
Bachelor's in Computer Science

SKILLS

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

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