|  |  |
| --- | --- |
| **GESER DUGAROV, Ph.D. | Java Developer, Big Data Engineer** | |
| email:  profiles: | [geserdugarov@gmail.com](mailto:geserdugarov@gmail.com)  [LinkedIn](https://www.linkedin.com/in/geserdugarov/), [GitHub](https://github.com/geserdugarov) |

SUMMARY

Software Engineer developing core functionality in a Data Lakehouse platform to extract value from PB-scale data. Also contributing to open source, [Apache Hudi project](https://github.com/apache/hudi/pulls?q=is%3Apr+author%3Ageserdugarov), with a focus on stream processing performance and usability of solution.

Wide experience in research and data analysis, PhD. Made a transition from fundamental research in Oil & Gas to research and development in IT.

Passionate about Big Data and Distributed Systems. Personal mission: "Living a balanced life. Helping professionals to work smarter, not harder by creating automatic systems for their routine."

TECHNICAL SKILLS

***Java, Python, Maven, PostgreSQL, Docker, Hadoop Ecosystem***

WORK EXPERIENCE

|  |  |
| --- | --- |
| May 23 - Current  (**~**2 yrs) | **Java Developer / Big Data Engineer**,  [**Huawei Cloud**](https://www.huaweicloud.com/intl/en-us/product/mrs.html) |

• Development of core functionality of a Data Lakehouse platform for Big Data processing on enterprise-level scalable clusters.

|  |  |
| --- | --- |
| Jan 24 - Current  (1+ yr) | **Apache Hudi Contributor**,  **The Apache Software Foundation** |

[Apache Hudi](https://hudi.apache.org/) is a Data Lakehouse platform, which brought database functionality into Data Lake, and provided incremental processing for low latency analytics.

• Optimized serialization/deserialization of data stream records in Flink stream writing, which accelerated processing speed by 30%, and reduced amount of allocated memory twice ([design doc](https://github.com/apache/hudi/pull/12697), [main changes](https://github.com/apache/hudi/pull/12796), [umbrella ticket](https://issues.apache.org/jira/browse/HUDI-8920)). Released in Apache Hudi 1.1.

• 40+ [merged PRs](https://github.com/apache/hudi/pulls?q=is%3Apr+author%3Ageserdugarov+is%3Amerged).

|  |  |
| --- | --- |
| Feb 22 - May 23 (1+ yr) | **Python Developer**,  **Digital Research** (computer vision startup) |

• Created an event-based architecture for a system for trucks monitoring. Developed server-side image processing (~20K images/day). In production, the system reduced fleet idle time by 12%.

• Created a customer-facing web UI with reports and data visualization. Also created an internal web UI for system monitoring.

EDUCATION

|  |
| --- |
| **PhD, Geophysics,** Trofimuk Institute of Petroleum Geology and Geophysics SB RAS |
| **MSc, Computational and Applied Mathematics,** Novosibirsk State University |

[Recent version of CV](https://geserdugarov.github.io/file/CV_Geser_Dugarov.pdf)