



Алиев Айдар

Мужчина, 37 лет, родился 12 мая 1988

+7 (926) 1783466 — предпочитаемый способ связи • tg: AjdarA1iev, whatsapp: +7 926 178-34-66

aidaraliev@gmail.com

Другой сайт: <https://github.com/aydaraliev/CV>

Проживает: Москва

Гражданство: Россия, есть разрешение на работу: Россия

Готов к переезду, готов к командировкам

Желаемая должность и зарплата

Data Engineer

Специализации:

- BI-аналитик, аналитик данных
- Программист, разработчик

Занятость: полная занятость, проектная работа

График работы: полный день, гибкий график, удаленная работа

Желательное время в пути до работы: не более часа

Опыт работы — 14 лет

Ноябрь 2022 —
настоящее время
2 года 9 месяцев

Иннотех, Группа компаний

Москва

Информационные технологии, системная интеграция, интернет

- Разработка программного обеспечения
- Системная интеграция, автоматизации технологических и бизнес-процессов предприятия, ИТ-консалтинг

Разработчик ETL-процессов (Data Engineer)

Ведущий российский интегратор IT-решений для корпоративного сектора

Разработчик ETL-процессов (Data Engineer)

Зона ответственности:

- * Автоматизация ETL-процессов для проекта ВТБ на стеке Airflow + PostgreSQL.
- * Переписывание и доработка Python/SQL модулей внутренних фреймворков.
- * Разработка DAG'ов для очистки и обслуживания таблиц.
- * Использование git и CI/CD на импортозамещённых решениях.

Ключевые достижения:

- ✓ Автоматизировал обслуживание 12+ таблиц, устранив необходимость ручного запуска скриптов, что сократило ~40 часов рутинной работы в месяц.
- ✓ Разрабатывал DAG'и для обработки таблиц объёмом до 500 млн строк / 100+ GB в рамках DWH ВТБ.
- ✓ Внедрил улучшенное логирование, повысив прозрачность и удобство поддержки процессов. Модуль стал обязательным для использования всей командой разработчиков.
- ✓ Написал модуль который используется в 100+ DAGax Airflow нашей команды. Отрефакторил что было, привёл к читаемому виду, применил подход DRY (вынес общие функции и константы

в отдельный модуль)

Стек: Python, SQL, Airflow, PostgreSQL, Spark, git

Сентябрь 2020 —
настоящее время
4 года 11
месяцев

Институт Молекулярной Генетики, НИЦ Курчатовский институт

Москва, prski.ru/

Образовательные учреждения

- Научно-исследовательская, научная, академическая деятельность

Аналитик данных/Биоинформатик

Ведущий российский научный центр в области молекулярной генетики, осуществляющий фундаментальные и прикладные исследования данных о мутациях и биомедицинских образцах.

Аналитик данных / Биоинформатик

Зона ответственности:

- * Обработка больших массивов биомедицинских данных (50–60 млн строк на 3–6 тыс. образцов) в Linux-среде.
- * Проведение прунинга, импутации и анализа данных о мутациях с использованием Python и специализированных утилит.
- * Интеграция C/C++ инструментов (tabix, vcftools, bcftools etc.) в автоматизированные пайплайны на Python и R.
- * Автоматизация процессов выявления мутаций, связанных с риском ОНМК, и подготовка аналитических отчетов.

Ключевые достижения:

- ✓ Проектировал и оптимизировал пайплайны для анализа генетических данных объемом 50+ млн записей, используя Python и bash-скрипты в распределённой среде Linux.
- ✓ Интегрировал C++ утилиты в пайплайны Python и R, повысив производительность вычислений с 36 часов до 3 часов при обработке VCF-файлов объемом до 150 GB.
- ✓ Автоматизировал процесс выявления и анализа ключевых мутаций.
- ✓ Разработал систему автоматизированных отчетов, которые позволяют выявить ошибки первого рода при поиске ассоциаций SNP-полиморфизмов с ОНМК.
- ✓ Обеспечил поддержку и масштабируемость процессов обработки данных для новых научных задач.

Стек: Python, R, Bash, Linux, tabix, CLI-программы (vcftools, bcftools, plink2 etc.)

Август 2021 —
Ноябрь 2022
1 год 4 месяца

Департамент здравоохранения Москвы

Москва, mosgorzdrav.ru/

Государственные организации

- Государственные организации

Аналитик данных

Крупнейшая медицинская организация, управляющая цифровыми сервисами и данными здравоохранения для мегаполиса.

Аналитик данных

Зона ответственности:

- * Автоматизация загрузки и обработки данных из ЕМИАС и S3 в локальное хранилище DWH на ClickHouse.
- * Разработка управленческих отчётов и аналитических дашбордов для подразделений департамента.
- * Расчет метрик эффективности для медицинских учреждений.
- * Автоматизация своевременного справочников (например адрес -> поликлиника).
- * Выполнение ad-hoc аналитики и подготовка данных по индивидуальным запросам.
- * Помощь с автоматизацией рутинных задач оперативному отделу.

Ключевые достижения:

- ✓ Автоматизировал загрузку и обработку данных (плоские файлы до 20 Гб.) в ClickHouse, обеспечив стабильность и скорость обновлений.
- ✓ Разработал отчёты, сократив подготовку управленческих данных с 2 дней до 3 часов.
- ✓ Автоматизировал интеграцию и ведение справочников адресов, устранив необходимость ручной работы 2 сотрудников.
- ✓ Разработал комплексный пайплайн для формирования выгрузок, ускорив процесс получения аналитики для оперативных решений.

Стек: Python, SQL, ClickHouse, Airflow, pandas, openpyxl, requests

Август 2011 —
Сентябрь 2020
9 лет 2 месяца

Работа в лабораториях, преподавание

Москва, massey.ac.nz

Образовательные учреждения

- Научно-исследовательская, научная, академическая деятельность

Исследователь / Преподаватель Python / Разработчик пайплайнов

Зона ответственности:

- * Моделирование и анализ биологических данных с использованием Python, Java, R и Bash.
- * Разработка, оптимизация и поддержка аналитических пайплайнов для исследований и учебных задач.
- * Преподавание основ Python и анализа данных студентам.
- * Интеграция CLI-утилит (C/C++/Fortran/R) в существующие пайплайны.

Ключевые достижения:

- ✓ Смоделировал изменения в человеческом геноме при расселении людей по островам тихого океана от 50 до 20 тысяч лет назад.
- ✓ Обучил более 50 студентов основам Python, сформировав фундаментальные навыки у начинающих специалистов.
- ✓ Автоматизировал обработку сложных данных с помощью CLI утилит и скриптов, снизив трудозатраты исследовательских команд.
- ✓ Участвовал в международных коллаборациях, применяя современные методы статистики и визуализации данных.

Стек: Python, Bash, R, Java, Linux, C/C++ tools

Образование

Магистр

2014	Massey University, New Zealand Institute of Fundamental Sciences, Master in Computational Biology
2011	Московский государственный университет им. М.В. Ломоносова, Москва Факультет вычислительной математики и кибернетики, Разработчик (доп. квалификация)
2011	Московский государственный университет им. М.В. Ломоносова, Москва Фундаментальной Медицины, Лечебное дело (врач)

Электронные сертификаты

2024	SQL для разработчиков, Яндекс Практикум
2023	Специалист по Data Science, Яндекс Практикум
2020	Python and Flask Bootcamp: Create Websites using Flask!
2018	CalcAPL1x: Calculus Applied! Calculus 1 Calculus 2 Calculus 3 DS101x: Data Science Online Readiness Assessment Data Science Math Skills Nginx Fundamentals: High Performance Servers from Scratch Object Oriented Programming in Java REST APIs with Flask and Python Statistics with R UT.5.05x: LAFF: Linear Algebra - Foundations to Frontiers

Навыки

Знание языков	Русский — Родной Английский — C1 — Продвинутый Голландский — A2 — Элементарный Кыргызский — A2 — Элементарный
---------------	--

Навыки	Python R Linux Java pandas Airflow Bash SQL Apache Airflow Matplotlib Jupyter Notebook REST API Seaborn PyCharm Clickhouse Scikit-learn Flask Git PostgreSQL Docker
--------	---

Дополнительная информация

Обо мне	<ul style="list-style-type: none">- Учился программировать на ВМиК МГУ параллельно с дипломом врача (ФФМ МГУ).- После выпуска занимался сначала биоинформатикой, затем полностью перешёл в ETL и Data Engineering.- Уверенно знаю Python и SQL и стремлюсь углубить свои знания.- Имею 3 года промышленного опыта разработки ETL процессов.- Сторонник непрерывного обучения. Сейчас прохожу курс по инженерии данных на Яндекс Практикум.- Интересуюсь достижениями искусственного интеллекта (и Data Science как отрасли), посещал лекции и семинары по теории игр вместе с магистрами в университете Амстердама в течение семестра.
---------	---

- Всю карьеру так или иначе работал со структурированными (табличными) данными.
- Преподавал Python и основы анализа данных студентам.
- Уверенно работаю в Linux, люблю автоматизацию процессов.
- Открыт к предложениям Data Engineer, ML Engineer и Data Analyst.