

Кирилл Дикалин

dikalinkirill@gmail.com ♦ github.com/kirilldikalin ♦ [@kirilldikalin](https://twitter.com/kirilldikalin)

Имею опыт проектирования аналитических систем хранения и обработки больших данных, знаю принципы работы БД и методологии моделирования.

Уверенно пишу код на Python и SQL

Понимаю как работать со стеком продуктов на платформе Hadoop

Образование

МГТУ им. Баумана

2018-2023

Кафедра информационной аналитики и политических технологий

Специальность: Прикладная информатика

Опыт работы

ПАО "СберБанк"

Апрель 2022-по н.в.

Middle Data Engineer блока *SberData*

- Формирование файлов конфигураций/инвентори. Создание и разворачивание дистрибутивов и плагинов для загрузки данных реплик на различные контуры (ИФТ, ПСИ, ПРОМ) с последующими запусками и выполнением тестирования работоспособности. Проведение совместного интеграционного тестирования с Источником в части загрузок реплик. Разбор и маршрутизация тикетов и дефектов, возникающих на тракте тиражирования данных. Ведение отдельных фичей, от оформления релизных артефактов до проведения ПСИ и непосредственного внедрения в ПРОМ. Написание запросов к данным в kafka (scala), hbase, hive.

ООО "Яндекс"

Июнь 2022-по н.в.

Ревьюер курса Data Engineer в Яндекс.Практикуме

- Проверка написанного студентом проекта и дача развивающей обратной связи.

ПАО "Трансконтейнер"

Ноябрь 2021-Март 2022

Data Engineer

- Проектирование и создание корпоративного хранилища данных компании (DWH). Разработка конвейеров обработки данных. Осуществлена интеграция разработок в существующие решения, а также интеграция с системами источниками данных. Разработка и оркестрация ETL и ELT в airflow и небольших сервисов, работающих в docker. Например, настройка оповещений о выполнении дагов в телеграмм, и автоматическое обновление внешних словарей в ClickHouse. Оптимизированы существующие sql скрипты, настроены CI/CD. Отладка, тестирование и исправление ошибок в своём и чужом коде. Работа велась в облачной платформе управления данными (Yandex.Cloud).

Центр хранения и анализа больших данных МГУ им. Ломоносова

Октябрь 2020-Ноябрь 2021

Data Analyst

- Решение проблем проектов, в которых необходимо было использовать способы работы с данными для получения результатов. Использование SQL (включая оконные и триггерные функции, хранимые процедуры, шардирование и партиционирование данных в БД) и python для работы с документами (Pandas, NumPy). Участие в построении корпоративного хранилища данных (Greenplum).

Лаборатория информатизации и социальной аналитики МГТУ им. Баумана

Январь 2021-Февраль 2022

Аналитик

- Общение с разработчиками и людьми, работающими над электронным университетом, определение необходимых данных и вариантов их анализа. Получена и проанализирована выгрузка данных ЭУ, обнаружены аномалии в данных, выявлены проблемы их достоверности. На основании полученных данных сформированы рекомендации для принятия управленческих решений и исправлению проблемы достоверности данных.

Технические навыки

Языки программирования

Python, Scala, SQL. little: Haskell, Java

Big data

Kafka, HBase, Hive, ClickHouse, PostgreSQL, AirFlow

Spark, MongoDB, Greenplum, Cassandra, RabbitMQ

Навыки в ...

Linux, Docker, git, bash, LaTeX