

АБ

DATA PLATFORMS

ПРОЕКТНОЕ РЕЗЮМЕ

Алексей Булыгин

АРХИТЕКТОР / ТЕХНИЧЕСКИЙ РУКОВОДИТЕЛЬ

Проектирую платформы данных
и руковожу их разработкой

КОНТАКТЫ

Telegram
@a_buligin

Почта
aleksej_bulygin@internet.ru

КАРЬЕРНЫЙ ТРЕК

С 2025

Архитектор платформ данных
и технический руководитель
проектов

2021–2025

Лид команды → Head of Data
→ Head of Data Platform

2017–2021

Аналитика, Data Science и ML-
инженерия

РУКОВОДСТВО

Смешанные команды
Технические решения
Ресурсы разработки
Оценка бюджета разработки
Развитие сотрудников
Работа с заказчиками

ПРОФЕССИОНАЛЬНЫЙ ПРОФИЛЬ

Совмещаю архитектуру DWH, Data Lake и Lakehouse с
руководством инженерами данных, аналитиками и специалистами
по ML. Основной опыт — on-premise, также работаю с облачной
инфраструктурой.

В инхаусе вырос от лида до Head of Data и Head of Data Platform;
отвечал за найм, оценку и продвижение сотрудников. С 2025 года
работаю в консалтинге, веду несколько проектов параллельно и
развиваю сотрудников как ментор.

ЭКСПЕРТИЗА

Архитектура и модели

DWH, Data Lake, Lakehouse; batch, NRT, real-time; Lambda, Kappa,
медальонная архитектура. Kimball, Inmon, Data Vault 2.0, Anchor
Modeling; ER, UML, C4.

Хранение и вычисления

PostgreSQL, MySQL, Greenplum, ClickHouse, BigQuery, Redshift,
MongoDB. HDFS, Ozone, MinIO, S3, Cloud Storage; Hadoop, Spark,
Spark Streaming, Trino, Hive, YARN. Iceberg; Polaris, Hive Metastore.

Интеграции и управление данными

Python, SQL, Airflow, dbt, NiFi, Dataflow, Kafka, Pub/Sub. Data
Governance, Data Quality, MDM, метаданные; Great Expectations,
DataHub.

Надёжность и инфраструктура

HA, disaster recovery, резервное копирование и мониторинг.
Kubernetes, Docker; IaC и GitOps: Ansible, Helm, Argo CD; GitLab
CI/CD. AWS, GCP, Yandex Cloud, VK Cloud.

Аналитика и ML

Tableau, Superset, DataLens, Looker Studio, Redash. NLP, скоринг,
классификация, ранжирование, аномалии; scikit-learn, PyTorch,
TensorFlow, Spark MLlib, MLflow.

01

2025 — СЕЙЧАС

ПРОЕКТНЫЙ ОПЫТ

01 / DWH

PostgreSQL, Patroni, etcd,
HAProxy
ClickHouse
Kafka, NiFi
Ansible, GitLab CI/CD

02 / DWH

Greenplum
Airflow
GitLab CI/CD

03 / РИТЕЙЛ

Kimball, Inmon
Greenplum, ClickHouse
Airflow, VK Cloud
GitLab CI/CD

04 / LAKEHOUSE

Хранение — Ozone
Таблицы — Iceberg
Каталог — Polaris
Обработка — Spark, Spark
Streaming
SQL — Trino
Витрины — ClickHouse

Архитектура платформ

01 Корпоративный DWH на открытом стеке

Апрель 2026 — настоящее время · Фармацевтическая компания

Архитектор DWH

Проектирую платформу с отказоустойчивым PostgreSQL и ClickHouse для аналитических нагрузок. Внедрение включает NRT-доставку данных и автоматизацию инфраструктуры через IaC.

02 Целевая архитектура корпоративного DWH

Ноябрь 2025 — настоящее время · Энергетическая компания

Архитектор DWH

Для хранилища объёмом около 5 ТБ разработал и описал целевую архитектуру: сеть, хранение, загрузку и интеграцию с системами-источниками. Представил и защитил архитектурное решение перед заказчиком.

03 Развитие действующей платформы данных

Ноябрь 2025 — настоящее время · Ритейл

Архитектор DWH

Отвечаю за технические решения и качество реализации. Подключаю источники, проектирую витрины, оптимизирую запросы и пайплайны. Провожу ревью кода, развиваю production-контур и работаю с заказчиком. Модель данных сочетает подходы Kimball и Inmon.

04 Запуск Lakehouse

2025 · Модернизация платформы данных

Архитектор и руководитель технической команды

Спроектировал платформу на open-source-стеке для замены двух устаревших систем: DWH на Microsoft SQL Server и Data Lake на Hadoop. Руководил командой по двум направлениям: развёртыванию платформы и разработке.

Применил медальонную архитектуру. Разделил хранение, обработку и SQL-доступ; аналитические витрины потребителей вынес в ClickHouse. Обосновывал перед бизнесом выделение оборудования по расчётному сайзингу.

Lakehouse введён в промышленную эксплуатацию

02

HEAD OF DATA

ПРОЕКТНЫЙ ОПЫТ

05 / МАРКЕТИНГОВЫЙ SAAS

GCP, BigQuery, Cloud Storage
Python, Dataflow, Pub/Sub
Cloud Composer, Cloud
Functions
Great Expectations
Looker Studio, Grafana
Flask, Docker, Ansible
GitLab CI/CD, Secret Manager

06 / ПРОДУКТОВАЯ АНАЛИТИКА

Greenplum, PostgreSQL,
ClickHouse, S3
Python, Spark, Kafka
Airflow, dbt, RudderStack
Great Expectations
Tableau, Redash, Grafana
Ansible, GitLab CI/CD

МОДЕЛИРОВАНИЕ

Data Vault 2.0
Сквозная идентификация
Golden Record

Платформы и команды

05 Маркетинговая платформа с нуля

Апрель 2023 — июнь 2025 • Маркетинговая SaaS-компания

Head of Data Platform и архитектор платформы

Задача. Объединить более 20 внешних и собственных источников в единую аналитическую платформу. Источники различались способами авторизации, форматами, структурой и гранулярностью данных и оставались вне моей зоны ответственности.

Моя ответственность. Руководил командой разработки, отвечал за архитектуру и техническое развитие платформы. Согласовывал работу над интеграциями, моделью данных и аналитическими сервисами.

Решения. Спроектировал DWH на GCP по Data Vault 2.0. Организовал API-интеграции, согласование данных разной гранулярности и обработку поздно поступающих данных. Платформа включала no-code-подключение источников, автоматизированную отчётность, алертинг и мониторинг.

Подготовка типовых отчётов сократилась с дней до минут

Более 20 источников объединены в SaaS-платформу с автоматизированной отчётностью.

06 Сквозная аналитика нескольких продуктов

Март 2022 — апрель 2023 • Продуктовая IT-компания

Head of Data и архитектор платформы

Задача. Объединить данные нескольких цифровых продуктов для сквозной и управленческой отчётности. Платформа обрабатывала миллионы событий с приростом данных около 200 ГБ в месяц.

Моя ответственность. Руководил направлением данных и развитием платформы. Совмещал управление командой с проектированием хранилища, событийной обработки и аналитических витрин.

Решения. Спроектировал DWH по Data Vault 2.0 и развивал событийную платформу. Организовал сквозную идентификацию пользователей и формирование Golden Record.

Golden Record и сквозная отчётность по нескольким продуктам

03

2017 — 2022

ПРОЕКТНЫЙ ОПЫТ

07 / EDTECH

Python, SQL
Greenplum, Amazon Redshift
Airflow, scikit-learn

08 / FINTECH

Python, PySpark
scikit-learn, MLflow
AWS, PostgreSQL
Elasticsearch, Kibana
Git

09 / DIGITAL

Python, SQL, PostgreSQL
pandas, NumPy
scikit-learn, TensorFlow, Keras
Ruby on Rails

ЗАДАЧИ ML

Скоринг и ранжирование
Выявление аномалий
NLP и временные ряды
Применение моделей в
реальном времени

Инженерия данных и ML

07 Маркетинговое DWH и аналитические модели

Июнь 2021 — март 2022 • EdTech-компания

Лид команды развития DWH

Руководил командой, развивающей хранилище маркетинговых данных. Проектировал модель данных, разрабатывал и перерабатывал ETL/ELT-процессы и DAG в Airflow. Работал над стабильностью пайплайнов и развитием витрин.

Результат. На сформированном контуре данных построены мультиканальная модель атрибуции и модели прогнозирования ключевых бизнес-метрик.

08 Данные и модели для решений в реальном времени

Январь 2020 — март 2021 • FinTech-компания

Инженер платформы данных и ML-решений

Развивал контур данных для клиентского скоринга, ранжирования лидов и фрод-мониторинга. Готовил признаки, строил пайплайны обработки больших объёмов данных и разделял Dev- и Prod-окружения. Разрабатывал модели и участвовал в их интеграции в рабочие системы.

Результат. На платформе разработаны и внедрены ML-модели принятия решений в реальном времени.

09 Аналитика и ML для цифровых продуктов

Март 2017 — январь 2020 • Digital-компания

Data Scientist и ML Engineer

Работал с продуктовой и маркетинговой аналитикой для медиа, электронной коммерции и контентных продуктов. Разрабатывал продуктовые метрики, модели классификации и ранжирования, решения для временных рядов и NLP, а также backend-сервисы.

В числе инженерных задач разработал мигратор данных из 1С-Битрикс в PostgreSQL.