

Data Reliability Engineering: A new approach to data quality

01 Introduction

02 How we got here

03 What is DRE?

04 DRE in the real world



Egor is co-founder and Chief Technology Officer at Bigeye. Before starting Bigeye, Egor was a staff engineer at Uber where he scaled the company's first data warehouse, supporting thousands of internal users and mission-critical workloads.

Egor Gryaznov Co-founder and CTO



The cost of poor data quality

Every year, poor data quality costs organizations an average of

\$12.9 million

Revenue risk

- 01 Had a \$1M revenue loss due to undetected billing errors.
- 02 Didn't detect that a high % of shoppers couldn't log in on Black Friday.
- 03 Fined after missing a compliance reporting deadline, due to poor data quality.

Compliance risk

- 01 A broken recommendation engine shared incorrect product information with shoppers.
- 02 Lost two days' worth of data in monthly report before realizing it.
- 03 Software bug stopped sales tax charges in a state.

Wasted engineering time

- 01 Reporting is manually checked because data quality is not trusted.
- 02 Data scientists spend 75% of their time on data quality issues.
- 03 End users often find data issues before the data team does.

01 Introduction

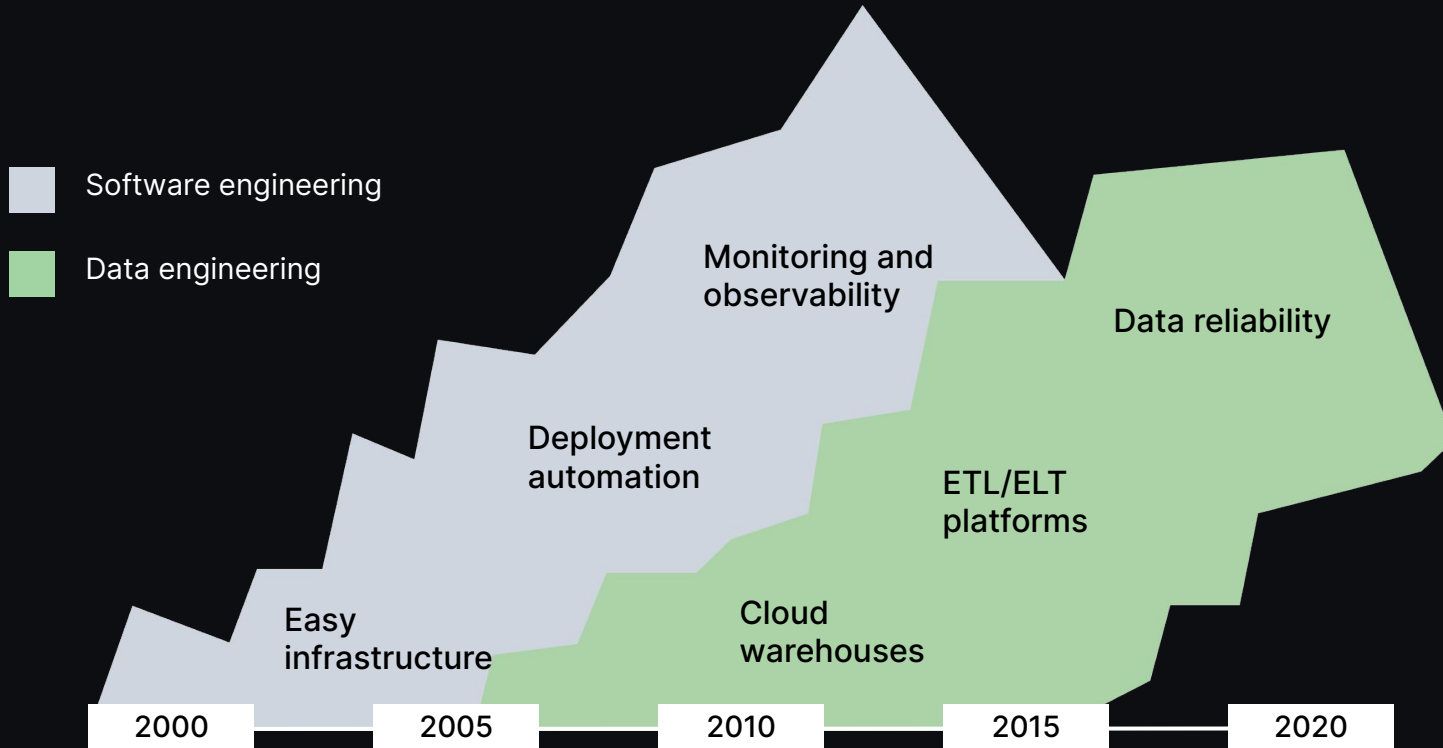
02 How we got here

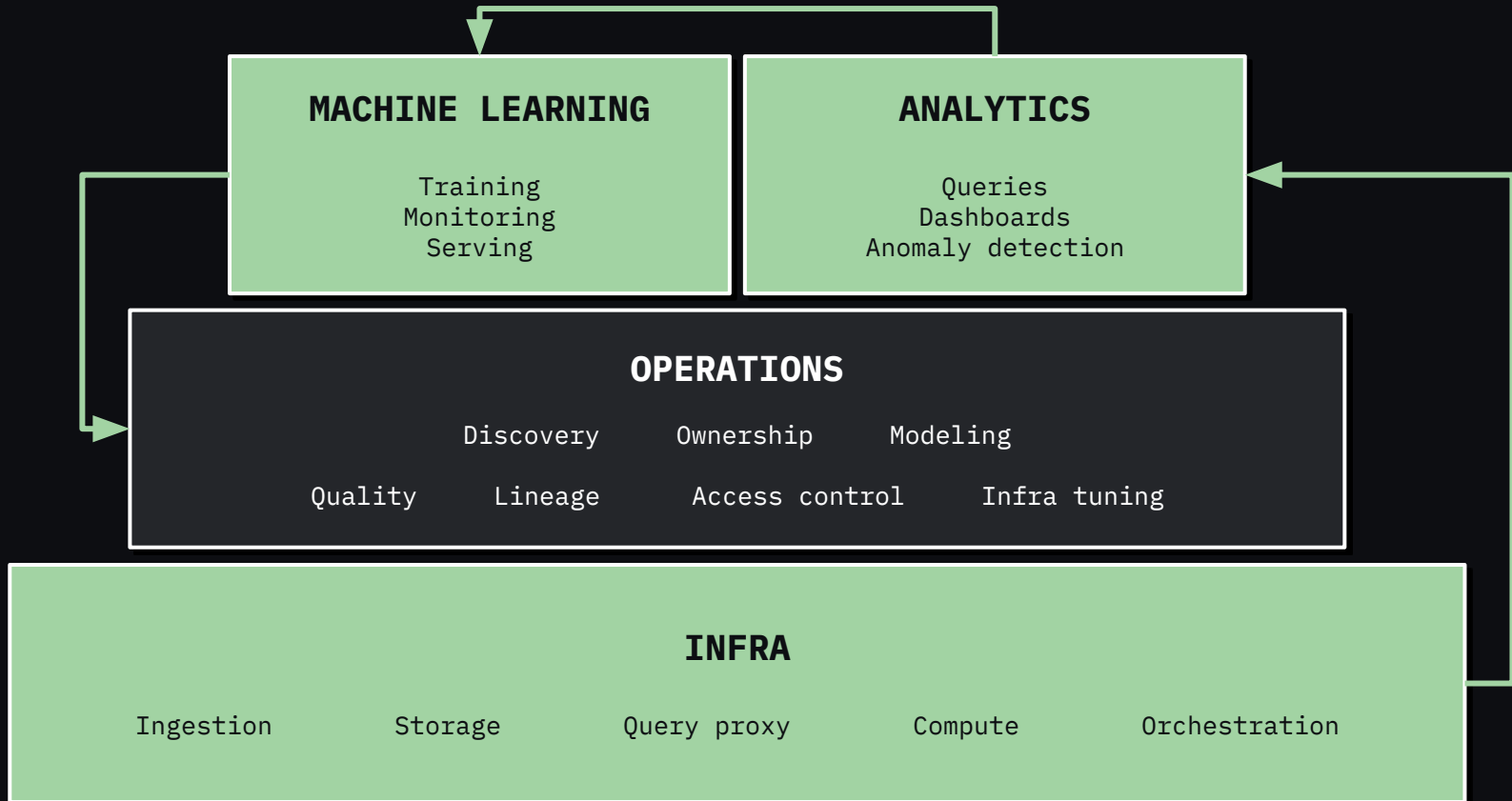
03 What is DRE?

04 DRE in the real world



Data engineering is evolving





01 Introduction

02 How we got here

03 What is DRE?

04 DRE in the real world

Site Reliability Engineering

Site reliability engineering (SRE) is a set of principles and practices that incorporates aspects of software engineering and applies them to infrastructure and operations problems.

The main goals are to create **scalable** and **highly reliable software systems**.

Principles

- 01 Embrace risk
- 02 Set standards
- 03 Reduce toil
- 04 Monitor everything
- 05 Use automation
- 06 Control releases
- 07 Favor simplicity

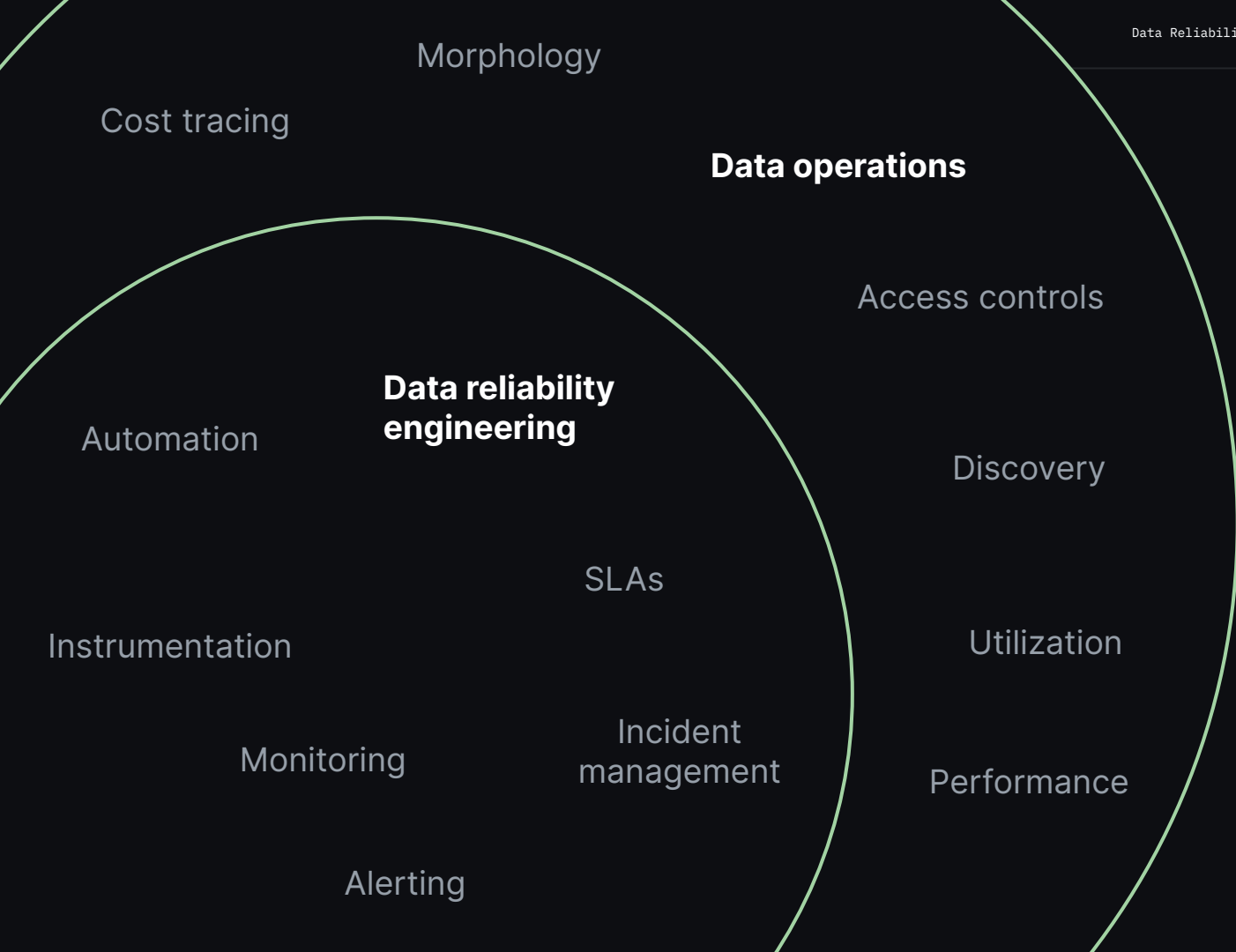
Data Reliability Engineering

Treating data quality like an engineering problem.

The key practices data teams must employ to ensure that data stays fit for use in key applications—decision making, AI/ML projects, and embedded analytics—without losing iteration velocity of their data environment.

Principles

- 01 Embrace risk
- 02 Set standards
- 03 Reduce toil
- 04 Monitor everything
- 05 Use automation
- 06 Control releases
- 07 Favor simplicity



01 Introduction

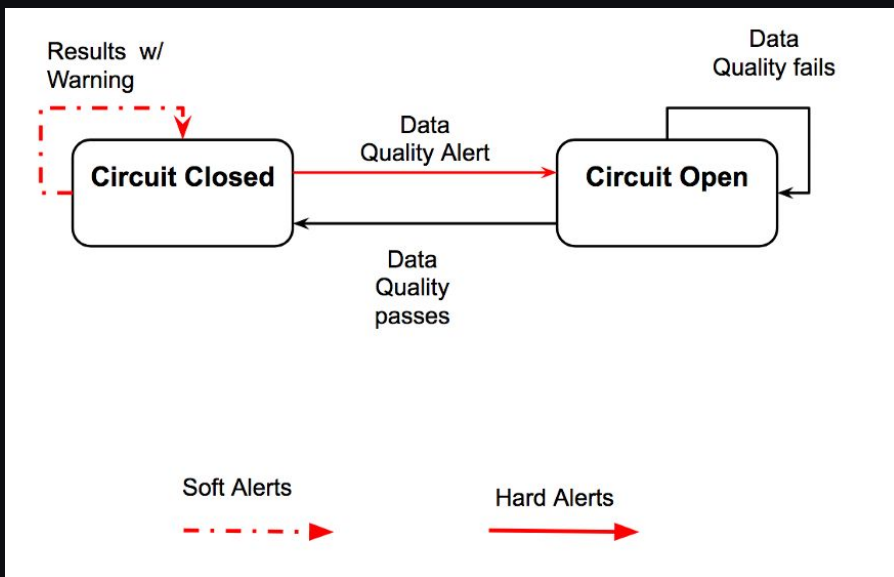
02 How we got here

03 What is DRE?

04 DRE in the real world



Circuit Breakers - Intuit



Principles

- 01 Embrace risk
- 02 Set standards
- 03 Reduce toil
- 04 Monitor everything
- 05 Use automation
- 06 Control releases
- 07 Favor simplicity

Midas - Airbnb



Certified by Midas

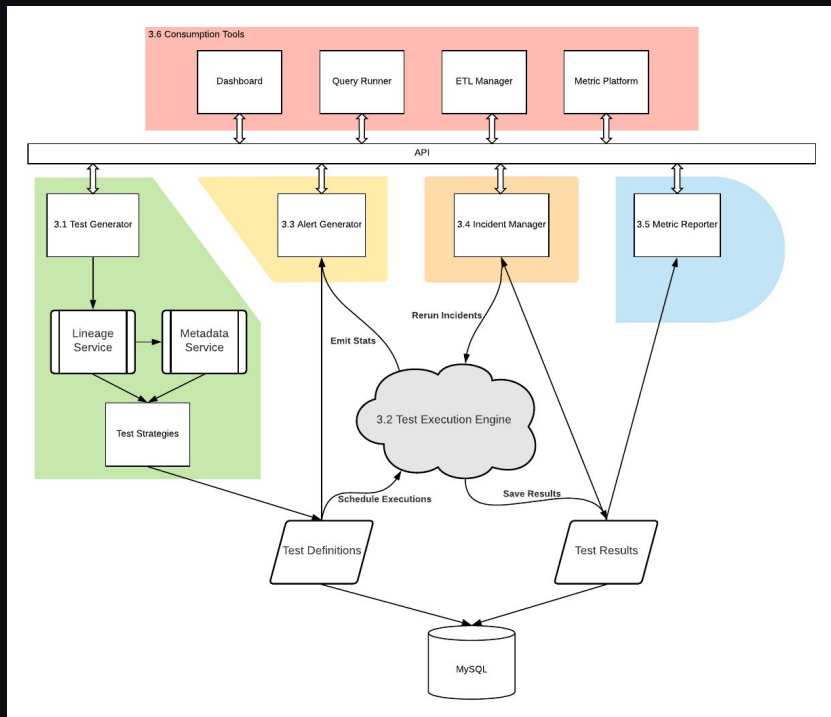
This certifies that the data represents the standardized definition for that business concept and employs best practices in data engineering that guarantee its reliability and accuracy over time.

[Learn more](#) >

Principles

- 01 Embrace risk
- 02 **Set standards**
- 03 Reduce toil
- 04 Monitor everything
- 05 Use automation
- 06 Control releases
- 07 Favor simplicity

UDQ - Uber



Principles

- 01 Embrace risk
- 02 Set standards
- 03 Reduce toil
- 04 Monitor everything
- 05 Use automation
- 06 Control releases
- 07 Favor simplicity

Thank you.



Egor Gryaznov
egor@bigeye.com