



What Metrics Make Engineering Teams Elite

An exclusive report on the metrics and workflows used by the highest performing dev teams in the world

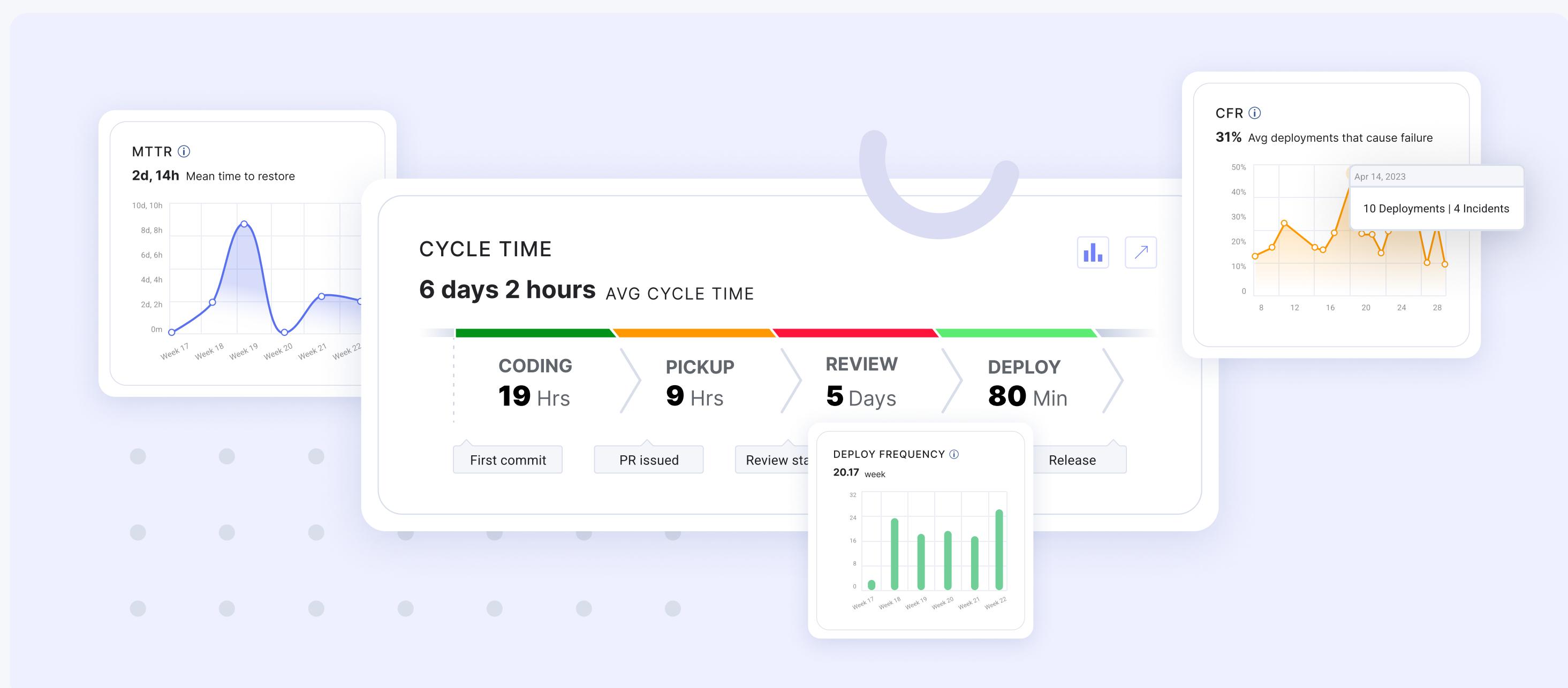


Table of Contents

- 1.** Introduction: How Metrics Shaped DevOps
- 4.** Beyond DORA: Process, Goals & Takeaways from LinearB's Engineering Benchmarks Study
- 7.** The New Metrics: 10 Benchmarks, 3 Categories and a Big Shift
- 11.** Tools of Elite Teams
- 12.** Conclusion

INTRODUCTION

How Metrics Shaped DevOps



LinearB is all about productivity — not just for enabling and improving, but also finding the best ways to measure output. **LinearB's Engineering Benchmarks Report** is the first of its kind: an examination of the key performance metrics that make engineering organizations elite, average or underwhelming as compared to peers in the industry.

In 2014, a group of behavioral scientists set out on a years-long journey to create the first set of industry standards for DevOps organizations. This team of individuals began as a collaboration between Nicole Forsgren, Gene Kim, Jez Humble and the automation software company Puppet, calling themselves the DevOps Research and Assessment (DORA) team.

In 2018, Forsgren, Humble and Kim co-authored the book *Accelerate: The Science of Lean Software and DevOps: Building and Scaling High Performing Technology Organizations*, which won the prestigious Shingo Publication Award.

The seven-year study presented in *Accelerate*, performed in collaboration with Google, analyzed 23,000 data points from more than 32,000 international professionals in the field. The development teams DORA studied included those from start-ups and large enterprises; nonprofit and for-profit organizations; and born-digital as well as legacy companies.

INTRODUCTION

How Metrics Shaped DevOps

Later that year, Google acquired DORA and folded the company into its Google Cloud division.

In *Accelerate*, the co-authors contend that the “highest performers are twice as likely to meet or exceed their organizational performance goals.”

They identified four key metrics that cut to the heart of how to measure software development and delivery success performance:

- ✓ Deployment frequency
- ✓ Lead time for changes
- ✓ Mean time to recovery
- ✓ Change failure rate

These DORA metrics create visibility and clarity into day-to-day operations of software development teams and offer data that can help leaders make better decisions. Growing awareness of DORA’s work began to have real impacts on teams across industries. The metrics helped foster healthy competition between developers — “gamifying the system,” making work like team sports, in which collaboration is essential for improvement.

It’s not hyperbole to say these metrics revolutionized the way DevOps is done.

Today, the DORA project is still the longest-running academically rigorous research investigation of its kind.

However, the tech landscape has changed dramatically since *Accelerate* was published in 2018. Visibility and metrics aren’t enough on their own. Engineering leaders require context and meaning to guide their teams to greatness.

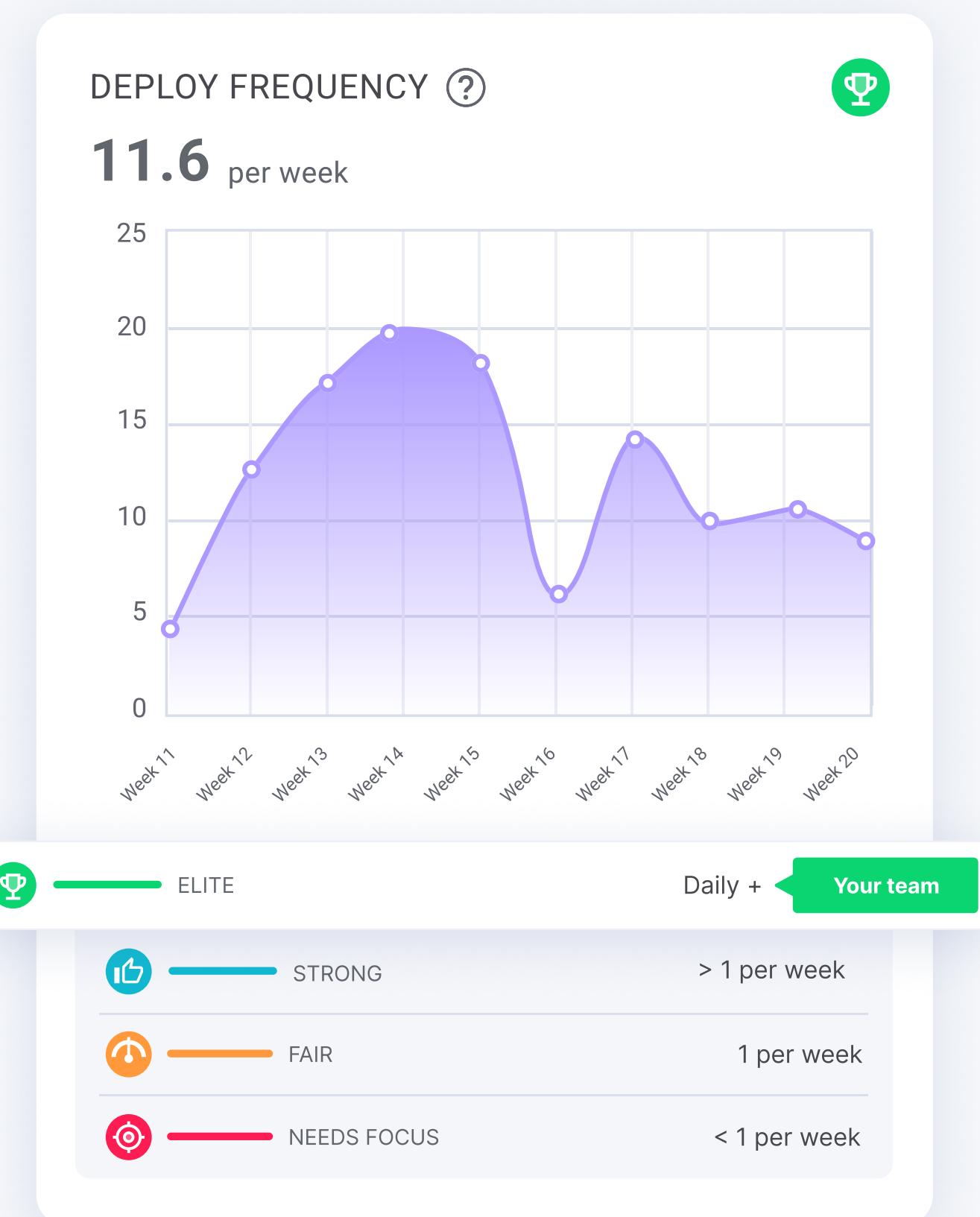
“

“**DevOps is not a goal, but a never-ending process of continual improvement.**”

Jez Humble
Author | *Accelerate*

INTRODUCTION

How Metrics Shaped DevOps



In an effort to support and continue DORA's legacy of groundbreaking research, LinearB published the *Engineering Benchmarks Report* based on the study of 1,971 dev teams and more than 4.5 million branches. It is a first-of-its-kind examination of the key performance metrics that make engineering organizations elite, average or underwhelming as compared to peers in the industry.



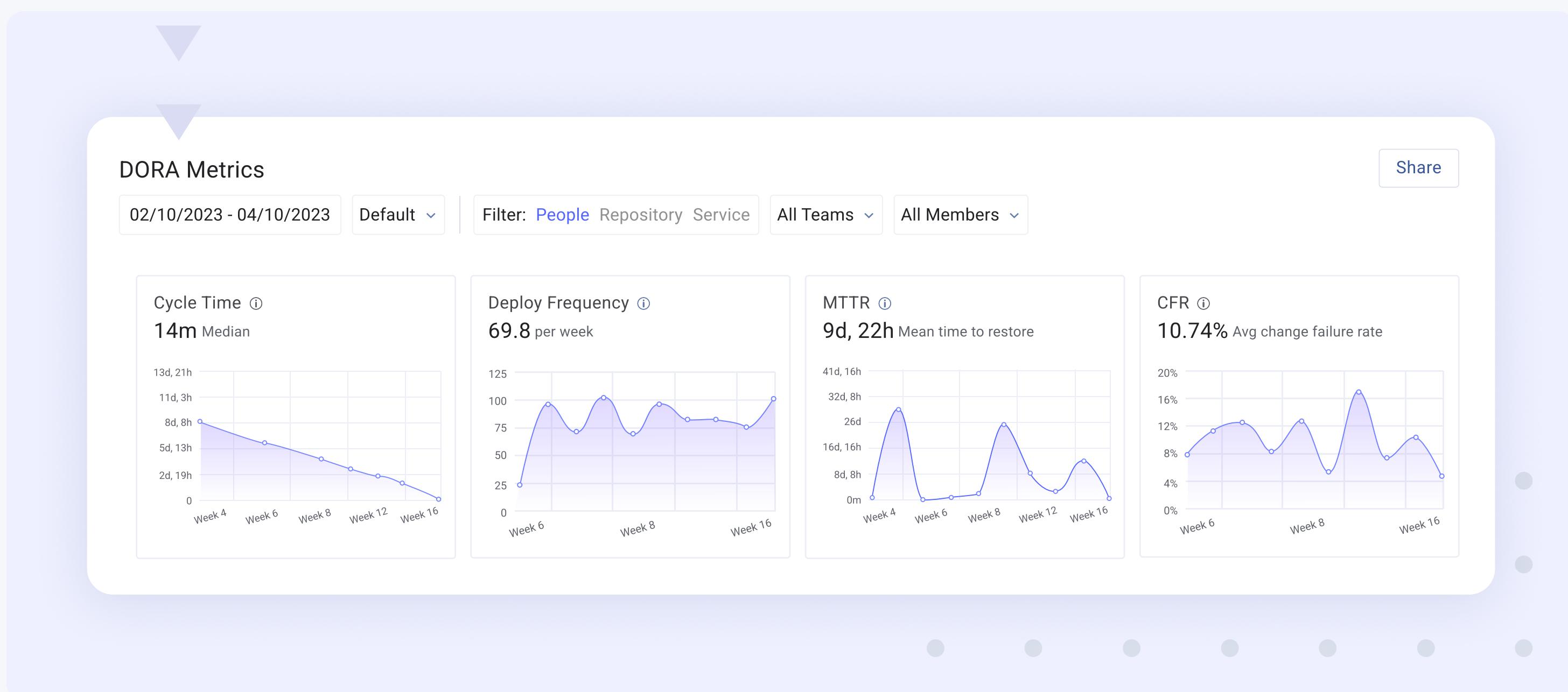
**"The biggest impact of metrics?
Going from 'darkness to visibility.'"**

Ori Keren
Co-Founder & CEO | LinearB

LinearB co-founder and CEO Ori Keren says DORA helped DevOps move from "darkness to visibility." But once there's full illumination, leaders need a full understanding of what makes one team more effective than another — and what characterizes the top tier of development teams that are truly elite.

BEYOND DORA

Process, Goals & Takeaways from LinearB's Engineering Benchmarks Study



Since the publication of *Accelerate*, the LinearB team saw how DORA's work transformed the DevOps space. But it soon became clear that DORA's four key metrics alone weren't sufficient to measure the wide variety and growing contingent of organizations practicing DevOps.

After adopting and measuring DORA metrics for several years within the LinearB R&D organization, we identified three categories where the DORA measures alone were lacking.

1. DORA Metrics are Lagging Indicators

Managing an engineering team is not so different from running a business — with stats like annual recurring revenue (ARR) standing in for DORA's four key metrics. Simply waiting for ARR to improve doesn't work; nothing will happen without action. Acting on leading indicators and KPIs is more effective than missing an ARR goal and backtracking to correct problems. It's better to identify opportunities early on to attract more leads, increase demand and optimize development execution.

BEYOND DORA

Process, Goals & Takeaways from LinearB's Engineering Benchmarks Study

It's the same in the engineering world; it's easier and more effective to identify trends and take proactive action, rather than scrambling after missing the mark iteration after iteration.

DORA metrics measure the velocity and stability of a team and provide an overall view of its effectiveness.

However, DORA metrics are lagging, rather than leading indicators — a key aspect of DORA that often gets overlooked.

Lagging indicators provide an understanding of a team's past and/or current performance, which is still crucial to fostering a culture of continuous improvement. And while they're relatively easy to measure, they're more difficult to change.

2. DORA Metrics Don't Accurately Measure Business Impact

It's possible to knock all four DORA metrics out of the park but still miss the mark on delivering on promises if DevOps isn't aligned with business goals and needs. Whatever a team builds isn't truly successful unless it fits within an organization's overall business context.

Think about an engine firing on all cylinders that's powering a car headed in the wrong direction.

Some metrics, like cycle time and deployment frequency, are important to an internal team but aren't meaningful to the business at large. Devs care about the relative speed of sections of code (however small or large) to keep their pipeline running smoothly, but business leaders want to see the entire feature(s) — consistently, and as often as possible. Velocity doesn't matter if there's no alignment with the rest of the business and the team isn't delivering what will make customers happy.

“

“DevOps shows how we optimize the IT value stream, converting business needs into capabilities and services that provide value for our customers.”

Gene Kim

Author | Accelerate

BEYOND DORA

Process, Goals & Takeaways from LinearB's Engineering Benchmarks Study

3. Measurements (Themselves) Aren't Improvements

DORA metrics are vital markers of the status quo, but they're not blueprints for change; they're a starting point.

Building a product and improving a team's performance are substantially different projects, both requiring operational plans. (That includes both a team's ability to iterate on the product and its ability to do better on the next ones.) Plus, DORA metrics are abstract statistics; we need benchmarks that can extend the lessons of DORA and be leveraged as prescriptive goals for bottom-up change among engineering teams.

So LinearB expanded on this research with additional metrics that provide a more holistic view of software development teams and the ways they work. The Engineering Benchmarks Report built a new set of metrics with an eye to the second decade of the 20th century.

The 10 metrics and benchmarks included in this newly published report help teams:

1. Position their performance against the industry;
2. Identify specific areas they can improve to boost performance; and
3. Map a strategy to achieve elite performers.

“

[The] highest performers are twice as likely to meet or exceed their organizational performance goals.”

DORA authors
Accelerate

THE NEW METRICS

10 Benchmarks, 3 Categories and a Big Shift

To drive change in engineering teams, LinearB's Engineering Benchmarks Study identified 10 metrics within 3 groupings that matter most in DevOps today.



**Delivery
lifecycle**



**Developer
workflow**



**Business
alignment**

THE NEW METRICS

10 Benchmarks, 3 Categories and a Big Shift

Delivery Lifecycle Metrics & How Elite Teams Perform

1. Cycle time

Cycle time (often synonymous with lead time for changes) is one of the four key DORA metrics. It represents the time elapsed from the first commit to when it's in customers' hands, encompassing coding time, pickup time, review time, and deployment time.

Ψ Elite cycle time is under 42 hours.

2. Coding time

This measures the time spent from the start of a first commit until a pull request (PR) is issued.

Ψ Elite coding time is less than half an hour.

3. Pickup time

This starts after the PR is issued and ends when the first comment is made.

Ψ Elite pickup time is under 1 hour.

4. Review time

This covers the code review and PR merge process, from first comment until merged into the main branch.

Ψ Elite review time is less than 1 hour.

5. Deploy time

This is the measurement from branch merge to code release.

Ψ Elite deploy time is less than 1 hour.

“

“Cycle time is an engineering super metric.”

Ori Keren

Co-Founder & CEO | LinearB

THE NEW METRICS

10 Benchmarks, 3 Categories and a Big Shift

Developer Workflow Metrics & How Elite Teams Perform

6. Deploy frequency

This metric measures how frequently code is released into production. It is higher when deploy time is lower. It can also be considered a function of cycle time.

Ψ Elite teams deploy on a daily basis.

7. Pull request (PR) size

This is arguably the most important leading indicator of a healthy development pipeline. Smaller PRs = faster pickup, faster and deeper reviews, faster merges, and higher deployment frequency.

Ψ Elite teams average fewer than 105 code changes in their PRs.

8. Rework rate

This measures code churn. Rework is categorized as changes to code 21 days old or less. Higher rework rates are good indicators of quality concerns.

Ψ An elite rework rate is under 2%.

Business Alignment Metrics & How Elite Teams Perform

9. Planning accuracy

This metric defines what was committed versus what was actually delivered.

Ψ Elite teams have a planning accuracy of 80% or more.

10. Capacity Accuracy

Capacity accuracy tells you if your teams are taking on the “correct” amount of work. It measures all completed (planned and unplanned) work as a ratio of planned work.

Ψ Elite teams capacity accuracy above 85%.

THE NEW METRICS

10 Benchmarks, 3 Categories and a Big Shift

Complete Performance Data

	ELITE	STRONG	FAIR	NEEDS FOCUS
 CYCLE TIME	< 42 hours	42 - 95 hours	95 - 188 hours	188+ hours
 CODING TIME	< .5 hours	.5 - 1 hours	1 - 4.5 hours	4.5+ hours
 PICKUP TIME	< 1 hours	1 - 3 hours	3 - 14 hours	14+ hours
 REVIEW TIME	< 1 hours	1 - 5 hours	5 - 21 hours	21+ hours
 DEPLOYTIME	< 1 hours	1 - 20 hours	20 - 196 hours	196+ hours
 DEPLOY FREQUENCY	Daily +	> 1/ week	1/ week	< 1/ week
 PR SIZE	< 105 code changes	105 - 155 code changes	155 - 229 code changes	229+ code changes
 REWORK RATE	< 2%	2% - 6%	6% - 12%	12%+
 PLANNING ACCURACY	> 80%	79% - 65%	64 % - 40%	< 40%

*Capacity-accuracy metrics do not fall easily along the categories above but will be included in future charts.

DATA
SOURCED FROM

 +1,971 TEAMS

 +4.5M BRANCHES

Tools of Elite Teams

The LinearB *Engineering Benchmarks Report* helps engineering leaders identify where they can impact improvement and set the goals to achieve increased performance. The best way to ensure every goal is tied to meaningful action? Embed those goals in a team's workflow with software built for this purpose and anchor check-ins and recurring ceremonies around these goals.

The right tools can optimize that workflow and provide granular details about the status of each metric. With data-driven retrospectives, teams can make adjustments constantly — not just iterations of code and features, but of the day-to-day processes themselves. In turn, teams can improve their planning accuracy ... and the domino effect of better results leading to even-better results goes on and on.

LinearB's platform can jumpstart any software development team's journey to elite performance in engineering efficiency, quality and job satisfaction.

DORA Metrics

LinearB's free tool integrates with an organization's Git and project management system to provide team-level and organization-level DORA metrics within minutes.

Engineering Impact Metrics

The LinearB software delivery management platform goes beyond the four DORA metrics to offer engineering leaders unique views into resource allocation, cost capitalization, and delivery forecasting.

Workflow Automation

When engineering teams have gained the visibility they need in order to understand their performance, they start their improvement journey with LinearB's workflow automation tools gitStream and WorkerB.



WorkerB

A pull request had 465 changes and was merged without review

PR: [linweb/Projects top bar refactor](#)

Related Issue: [LINB-2178- Admin can self configure projects](#)

- ✓ Impact Metrics: Build a robust metrics program
- ✓ Programmable Workflows: Streamline code merge process
- ✓ Resource Allocation: R&D resources aligned with company goals
- ✓ Predictable Project Delivery: On-time forecasting for project delivery
- ✓ Policy-as-code: Standard merge policy-as-code across all dev teams

Conclusion

The goal of the *Engineering Benchmarks Report* is to give teams a better understanding of their performance today and build a strategy that helps them drive the improvement they need in order to impact business goals.

Teams can make the biggest improvements when they know what to expect. A predictable pipeline is one that allows for incremental shifts in the right direction. And because all the metrics are connected, even small changes can produce big wins.

