Drive AI Adoption Through Experimentation

Meet Your Presenters

Deanna Ballew

Chief Product Officer, Listrak

20+ years of martech leadership. Previously SVP of Innovation, AI & DXP at Acquia and Chief Innovation Officer at Widen.

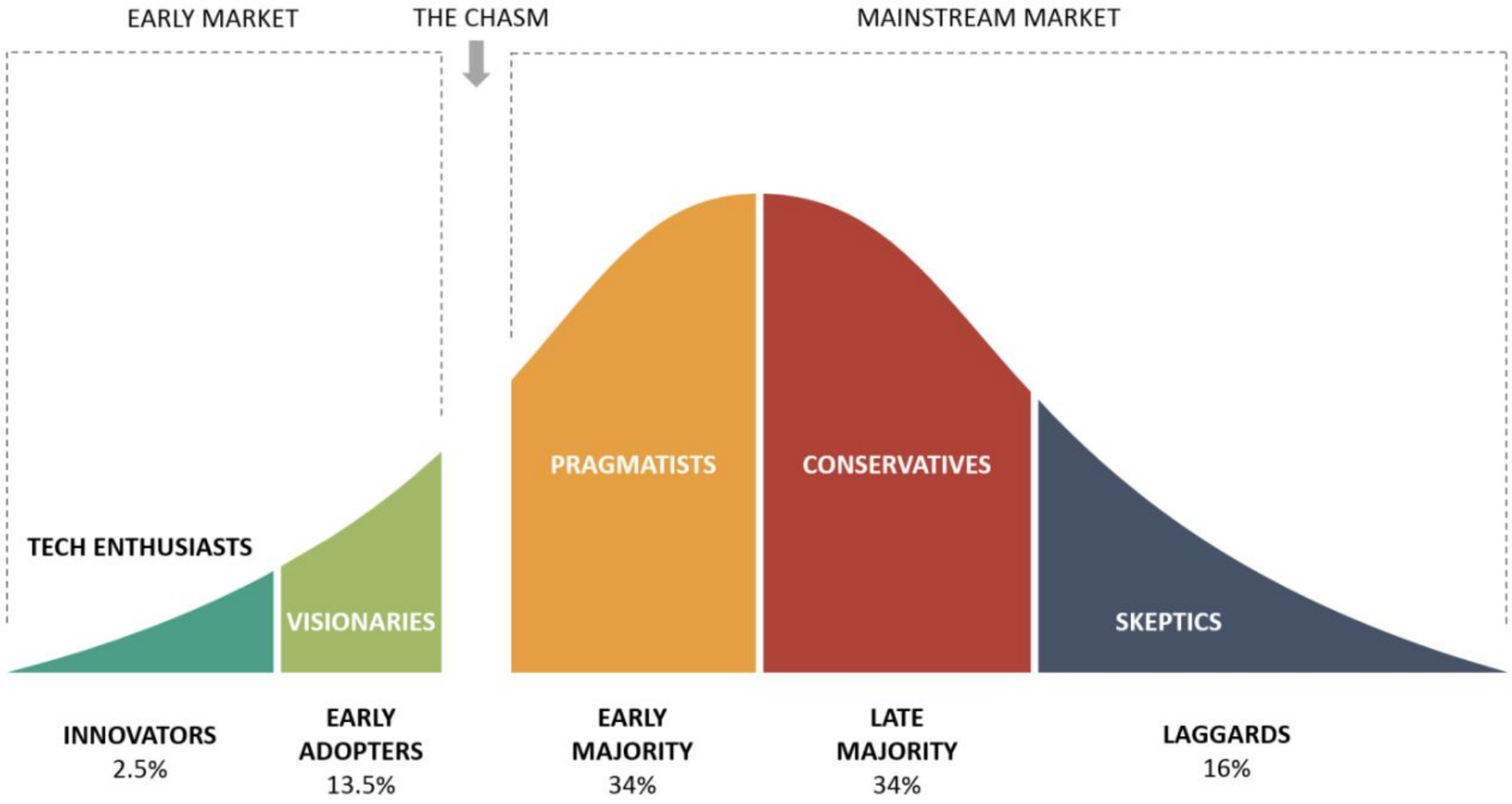
LinkedIn Profile

Ben Dotte

VP of Engineering, Acquia

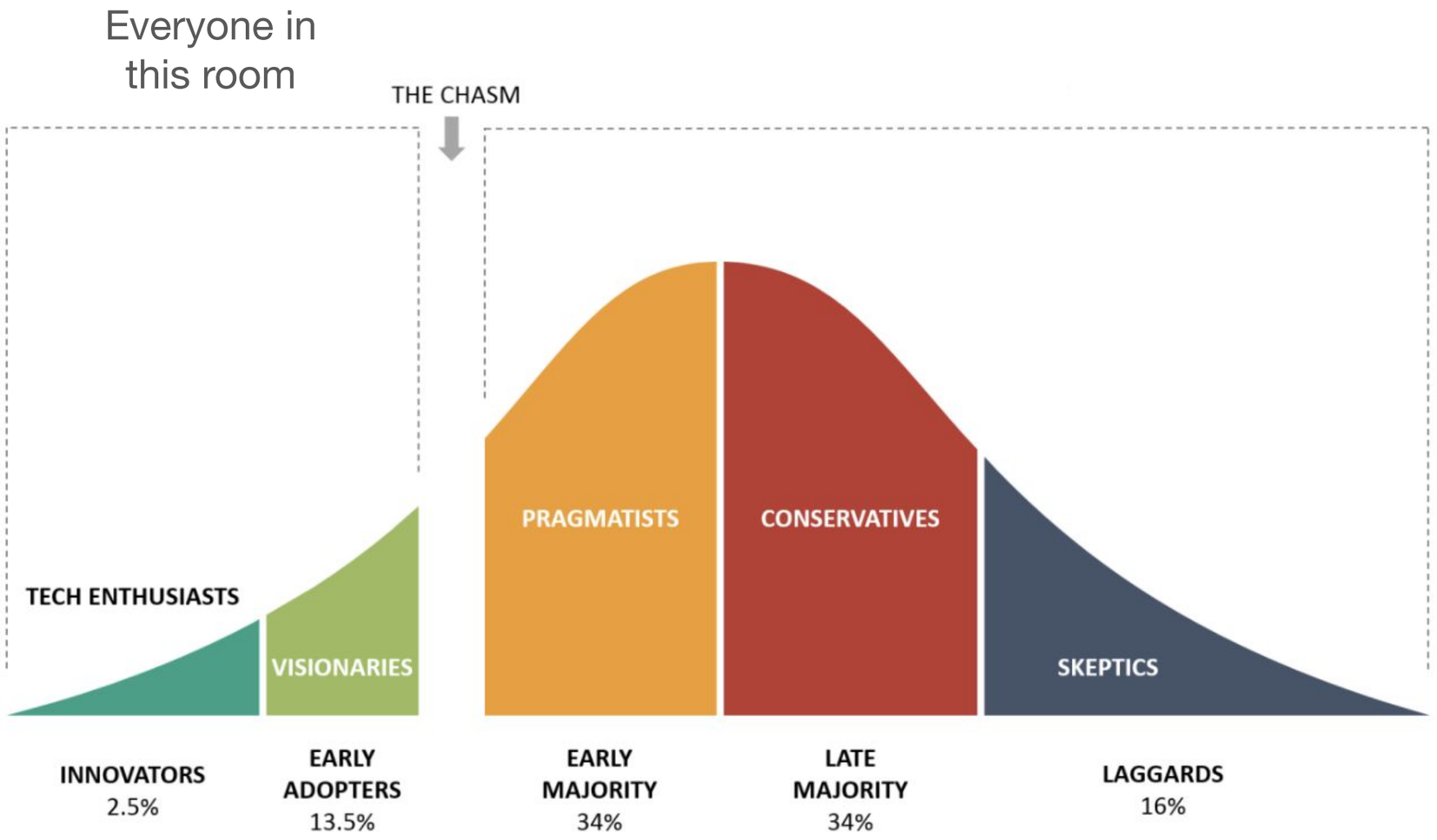
20+ years of martech engineering leadership. Previously Senior Director of Engineering at Widen.

LinkedIn Profile



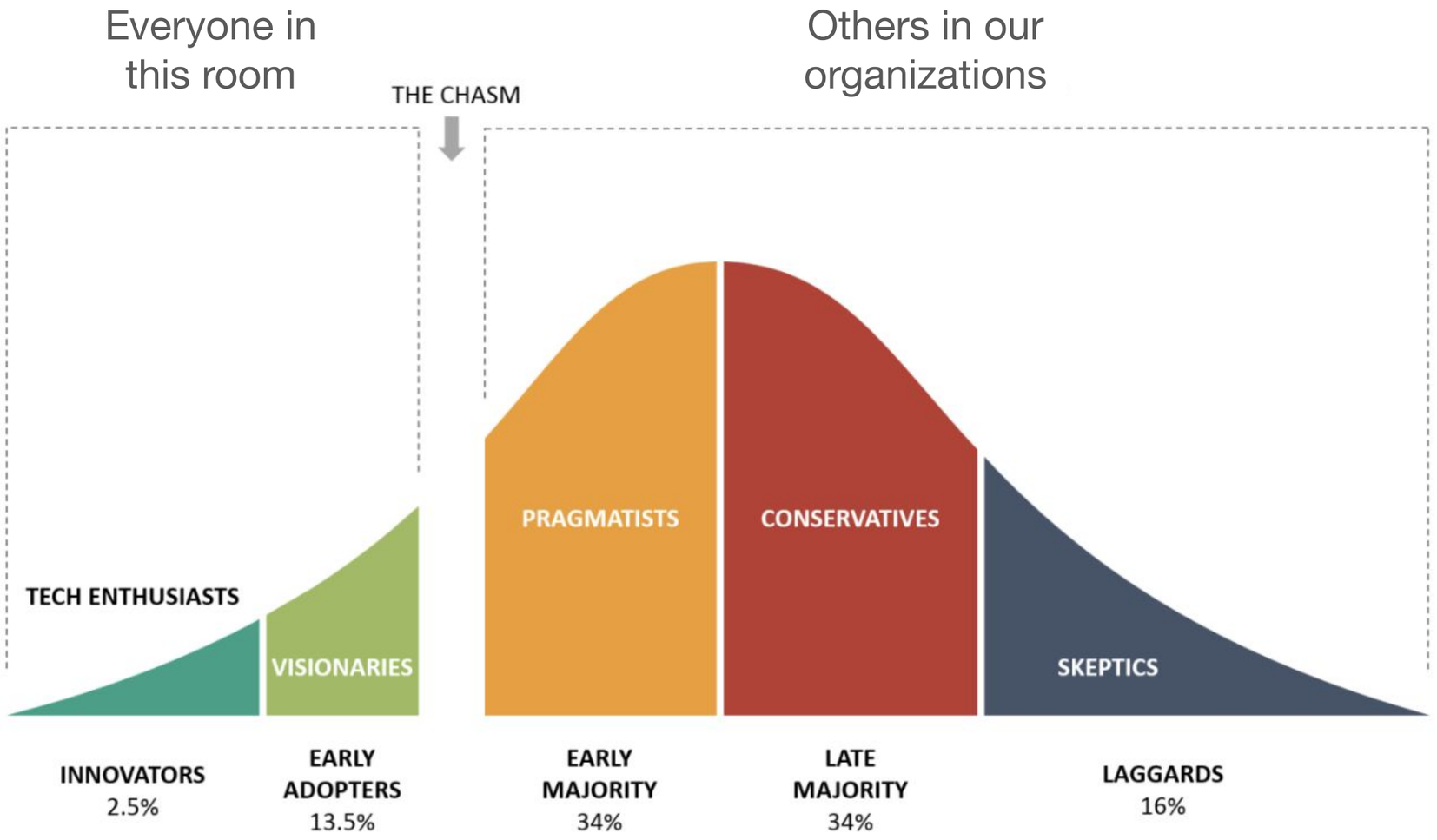
Technology Adoption Curve

Crossing the Chasm, Geoffery A. Moore



Technology Adoption Curve

Crossing the Chasm, Geoffery A. Moore



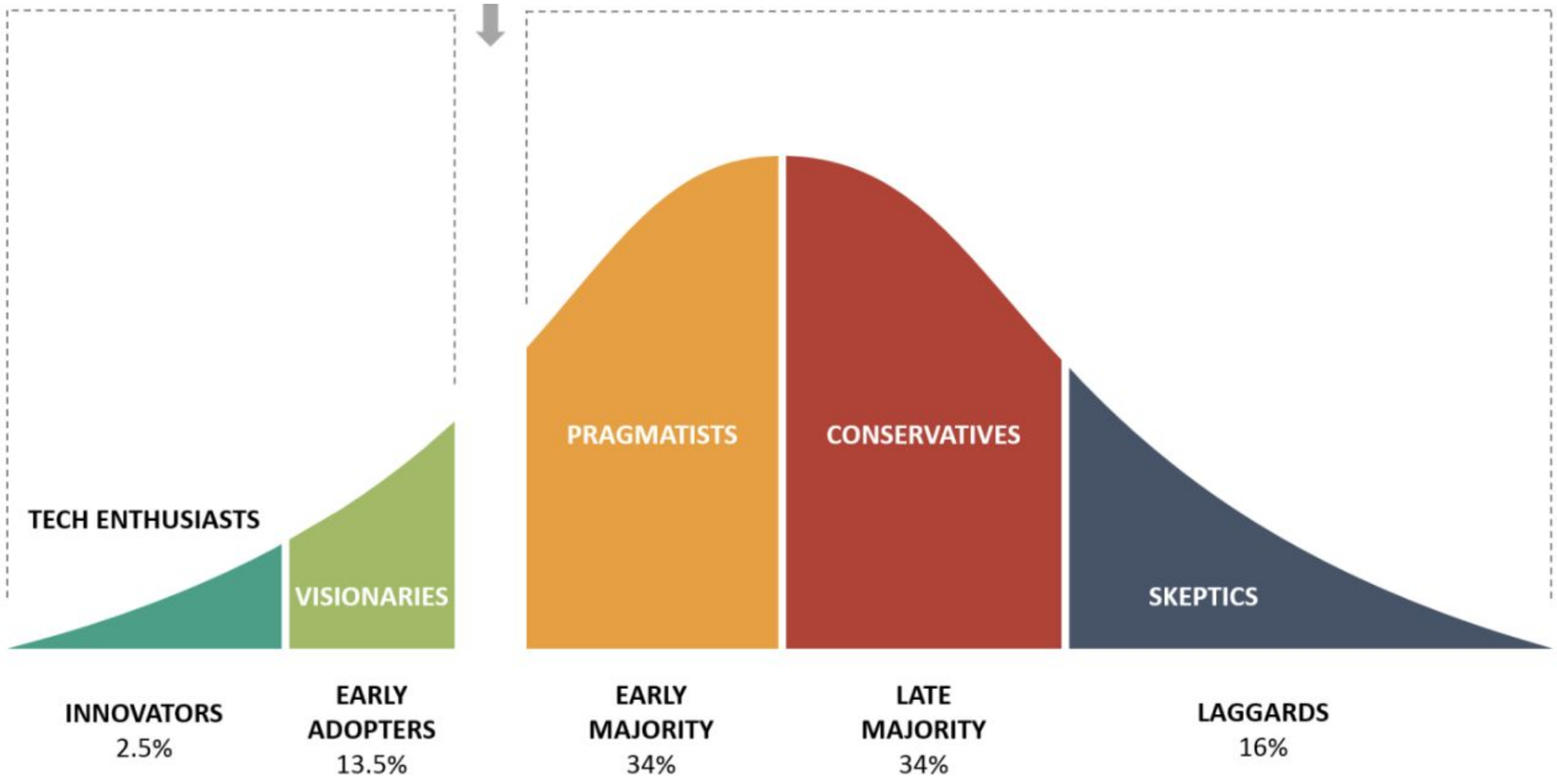
Technology Adoption Curve

Crossing the Chasm, Geoffery A. Moore

Everyone in this room

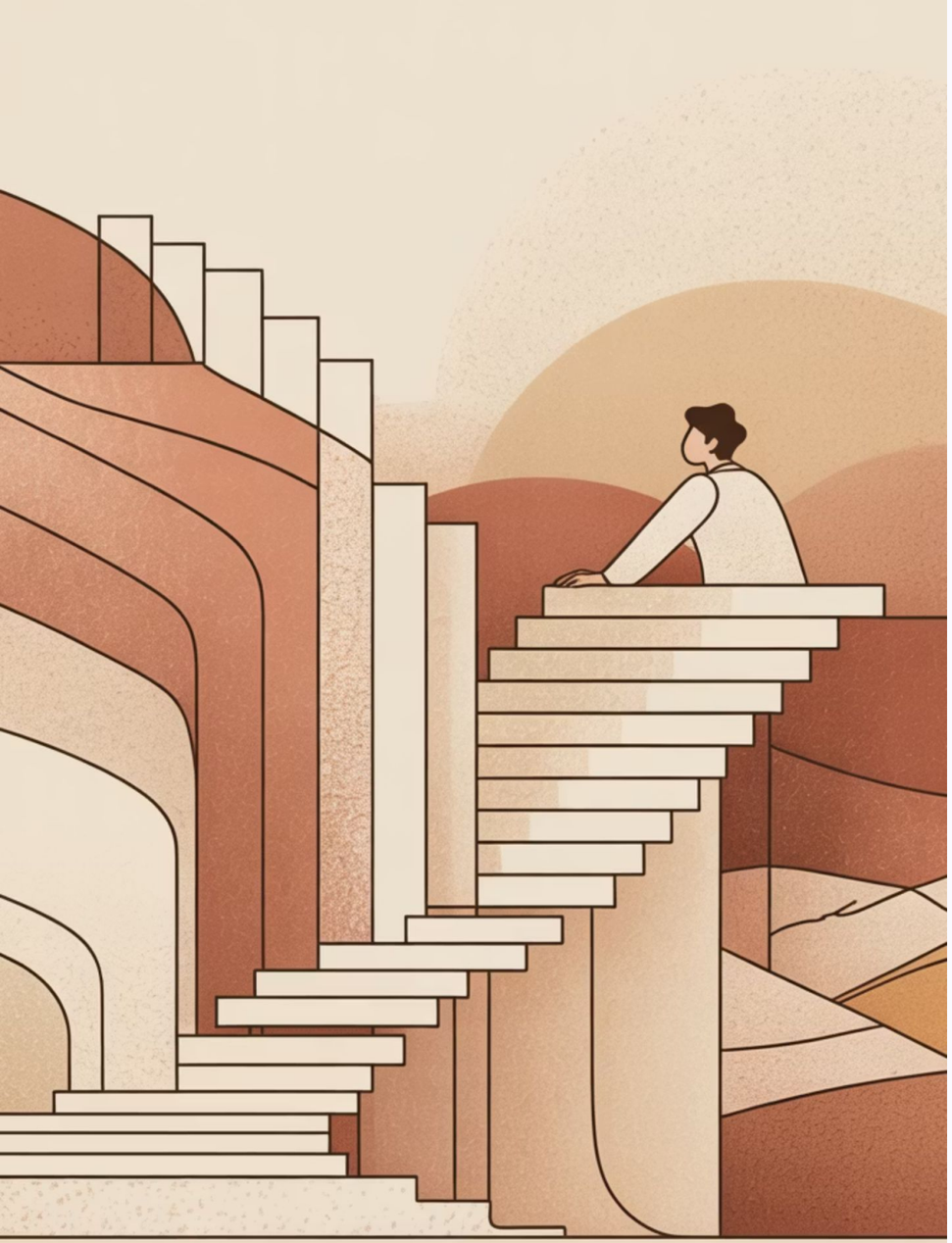
Experiment to cross the Chasm

Others in our organizations



Technology Adoption Curve

Crossing the Chasm, Geoffery A. Moore

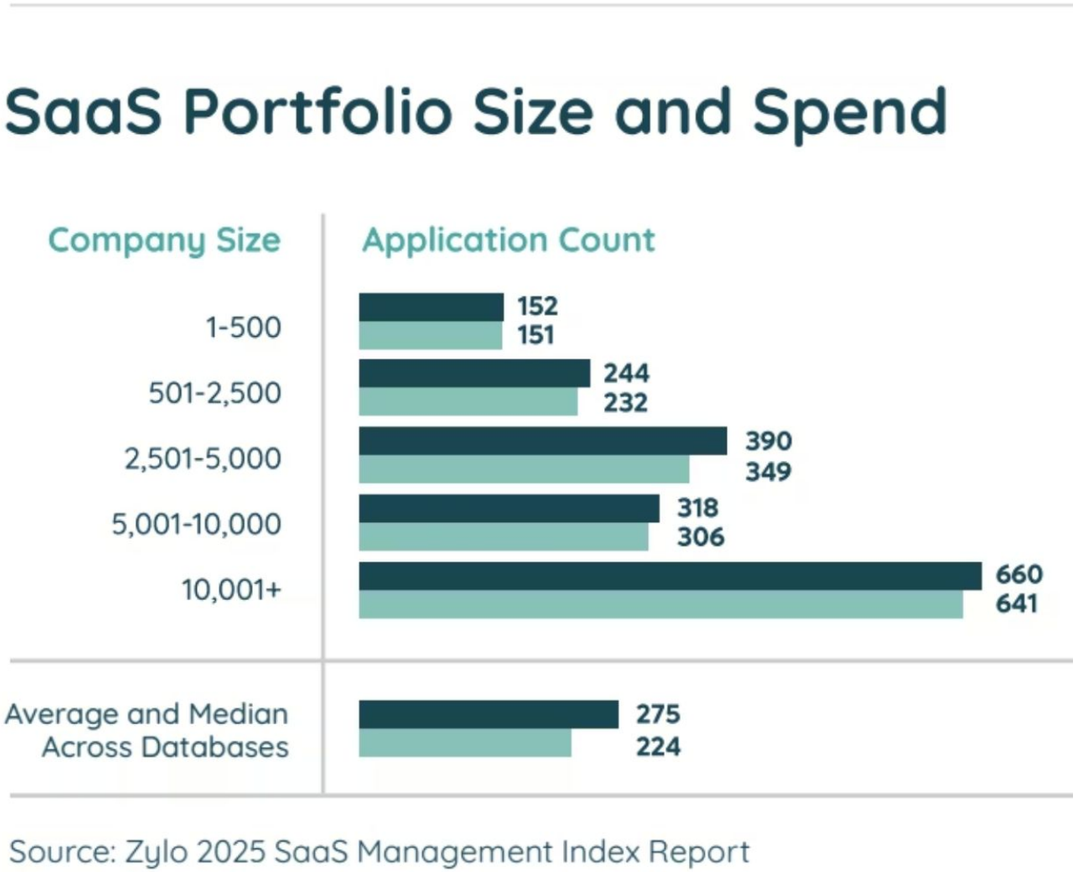


Three Challenges to GenAI Adoption

Challenge #1
Plethora of Choices

- Which features solve real problems?
- Which tools fit your risk profile?
- How do you calculate total cost of ownership?

275 SaaS applications
On average



Source: Zylo 2025 SaaS Management Index Report

The True Cost of AI Tools

$360K

Microsoft Copilot

Annual cost for 1,000 employees at $30/user/month

$180K

Slack + Google Gemini

Combined annual cost for 1,000 employees*

$600K

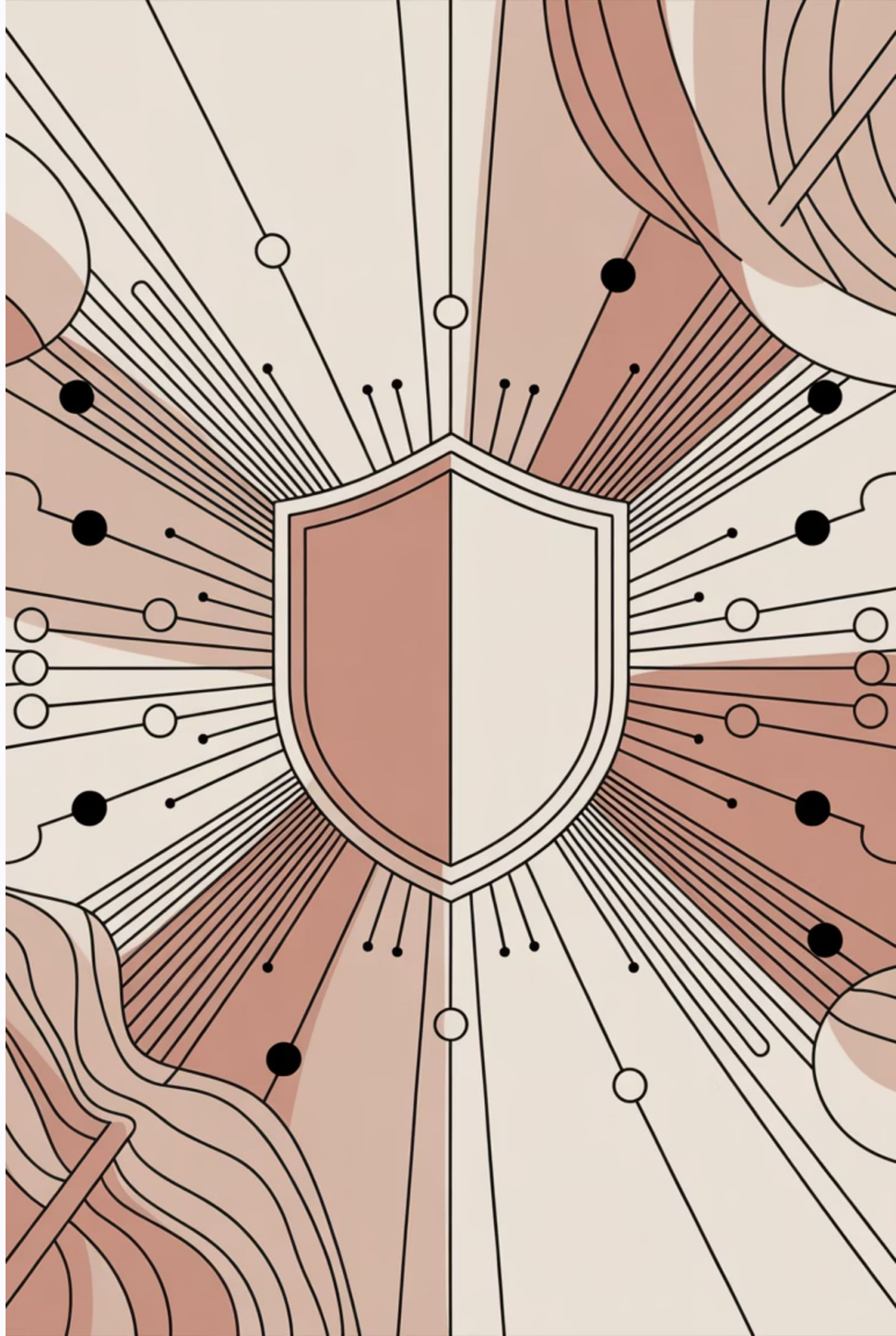
Salesforce Einstein GPT

Annual investment for 1,000 employees at $50/user/month*

* Pricing in 2023

Challenge #2

Security Concerns & Regulatory Compliance



Navigating the Regulatory Landscape

EU AI Act

Comprehensive regulatory framework establishing requirements for AI systems based on risk levels, with strict compliance mandates for high-risk applications.

NIST AI Framework

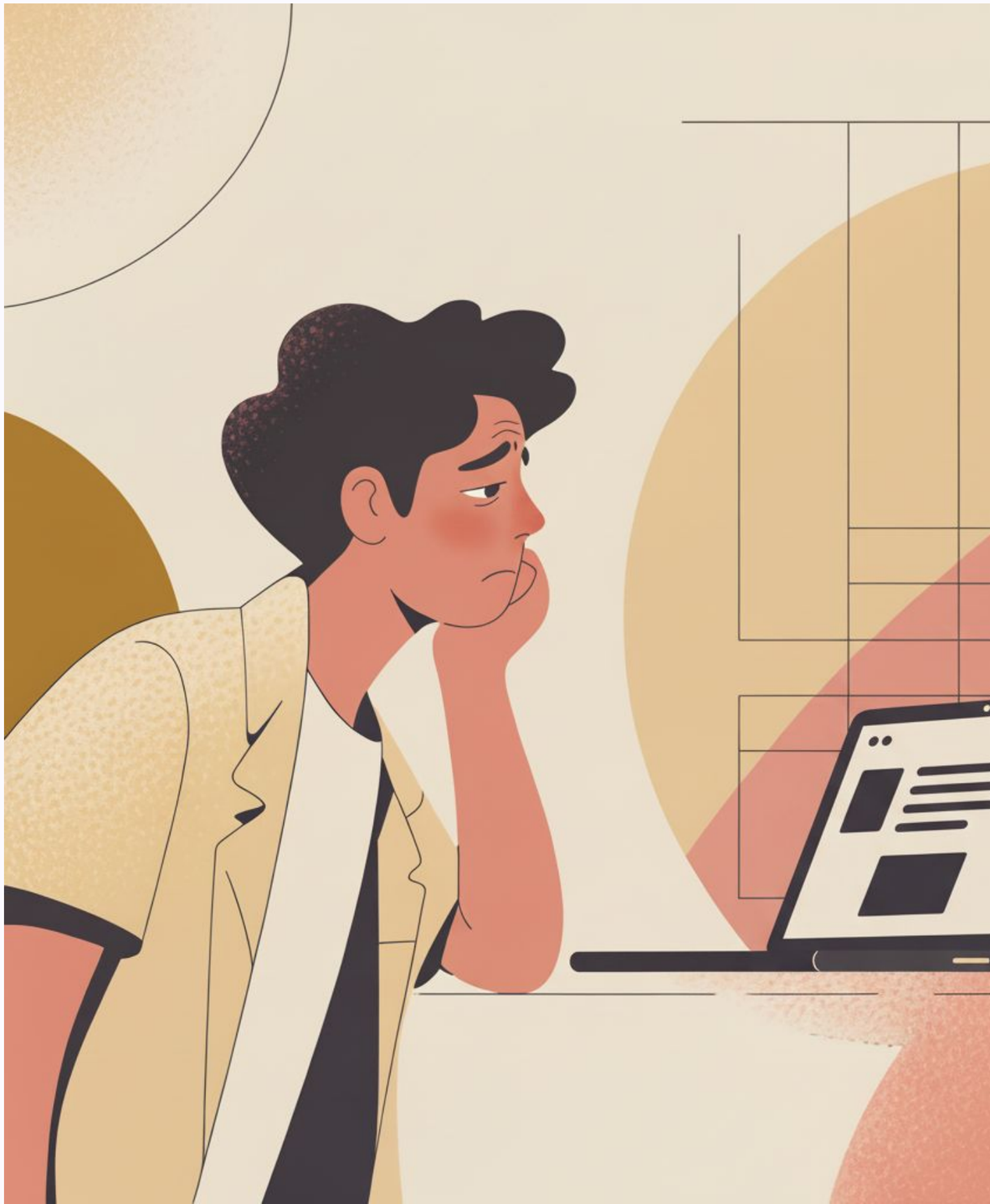
U.S. National Institute of Standards and Technology provides guidance for managing AI risks, trustworthiness, and responsible development practices.

ISO/IEC 42001

International standard for AI management systems, providing organizations with a structured approach to responsible AI governance and deployment.

Challenge #3

Humans Resist Change



The Three Dimensions of Human Resistance

Resistance to Change

Humans prefer predictable routines. New AI tools disrupt familiar workflows, creating psychological discomfort that manifests as pushback.

Fear of the Unknown

AI's "black box" nature creates anxiety. Without understanding how systems work, users worry about reliability, accuracy, and potential consequences.

Trust Deficit

Trust must be earned through consistent, positive experiences. Initial skepticism only dissolves when users see reliable results firsthand.

The Solution

Run Experiments

Why Experiments Overcome Barriers

01

Evaluate features and determine ROI

Test specific use cases with real data to understand which AI capabilities deliver measurable business value, providing budget justification.

02

Minimize security risks

Limit exposure by experimenting with controlled data sets and specific use cases, demonstrating safe implementation before scaling.

03

Build employee trust and buy-in

Early adopters become champions who influence peers. Hands-on experience replaces skepticism with confidence and enthusiasm.

Designing Effective Experiments

1

Define the Problem

Clearly articulate the challenge with validated qualitative or quantitative evidence. Vague problems yield vague results.

2

Propose a Solution

Describe how your chosen approach addresses the problem. Consider buy, build, or partner options.

3

Establish Success Metrics

Define measurable criteria for success and failure before beginning. What outcomes prove value?

Generate Your Hypothesis

> We believe [the proposed solution] will produce [expected impact] for [who/what will be impacted] by **[how much]** during/over **[how long]**.

A well-structured hypothesis provides clarity and focus, ensuring everyone understands what you're testing and how success will be measured.

Test & Analyze Results

Create the Test Plan

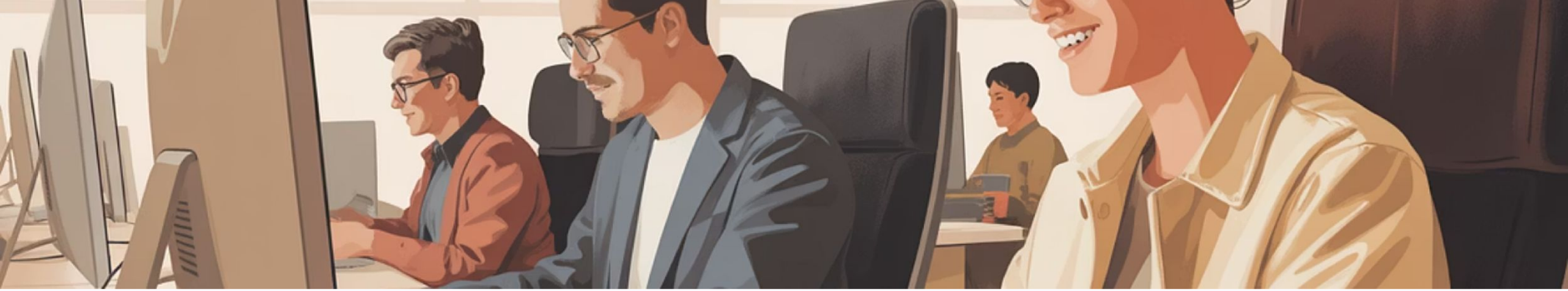
Design appropriate testing methods based on your hypothesis:

- User surveys and feedback
- Prototype testing sessions
- Proof of concept implementations
- A/B testing with control groups
- Pilot programs with limited scope

Analysis & Communication

Compare results against your initial success criteria:

- Document quantitative outcomes
- Capture qualitative learnings
- Identify unexpected insights
- Share findings transparently
- Make go/no-go recommendations



Case Study

Engineering Team Example

GitHub Copilot Pilot Program

GitHub Copilot Experiment Design

Problem Statement

Engineers spend excessive hours on repetitive coding tasks, reducing time available for complex problem-solving and innovation.

Proposed Solution

GitHub Copilot offers an AI-powered code assistant to automate repetitive tasks, stub out initial code blocks, and provide debugging guidance.

Why We Believe It Will Work

Senior engineers who can properly evaluate AI-generated code responses will gain significant productivity improvements by having an intelligent coding assistant.

Success Metrics

- Increased commit volume from pilot group engineers (measured via Jellyfish metrics)
- Improved engineer satisfaction (qualitative surveys)
- Reduction in time spent on repetitive coding tasks

GitHub Copilot Hypothesis

We believe [**GitHub Copilot**]

will produce [**higher productivity, faster task completion, higher satisfaction**]

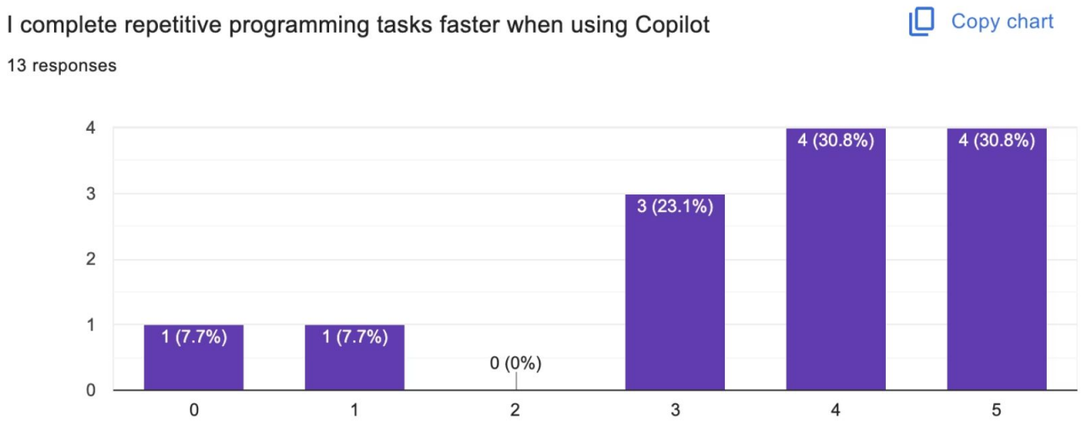
for [**engineers**]

by [**increasing the number of commits, 10% increased task efficiency, increased satisfaction**]

during/over [**30 days**]

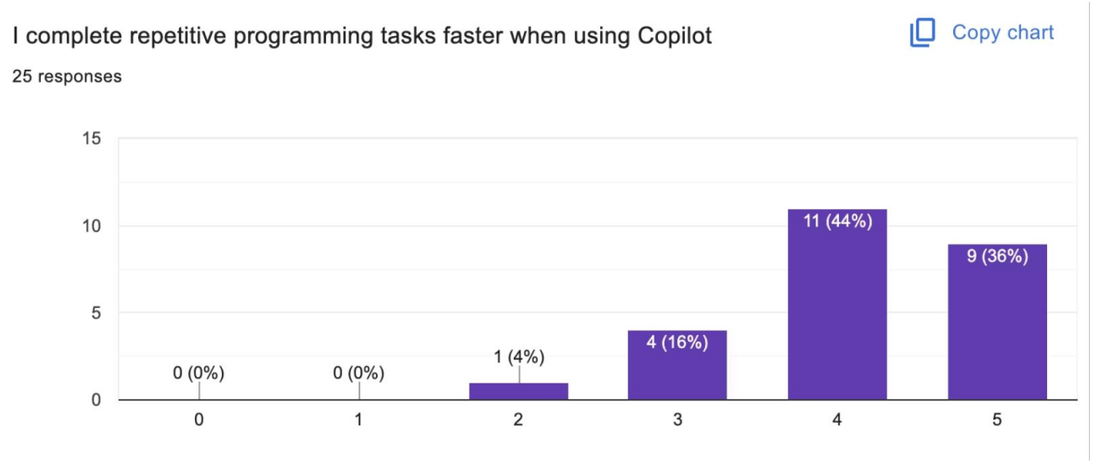
Weekly Survey Results: Satisfaction

Week 1 Response



Initial reactions showed cautious optimism with engineers still learning the tool's capabilities.

Week 4 Response

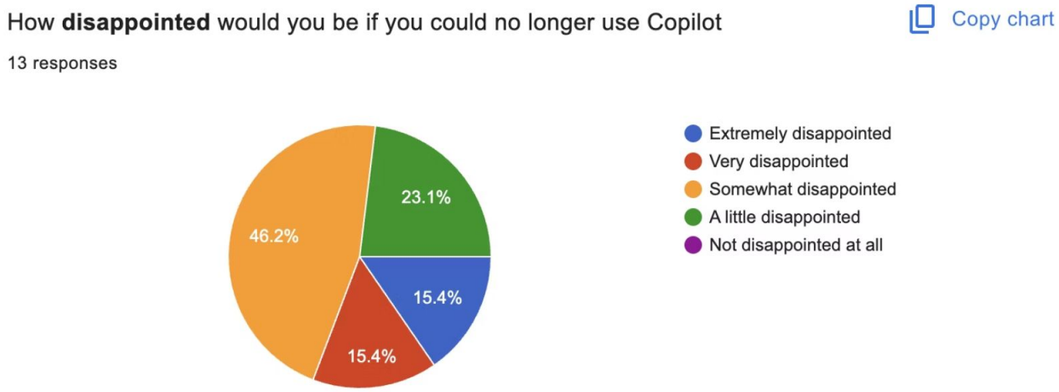


Satisfaction increased significantly as engineers gained confidence and discovered productivity gains.

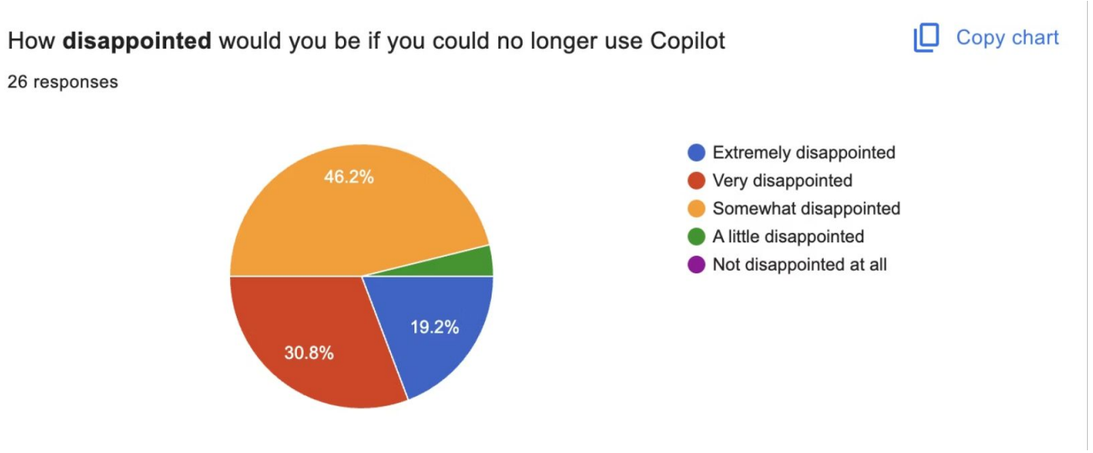
We surveyed 20+ engineers weekly throughout the four-week evaluation period to track sentiment as they gained comfort with the tool.

Weekly Survey Results: Productivity

Week 1 Sentiment



Week 4 Sentiment



Survey questions were framed using research frameworks such as Kano Analysis to understand feature satisfaction and perceived productivity impact.

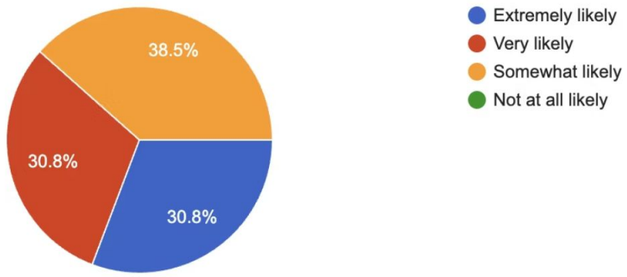
Experiment Outcomes & Rollout Decision

Week 1 Baseline

How likely are you to **recommend** that your team continues to use Copilot?

Copy chart

13 responses

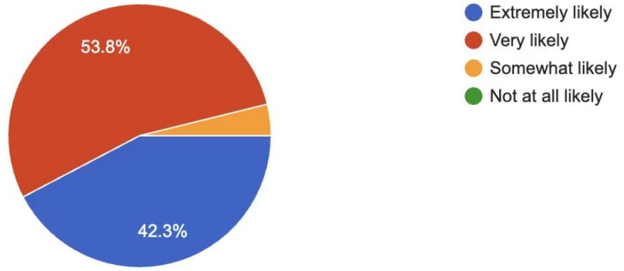


Week 4 Results

How likely are you to **recommend** that your team continues to use Copilot?

Copy chart

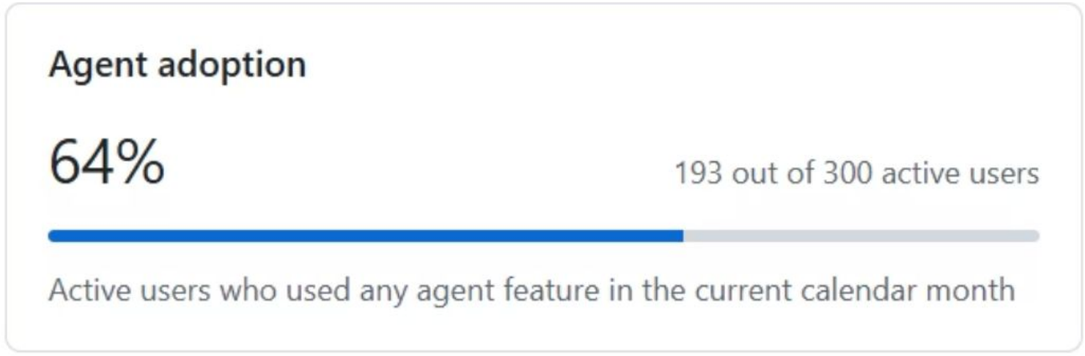
26 responses

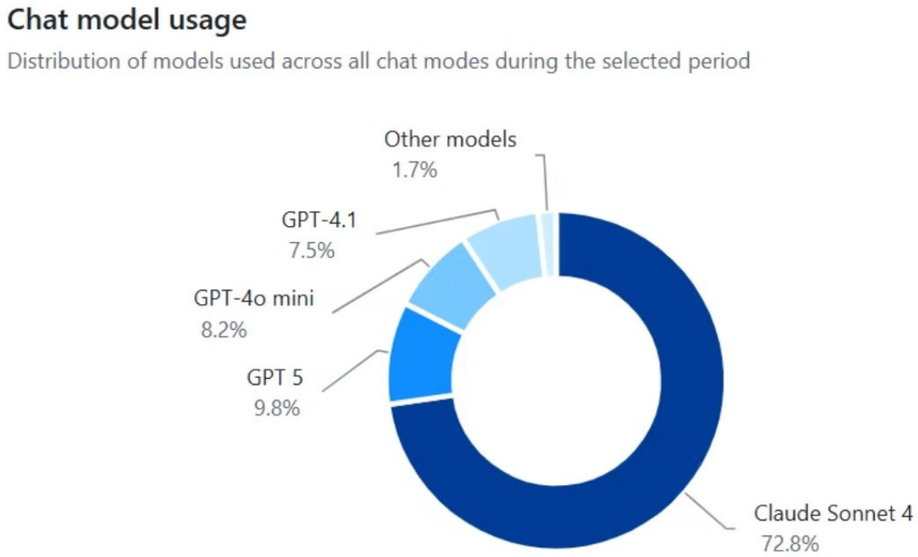


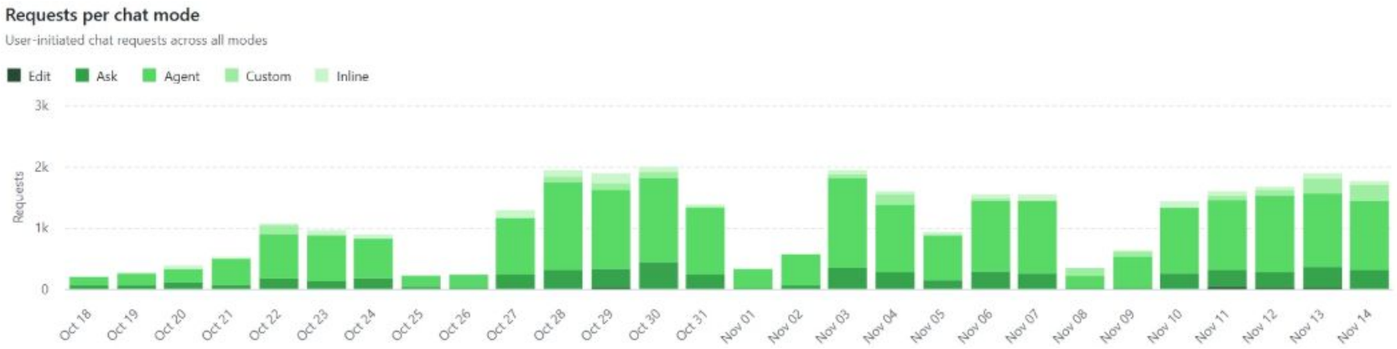
Hypothesis Validated

Senior engineers demonstrated measurably higher productivity and satisfaction. Budget was secured across all 3 Business Units mid-cycle, enabling rollout to 300+ engineers company-wide.

Copilot Usage Tracking Dashboard







Ongoing usage metrics provide visibility into adoption patterns, helping identify power users and opportunities for additional training.

Case Study

AI Hackathon

Organization-Wide Learning Event

AI Hackathon Experiment Design

Problem Statement

Employees needed direct, hands-on experience with generative AI technologies to understand capabilities and build practical skills.

Proposed Solution

Host a company-wide hackathon open to all employees, featuring training resources, dedicated experimentation days, and demonstration opportunities.

Why We Believe It Will Work

Providing dedicated time and space away from daily responsibilities allows employees to focus on learning the tools that will shape our collective future.

Success Metrics

- Cross-functional participation across teams and regions
- Number and quality of project demonstrations
- Projects advanced to product roadmaps
- Employee feedback on learning experience

Hackathon Structure

1

Pre-Event: Training Requirements

All participants completed free AWS generative AI courses (1-4 hours) to establish baseline knowledge before the event.

2

Registration Phase

Individual and team sign-ups accommodated different working styles. Fully remote format enabled global participation.

3

Day 1: Project Development

Teams focused on building functional prototypes, testing hypotheses, and iterating on their AI-powered solutions.

4

Day 2: Completion & Demos

Final development work followed by demonstrations. Participants could present via live demo, documentation, slides, or discussion format.

5

Evaluation & Awards

Hosts and judges attended each demo. All participants received monetary awards, with larger prizes for top projects.

Hackathon Outcomes

131

Total Participants

Strong cross-functional engagement across the organization

3

Regional Demo Events

Americas, Europe, and India sessions accommodated time zones

41

Projects Demonstrated

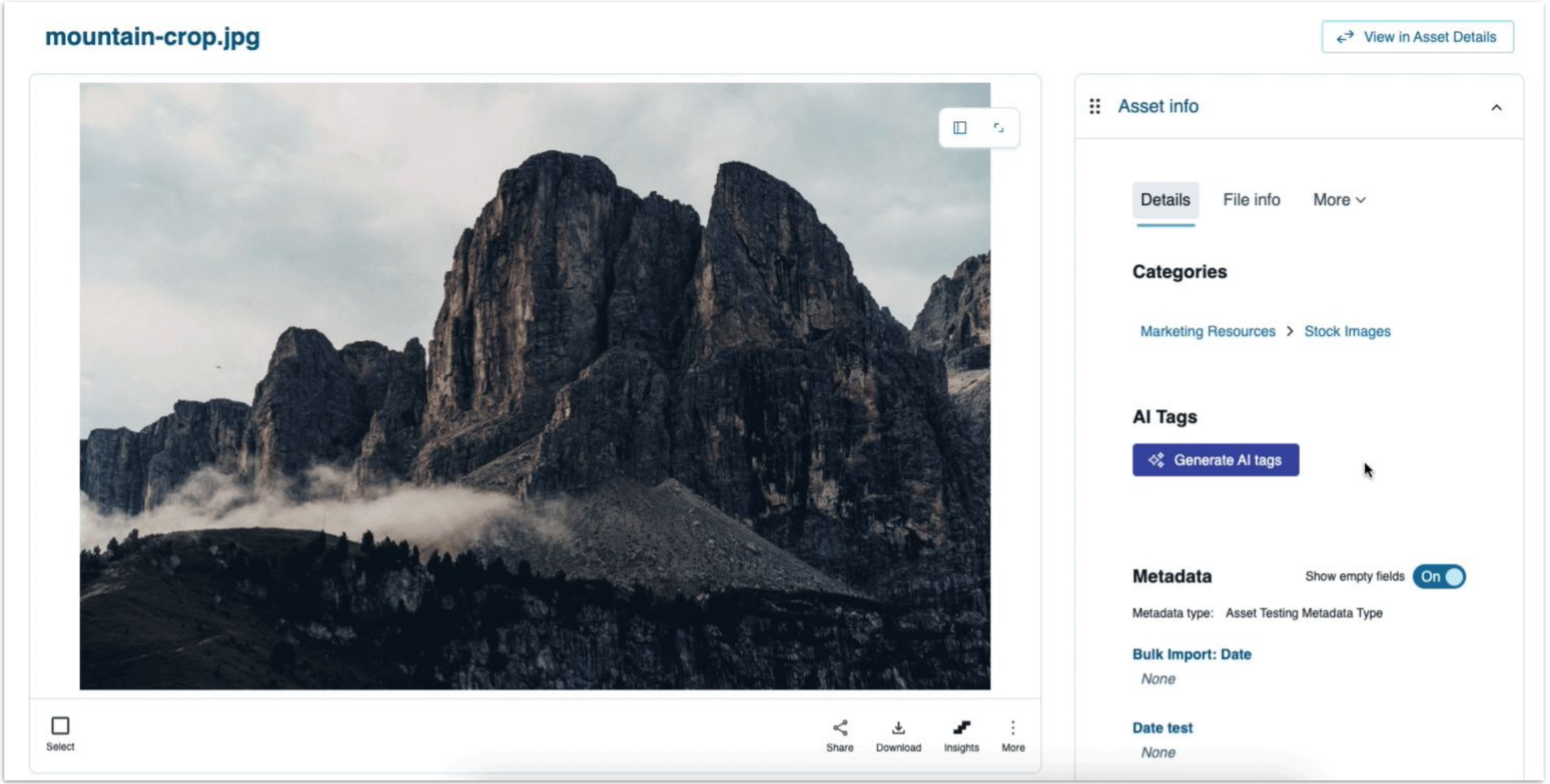
Diverse solutions ranging from automation to customer insights

5+

Roadmap Candidates

Projects advanced to product roadmap consideration

AI Hackathon - Released Feature



AI Hackathon - Analysis

Key Learnings & Outcomes

High participation across the organization

Training by AWS prior to the hackathon got great feedback

Low # of demos vs. # of participants - expected as the focus was learning

Overall a success and we are now hosting 2 org-wide hackathons per year

Overall Satisfaction Score

Case Study

AI Playground

Secure Internal AI Access

AI Playground Experiment Design

Problem Statement

Employees needed cost-effective access to generative AI without risking confidential information leakage to public models.

Proposed Solution

Deploy an open-source ChatGPT-like interface within our private AWS environment, routing requests through the OpenAI API with enterprise security controls.

Why We Believe It Will Work

An approved, secure environment configured with no model training enables safe AI experimentation while addressing data privacy concerns.

Success Metrics

- Usage metrics across departments (request volume)
- Employee understanding of generative AI capabilities (surveys)
- Zero security incidents or data leaks
- Cost per request vs. productivity gained

AI Playground Usage Metrics

Week of June 22, 2025

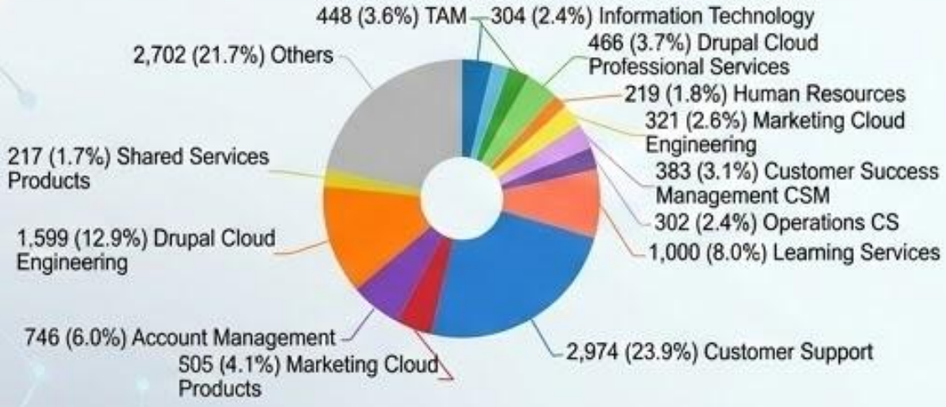
12,432
Total Prompt Count

1,554
Average Prompts Per Day

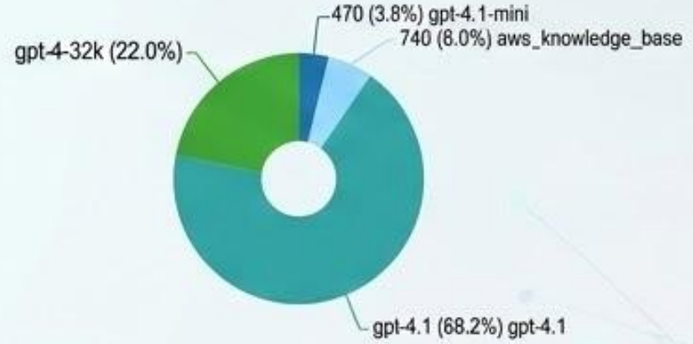
553
User Count

38
Department Count

Department-wise Usage Metrics



Model-wise Usage Metrics

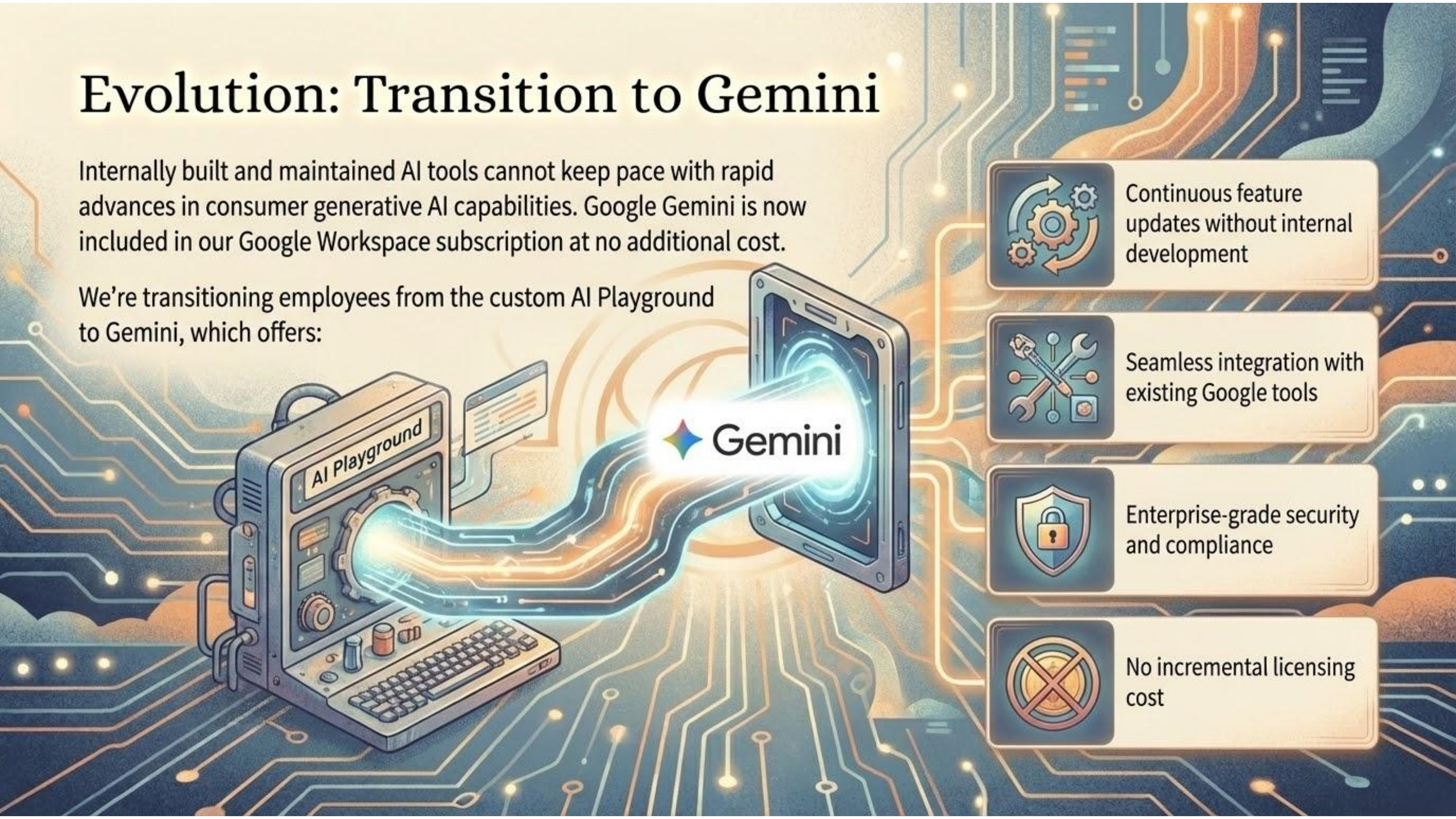


Usage data revealed consistent engagement across multiple departments, validating the need for secure, internal AI access.

Evolution: Transition to Gemini

Internally built and maintained AI tools cannot keep pace with rapid advances in consumer generative AI capabilities. Google Gemini is now included in our Google Workspace subscription at no additional cost.

We're transitioning employees from the custom AI Playground to Gemini, which offers:



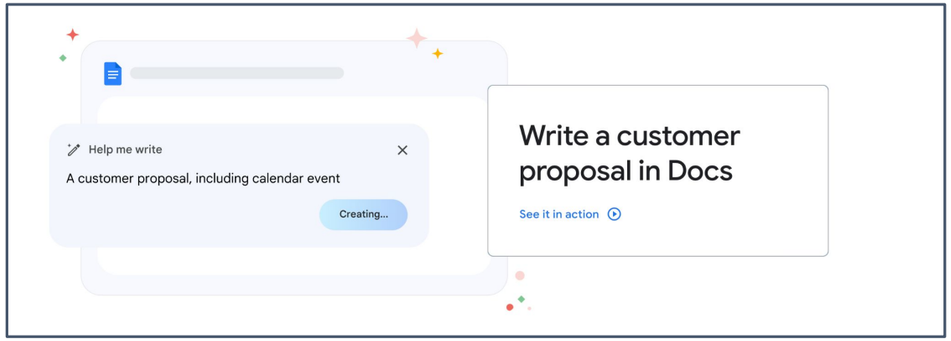
- Continuous feature updates without internal development
- Seamless integration with existing Google tools
- Enterprise-grade security and compliance
- No incremental licensing cost

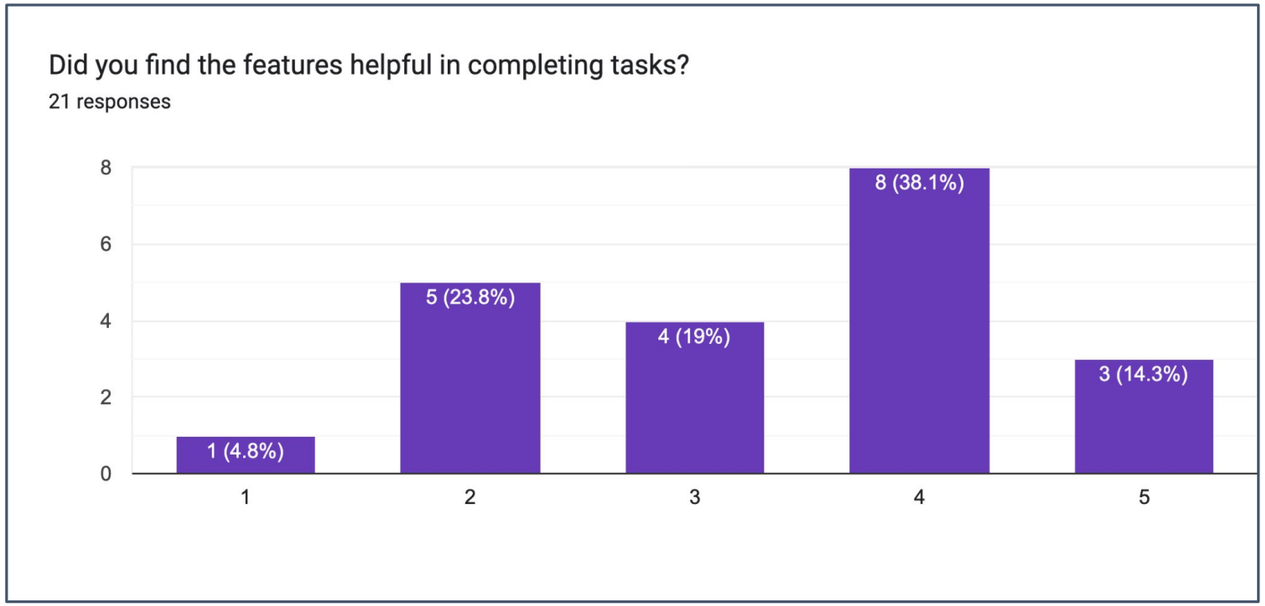
Confidently Say No

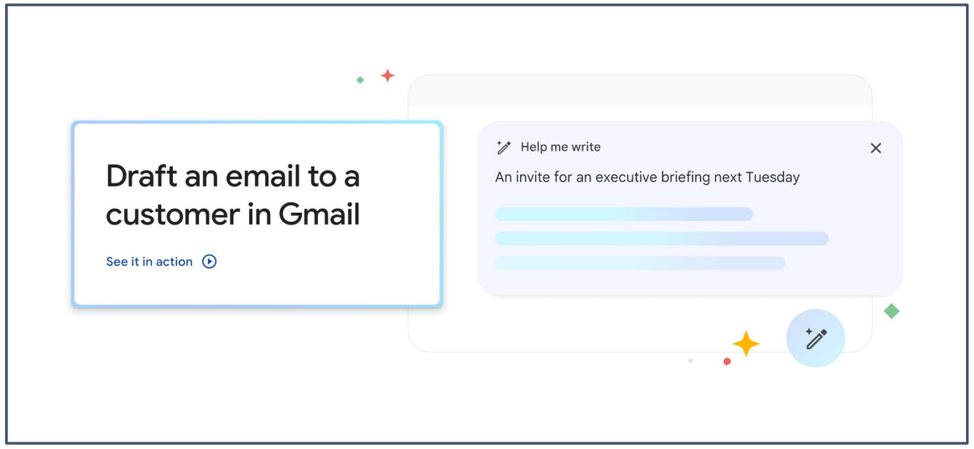
With Data-Driven Decisions

Failed Experiment: Google Workspace

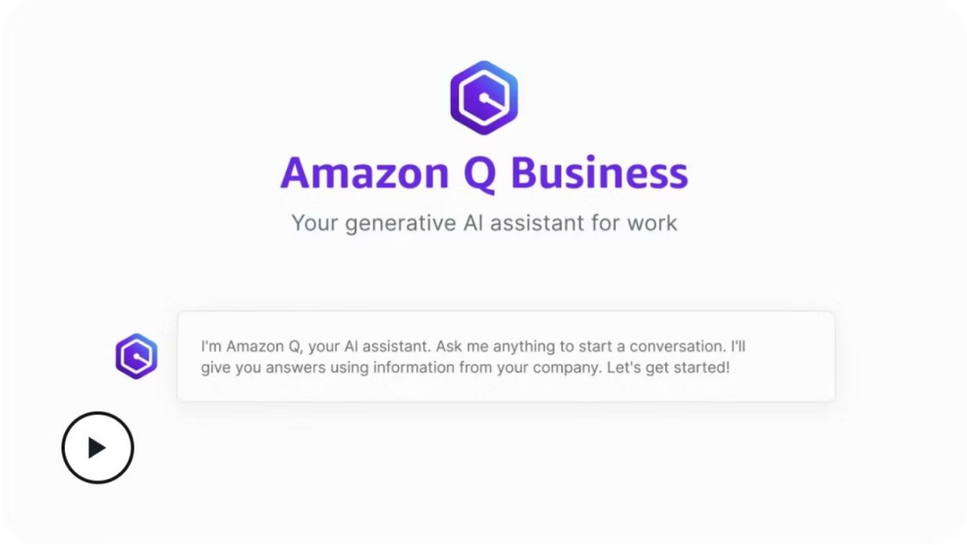
Trialed by 50 users for 30 days
Department budget approval was needed $360/user/yr
Leaders didn’t have enough ROI to make tradeoffs







Failed Experiment: Amazon Q



We invested three weeks implementing Amazon Q with augmented data sources from our public support center and private Confluence documentation.

The outcome: The system failed user acceptance testing due to inaccurate responses. Data sources required significantly more refinement than anticipated.

Rather than invest additional months, we paused the project with clear documentation for potential future revisit.

Measure What Matters

Start at the Top: Organizational Goals

Effective AI adoption requires alignment with executive priorities. Establish organization-wide annual goals and objectives focused on:

- Operational efficiency improvements
- Technology adoption rates
- Innovation metrics
- Cost optimization

These goals cascade down through leadership, creating accountability and focus at every organizational level.

Leading & Lagging Indicators

Lagging Indicators

- Time efficiency improvements
- Cost savings achieved
- Customer satisfaction scores

Leading Indicators

- Number of experiments conducted
- Tools evaluated
- Employee training completion

Long-Term Progress Tracking

Sustainable AI adoption requires multi-year perspective. Establish long-term goals that are:

Measurable: Quantifiable with clear success criteria

Time-bound: Multi-year targets with milestone checkpoints

Externally validated: Benchmarked against industry standards

Publicly committed: Shared transparently to drive accountability

Consider external assessment frameworks to provide objective measurement against industry standards.

TDWI Maturity Assessment

The TDWI (Transforming Data with Intelligence) organization offers comprehensive maturity assessments for data and analytics capabilities.

Access free assessments: https://tdwi.org/assessments

NEW! TDWI BI and AI Maturity Model Assessment

The BI and AI Maturity Model can help organizations improve their analytics programs. It provides a framework for companies to understand where they are, where they've been, and where they still need to go to support excellent data practices.

NEW! Data Ecosystems Maturity Model Assessment

TDWI created the Data Ecosystems Maturity Model and Assessment to help organizations understand the maturity of their data ecosystems and prioritize their next steps to leverage data for better operational efficiency, strategic decisions, and innovation.

Data Governance Maturity Model Assessment

TDWI created the Data Governance Maturity Model and Assessment to help organizations understand the maturity of their data governance efforts and prioritize their next steps to achieve greater success with data governance.

AI Readiness Assessment

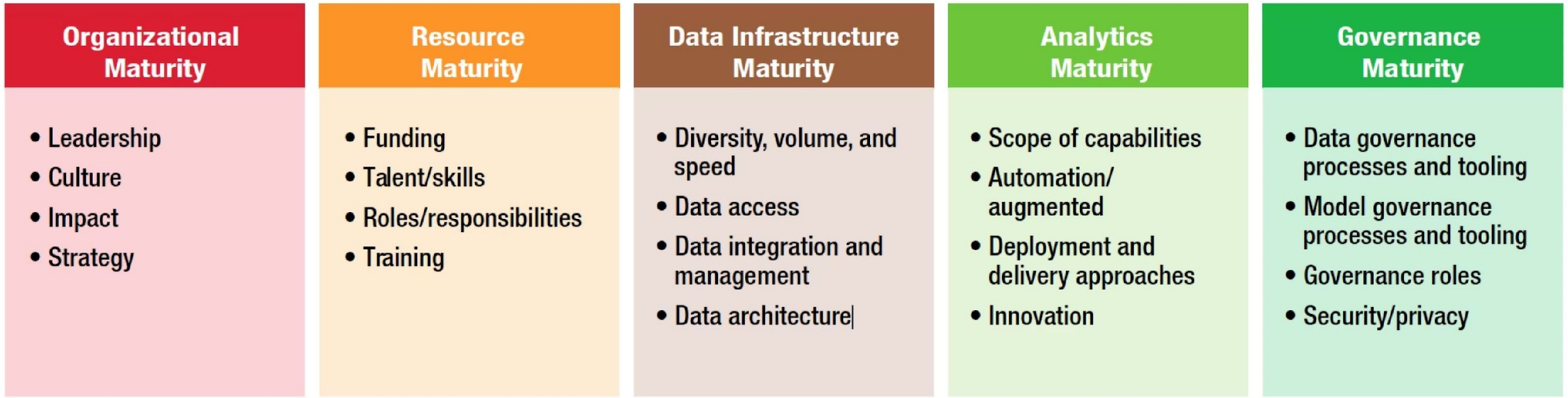
TDWI has created a new AI Readiness Assessment to help organizations better understand what it takes to support a successful AI program and how to advance their analytics efforts to sustain AI. The assessment measures an organization's AI readiness in an objective way across five dimensions.

Data Quality Maturity Model Assessment

TDWI has created a new Data Quality Maturity Model and Assessment to help organizations better understand the maturity of their data quality efforts and provide current, best-in-class insight and support. The assessment measures the maturity of a data quality program in an objective way across five dimensions.

TDWI Model Dimensions

The TDWI framework evaluates organizational maturity across five critical dimensions, each scored independently to identify specific strengths and improvement opportunities.



Assessment Progress & Targets

2024 target
72/100

Dec 2022
48/100
Stage: Early

Organization (Score: 11/20)
Executive sponsor emerges, pocket of analytic advocacy, organizational interest increasing

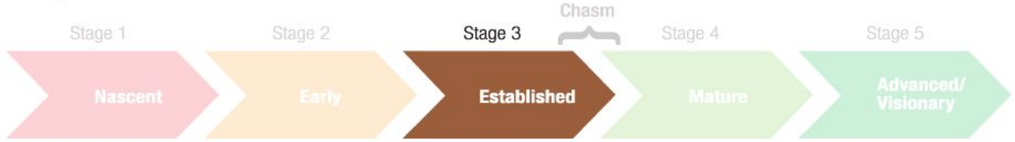
Resources (Score: 10/20)
Still often scattered or disperate

Data infrastructure (Score: 10/20)
Emerging vision for central infrastructure, data warehousing begins small

Analytics (Score: 10/20)
Reports and dashboards being created, BI tools purchased, more questions being asked

Governance (Score: 6/20)
None

Dec 2023
65/100
Stage: Established



Organization (Score: 14/20)
Data strategy created, business unity & data partnerships, general excitement about future

Resources (Score: 11/20)
Analytics group or COE, data literacy programs, dedicated data specific hires prioritized

Data infrastructure (Score: 15/20)
Warehouse in place, plan to move beyond structured data, data volume growing rapidly

Analytics (Score: 14/20)
Reports, dashboards, self-service visualization tools in place, predictive analytics being tested

Governance (Score: 11/20)
Organizational governance team or committee established but often still struggles to make progress

The Experimentation Flywheel

Run Experiments

Test hypotheses with controlled scope and clear success metrics

The Experimentation Flywheel

Run Experiments

Test hypotheses with controlled scope and clear success metrics

Prove Value

Demonstrate measurable ROI that justifies investment decisions

The Experimentation Flywheel

Run Experiments

Test hypotheses with controlled scope and clear success metrics

Prove Value

Demonstrate measurable ROI that justifies investment decisions

Reduce Risks

Identify and mitigate concerns through limited, safe pilots

The Experimentation Flywheel

Run Experiments

Test hypotheses with controlled scope and clear success metrics

Prove Value

Demonstrate measurable ROI that justifies investment decisions

Build Trust with Data

Create evidence-based confidence among stakeholders and users

Reduce Risks

Identify and mitigate concerns through limited, safe pilots

Each successful experiment fuels the next, creating self-reinforcing momentum toward organization-wide AI adoption.

Thank You

Deanna Ballew

Chief Product Officer at Listrak

deannaballew@gmail.com

LinkedIn.com/in/deannaballew/

Ben Dotte

VP of Engineering at Acquia

ben.dotte@acquia.com

linkedin.com/in/benjamin-dotte/