Fivetran

The AI tax: The hidden cost when AI isn't connected to your data

84% of organizations lose at least 10% of AI spend to avoidable work — and nearly 4 in 10 lose a quarter or more.

Research by Redpoint INSIGHTS

Executive Summary

83% of organizations spent at least $1 million on AI/ML in the past year. However, **84% of organizations lose at least 10% of their AI/ML spending to avoidable work and nearly 4 in 10 lose a quarter or more**. We call this the AI tax, and only 16% of organizations keep it under 10%.

> **The AI tax:** the share of AI/ML spending lost to avoidable work, such as fixing data problems, rebuilding integrations, reprocessing data, duplicating effort, and retraining models.

A high AI tax goes hand in hand with more stalled projects

Organizations with a High or Very High AI tax are **5x as likely** as Low-tax peers to have a quarter or more of their AI initiatives delayed or stalled due to data foundation gaps. Respondents also see potential to recover that time. **41% estimate that eliminating avoidable data** work could help AI initiatives deliver business value more than **25% faster**.

Reliance on custom integration is associated with a higher AI tax

53% of organizations relying mostly on custom, project-specific integrations report a High or Very High AI tax, compared with 40% of those using a standardized integration platform.

Lower-tax organizations report stronger AI outcomes

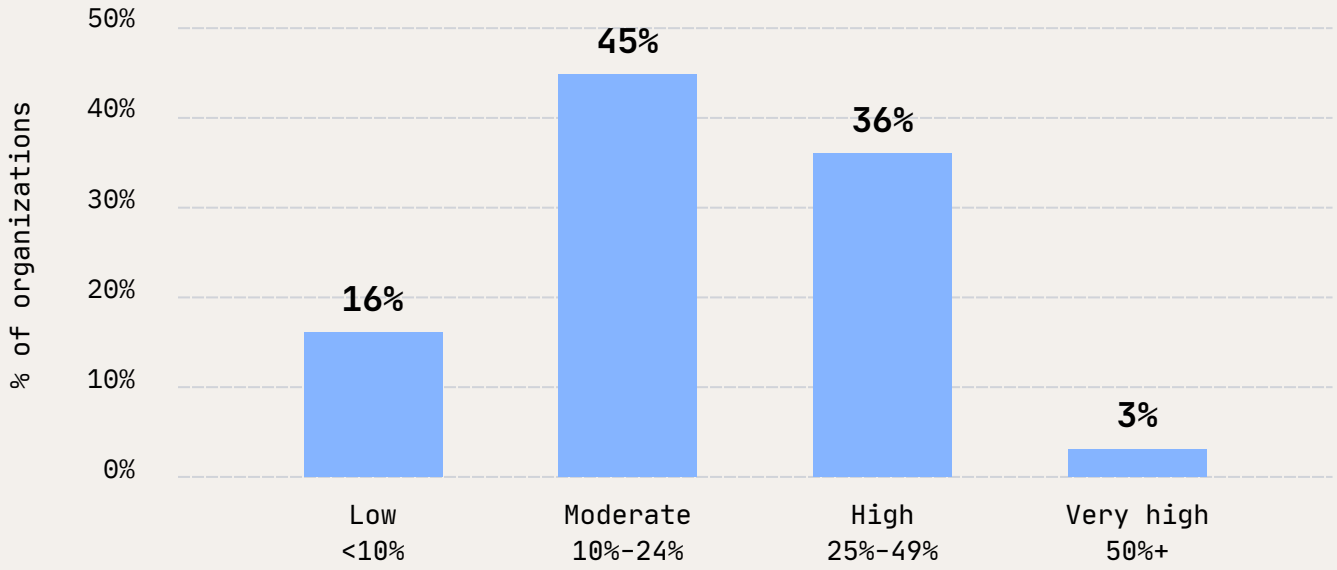
60% of Low-tax organizations rate their performance at moving AI initiatives into production over the past year as exceeding expectations, compared with 41% of those with a High or Very High tax. Low-tax organizations also more often report exceeding expectations for ROI and scaling AI initiatives across the organization: 62% and 56%, respectively, compared with just 40% for each among Very High-tax organizations.

Data foundations lead priorities for additional AI investment

Reducing the AI tax starts with the data foundation. Asked where they would invest additional AI/ML budget over the next 12 months, nearly one-third of respondents (31%) chose the data foundation, ahead of talent (17%) and AI models (16%).
Based on Fivetran's global survey of 500 senior IT, data, and AI/ML leaders, this report examines where the AI tax accumulates and how it relates to AI outcomes.

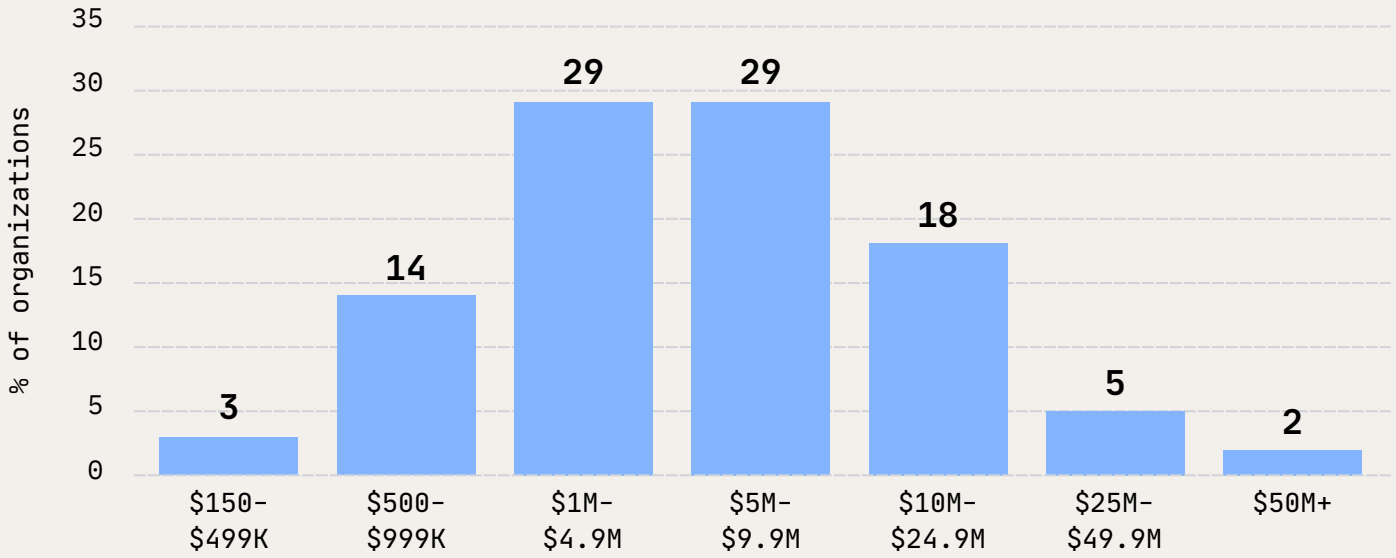
How large is the AI tax bill?

Share of AI/ML spend in the past 12 months that went to avoidable work



39% of organizations fall into the High or Very high AI tax tiers.
Based on 500 organizations.

How much did your organization spend on AI/ML initiatives during the past 12 months, in U.S. dollars?

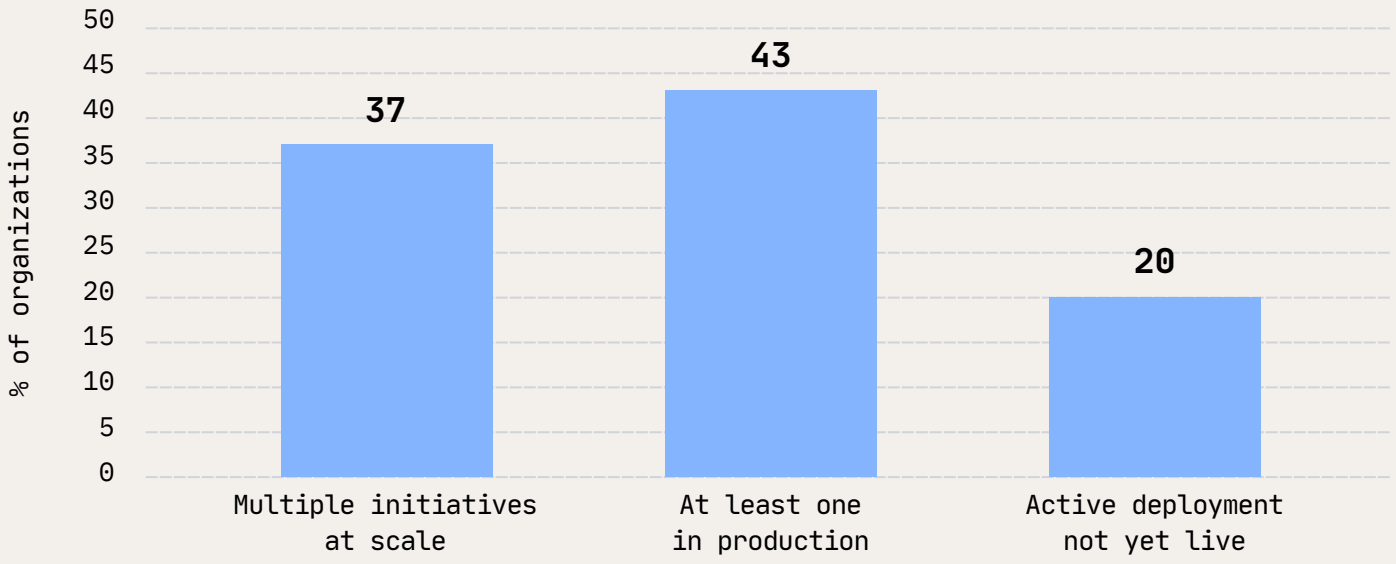


Big AI budgets have yet to translate into widespread scale

AI is now a seven-figure investment for most organizations. In the past 12 months, 83% spent $1 million or more on AI/ML initiatives, and more than half (54%) spent $5 million or more. Yet deployment has been uneven. While 80% of organizations have at least one AI/ML initiative live in production, only 37% report multiple initiatives at scale, meaning AI projects running in production and in sustained operational use.

Moving AI into production does not eliminate the cost of fixing data, rebuilding integrations, and reprocessing. Even organizations at scale face substantial rework. Among this group, nearly half (43%) say at least a quarter of AI spending goes to avoidable work.

Which best describes your organization's current stage of AI/ML deployment?



The AI tax starts in the data foundation

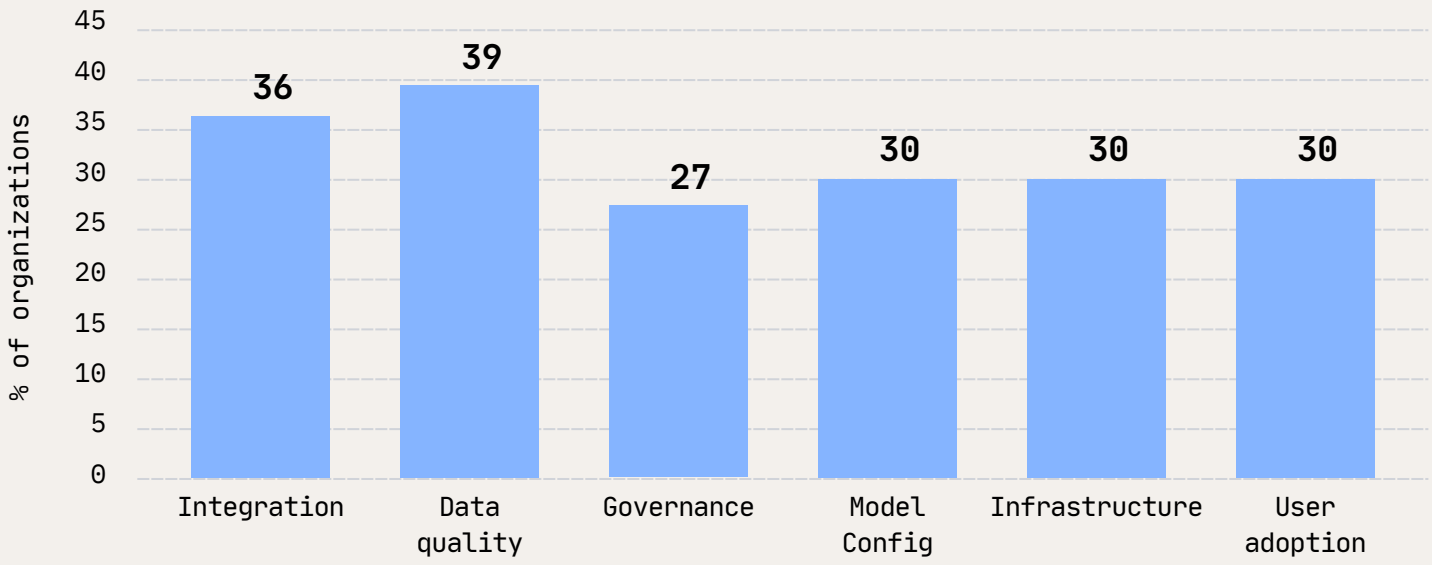
The biggest sources of unexpected AI work sit in the data foundation. Two-thirds (66%) cite data quality or integration as a top source of unexpected work, delay, or cost. Data quality and preparation lead at 39%, with integration close behind at 36%.

For 40% of organizations, the data foundation already accounts for more than 30% of total AI/ML spend. Fixing data quality is also the single biggest drain on engineering resources (38%), ahead of retraining models (33%), rebuilding integrations (32%), and reprocessing data (32%), from up to two choices.

Compounding these challenges, most organizations lack a complete view of what data work is costing them. Only 38% track these costs centrally, leaving nearly two-thirds with fragmented or incomplete visibility. Without a consolidated picture of spending on data preparation, integration, and remediation, organizations may struggle to identify where avoidable costs are accumulating.

Data problems also consume engineering time and delay AI initiatives once projects are underway.

Where do AI initiatives most commonly encounter unexpected work, delay, or cost?



A high AI tax is associated with more stalled initiatives

Organizations with a High or Very High AI tax are 5x as likely as Low-tax peers to have at least a quarter of their AI initiatives delayed or stalled by data foundation gaps.

Share of organizations with at least 25% of AI initiatives delayed or stalled by data foundation gaps, by AI tax tier

AI Tax Tier	% with 25% of initiatives delayed
Low AI Tax	17%
Moderate AI Tax	50%
High AI Tax	**86% ↑**
Very High AI Tax	**87% ↑**

Higher-tax organizations also face longer delays. When data foundations aren't ready, 75% of High- and Very High-tax organizations report typical delays of 3 weeks or more, compared with 46% of Low-tax organizations.

The work behind those delays often exceeds expectations. 66% of organizations needed moderately or significantly more data work than planned once projects began, with data preparation the most commonly underestimated cost, cited by 33%.

Poor-quality data adds further work and undermines confidence in AI. 98% of organizations have experienced consequences from models trained or run on incomplete, stale, or poor-quality data. 57% report inaccurate outputs or predictions, 57% have had to retrain or recalibrate models, and 52% say stakeholders lost trust in AI results.

Ownership and alignment also matter. 42% also cite unclear ownership or cross-team misalignment as a reason AI initiatives stall or underperform.

Avoidable work consumes engineering capacity as well as budget

58% of organizations say technical teams spend at least a quarter of their time supporting AI/ML on avoidable data work, including remediation, reprocessing, retraining, and rebuilding integrations.

For a 10-person data engineering team dedicated to AI, losing a quarter of its time to avoidable work represents approximately $315,000 a year in salary costs — capacity that could support AI deployment and expansion[1].

U.S. teams report the greatest engineering burden

22% of U.S. technical teams spend half or more of their time supporting AI/ML on avoidable work, compared with 14% in EMEA and 11% in APAC.

[1] *Illustrative estimate based on Built In's average U.S. data-engineer base salary. Assumes full-time AI support and excludes benefits and overhead.*

Over half of organizations relying mostly on custom integrations report a high AI tax

53% of organizations relying mostly on custom, project-specific integrations lose at least a quarter of their AI/ML spending to avoidable work, compared with 40% of those using a standardized integration platform.

Selective custom work does not show the same burden. Among organizations using standardized approaches with some custom builds, 35% report a High or Very High AI tax. The distinction is whether custom integration supplements a standardized foundation or serves as the primary approach.

U.S. organizations rely more heavily on custom integration

U.S. organizations are more than twice as likely as those in EMEA or APAC to rely mostly on custom integration. They also report the highest AI tax, with 45% in the High or Very High tiers versus 34% in EMEA and 37% in APAC.

Organizations relying mostly on custom integrations are twice as likely to face data-related delays of more than a month: 40% vs. 20% of those using a standardized platform.

Reducing avoidable data work could accelerate AI time to value

41% of respondents estimate that eliminating avoidable data work could help their AI/ML initiatives deliver value more than 25% faster, and another 49% expect gains of 11% to 25%.

Lower-tax organizations also report stronger production outcomes. 60% of Low-tax organizations rate their performance at moving AI initiatives into production as exceeding expectations, compared with 41% of those with a High or Very High tax.

The relationship extends to financial returns. Nearly two-thirds (62%) of Low-tax organizations say AI ROI exceeded expectations, compared with 40% of organizations with a Very High AI Tax. While the findings don't establish causation, they suggest that organizations spending less on avoidable data work are also more likely to realize stronger returns on their AI investments.

Data foundations lead priorities for additional investment

Asked where they would put additional AI/ML budget over the next year, 31% chose the data foundation, ahead of talent (17%) and models (16%).

---

9 in 10 organizations expect removing avoidable data work to speed up AI time to value by more than 10%.

Budget influencers report a higher AI tax than budget owners

47% of respondents who jointly influence or recommend AI budget allocations report a High or Very High tax, compared with 36% of those who directly own or approve budgets. Only 4% of influencers report a Low tax, compared with 20% of budget owners.

U.S. organizations expect the greatest time-to-value gains

50% of U.S. organizations expect AI time to value to improve by more than 25% if avoidable data work is eliminated, compared with 35% in EMEA and 38% in APAC.

What data leaders should do to lower the AI tax and improve AI returns

The AI tax is widespread, predictable, and costly. 84% of organizations lose at least 10% of AI spend to avoidable work, and a higher tax goes hand in hand with more stalled initiatives and weaker returns. The tax grows when every AI project rebuilds its own connection to data instead of relying on a shared, trusted foundation. But it isn’t an inevitable cost of adopting AI. Three priorities can help organizations reduce avoidable work and improve returns.

Priority 1: Standardize data integration

Standardize the integration work that recurs across initiatives (the pipelines, connections, and maintenance processes teams otherwise rebuild for each project), and reserve custom builds for requirements the shared foundation can't meet. Organizations that standardize most of their integration, whether on a single platform or with selective custom builds, report a High or Very High AI tax far less often than those relying mostly on custom integrations: 37% vs. 53%.

Priority 2: Track data costs and plan for data work upfront

Make data work a planned cost, not a surprise. Build data preparation and integration into every AI business case as explicit line items, since data preparation is the most commonly underestimated cost (33%). Then track the AI tax as a standing metric: measure the share of AI spend lost to avoidable work by team and initiative, and review it alongside ROI.

Priority 3: Align business and technical leaders on success

Before an initiative launches, have business and technical leaders agree on what success looks like and name one owner accountable for it. Unclear ownership or misalignment is the most common reason AI initiatives stall (42%), and organizations with misaligned leaders are half as likely to say AI ROI exceeded expectations (27% vs. 57% among very aligned organizations).

How organizations measure AI success

Organizations' top technical metrics for judging AI are data pipeline reliability (42%) and data quality (40%). Their top business metrics are adoption or usage (49%) and revenue impact (48%).

AI returns are shaped long before an initiative reaches scale. Connect AI to a shared, trusted foundation once, and every project starts with the data and context it needs instead of starting from scratch.

See what a trusted data foundation looks like in practice

Learn how organizations are using Fivetran to connect enterprise data and build the foundation for AI.

Methodology

This report is based on a global survey of 500 senior IT, data, and AI/ML leaders conducted from August 24 to September 7, 2026, with statistical significance tested at a 95% confidence level across the U.S. (40%); France (11%), Ireland (10%), Germany (10%), and the UK (10%); and Japan (9%), Australia (5%), and Singapore (5%).

The survey examined 4 core areas of AI/ML operations: investment and deployment, data foundations and integration, avoidable work and delays, and alignment, performance, and ROI. Findings were analyzed to identify global patterns in AI spending, engineering effort, integration approaches, and the relationship between avoidable work and AI outcomes. Company size skewed toward organizations with fewer than 2,000 employees, including 200–499 employees (15%), 500–999 (34%), 1,000–1,999 (37%), 2,000–9,999 (14%), 10,000–49,999 (0.6%), and 50,000+ (0.4%).

The survey captured a broad mix of industries, including manufacturing (24%), financial services (22%), technology (18%), retail/CPG (17%), healthcare (11%), and hospitality (6%). Respondents held senior roles such as CTO/VP of Engineering (44%), CEO/President/Co-founder (15%), VP/Head of AI/ML (13%), CIO/VP of IT (11%), Chief Data Officer/Chief Data & Analytics Officer (7%), and VP/Head of Data (6%).

The AI tax is the self-reported share of AI/ML spending over the past 12 months consumed by avoidable work. Organizations were grouped into 4 tiers: Low (less than 10%), Moderate (10–24%), High (25–49%), and Very High (50% or more).

About Fivetran

Fivetran is the company that connects your AI to your data, so customers can act with confidence. Fivetran moves data from more than 750 sources and gives it the definitions, logic and lineage that make it trustworthy. It connects that data to any agent, tool or team through open infrastructure, giving customers the freedom to choose their clouds, data platforms and AI tools while maintaining control of their data. Nearly 10,000 customers, including Anthropic, Canva, OpenAI, Perplexity, Siemens and Verizon, rely on Fivetran, while more than 100,000 data teams use dbt. Fivetran's offerings include Fivetran data movement, dbt, Managed Data Lake and dbt Wizard, with dbt Core and its open-source community remaining central to the company's commitment to open infrastructure.

Learn more at Fivetran.com and getdbt.com, or follow Fivetran on LinkedIn and X, and both Fivetran and dbt on YouTube. We’re just getting started.