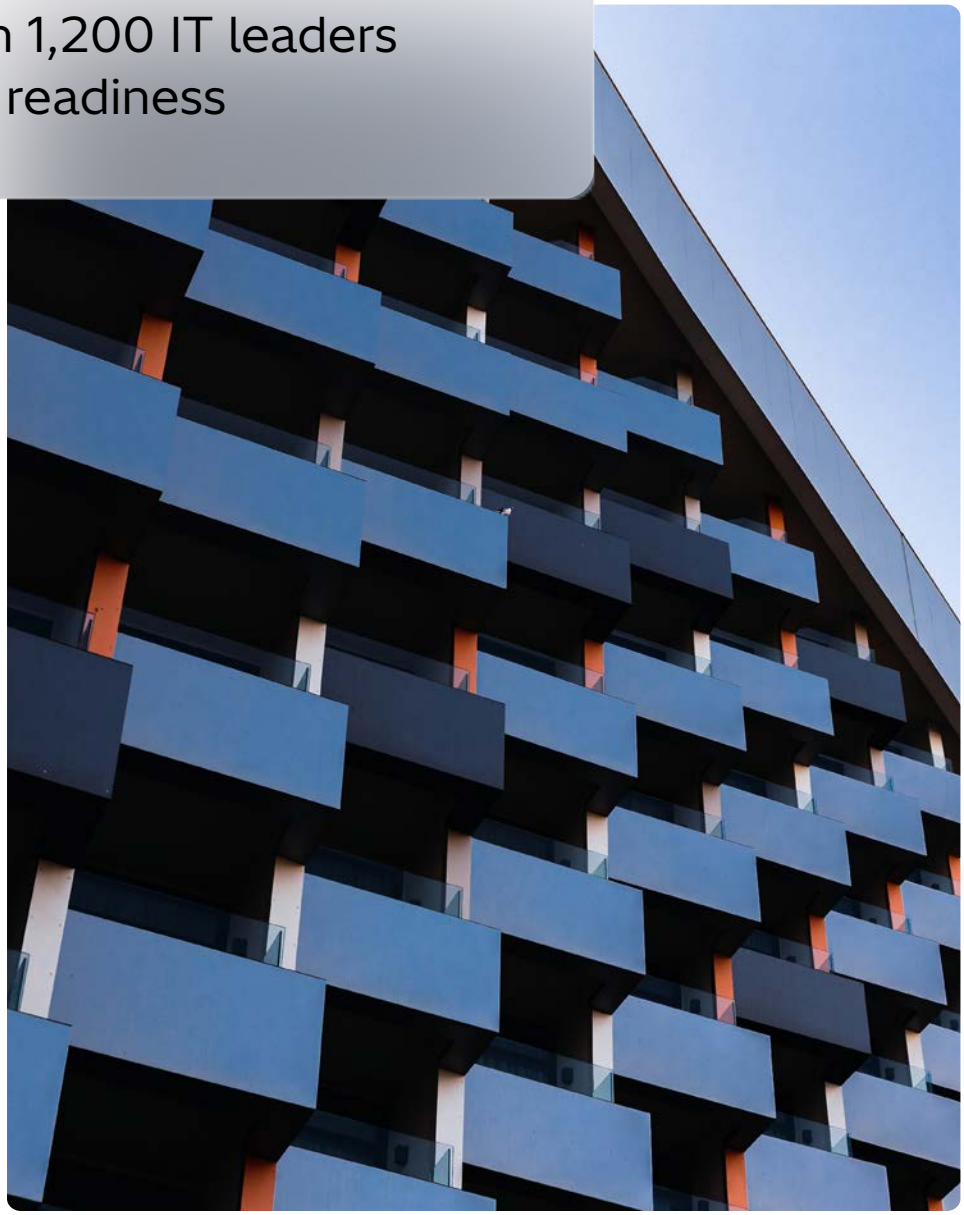
Hitachi Vantara

State of Data Infrastructure Global Report 2025

From Fragile to Optimized

What more than 1,200 IT leaders reveal about AI readiness



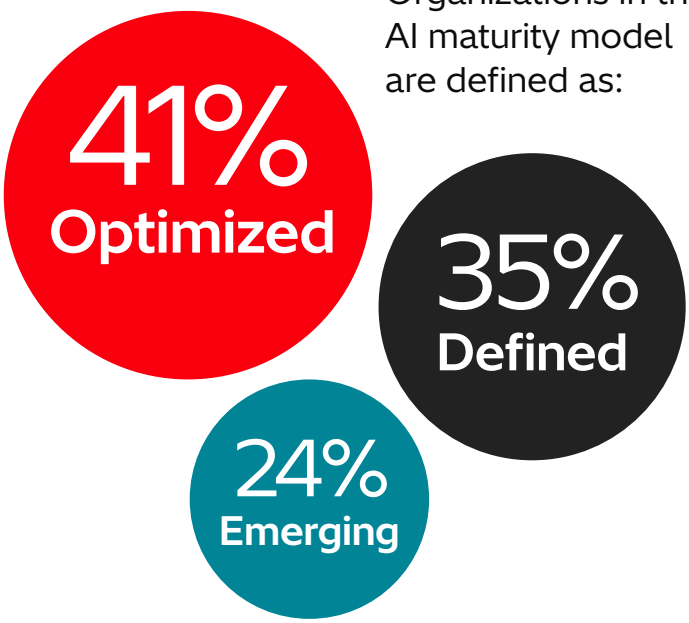
HITACHI

Global enterprises are drowning in petabytes — juggling hybrid, private, and public clouds while complexity skyrockets. The result? A fragile ecosystem that keeps executives awake at night.

At the same time, the transition to AI-first everything is relentless. Over the next two years, IT leaders expect to increase AI investment by 70% and expect to increase hiring of skilled workers by 68%. Yet, 37% of organizations still can't calculate ROI on AI initiatives — though 58% say they have established ROI, and 71% are treating AI like R&D expecting returns over the long term. Meanwhile, internal AI threats have surged: concern about breaches from internal AI jumped from 31% to 41% year-over-year now nearly matching external AI-enabled attacks (43%). All while in the midst of one of the most threatening and polarizing market and geo-political environments since the Cold War.

In this year's *State of Data Infrastructure Global Report*, Hitachi Vantara's proprietary research draws a sharp line between enterprises that will thrive in this unique market, and those that will lag, with a model defining AI maturity. The research reveals 24% of organizations are **Emerging**, 35% are **Defined**, and 41% are **Optimized**. **Optimized** organizations have resilient infrastructure with clean data powering AI-driven operations and measurable returns. **Emerging** firms are risk-averse, skill-starved, and clinging to manual processes that leave them unable to scale. **Defined** organizations linger in the gray area between the two, risking irrelevance with marginal progress, but lack the talent and strategy to execute.

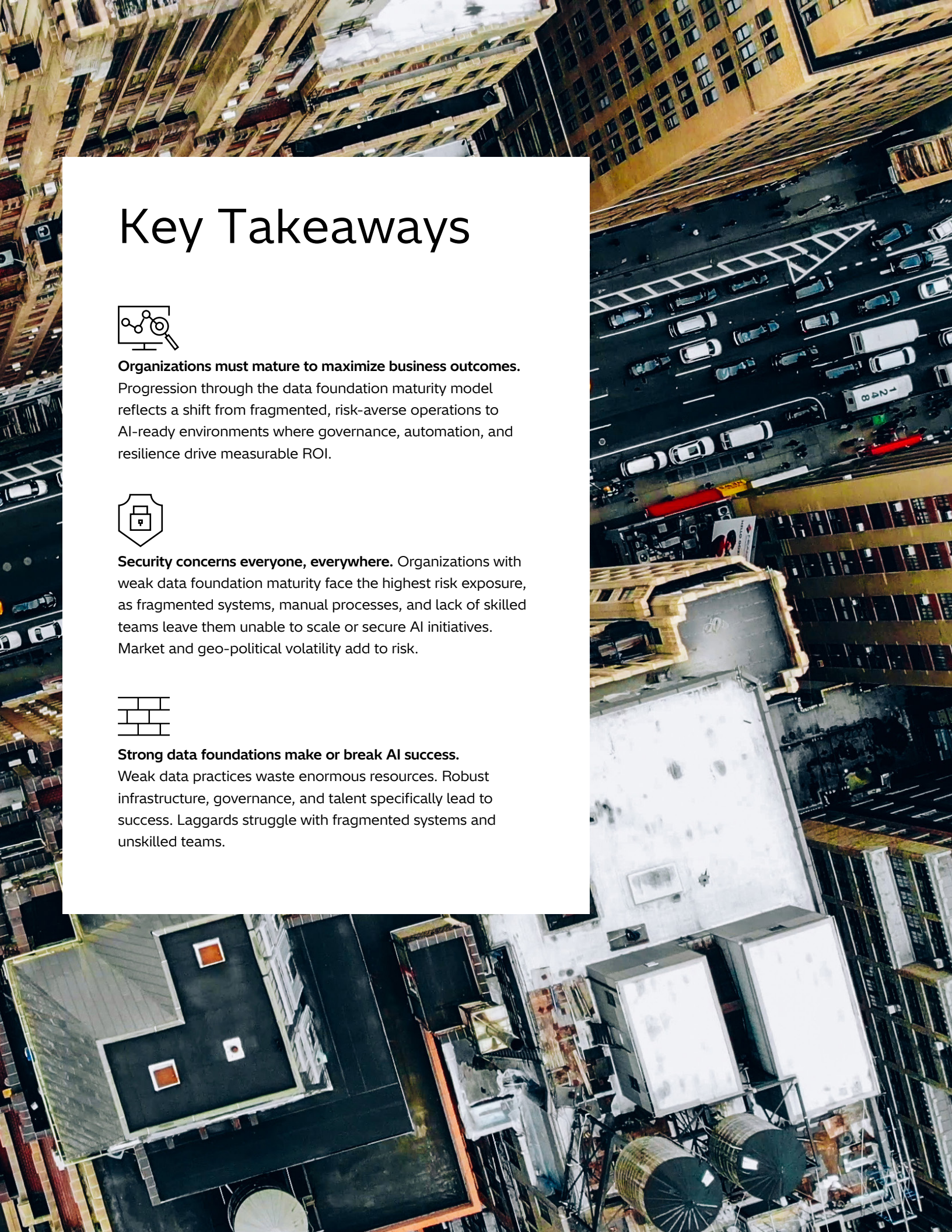
Organizations in the AI maturity model are defined as:



The difference among these organizations is clear and falls on the shoulders of leadership. Critically, this distribution is not correlated with budget, company size, or data volume — maturity is a strategic choice driven by leadership prioritization.

Those who invest in infrastructure, governance, and trusted partnerships win big.

The rest will be left counting losses, while competitors turn insight into dollars.



Key Takeaways

Organizations must mature to maximize business outcomes. Progression through the data foundation maturity model reflects a shift from fragmented, risk-averse operations to AI-ready environments where governance, automation, and resilience drive measurable ROI.

Security concerns everyone, everywhere. Organizations with weak data foundation maturity face the highest risk exposure, as fragmented systems, manual processes, and lack of skilled teams leave them unable to scale or secure AI initiatives. Market and geo-political volatility add to risk.

Strong data foundations make or break AI success. Weak data practices waste enormous resources. Robust infrastructure, governance, and talent specifically lead to success. Laggards struggle with fragmented systems and unskilled teams.

Table of Contents

The Data Infrastructure Maturity Model

From Emerging to Optimized

Hitachi Vantara's proprietary research introduces a comprehensive framework for assessing data infrastructure maturity across six critical dimensions: Scalability, Reliability, Security, Governance, Sovereignty, and Sustainability. Organizations fall into three distinct stages — **Emerging**, **Defined**, and **Optimized**.

Emerging firms are typically risk-averse, constrained by limited skills, and reluctant to form partnerships; they often restrict AI adoption and struggle to quantify ROI. **Defined firms** have established foundational capabilities but advance slowly due to skill gaps and inconsistent strategies. In contrast, **Optimized firms** prioritize investment in talent and capacity, embed governance and reliability into operations, emphasize data quality, and maintain a clear executive vision for AI, supported by dedicated teams.

Six critical dimensions for assessing data infrastructure maturity:

Organizations fall into three distinct stages of maturity:

- Emerging
- Defined
- Optimized

Ready to find out where your organization stands — and how to move to the next level?

Dive into the full model and start building a data infrastructure that powers innovation and competitive advantage.

Complexity and Risk Are Out of Hand

Complexity is growing at a pace that feels unmanageable, according to U.S. business leaders. The rapid storage growth, data volume explosion, and platform proliferation are all converging to increase risk exposure. The top security concerns across the board are:

- **Data breaches from human error (~43%)**
- **Loss of customer or investor confidence (~37%)**
- **Regulatory penalties (~34%)**

Adding to the literal complexity of technology is the abstract complexity with which U.S. companies perceive their own capabilities. Case in point: While U.S. respondents trust in their own ability to quickly detect breaches, they also believe AI might assist attackers more than hinder them. Furthermore, these same organizations reported strong trust in their security policies, yet believe complexity will get in the way of threat detection. Adding to the challenge is that 55% of AI experts have little or no confidence in U.S. companies to develop and use AI responsibly, according to Pew Research — presenting reputation risks to firms with little visibility into their data, or a lack of skill to explain AI output.

55%

of AI experts

have little or no confidence in U.S. companies to develop and use AI responsibly.

Pew Research

Across the pond, legal and ethical concerns are more acute, with 45% of UK respondents citing them as top sticking points to AI adoption compared to the global average of 32%.

The anxiety is widespread: 57% of respondents say data loss would be catastrophic to their business. Fifty-five percent say complexity makes breach detection harder, and 52% believe AI benefits attackers more than defenders. What's more, confidence in employee AI safety dropped significantly: only 65% of leaders are confident employees use AI safely and only for business purposes, down from 77% in 2024. Data security as an AI implementation concern runs above the global average (56%) in the U.S. (63%), UK (67%), and India (67%).

As a result of this risk and complexity, organizations have a responsibility to be transparent about their abilities to prevent attacks and recognize that the experience may be humbling.

57% say data loss would be catastrophic to their business

52% believe AI benefits attackers more than defenders

46%

of leaders say that if executives saw the **true fragility** of their current environment, **it would keep them up at night.**

The Billion Dollar Wake-Up Call

MIT research reveals that 95% of organizations are getting zero return on $30–40 billion in GenAI investment — not because the models are bad, but because the underlying infrastructure can't support them. This mirrors our own findings: organizations with weak data foundations struggle to scale AI, while those with mature, integrated systems extract real value. The divide isn't about who adopts AI first. It's about who's ready.

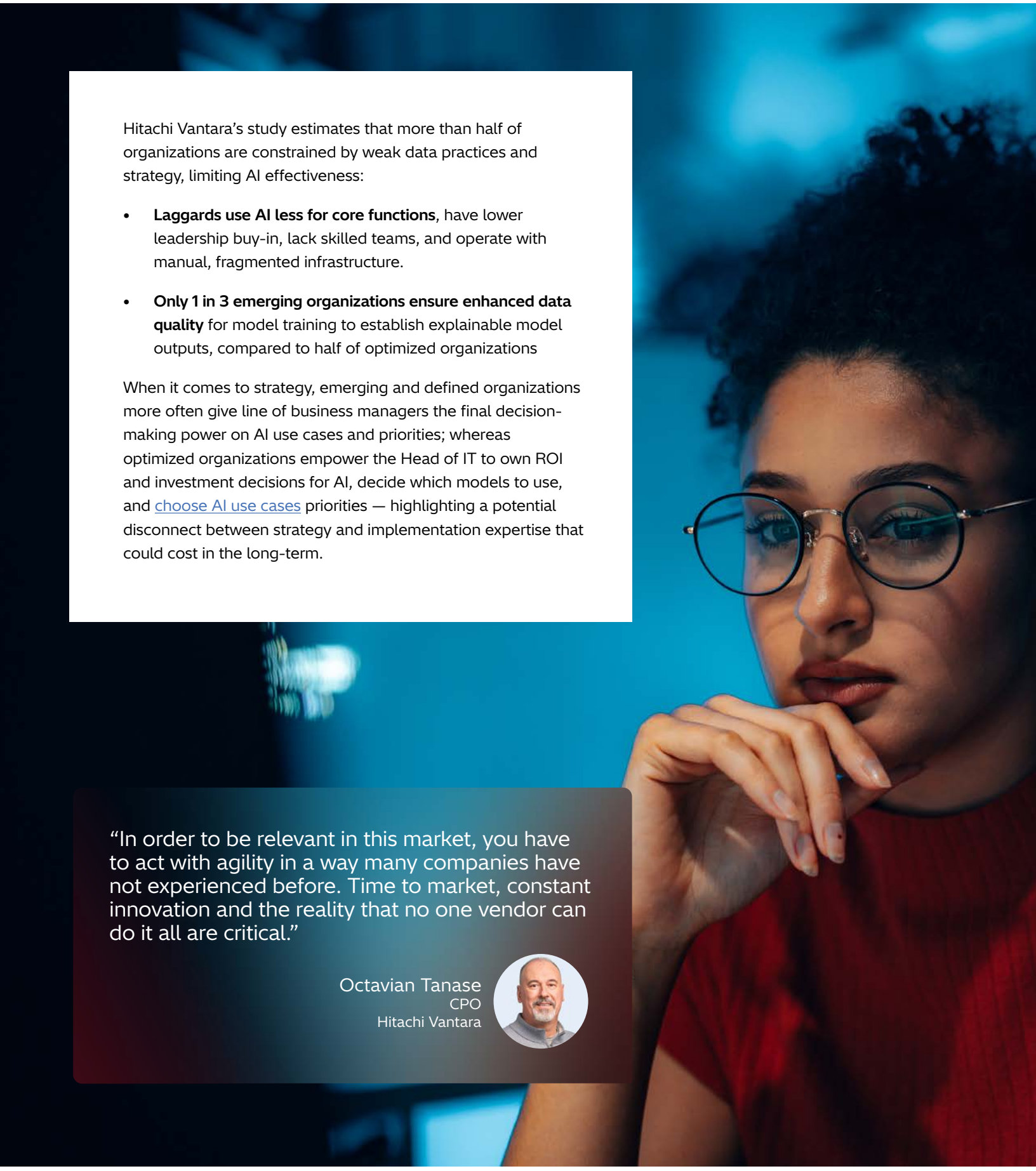
When asked why AI projects were successful, optimized organizations in Hitachi Vantara's study were twice as likely to cite data quality as a key component than emerging organizations (48% vs. ~25% for Emerging organizations).

Quick-win use cases (generative tools and AI embedded in existing software) tend to be most successful when calculating ROI, whereas owned or rented LLMs deliver mixed results. Inaccurate outputs and insufficient data quality are still among the leading reasons AI projects fail – consistent with last year's failure drivers.

Hitachi Vantara's study estimates that more than half of organizations are constrained by weak data practices and strategy, limiting AI effectiveness:

- **Laggards use AI less for core functions**, have lower leadership buy-in, lack skilled teams, and operate with manual, fragmented infrastructure.
- **Only 1 in 3 emerging organizations ensure enhanced data quality** for model training to establish explainable model outputs, compared to half of optimized organizations

When it comes to strategy, emerging and defined organizations more often give line of business managers the final decision-making power on AI use cases and priorities; whereas optimized organizations empower the Head of IT to own ROI and investment decisions for AI, decide which models to use, and choose AI use cases priorities — highlighting a potential disconnect between strategy and implementation expertise that could cost in the long-term.



"In order to be relevant in this market, you have to act with agility in a way many companies have not experienced before. Time to market, constant innovation and the reality that no one vendor can do it all are critical."

Octavian Tanase
CPO
Hitachi Vantara

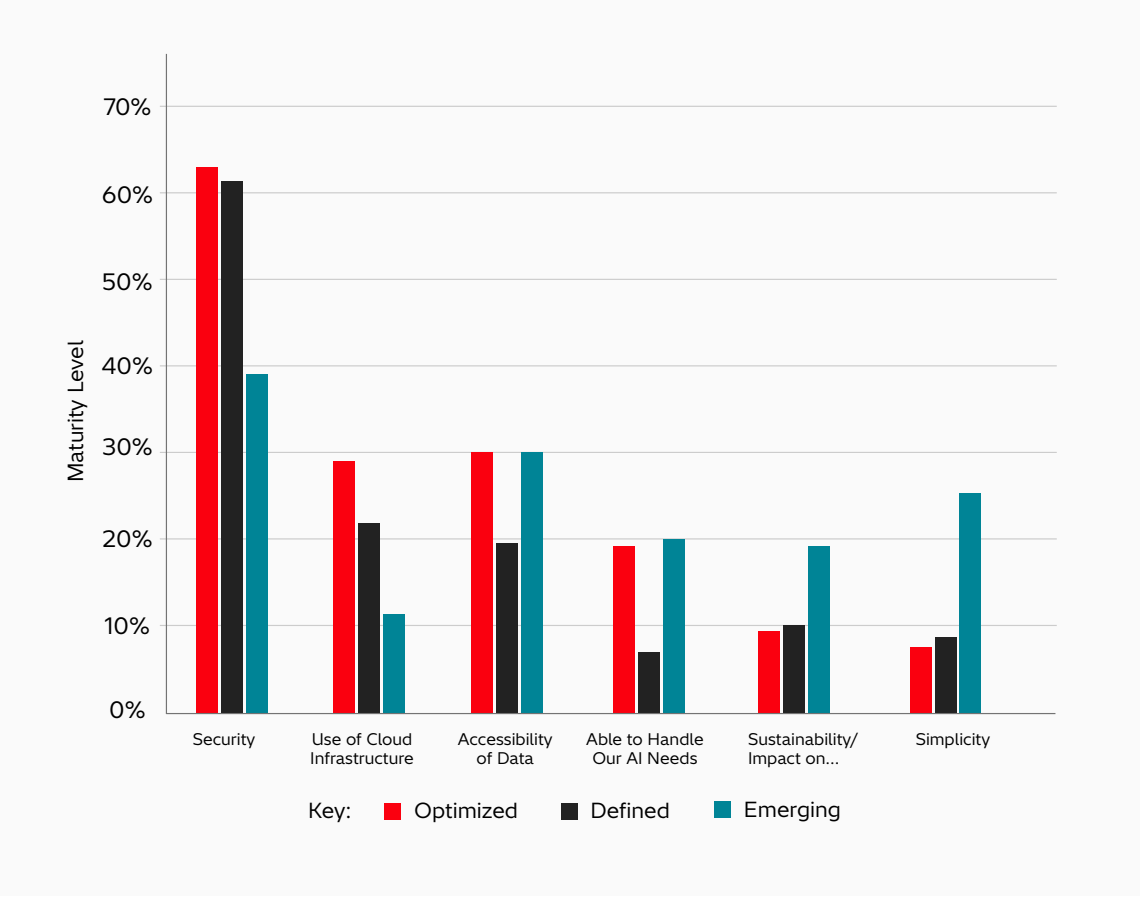
What Data Infrastructure Leaders Do Differently

The Non-Negotiables of Data Infrastructure

At all levels of maturity, organizations prioritize security and accessibility of data when it comes to data infrastructure. However, organizations differ when it comes to the use of cloud infrastructure, with defined and optimized organizations favoring and empowered to adopt sophisticated infrastructure over simplicity.

Because emerging organizations rely on simplicity to operate, opportunities for global expansion and advanced analytics are limited. These organizations also pay higher, long-term costs due to less sophisticated legacy systems.

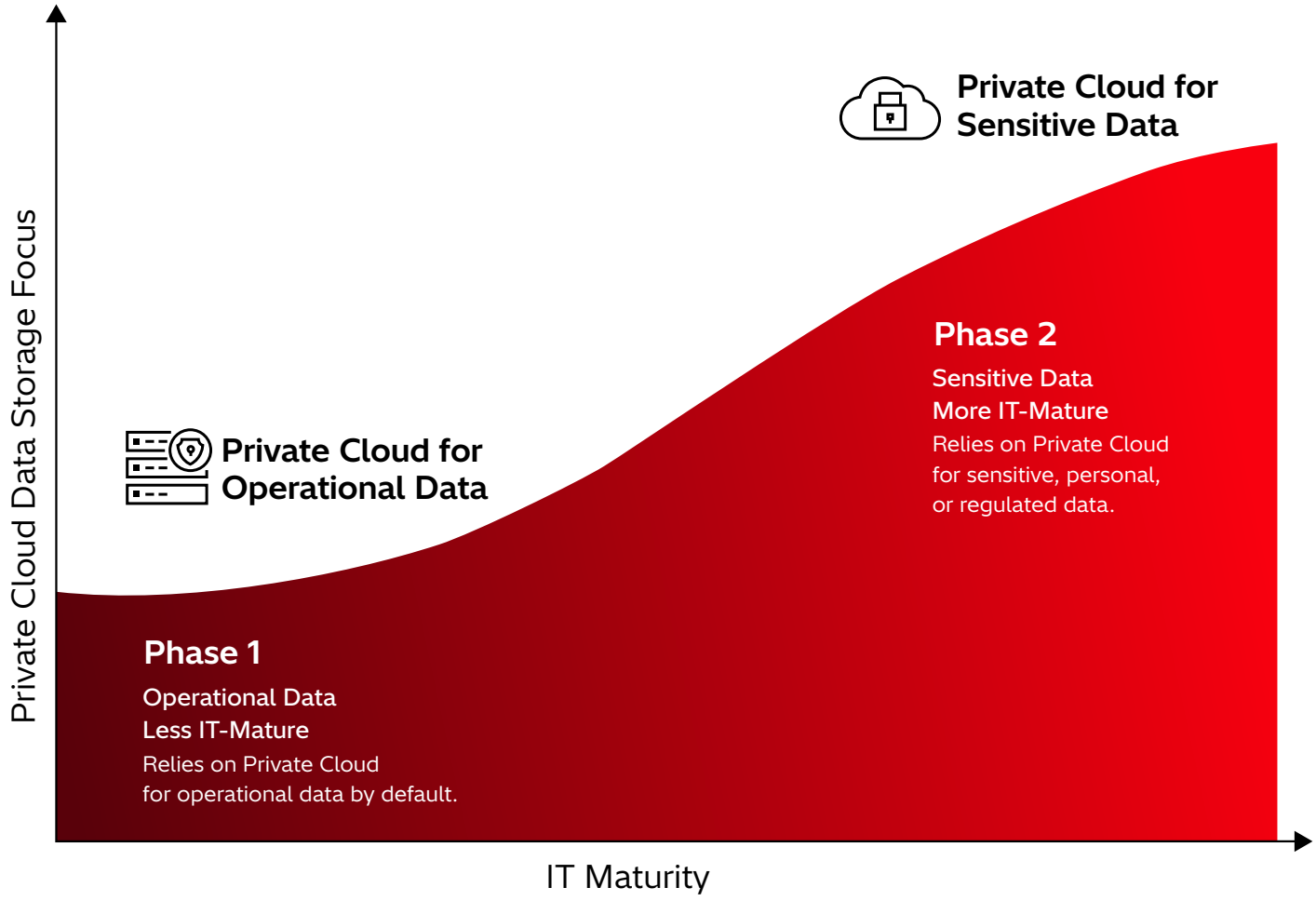
Most Important Data Infrastructure Aspects by Maturity



Maturity Narrows Private Cloud Usage

For less mature organizations, private cloud is often the default because data isn't rigorously classified and it feels safer. As organizations mature, they shift from storing by necessity to storing by strategy — enabling risk-based storage policies and increase confidence in public cloud security controls.

As organizations mature, they stop thinking in terms of systems and start thinking in terms of data domains. With maturity, organizations enable formal data classification (operational vs. sensitive vs. regulated), establish risk-based storage policies including strong data governance, and have increased confidence and trust in public cloud and its security controls. This is the key that unlocks speed, scale, and analytics for operational data. Add to that control, isolation, and auditability for sensitive data.

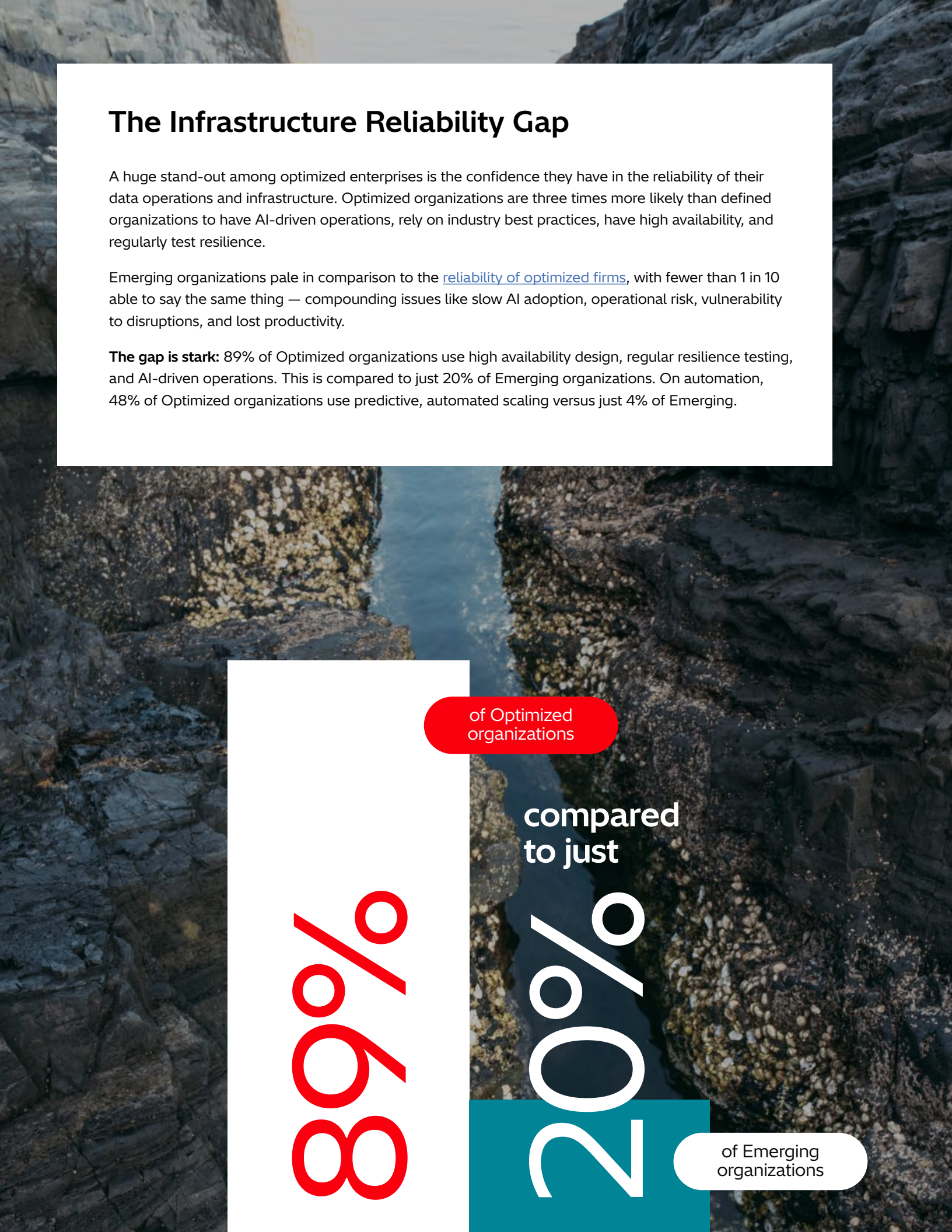


The Infrastructure Reliability Gap

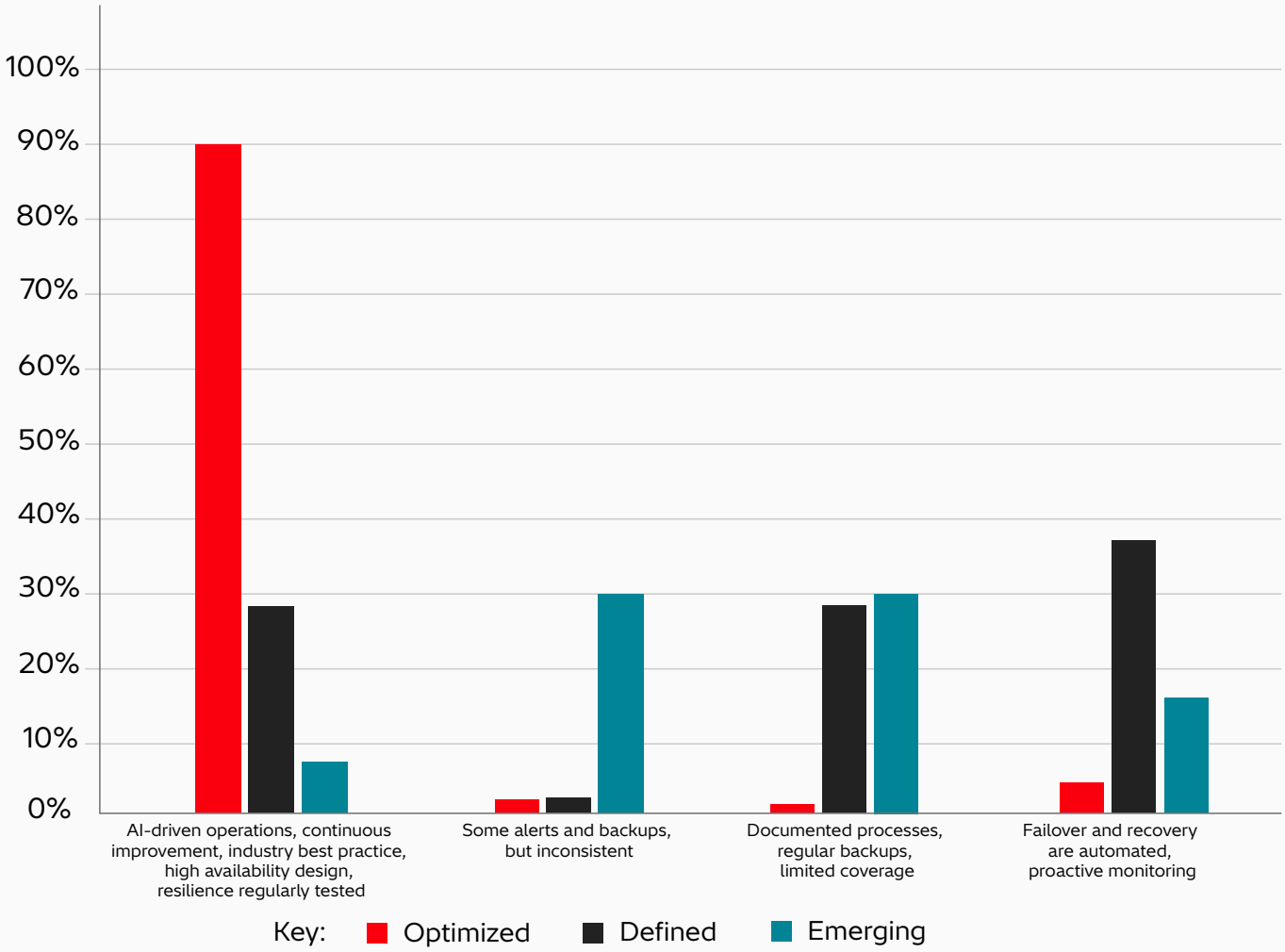
A huge stand-out among optimized enterprises is the confidence they have in the reliability of their data operations and infrastructure. Optimized organizations are three times more likely than defined organizations to have AI-driven operations, rely on industry best practices, have high availability, and regularly test resilience.

Emerging organizations pale in comparison to the reliability of optimized firms, with fewer than 1 in 10 able to say the same thing — compounding issues like slow AI adoption, operational risk, vulnerability to disruptions, and lost productivity.

The gap is stark: 89% of Optimized organizations use high availability design, regular resilience testing, and AI-driven operations. This is compared to just 20% of Emerging organizations. On automation, 48% of Optimized organizations use predictive, automated scaling versus just 4% of Emerging.



Reliability of Data Operations by Maturity



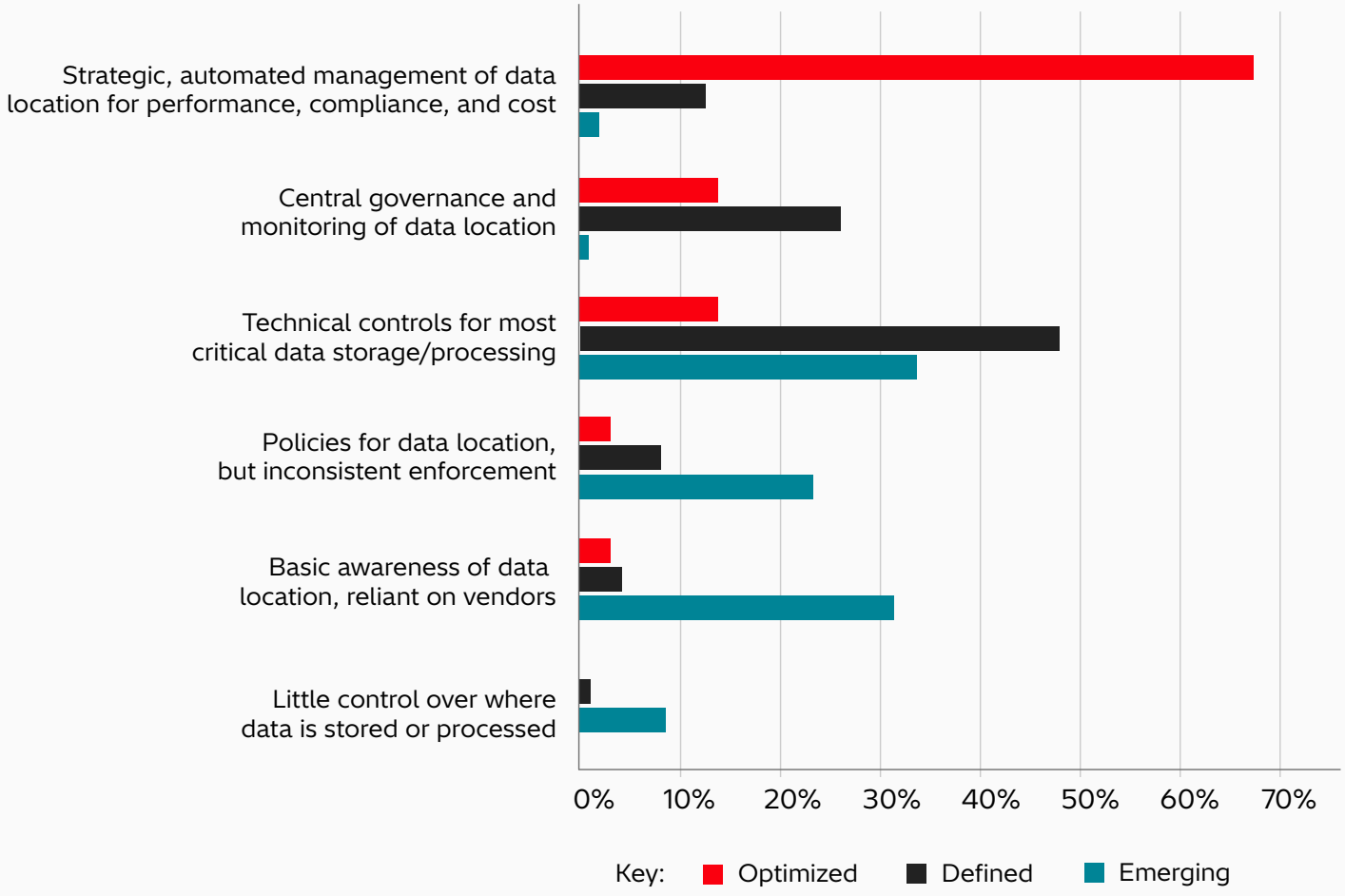
“There is no such thing as cyber resilience without infrastructure resilience.”

Chris Millington
Global Solutions Lead, Data and Cyber Resilience
Hitachi Vantara

Storage Strategy Defines AI Success

Maturity doesn't just mean more storage — it's a shift of the control plane upward: from ad hoc awareness and technical controls to centralized governance and automation. This means standardizing data location policies, automating their enforcement, and instrumenting observability across AI and data (often with partners, given that 94% of companies say they need third party help on data infrastructure) are key to maturing as an organization.

Data Storage and Processing by Maturity

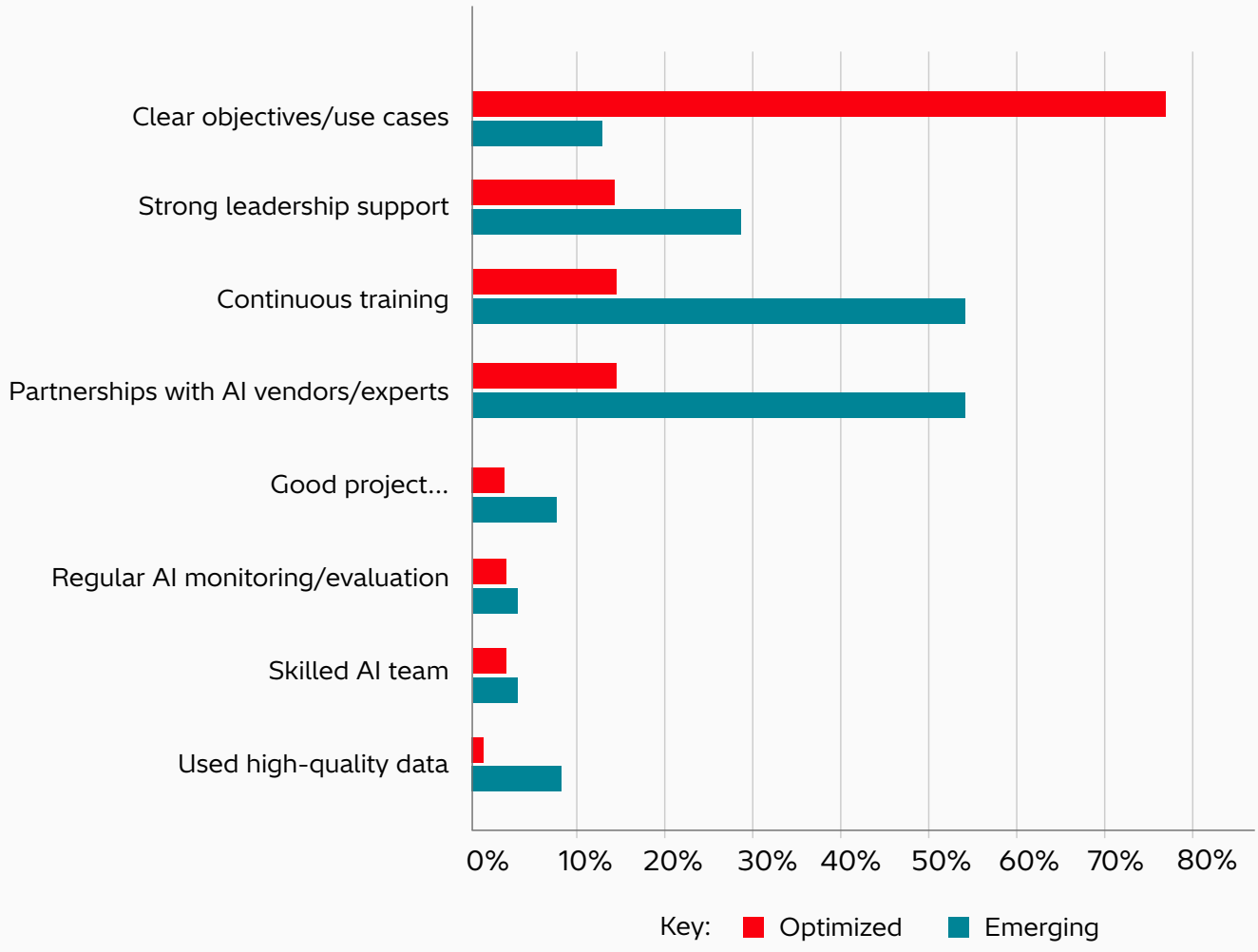


Optimized Organizations Prioritize Data Infrastructure

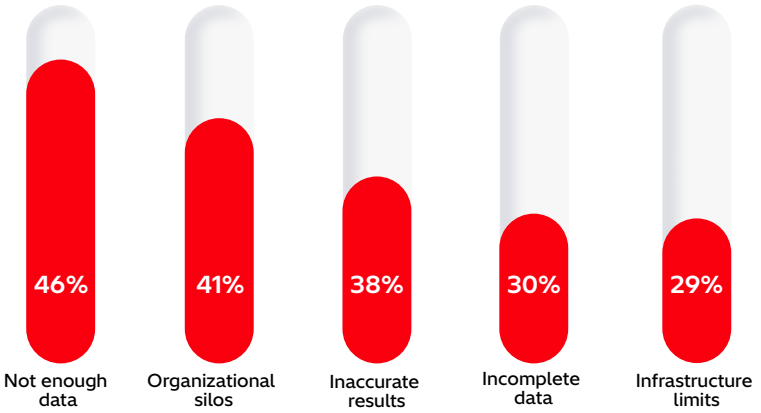
Optimized organizations prioritize their data foundation, with almost all citing that the right infrastructure to support AI workloads is the key to business success. Leaders also hire and upskill aggressively and operate with a clearly defined AI vision. Sustainability is embedded as an outcome of efficiency and cost control, with continuous measurement and optimization.

Regardless of maturity, company size, or industry most organizations need help from third parties on data infrastructure — though the reasons vary across organizations. Optimized organizations are more likely to need help building AI models/LLMs, sustainable implementation, and training users; at all stages of maturity, organizations need help with data processing and data preparation.

Reason for AI Project Success by Maturity



Conversely, failed projects more often link to a lack of data or incomplete data, as well as infrastructure limits, highlighting persistent data foundation and change-management gaps.



Failure Factors

“The true constraint is [no longer the model or hardware, but] the data pipeline, storage, and network architecture that support it, and the operational discipline required to run these systems reliably.”

Lawrence Yeo
ASEAN Solutions Director
Hitachi Vantara

What's Different From 2024?

With AI adoption steady, and no major improvement in data trust or model quality, the standout difference in the state of data this year since last year is the intensity of security.

- **Security has intensified** as the dominant theme—especially in the U.S.—moving beyond last year's already high concern.
- **Hallucinations are down,** but **accuracy/trust remain at just ~40%**, echoing last year's "data quality gap" narrative.
- **Sustainability still lags** when it comes to actual priorities—despite the annual increase of extreme heat and severe storms on data centers.

Security Concerns by Type (Year-over-Year)

Metric	2024	2025	Change
Data security as top AI implementation concern	37%	56%	+19
Internal AI breach concern	31%	41%	+10
External AI-enabled attack concern	41%	43%	+2
Regulatory compliance concern	31%	37%	+6

AI Success Rates by Model Type (Year-over-Year)

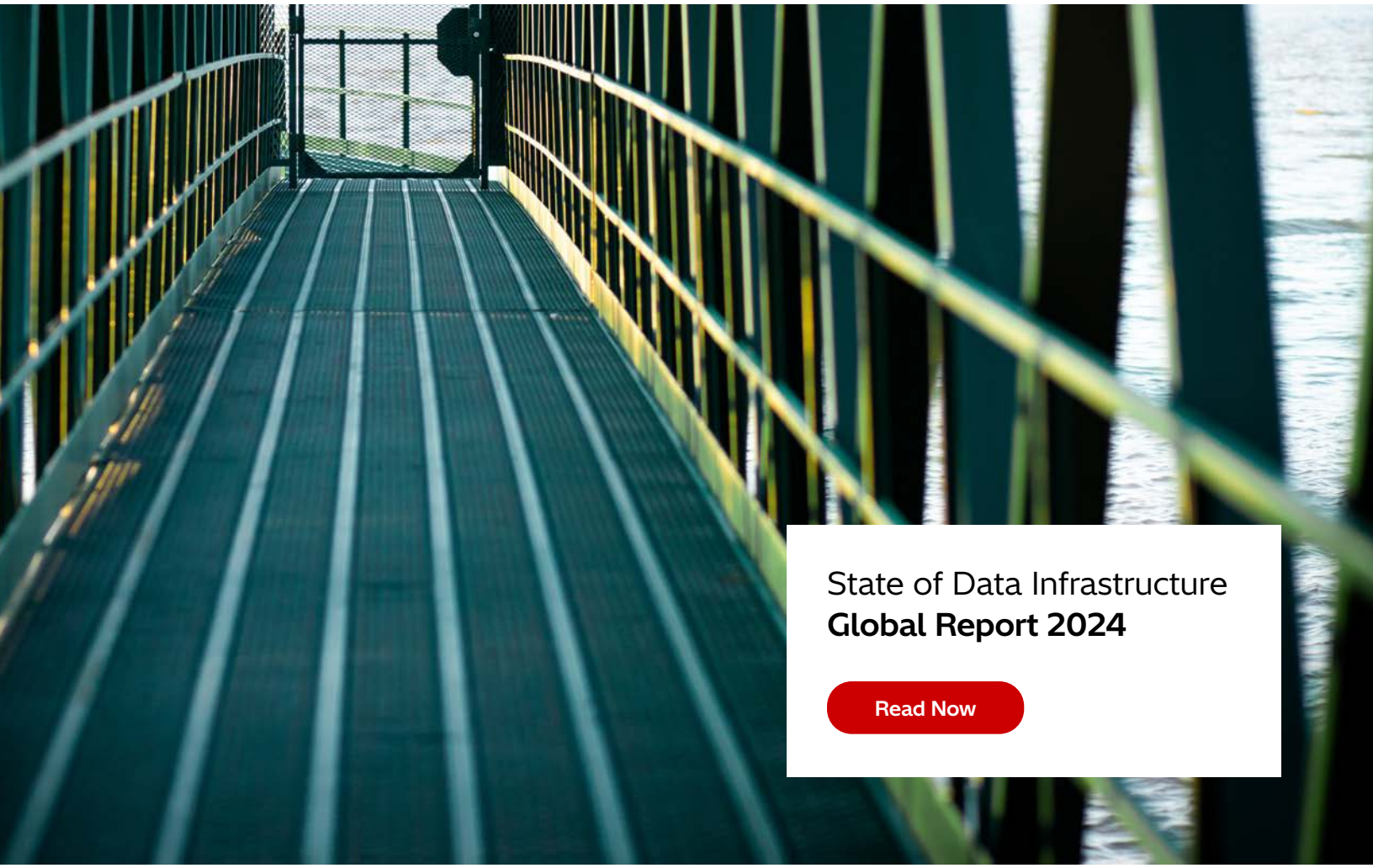
Metric	2024	2025	Change
Generative AI tools (ChatGPT, Copilot)	95%	81%	-14
AI tools integrated into existing software	85%	78%	-7
Purchased LLM model	85%	73%	-12
Company-built LLM model	84%	70%	-14
Free/open-source model	72%	70%	-2

Explainability Strategies (Year-over-Year)

Metric	2024	2025	Change
No strategy to explain model outputs	42%	24%	-18
Enhanced data quality for model training	39%	49%	+10
AI governance frameworks	32%	45%	+13
Regular AI performance audits	36%	44%	+8

Employee Confidence (Year-over-Year)

Metric	2024	2025	Change
Confident employees use AI safely/only for business	77%	65%	-12



A Blueprint for Action

Tailored Moves for Every Maturity Stage

Ninety-four percent of companies say they need third party help on data infrastructure. Priority gaps include training IT staff, building models, secure and strategic implementation, data processing and preparation, and end-user training. Many organizations manage gaps with a mix of technology, hiring, upskilling, and partnering. Hiring skilled AI workers remains a particular challenge in India and Canada, with 54% and 60% of respondents, respectively, citing it as the top concern. That's well above the global average of 37%.

The research shows who actually owns AI decisions: Head of IT leads across all decision categories (27-39% depending on decision type), with Head of AI second (17-29%). This supports the finding that Optimized organizations empower IT leadership, while less mature organizations diffuse decision-making to line-of-business managers.

"Data governance is the foundation of operational AI."

Octavian Tanase
CPO
Hitachi Vantara

Here's a checklist for organizations to accelerate their maturity:

Emerging:

- Normalize policies
- Centralize data-location controls
- Set measurable KPIs and create dedicated teams.
- Prioritize data-quality pipelines (preparation, accessibility, accuracy)

Defined:

- Consolidate and govern the data landscape
- Establish high-availability and resilience patterns with automated monitoring and failover
- Centralize data-location governance to meet performance, compliance, and cost goals
- Raise data quality through rigorous preparation, accessibility, and accuracy controls
- Define an executive AI vision linked to revenue, savings, and experience outcomes
- Invest in talent and continuous training; institute ongoing audits and model monitoring
- Use partnerships to accelerate capability while maintaining sovereignty and security

Optimized:

- Prioritize strategic and secure implementation, build models, make data available, train IT/users, and provide scalable solutions
- Partner and outsource to accelerate capability while preserving sovereignty and compliance
- Scale predictive automation
- Embed sustainability measurements

Partner for Progress

How Hitachi Vantara Accelerates Your AI Journey

For organizations with strong foundations, Hitachi Vantara can coach on AI vision, strategy, and best-practice adoption. For those with ambitious AI goals but weaker foundations, Hitachi can provide unified, secure, and sustainable infrastructure solutions that unlock ROI. The proprietary maturity model helps identify where solutions and services are most appropriate, enabling a step-change from Emerging to Defined to Optimized. Any organization can make the shift—but senior leaders must prioritize data infrastructure and commit to a long-term partnership that integrates strategy, services, and solutions.

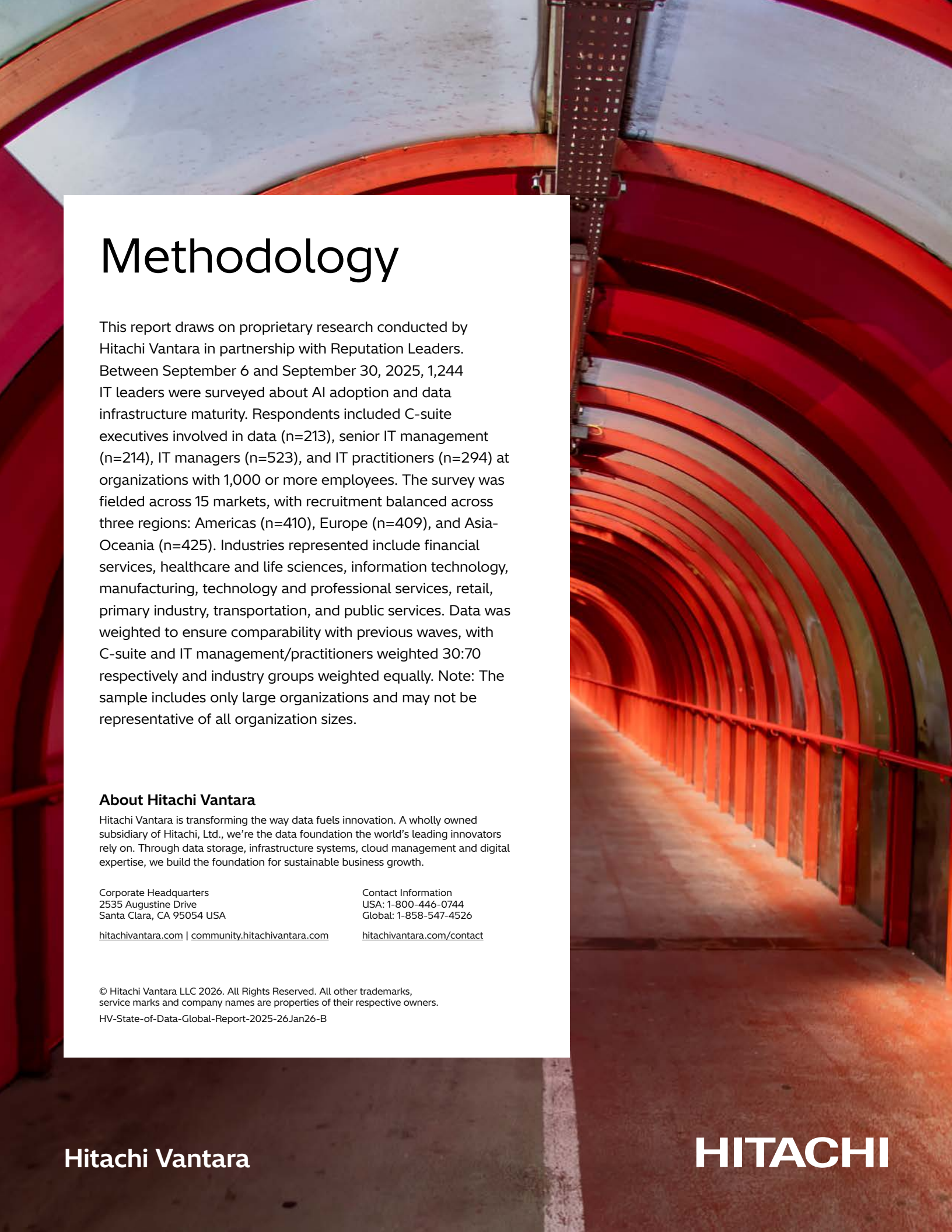
"Hitachi Vantara's VSP One platform exemplifies a unified strategy for modern storage.

It delivers a single, resilient platform spanning edge, core and cloud, while supporting workloads ranging from traditional enterprise systems to AI-driven use cases."

Whit Walters, Field CTO, GigaOm[1]

Learn more about Hitachi Vantara's data foundation.

1 GigaOm Radar Recognizes Hitachi Vantara as a Leader and Outperformer in Primary Storage for Second Consecutive Year



Methodology

This report draws on proprietary research conducted by Hitachi Vantara in partnership with Reputation Leaders. Between September 6 and September 30, 2025, 1,244 IT leaders were surveyed about AI adoption and data infrastructure maturity. Respondents included C-suite executives involved in data (n=213), senior IT management (n=214), IT managers (n=523), and IT practitioners (n=294) at organizations with 1,000 or more employees. The survey was fielded across 15 markets, with recruitment balanced across three regions: Americas (n=410), Europe (n=409), and Asia-Oceania (n=425). Industries represented include financial services, healthcare and life sciences, information technology, manufacturing, technology and professional services, retail, primary industry, transportation, and public services. Data was weighted to ensure comparability with previous waves, with C-suite and IT management/practitioners weighted 30:70 respectively and industry groups weighted equally. Note: The sample includes only large organizations and may not be representative of all organization sizes.

About Hitachi Vantara

Hitachi Vantara is transforming the way data fuels innovation. A wholly owned subsidiary of Hitachi, Ltd., we're the data foundation the world's leading innovators rely on. Through data storage, infrastructure systems, cloud management and digital expertise, we build the foundation for sustainable business growth.

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA

hitachivantara.com | community.hitachivantara.com

Contact Information
USA: 1-800-446-0744
Global: 1-858-547-4526

hitachivantara.com/contact

HV-State-of-Data-Global-Report-2025-26Jan26-B

Hitachi Vantara

HITACHI