CLOUDERA

CLOUDERA AND NATIONAL TECHNOLOGY NEWS RESEARCH REPORT

The state of hybrid data architectures in the AI era

In collaboration with

NTN
National Technology **News**

Introduction

As organisations contend with a surge in new technologies and expanding data sources, the challenge of harnessing information for strategic advantage has never been greater. From cloud-native applications and legacy on-premises systems to edge devices, the sheer diversity of data origins is complicating how businesses manage, integrate and govern their information. This complexity is prompting many to rethink their data strategies, aiming to meet operational needs, comply with evolving regulations and unlock opportunities for innovation.

Rapid advances in artificial intelligence are reshaping business priorities. The adoption of generative and agentic AI models is driving demand for platforms that support diverse workloads and enable real-time decision-making. At the same time, regulations such as the EU AI Act, GDPR and industry-specific mandates require organisations to demonstrate transparency, control and accountability across distributed data environments.

This environment compels organisations to balance innovation with risk management. Compliance is no longer a procedural formality; it serves as a catalyst for responsible AI adoption and scalable data operations. Integrating legacy systems, managing fragmented data silos and meeting stricter regulatory requirements all point to the need for a unified approach to data architecture.

Hybrid data architectures, which unify cloud and on-premises infrastructure, are emerging as a practical response. These approaches offer flexibility to support legacy systems, improve scalability and address data residency requirements. However, they also present new challenges, especially in governance, security and regulatory compliance.

To understand how organisations are navigating these issues, National Technology News and Cloudera conducted a survey of professionals in data, technology and operations roles. This report presents the survey's key findings, analyses the drivers behind hybrid adoption and examines the relationship between AI, compliance and operational resilience in today's evolving data environment.

Methodology

National Technology News and Cloudera surveyed 100 professionals working across data, technology and operations to understand how organisations are evolving their data architecture to support critical AI-based operational use cases, what kind of approach they are taking to data architectures for operational use cases, the main factors determining the choice of infrastructure type, and what role AI and analytical use cases play in their strategy.

Key Findings:

1. **Lifecycle complexity is the top obstacle:** 87 per cent struggle with managing distributed AI/ML lifecycles, making it the single most cited challenge.
2. **Security leads near-term priorities:** 63 per cent rank real-time anomaly detection as their number-one strategic use-case, with 45 per cent planning automated cybersecurity response.
3. **AI investment accelerates:** 76 per cent will expand AI and machine-learning capabilities in the next 12-24 months, signalling rapid budget growth.
4. **Hybrid solves the legacy lock-in:** 72 per cent adopt hybrid architecture primarily to integrate legacy systems that cannot move to the cloud.
5. **Governance maturity rises, gap remains:** 68 per cent have enterprise-wide governance, yet a full 32 per cent still rely on partial or ad-hoc policies.

Table of contents

1. Best approaches to data architecture for operational use cases 4
2. Key factors driving hybrid data architecture adoption 5
3. What makes a hybrid architecture strategy optimal? 6
4. Operational use cases for AI and data analytics 7
5. Strategic use cases in the next 12–24 months 8
6. Approaches to data governance in heterogeneous environments 9
7. Preferred AI/ML models: An overview 10
8. Challenges in implementing AI/ML in hybrid environments 11
9. Managing compliance and regulations in hybrid environments 12
10. Investment priorities for hybrid data in the next 12–24 months 13

Conclusion 14

1. Best approaches to data architecture for operational use cases

The survey results indicate that most organisations are moving towards a balanced hybrid approach for operational data architecture. More than half of respondents (54 per cent) describe their strategy as a near-equal mix of cloud and on-premises infrastructure. This preference reflects a practical recognition that hybrid models provide the adaptability needed to integrate legacy systems, optimise performance and maintain operational resilience.

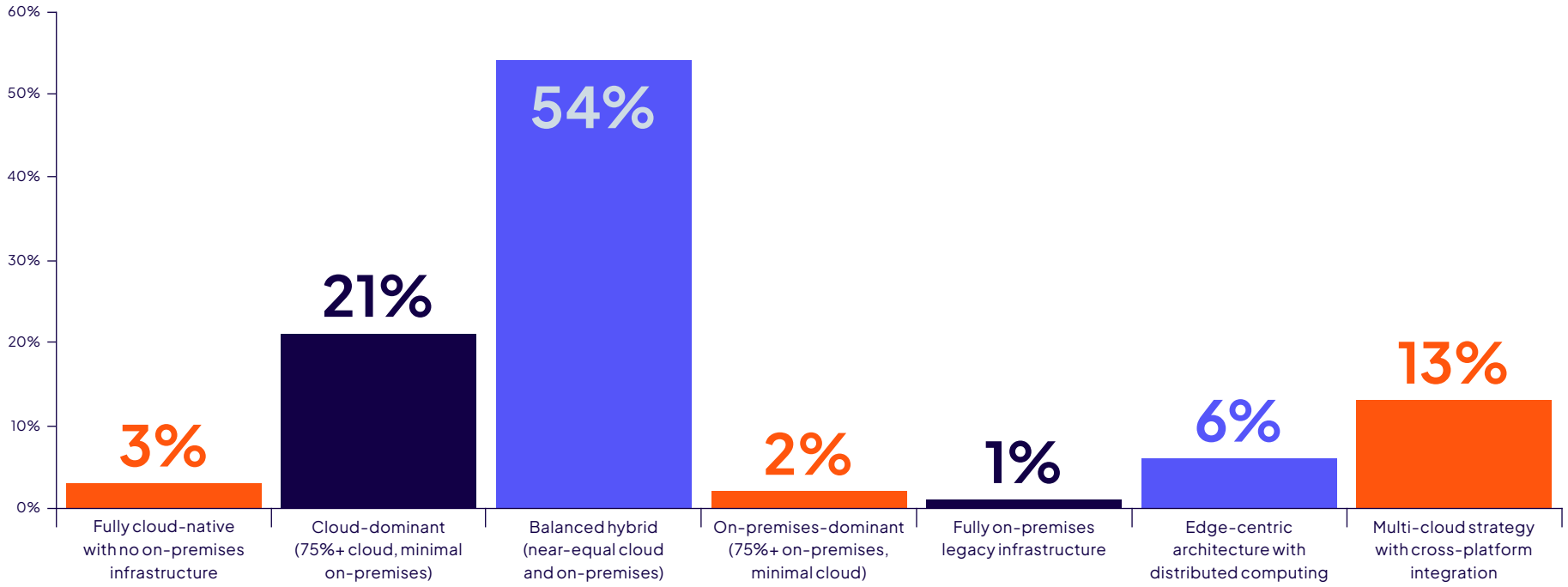
Hybrid architectures are valued for their ability to bridge the gap between existing on-premises investments and the opportunities presented by cloud platforms. Many organisations continue to rely on critical legacy systems that are not easily migrated to the cloud. By adopting a hybrid approach, organisations can modernise at their own pace, reducing risk and avoiding disruption to essential operations.

The survey also highlights the rise of cloud-dominant and multi-cloud approaches. Around one-fifth of organisations (21 per cent) favour a model where cloud infrastructure is the primary environment, with minimal reliance on on-premises resources. A further 13 per cent are adopting multi-cloud strategies, selecting services from multiple providers to optimise workload distribution and reduce dependence on any single vendor.

Few organisations are committing to either extreme. Only one per cent prefer a fully on-premises legacy infrastructure, and just three per cent have adopted a fully cloud-native approach. This suggests that most organisations see value in retaining a mix of environments, rather than pursuing a complete transition in either direction.

These findings underscore the importance of adaptability and risk management. By maintaining both cloud and on-premises capabilities, organisations can respond to changing business needs, regulatory requirements and technology developments without locking themselves into a single platform. This flexibility is especially relevant as regulatory and operational demands continue to evolve.

What best describes your organisation's current approach to data architecture for operational use cases? (Select one option)



2. Key factors driving hybrid data architecture adoption

Survey responses highlight that organisations are motivated by a combination of technical, operational and regulatory factors when adopting hybrid data architectures. Most respondents selected multiple drivers, reflecting the complexity of the decision-making process.

Integration with legacy systems was the most frequently cited reason, chosen by 72 per cent of respondents. Many organisations operate critical systems that cannot be easily migrated to the cloud due to technical constraints, cost, or risk. A hybrid approach allows these systems to remain operational while enabling the gradual adoption of new technologies and cloud services.

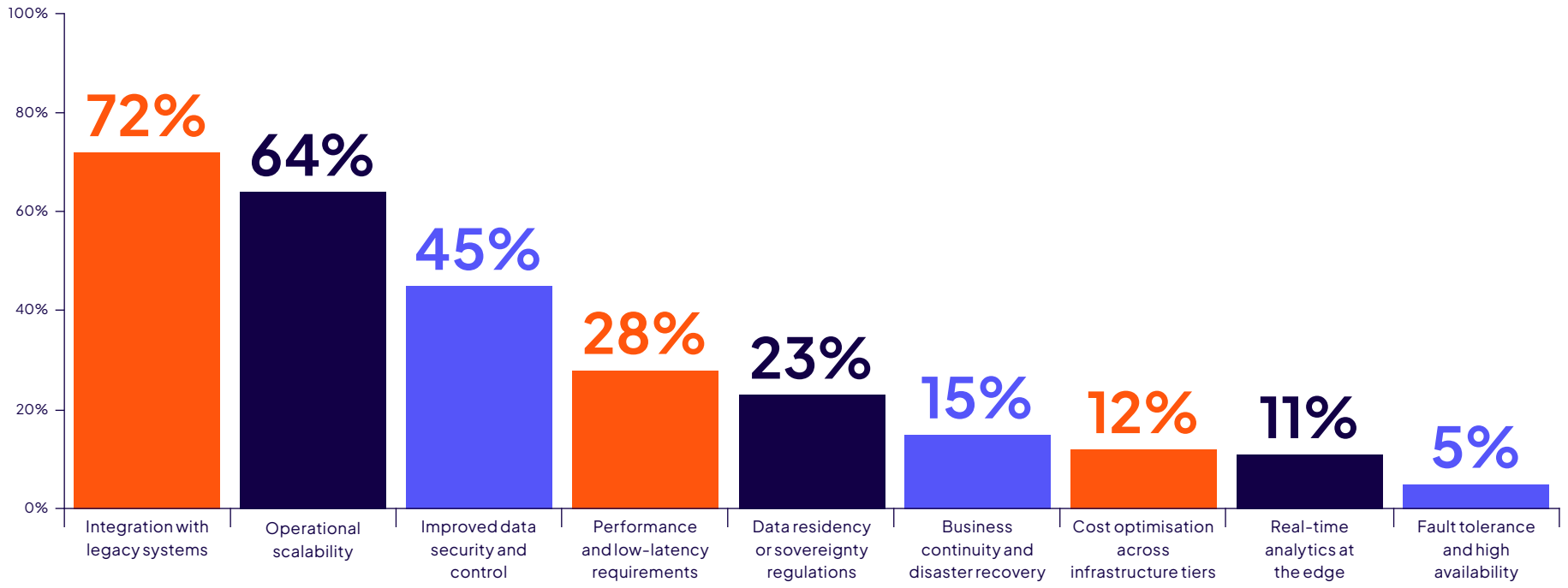
Operational scalability is another important factor, identified by 64 per cent of respondents. Hybrid architectures offer the ability to scale workloads efficiently, using cloud resources where appropriate while retaining control over sensitive or high-priority data and processes on-premises.

Improved data security and control (45 per cent), performance and low-latency requirements (28 per cent), and data residency or sovereignty regulations (23 per cent) were also significant considerations. These factors reflect the need to balance innovation with compliance and risk management, especially as regulations become more demanding and data privacy concerns grow.

Other considerations, such as business continuity, disaster recovery, and cost optimisation, were selected less frequently but still play a role in shaping hybrid strategies. The survey data shows that organisations are not driven by a single issue but by a blend of priorities that reflect their operational realities and regulatory obligations.

In summary, the adoption of hybrid data architectures is shaped by the need to integrate legacy systems, maintain operational flexibility, address regulatory requirements and manage risk. Organisations are seeking solutions that allow them to innovate without compromising security, compliance or performance.

What are the primary drivers for adopting a hybrid data architecture in your organisation? (Select up to three options)



3. What makes a hybrid architecture strategy optimal?

Maturity in hybrid data architecture is closely linked to the presence of robust governance frameworks. However, the survey reveals that most organisations are still on the path to achieving this goal. Although 68 per cent of organisations describe their hybrid data architecture strategy as aimed at fully operational hybrid architecture with enterprise-wide governance, the figures reflect a strong aspiration rather than a completed transformation. In reality, many organisations are still grappling with the complexities of integration, policy enforcement and infrastructure coordination between cloud and on-premises environments.

A well-developed governance framework serves as the cornerstone of a successful hybrid data architecture, elevating it beyond a patchwork of systems to a unified, strategic enterprise asset. This level of governance not only ensures compliance and risk management, but also enables organisations to realise the full value of their data investments, making it a critical driver of sustainable competitive advantage.

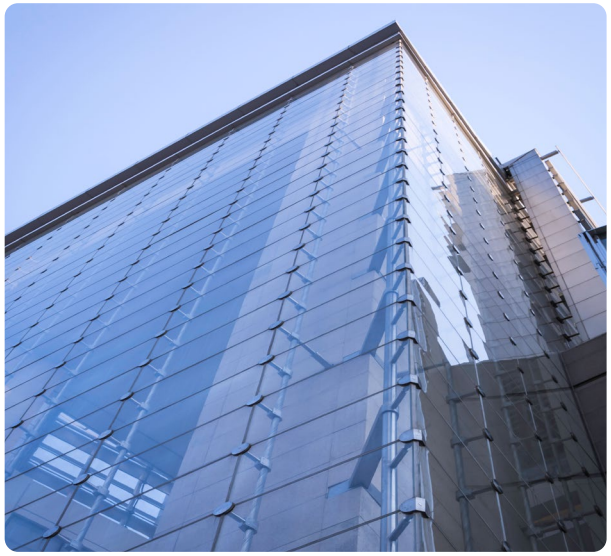
Organisations that implement unified policies and controls across cloud and on-premises environments are better positioned to manage data privacy, security and compliance. This approach also helps to reduce fragmentation, improve visibility and streamline operations, which are essential for scaling AI and analytics initiatives responsibly.

The regulatory landscape is a central factor shaping these strategies. Laws such as the EU AI Act, GDPR and sector-specific mandates impose strict requirements on transparency, data residency, risk management and accountability. These regulations not only govern how data is stored and processed, but also dictate where it can reside and how it must be protected. Non-compliance can result in significant financial penalties and reputational harm. As a result, organisations are under pressure to demonstrate that their data management practices are both robust and adaptable to evolving legal standards.

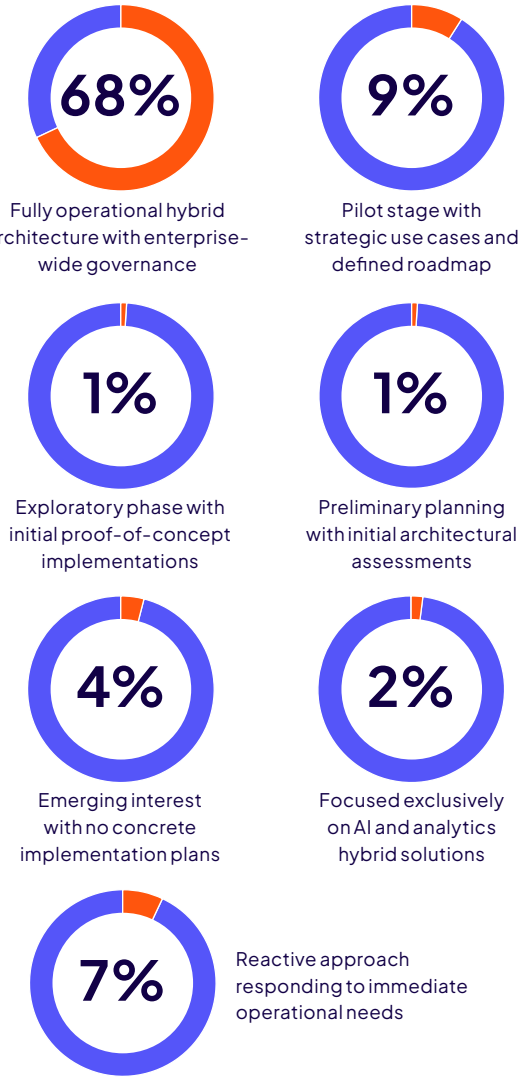
The survey also reveals that a smaller proportion of organisations are still in earlier stages of hybrid adoption. Nine per cent are in the pilot phase, while only a small minority are at the exploratory or planning stages. This distribution suggests that while progress has been made, some organisations are still working towards establishing mature, governed hybrid environments.

A hybrid strategy is most effective when it balances flexibility with control. Adaptability to evolving requirements, seamless integration of legacy and modern systems, and the enforcement of consistent governance are all crucial. Notably, the question arises: is a hybrid approach only viable once an enterprise-wide governance framework has been established, or can governance and hybrid adoption evolve in tandem?

The survey findings suggest that organisations prioritising governance and operational discipline are the ones best positioned to realise the full potential of hybrid data architectures, driving innovation while managing risk and meeting regulatory demands.



Which of the following best characterises your hybrid data architecture strategy?
(Select one option)



4. Operational use cases for AI and data analytics

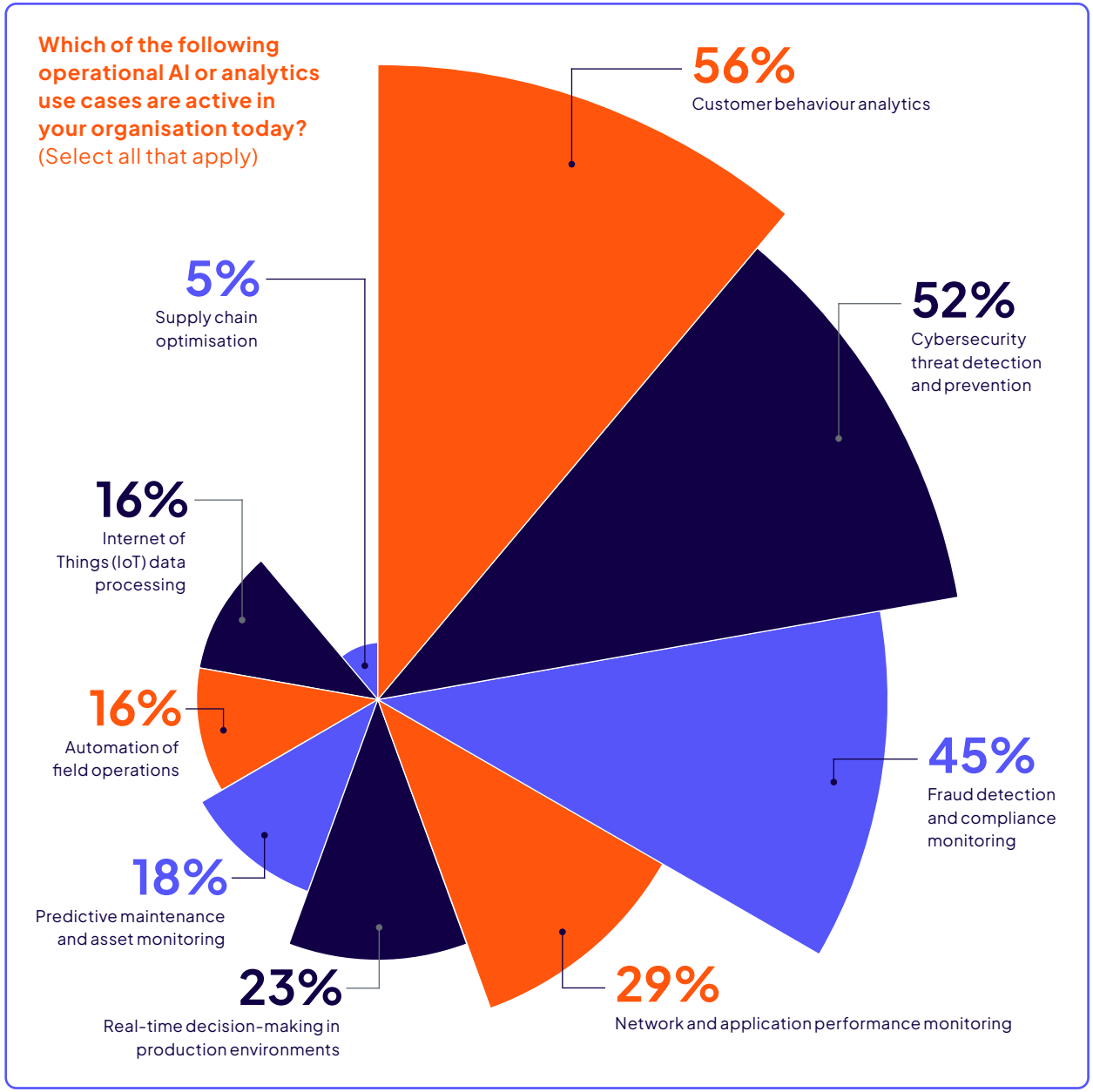
Organisations are embedding AI and analytics across a wide range of operational activities, signalling the technology's transition from isolated pilots to core business functions.

The most common application is customer behaviour analytics, with 56 per cent of organisations using AI to understand customer needs, personalise engagement and drive loyalty strategies. The rise of enterprise AI agents is accelerating this trend, as companies increasingly deploy these tools in customer-facing roles such as support and marketing, and plan to expand their use further.

A crucial requirement for these applications is the ability to use proprietary data securely, without exposing sensitive information outside the organisation. Systems that integrate both public and on-premises infrastructure, and that support seamless data and AI lifecycle management, are becoming essential for keeping models current while enforcing strict access controls and audit requirements.

Security-focused use cases are also prominent. Detection and prevention of cybersecurity threats (52 per cent) and fraud detection and compliance monitoring (45 per cent) are major priorities. These results highlight a strong focus on using AI to strengthen organisational resilience. Operational AI is now central to cybersecurity, enabling real-time anomaly detection, predictive risk assessment, automated compliance reporting and rapid response to potential regulatory breaches.

However, an excessive focus on protection can risk stifling innovation if it leads to overly restrictive practices. To address this, organisations are encouraged to treat security as an enabler of innovation, not a barrier. Hybrid data platforms that offer advanced data protection, such as tokenisation and masking, while enabling collaborative analytics and AI development, allow organisations to innovate with confidence, balancing risk management with operational agility.



5. Strategic use cases in the next 12–24 months

Looking ahead, organisations are prioritising hybrid data and AI use cases that directly address operational risk, resilience and personalisation. Real-time anomaly detection in operations is the top strategic priority, selected by 63 per cent of respondents. Automated cybersecurity response across distributed systems (45 per cent), operational forecasting and optimisation (41 per cent), and personalised service delivery using AI (41 per cent) follow closely behind.

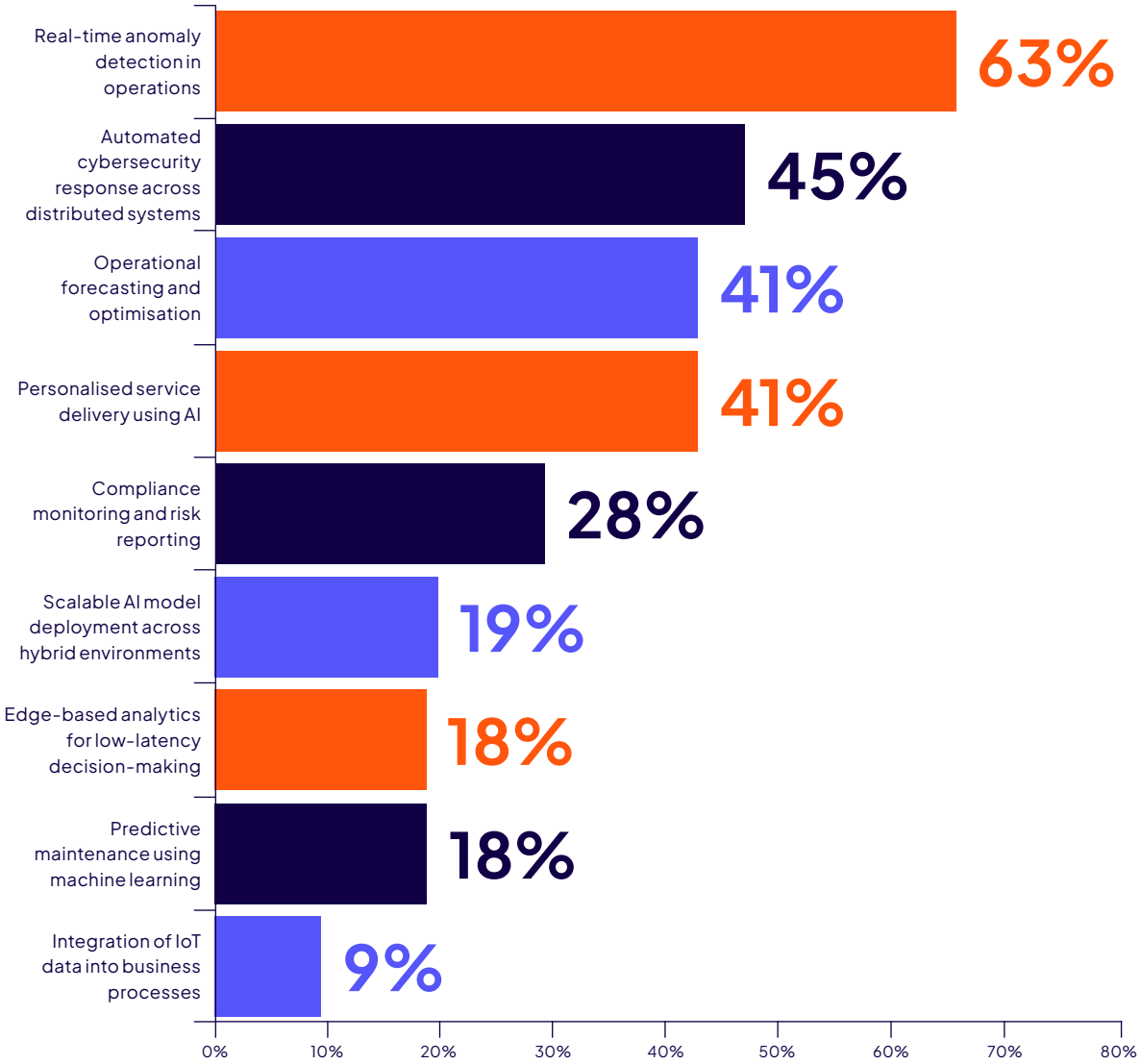
The focus on real-time anomaly detection and automated cybersecurity reflects the growing threat landscape and the challenge of managing large volumes of security data across fragmented environments. Many organisations struggle to process and analyse logs from sources such as DNS, firewalls and endpoints quickly enough to respond to threats. This has driven interest in open data lakehouse architectures and advanced AI models that can identify anomalies and support proactive threat mitigation.

Personalised service delivery and operational forecasting are also gaining traction, highlighting the demand for AI-driven insights that can improve customer experiences and optimise business processes. Compliance monitoring and risk reporting remain important, cited by over a quarter of organisations, reflecting the ongoing need to meet regulatory requirements and demonstrate accountability.

Implementing scalable AI models across hybrid environments is a key challenge for 19 per cent of respondents, underlining the importance of unified solutions that allow organisations to use proprietary data securely and deploy models across cloud, on-premises and edge environments.

Overall, the survey suggests that organisations are aligning their strategic investments with use cases that strengthen security, enhance resilience, and drive personalisation, while ensuring compliance and operational efficiency in an increasingly complex regulatory and threat environment.

Which hybrid use cases are considered strategically important to your organisation over the next 12–24 months? (Select all that apply)



6. Approaches to data governance in heterogeneous environments

The survey reveals a strong industry-wide consensus on the need for unified governance across cloud, on-premises and edge environments. All respondents indicated that their organisations have either a unified governance framework or common policies with environment-specific variations. This reflects a recognition that fragmented governance can create security risks, regulatory gaps and operational inefficiencies.

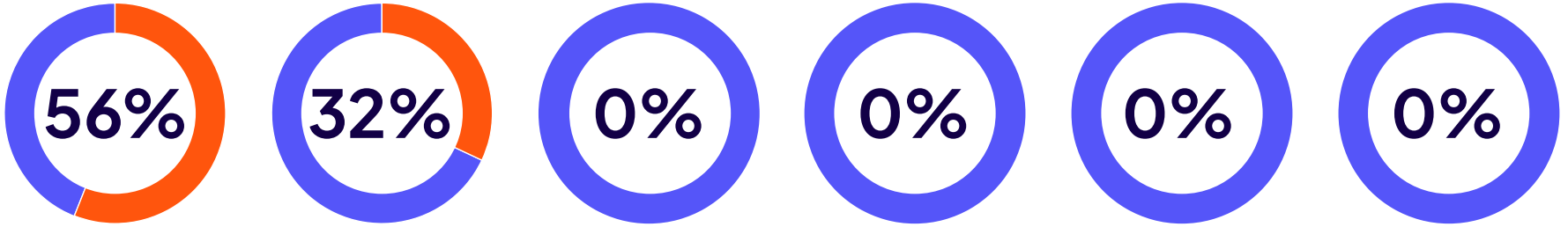
Unified governance frameworks are increasingly underpinned by technologies that enable consistent policy enforcement, metadata management, audit logging and federated data cataloguing. Many organisations are adopting solutions that allow them to apply governance controls regardless of where the data resides—whether in public cloud, private cloud, on-premises clusters or edge devices. This approach supports true hybrid cloud portability, enabling seamless movement of data and workloads while maintaining compliance and reducing latency.

Replication and synchronisation of metadata, classification tags and security policies are also seen as essential for business continuity and analytics. Enterprise-grade replication solutions help ensure that governance controls follow the data, supporting high availability, disaster recovery and regulatory obligations.

The regulatory environment continues to drive these developments. Laws such as the EU AI Act and GDPR require organisations to demonstrate strong data stewardship, transparency and the ability to enforce controls across distributed environments. Without a unified governance model, organisations risk data silos, inconsistent security, reduced data quality and increased compliance risks.

While unified frameworks are ideal, the survey acknowledges that some organisations still rely on common policies with environment-specific adaptations. However, those that can standardise governance across all environments are better positioned to support innovation, maintain compliance and respond quickly to regulatory changes.

How does your organisation approach data governance across heterogeneous environments (cloud, on-premises, edge)? (Select one option)



7. Preferred AI/ML models: An overview

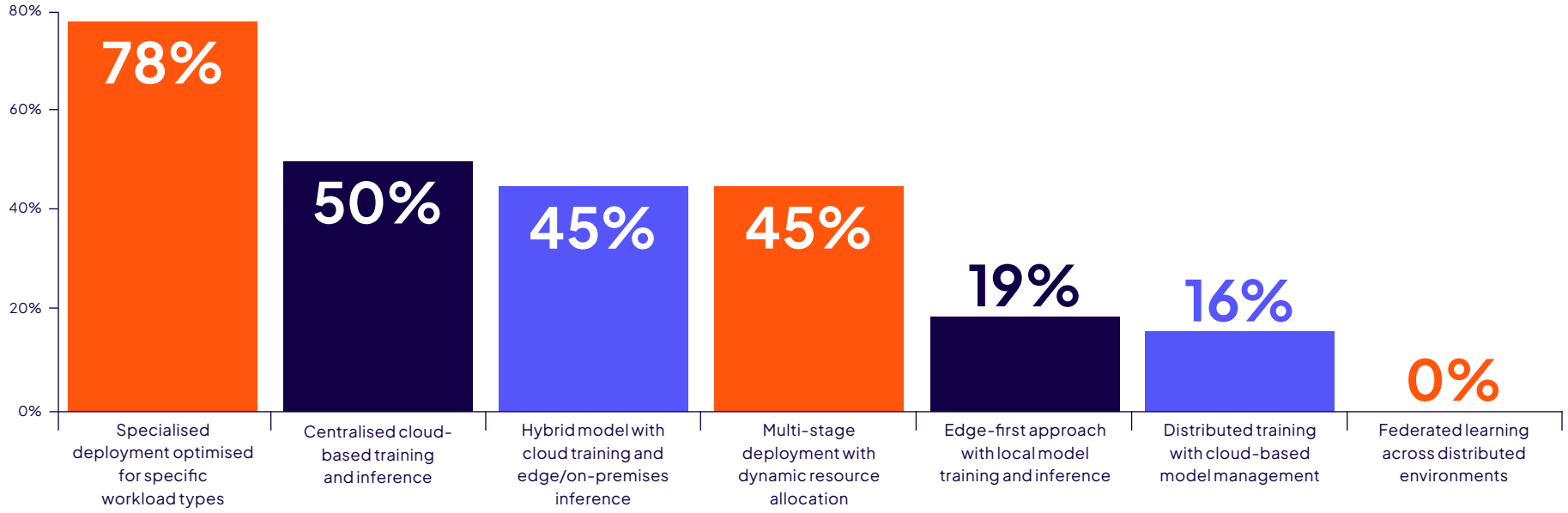
Survey responses show a clear trend towards deploying AI and machine learning models in environments optimised for specific workload types. Seventy-eight per cent of organisations report using specialised deployments, often relying on cloud-based infrastructure to support scalable training and inference. Half of respondents use centralised cloud-based training and inference, while 45 per cent use hybrid models that combine cloud training with edge or on-premises inference.

This shift is driven by the need for flexibility, scalability and access to high-performance computing resources, particularly as generative AI and large language models become more prevalent. Cloud-native platforms are evolving rapidly to support these requirements, providing tools for containerisation, dynamic resource allocation and integration with popular AI frameworks.

Some organisations continue to use edge-first and distributed approaches, especially where latency, data sovereignty or security concerns require local processing. However, the overall trend is towards environments that can support both centralised and distributed AI workloads. This approach allows organisations to optimise for performance, compliance and cost.

The survey findings highlight the importance of aligning AI and machine learning deployment strategies with operational needs and regulatory constraints. Organisations that can flexibly deploy models across cloud, on-premises and edge environments are better positioned to deliver value from AI initiatives while maintaining control and compliance.

Where are AI/ML models currently being trained and deployed in your organisation? (Select all that apply)



8. Challenges in implementing AI/ML in hybrid environments

Organisations report several persistent challenges when deploying AI and machine learning in hybrid environments.

The most significant is the complexity of managing distributed model lifecycles, cited by 87 per cent of respondents. This challenge is often compounded by a lack of integration between data and machine learning platforms, reported by 73 per cent, which can result in fragmented workflows, manual processes and delays in moving models from development to production.

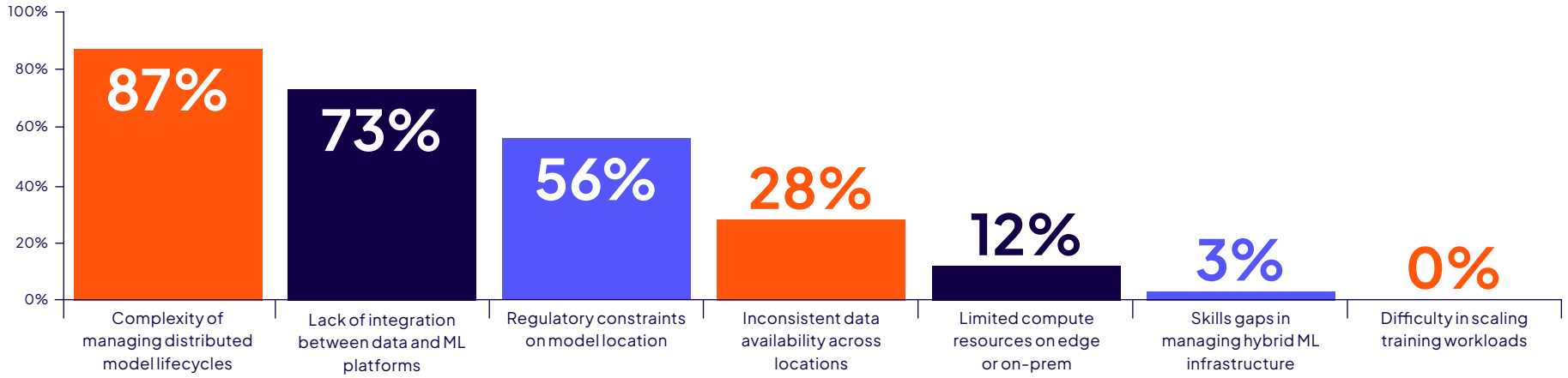
Regulatory constraints on model location are another major concern, with 56 per cent highlighting this issue. Laws such as the EU AI Act and GDPR, as well as similar regulations in other regions, may require that both data and the models trained on that data remain within specific jurisdictions. This creates additional hurdles for organisations seeking to leverage hybrid or multi-cloud deployments.

Other obstacles include inconsistent data availability across locations, limited compute resources on edge or on-premises infrastructure, and skills gaps in managing hybrid machine learning systems. While these issues are less frequently cited, they can still impede the effective scaling of AI initiatives.

To address these challenges, organisations are increasingly adopting unified hybrid architectures that specialise in MLOps, centralised model registries and automated deployment pipelines. These solutions support traceability, version control and secure, programmatic deployment of models across environments. Hybrid and multi-cloud capabilities also provide the flexibility needed to comply with regulatory requirements by allowing organisations to choose where to train, store and serve models.

The findings suggest that overcoming these challenges is critical for realising the full benefits of AI in hybrid environments. Organisations that invest in integrated platforms and robust governance are better positioned to accelerate innovation while maintaining compliance and operational control.

What are the most significant challenges you have encountered in deploying AI/ML in hybrid environments? (Select up to three options)



9. Managing compliance and regulations in hybrid environments

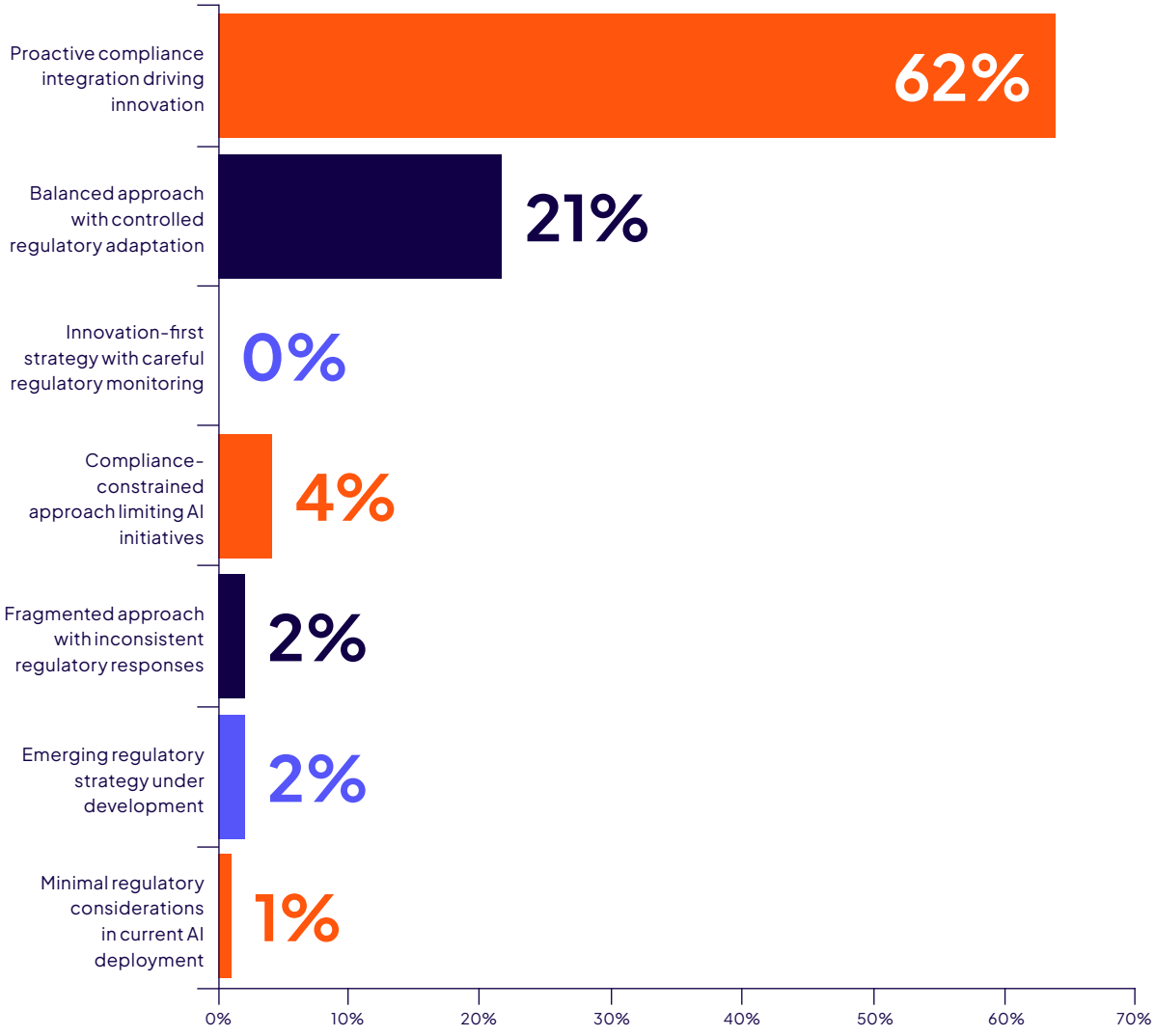
Compliance remains a central concern for organisations deploying AI across hybrid environments. The survey shows that most organisations view regulatory compliance not as a barrier, but as an enabler for responsible innovation. Sixty-two per cent report that proactive compliance integration is actively driving AI innovation, while a further 21 per cent take a balanced approach, adapting their strategies to meet evolving regulatory demands.

The regulatory landscape is evolving rapidly, with frameworks such as the EU AI Act and GDPR imposing strict requirements for transparency, risk management and data residency, especially for high-risk AI applications. Organisations must ensure that data and models remain within approved jurisdictions, maintain detailed audit trails and implement robust governance to avoid costly penalties and reputational damage.

A small minority of respondents still report that compliance constrains their AI initiatives, or that their approach to regulation is fragmented or underdeveloped. However, the prevailing trend is towards embedding compliance into AI workflows and infrastructure, enabling organisations to innovate confidently while meeting legal obligations.

Organisations that integrate compliance into the design of their AI systems are better able to manage risk, accelerate deployment and maintain trust with customers and regulators. This proactive stance not only reduces the risk of non-compliance but can also open up new opportunities by enabling responsible AI adoption in sensitive or highly regulated sectors.

How is your organisation managing the balance between regulatory compliance (e.g. EU AI Act, data sovereignty laws) and innovation in AI deployments across hybrid environments? (Select one option)



10. Investment priorities for hybrid data in the next 12–24 months

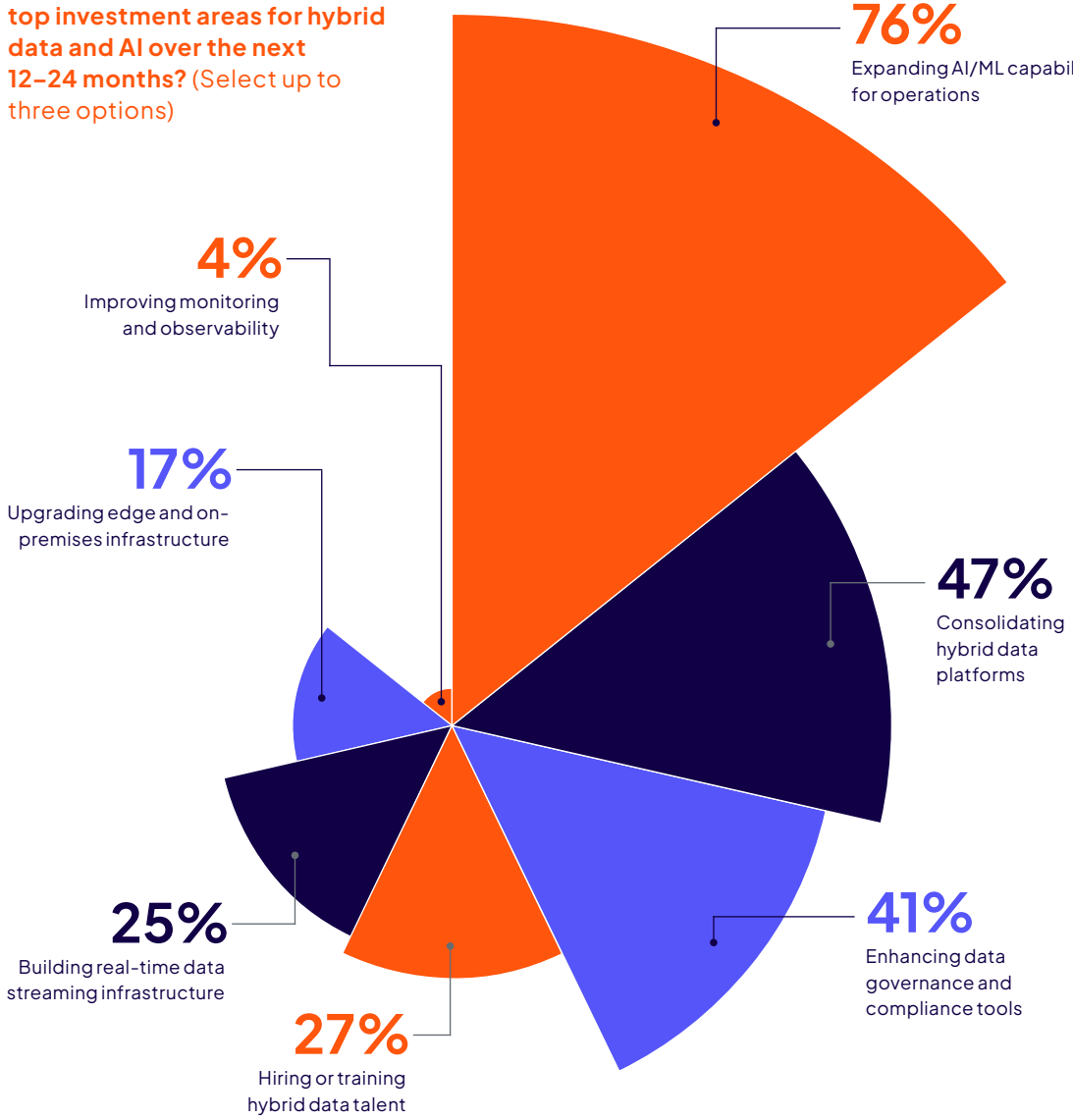
Over the coming one to two years, organisations are planning significant investments in hybrid data and AI capabilities. The top priority, chosen by 76 per cent of respondents, is the expansion of AI and machine learning capabilities within operations. This reflects the central role that AI now plays in driving efficiency, agility and competitive advantage.

Consolidating hybrid data platforms is the next most common investment, with 47 per cent seeking to simplify architectures, improve data accessibility and reduce costs by integrating disparate systems. Enhancing data governance and compliance tools is also a major focus, cited by 41 per cent, as organisations respond to increasing regulatory demands and the need to manage privacy and security risks.

Other key investment areas include hiring or training hybrid data talent (27 per cent), building real-time data streaming infrastructure (25 per cent), and upgrading edge and on-premises infrastructure (17 per cent). These priorities highlight the importance of developing skills, improving data agility and supporting low-latency use cases.

Overall, the survey suggests that organisations view AI expansion, platform consolidation and governance as interconnected priorities. The ability to innovate at scale depends on robust control, unified data management and the capacity to adapt quickly to regulatory and operational changes.

What are your organisation's top investment areas for hybrid data and AI over the next 12–24 months? (Select up to three options)



Conclusion

The findings of this report demonstrate a decisive shift towards hybrid data architectures as the foundation for operational resilience, innovation and regulatory compliance.

Most organisations now favour a balanced approach, combining the strengths of cloud and on-premises environments to achieve flexibility, scalability and integration with legacy systems. This strategy enables organisations to adapt to evolving business needs and regulatory requirements without being locked into a single platform.

Governance and compliance have moved to the centre of the data strategy agenda. The adoption of unified governance frameworks and proactive compliance integration is enabling organisations to manage risk, demonstrate transparency and accelerate responsible AI innovation. Rather than treating regulation as a barrier, leading organisations are embedding compliance into their workflows and infrastructure, unlocking new opportunities while maintaining trust.

AI and analytics are now operational realities, supporting a wide range of use cases from customer engagement to cybersecurity and risk management. The ability to deploy AI and machine learning models flexibly across hybrid environments is increasingly critical, as is the capacity to scale innovation while maintaining control and oversight.

However, significant challenges remain. Managing distributed model lifecycles, integrating data and machine learning platforms, and navigating complex regulatory landscapes require ongoing investment, robust governance and continual adaptation. Organisations that address these challenges directly—by consolidating platforms, investing in talent and prioritising unified data management—will be best positioned to harness the full value of their data.

The imperative is clear: organisations must continue to evolve their data architectures and governance strategies to keep pace with technological advances and regulatory change. By doing so, they can drive innovation, strengthen operational resilience and build a sustainable competitive advantage in an increasingly complex digital world.

CLOUDERA

In collaboration with

NTN

National Technology **News**