Enterprise Strategy Group™

TECHNICAL VALIDATION

Managing AI Project Lifecycle With Tata Communications Vayu AI Cloud

Unlocking Business Success With Effortless AI Transformation

By Tony Palmer, Practice Director
Enterprise Strategy Group

September 2025

Contents

Introduction

This Technical Validation by Enterprise Strategy Group details our evaluation of the Tata Communications Vayu AI Cloud platform. This validation explores how Vayu AI Cloud provides a secure, unified platform that minimizes complexity and addresses the challenges that impede the production deployment of enterprise AI initiatives.

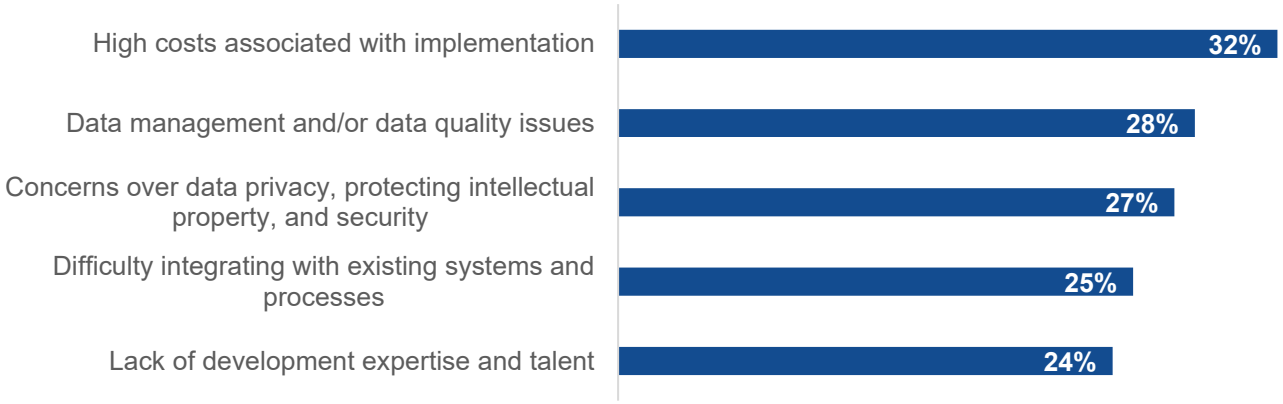
Background

In today's digital landscape, artificial intelligence (AI) is a must-have necessity for organizations that want to compete effectively in the marketplace. With increasing use of generative AI (GenAI) applications, the demand for advanced AI models capable of processing enormous amounts of information, generating insights, and driving innovation at scale is growing rapidly.

Organizations are ramping up their efforts to deploy AI-enabled applications to create competitive advantage. Research from Enterprise Strategy Group showed 58% of organizations have made significant investments in AI.[1] The interest in GenAI is high, as 87% of survey respondents are highly prioritizing GenAI in their efforts to integrate it into several areas of the business.[2]

Organizations seek to adopt GenAI to address several business objectives, such as increasing operational efficiency, improving product or service quality, and enhancing customer experience and engagement. Yet, many challenges are encountered when implementing AI, with 32% of organizations citing the high costs associated with implementation as one of their top three challenges (see Figure 1).

Figure 1. Top Challenges Organizations Have Faced When Implementing AI



Source: Enterprise Strategy Group, now part of Omdia

There is a generative AI skills gap present in many organizations today. Among surveyed organizations, we see that AI system deployment and cloud computing, two infrastructure-focused skills, top the list of skills needed to grow internal generative AI talent.[3] This implies that, while organizations might excel in software development, they frequently need specialized support to build the production-ready infrastructure that AI systems demand.

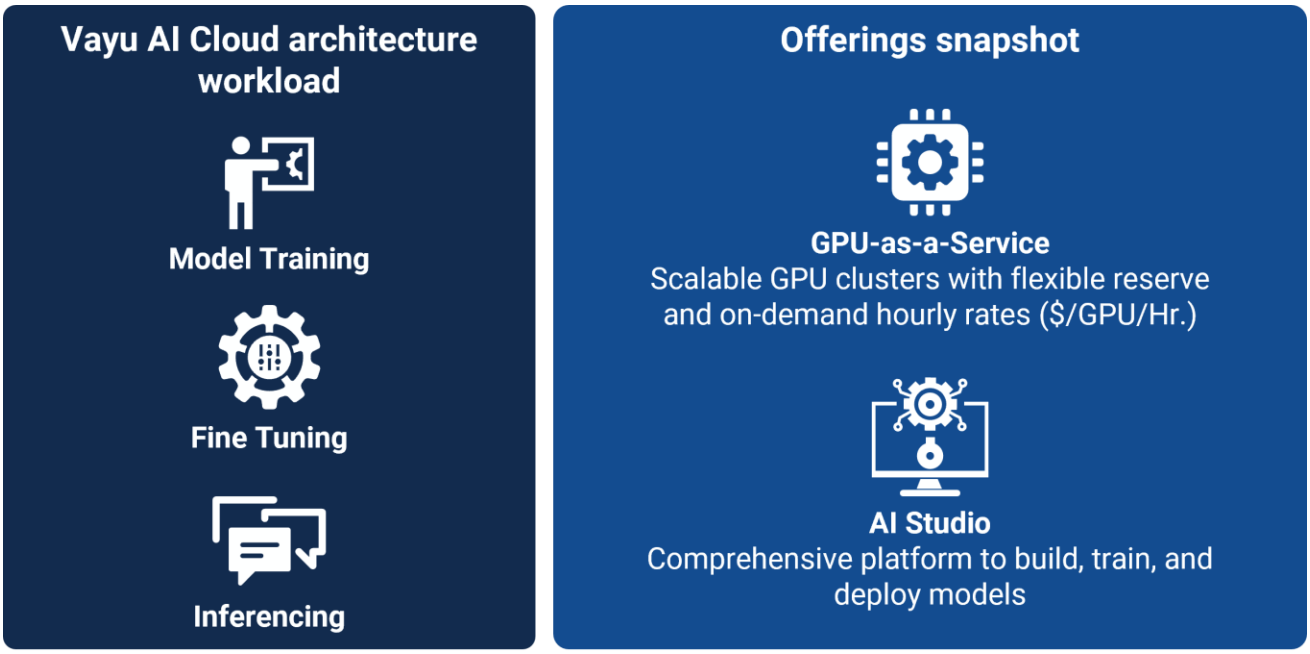
[1] Source: Enterprise Strategy Group Research Report, *Navigating Build-versus-buy Dynamics for Enterprise-ready AI*, January 2025. All Enterprise Strategy Group research references and charts in this technical validation are from this report unless otherwise noted.
[2] Source: Enterprise Strategy Group Research Report, *Navigating the Generative AI Partner and Alliance Landscape*, November 2024.
[3] Ibid.

Tata Communications Vayu AI Cloud

Tata Communications Vayu AI Cloud offers scalable GPU resources and state-of-the-art AI tools designed to give organizations the ability to develop, train, and deploy AI models at scale. Vayu AI Cloud is offered as a comprehensive platform to deliver a suite of capabilities, including large-scale GPU computing (GPU-as-a-service), large language model operations (LLMOps), and serverless functions that enterprises can use to build, train, and deploy models at scale with efficiency. AI Studio provides access to popular AI frameworks and tools, advanced data management, and an AI ecosystem designed to help organizations drive innovation and collaboration.

The Vayu AI Cloud platform is designed to be flexible to address a wide range of customer needs, enabling customers to utilize the platform exactly where they need it—from foundational model builders to enterprises looking to integrate AI into their business processes.

Figure 2. Tata Communications Vayu AI Cloud



Source: Enterprise Strategy Group, now part of Omdia

The key components of the Vayu AI Cloud platform include a network fabric that enables data movement, a scalable GPU infrastructure for compute, a data management platform with data connectors and processing capabilities, and a machine learning operations (MLOps) and LLMOps platform (AI Studio) for model development, experimentation, and deployment.

GPU-as-a-service includes bare-metal GPUs and high-performance networking and storage. GPU-as-a-service offers NVIDIA H100, H200,L40S, and L4 GPUs as of this writing. The AI Studio component offers open source and enterprise AI tools, with a focus on agility and ease of use for data scientists and business users. The AI Supermarket provides a marketplace of pre-built models, APIs, and tools to enable rapid integration of generative AI capabilities.

The MLOps and LLMOps services provide efficient model deployment, monitoring, and management, including features such as responsible AI guardrails, data sovereignty, governance, and compliance.

Tata Communications helps their customers avoid vendor lock-in by providing open source-based solutions and exposing stock open source APIs. In addition, the platform integrates with Tata Communications' broad portfolio of offerings, including customer experience and media services, to deliver end-to-end AI-powered solutions.

Enterprise Strategy Group Technical Validation

Enterprise Strategy Group validated Vayu AI Cloud through the lens of three real-world use cases, evaluating the challenges faced by organizations, how Vayu AI Cloud addresses those challenges, and the benefits that can be expected by organizations utilizing Vayu AI Cloud.

As organizations seek to improve competitiveness, reduce costs, and unlock new growth potential, AI is becoming a critical business enabler. And top drivers, like improving operational efficiency, enhancing product or service quality, and enhancing customer experiences, are unlocking even more investment. Yet, many challenges are encountered when implementing AI, with 32% of organizations citing the high costs associated with implementation as one of their top three challenges, 28% citing data management and data quality issues, and 27% calling out concerns over data privacy, sovereignty and security.

Intelligent Customer Service Automation

In this example, a global financial services firm needed to enhance customer service operations while reducing costs. Its existing customer support system struggled with high call volumes and extended customer wait times, inconsistent service and support quality across global operations, compliance risks from manual handling of sensitive financial data, and an inability to scale support during peak periods.

Tata Communications Vayu AI Cloud helps organizations overcome critical enterprise AI implementation challenges through integrated infrastructure services, simplified deployment, and comprehensive governance capabilities. We evaluated how this organization would implement Vayu AI Cloud to address these challenges.

Using Vayu AI Cloud, the firm deployed a comprehensive customer service AI solution, addressing these five key areas:

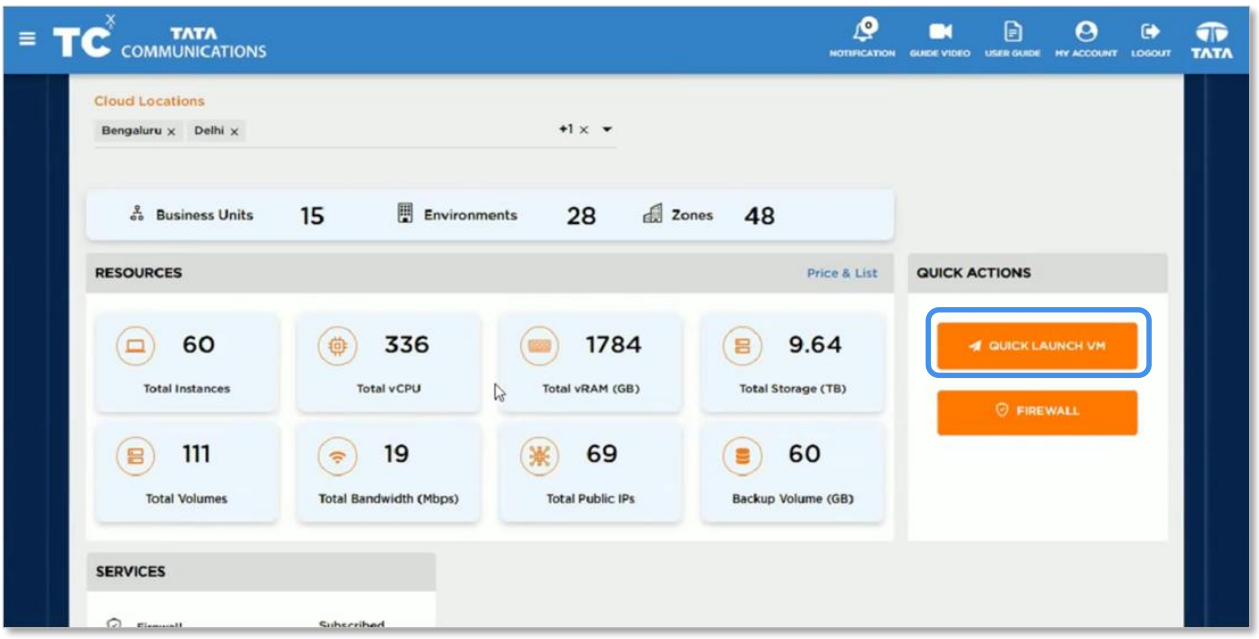
1. **Infrastructure deployment and database as a service**: They utilized Vayu Kubernetes cluster provisioning with automatic scaling to support fluctuating demand.
2. **Data integration, preparation, transformation, and validation**: They connected customer data sources within secure zones to maintain compliance with financial regulations.
3. **AI model development**: Experimentation, prototype development, and deployment of the LLM with GPUs through the unified portal enabled the firm to analyze customer inquiries more effectively.
4. **Backup protection**: Leveraging Vayu's integrated backup-as-a-service ensured business continuity for the firm's critical customer service application.
5. **Governance controls**: Vayu enabled the firm to implement AI ethics frameworks to ensure fair and transparent customer interactions.

Vayu AI Cloud Solution Benefits

Cost Optimization

- The Vayu AI Cloud unified portal eliminates the need for multiple management tools and reduces administrative overhead (Figure 3).

Figure 3. Tata Communications Cloud Console



Source: Tata Communications and Enterprise Strategy Group, now part of Omdia

- One-click provisioning using DevOps tools accelerates deployment and reduces implementation costs.
- Auto-scaling capabilities optimize resource utilization and reduce associated costs.

Data Management Enhancement

- Vayu AI Cloud's integrated catalog of databases, data connectors, and batch and stream frameworks simplify data pipeline creation.
- Logical zone segregation enables proper data handling across development, test, and production.
- Streamlined infrastructure supports improved data quality management.

Security and Compliance Framework

- Vayu AI Cloud employs purpose-built logical segregation for secure multi-environment deployments.
- Comprehensive governance controls enable AI ethics and responsibility.
- Enterprise-grade security is integrated throughout the platform.

Simplified Kubernetes Management

- Vayu AI Cloud enables teams to deploy Kubernetes clusters with just one click, using easily customizable configurations.
- Organizations can protect AI workloads with integrated, native backup-as-a-service.
- Streamlined and automated infrastructure scaling and training lifecycle management reduces operational complexity.

It's important to note that, as of this writing, Tata Communications' Kubernetes service is the first offering Enterprise Strategy Group has examined that offers an integrated, native, backup as-a-service. Once a customer has selected and reviewed all configurations, they can deploy the cluster with one click. Once the infrastructure is deployed, other apps, like databases and DevOps services can be added from the catalog with another click.

Why This Matters

Organizations seek to adopt GenAI to address important business objectives, such as increasing operational efficiency, improving product or service quality, and enhancing customer experience and engagement. Yet, many challenges are encountered when implementing AI, including the high costs associated with implementation (32%), data management complexities (28%), and concerns with security and compliance (27%).

Vayu AI Cloud enables organizations to rapidly deploy enterprise-grade AI applications that deliver measurable business outcomes while addressing critical implementation and operational challenges. As demonstrated in the customer service example, organizations can transform core business functions while maintaining governance, security, sovereignty, and cost efficiency.

Organizations leveraging Vayu AI Cloud could expect to reduce average customer resolution time, decrease operational costs through automation and significant reduction in the complexity of GPU infrastructure management, improve compliance with financial data regulations, and seamlessly scale during seasonal demand peaks, while guaranteeing consistent service quality across all global operations.

Enterprise Strategy Group's evaluation of Tata Communications Vayu AI Cloud revealed that Vayu AI Cloud is a remarkably complete AI solution designed for the enterprise. Vayu AI Cloud can deliver secure, trusted GenAI solutions that can provide real and tangible business benefits quickly and safely, regardless of an organization's level of expertise and experience.

Intelligent Knowledge Management System

Tata Communications' Vayu AI Cloud AI Studio provides a comprehensive end-to-end toolset designed to accelerate AI experimentation and deployment, addressing critical skills gaps and integration challenges that typically delay value realization.

In this example, A multinational retail business needed to transform its customer support operations to improve agent responsiveness and knowledge sharing across global facilities. Its existing customer support ecosystem was challenged with dispersed product documentation across multiple systems and formats, which led to inconsistent responses to similar customer questions. An inability to quickly incorporate product availability information into support systems meant agents were not able to reliably obtain consistent and up-to-date inventory data. Limited AI expertise within the technical support team was hampering progress to address these issues.

Using Vayu AI Cloud and AI Studio's capabilities, the company deployed an intelligent knowledge management system that addressed the requirements of the situation:

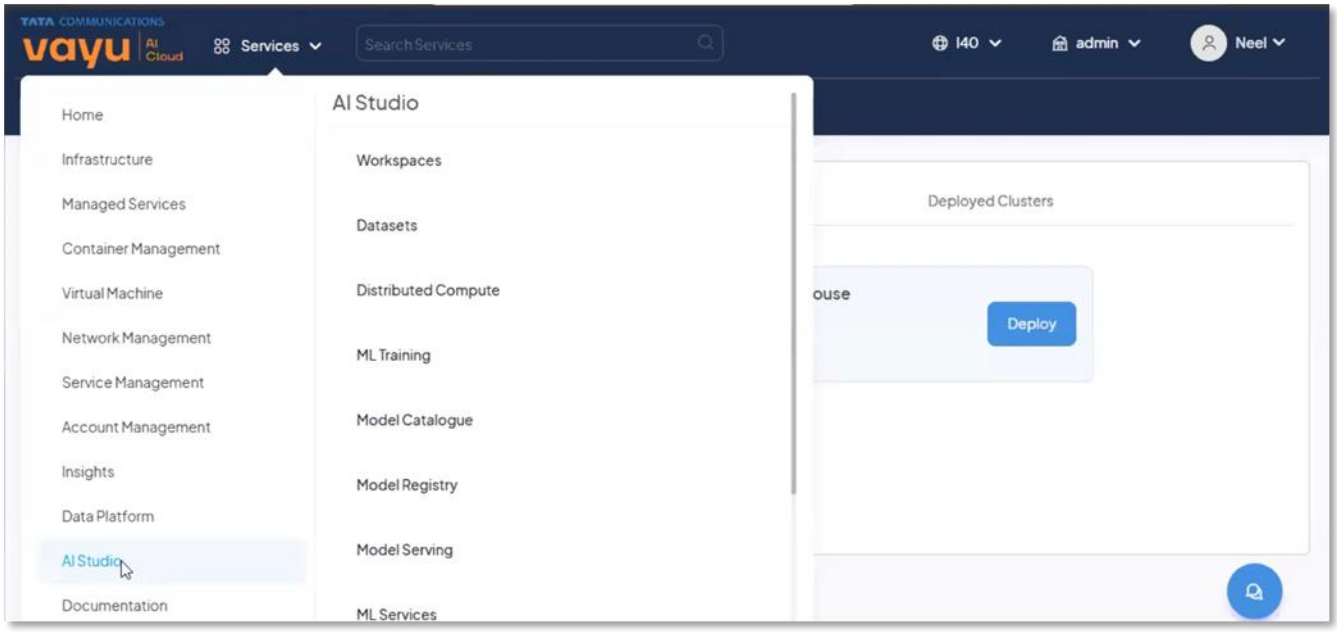
1. **Data integration**: Out-of-the-box connectors aggregated technical documentation, service manuals, and historical support tickets.
2. **Knowledge processing**: Dataset Manager was employed to standardize, label, and version all technical content.
3. **Model selection**: Appropriate foundational models were identified and selected through the AI Studio interface.
4. **Retrieval-augmented generation (RAG) implementation**: RAG connected the knowledge base with selected models using the no-code platform.
5. **Deployment**: The solution was deployed as an API service accessible through the company's existing support portal.

Key AI Studio Capabilities

Simplified Data Management

- AI Studio (Figure 4) provides multiple pre-built connectors for streamlined data pipeline creation.
- Dataset Manager simplifies labeling, versioning, and access control.
- Feature Store enables efficient data reuse across projects.

Figure 4. Tata Communications Vayu AI Cloud Dashboard—AI Studio



Source: Tata Communications and Enterprise Strategy Group, now part of Omdia

Flexible Development Environment

- Purpose-built workspaces enable creation of a development environment with pre-configured tools suited to the problem being solved (e.g., Jupyter, VS Code, and RStudio).
- Resources are easily customizable with guarantees, and reference notebooks are provided for common use cases.
- Customizable resource allocation enables organizations to build the flexible infrastructure that AI development teams demand.

Accelerated Model Deployment

- AI Studio RAG as-a-service is a low-code/no-code platform that enables organizations with limited GenAI resources—almost every organization implementing GenAI at this point—to use existing models and their own knowledge bases to deploy GenAI. Basically, a subject matter expert with no GenAI expertise can deploy critical GenAI use cases like chatbots, customer support systems, or recommendation systems with zero coding required across text, visual, embedding, and reranker models for no-code model deployment as API services.
- In addition to RAG, AI Studio also supports low-rank adaptation (LoRA), a different technique that adapts ML models to new contexts. LoRA can adapt models to specific use cases without changing the entire model by simply adding lightweight pieces to the original model.

Enterprise Integration

- AI Studio is designed to integrate with modern enterprise operations and processes, with comprehensive observability management, integration with existing tools and processes, and robust governance and access controls.

Why This Matters

As new AI use cases and approaches emerge, organizations appear to be struggling to recognize value quickly. Over a fifth (61%) of organizations have taken six months or more to see value from their AI initiatives. The top contributors to these delays include difficulty integrating AI with existing systems (41%), insufficient skilled personnel (40%), ineffective use of AI tools (33%), and an inability to scale AI solutions (32%). These factors can lead to skepticism about AI's potential and hinder further investment. This is a big reason why organizations are working hard to better streamline development and implementation, prioritize data quality, and establish clear metrics that demonstrate AI's value to the business.

AI Studio enables organizations to overcome the primary barriers to AI adoption by bridging the skills gap through no-code/low-code interfaces and pre-configured environments, accelerating time to value through streamlined development and deployment processes, simplifying integration with existing systems through flexible connectors and API services, and enabling subject matter experts to deploy AI solutions without specialized AI expertise.

Organizations leveraging Vayu AI Cloud could expect to reduce the time needed to answer complex technical queries, decrease training time for new support personnel, ensure consistent technical guidance across all global facilities, and increase first-contact resolution rates.

As demonstrated in the knowledge management example, AI Studio empowers organizations to rapidly implement enterprise-grade AI applications that deliver measurable business outcomes while addressing the most common implementation challenges.

Secure and Trusted AI Business Use Case

When selecting the right AI solution for enterprise use, teams most often look for assurance of AI governance, ethics, and responsibility; consistent and accurate results; and integration with existing tools and processes. To execute effective AI development, organizations need to focus on data, integration, and governance. Key attributes of an ideal solution include customization, control, and seamless integration with existing systems.

Vayu AI Cloud is designed to deliver a secure, ethical, and resilient environment, ensuring trust and compliance at every level. This includes a stateful firewall that provides basic security features; role based access control to ensure that users have access to just the resources and features that are appropriate to their roles; secure, segmented networking with InfiniBand; an SSL VPN client to encrypt all data in flight; and load balancer as-a-service to automate traffic distribution across multiple resources, improve resource utilization, facilitate transparent scaling, and enable high availability. Additional security is facilitated with RAID 6 data protection, available global hardware-level encryption, and checkpointing and snapshots on high-speed storage to ensure data protection, fault tolerance, and quick recovery for high-performance workloads.

In this example, a financial institution needed to modernize its loan approval process with AI while ensuring compliance with financial regulations and avoiding discriminatory lending practices. Their specific challenges included adhering to rigorous regulatory requirements for transparent and explainable lending decisions, integration of historical data containing potential socioeconomic and demographic biases, the need for consistent application of lending criteria across regions, maintaining data sovereignty for customer financial information, and implementing robust audit capabilities continuously to validate all of the above.

Using Vayu AI Cloud, the institution deployed a responsible AI-powered lending decision system using Vayu AI Cloud to meet the following requirements:

1. **Secure infrastructure setup**: Role-based access controls and segmented networking were implemented to protect sensitive financial data.
2. **Bias detection and mitigation**: Built-in explainability tools were used to identify and address potential biases in historical lending data.
3. **Ethical guardrails**: NeMo Guardrails were deployed to enforce lending policies and prevent discriminatory decisions.
4. **Transparent decision framework**: Explainability tools provided clear rationales for all lending decisions.
5. **Data sovereignty controls**: Vayu AI Cloud's data sovereignty features ensured compliance with regional financial regulations.

Key Vayu AI Cloud Security and Trust Capabilities

Comprehensive Security Infrastructure

- Vayu AI Cloud’s stateful firewall provides core network security features, and role-based access control ensures that users have access to just the resources and features that are appropriate to their roles.
- Secure, segmented networking with InfiniBand ensures high performance while keeping workloads secure.
- SSL VPN encryption secures data in transit.
- Load balancer as-a-service distributes tasks across resources, optimizing response time and providing high availability.
- RAID 6 data protection and hardware-level encryption protects and secures data at rest.

Responsible AI Framework

- NeMo Guardrails enforce ethical AI behavior, while enabling detection and mitigation of harmful content.
- Policy-based controls for AI responses enhance protection against inappropriate, biased, or misleading outputs.

AI Transparency and Explainability

- Vayu AI Cloud explainability tools help users and developers understand AI’s decision-making processes.
- Open source tools are used to identify factors that influence model predictions, provide visual and statistical explanations of outputs, and evaluate potential biases in predictions.
- Vayu AI Cloud explainability tools are critical for model debugging, complying with AI regulatory guidelines, and improving performance.

Data Sovereignty

- Vayu AI Cloud provides controls to maintain compliance with regional data regulations and secure data handling across jurisdictions. This ensures local control over data access, storage, and usage.

Why This Matters

When selecting the right AI solution for enterprise use, teams most often look for assurance of AI governance, ethics, and responsibility; consistent and accurate results; and integration with existing tools and processes. To execute effective AI development, organizations need to focus on data, integration, and governance. Key attributes of an ideal solution include customization, control, and seamless integration with existing systems.

Vayu AI Cloud is designed to enable organizations to implement AI solutions that are not only powerful but also responsible, transparent, and compliant with extant and emerging regulations.

As demonstrated in the financial lending example, Vayu AI Cloud helps organizations deploy AI in highly regulated industries with confidence, resulting in reduced loan processing times while maintaining compliance and mitigating the risks associated with biased or unethical AI decisions. Transparency and explainability required by stakeholders and regulators are maintained, along with a complete audit trail of all AI-assisted lending decisions for regulatory review. The solution maintains data sovereignty while leveraging advanced AI capabilities. The complete solution is a platform upon which organizations can build trust with customers through responsible AI practices.

By addressing the critical challenges of responsible AI implementation, Vayu AI Cloud helps organizations like our modeled financial institution realize the benefits of AI while minimizing associated risks and compliance concerns.

Conclusion

Despite the potential benefits of AI-driven applications and the appetite to use them, challenges abound in bringing them to fruition. According to Enterprise Strategy Group research, 62% of survey respondents indicated that they have faced moderate to extensive challenges moving AI models from development into production.

Enterprise Strategy Group validated that Vayu AI Cloud provides the tools and frameworks, along with the management software, necessary to help organizations build, deploy, and see immediate value from their AI-driven applications.

Key Takeaways

Tata Communications Vayu AI Cloud delivers a comprehensive AI platform that addresses four critical enterprise challenges:

Cost Reduction

- Vayu AI Cloud offers custom combinations of fully managed bare-metal NVIDIA H100, H200, L40S, and L4 GPUs at approximately one-third the cost of hyperscalers. Customers have the option to switch at any time without the burden of capital expenditure.
- Hidden expenses are eliminated with zero ingress and egress data transfer charges.
- Resource utilization is optimized through Kubernetes/SLURM orchestration with auto-scaling capabilities.
- Pricing models are transparent and predictable compared to complex hyperscaler billing structures.

Enhanced Data Management

- Vayu AI Cloud integrates high-speed parallel file systems and S3-compatible object storage for optimized AI workloads.
- Seamless multi-cloud connectivity missing from GPU-only service providers is provided for unified data access across environments.
- Comprehensive observability tools provide end-to-end visibility and data pipeline optimization.
- Efficient data workflows are supported through integrated storage solutions missing from GPU-only service providers.

Security and Governance Enhancement

- Vayu AI Cloud implements secure, segmented networking architecture for protected AI development.
- Self-service capabilities with granular controls and responsible AI tools maintain governance over sensitive data.
- Compliance with data sovereignty requirements is enabled through regional deployment options.
- Dedicated bare-metal infrastructure provides complete isolation of sensitive AI workloads.

Skills Gap Mitigation

- Vayu AI Cloud accelerates development with AI Workbench, featuring pre-configured environments for immediate productivity.
- Model access is simplified through AI Supermarket with ready-to-use premium models (Llama, Cohere, Mistral, etc.).
- Technical complexity is reduced through fully managed GPU services and operational support.

With Vayu AI Cloud, organizations can significantly reduce the time to market and time to value for deploying their AI-driven applications since they won't have to manage the infrastructure but instead can focus on developing the application and realizing business value.

This integrated approach enables organizations to deploy enterprise-grade AI solutions faster while maintaining control over infrastructure, costs, and security, addressing the key barriers to successful AI implementation that competitors with less-complete offerings fail to comprehensively solve.

If your organization is working toward seamlessly integrating impactful AI into every aspect of its business with minimal effort to speed time to market, provide flexibility to grow the business, and secure a competitive advantage, it's Enterprise Strategy Group's considered opinion that you would be smart to seriously examine Tata Communications Vayu AI Cloud.

About Enterprise Strategy Group
Enterprise Strategy Group, now part of Omdia, provides focused and actionable market intelligence, demand-side research, analyst advisory services, GTM strategy guidance, solution validations, and custom content supporting enterprise technology buying and selling.

contact@esg-global.com
www.esg-global.com