![aws FOR TELECOM]

WHITEPAPER

A Pragmatic Approach to Generative AI

aws FOR TELECOM

A Pragmatic Approach to Generative AI

Telecom executives face rising costs, margin pressures, fierce competition, and disruptive technologies. Communication service providers (CSPs) that can expand their networks, innovate business models, and adopt new technologies while controlling expenses will generate value for stakeholders. The emergence of generative AI presents an opportunity for forward-thinking telecom executives to progress toward these goals.

Generative AI: A Breakthrough in AI

Generative AI marks an advancement in AI and machine learning (ML), enabled by the affordability of large-scale computing, the availability of extensive data corpora, and the innovative Transformer Architecture. Unlike traditional approaches focused on prediction, classification, or optimization, generative AI, powered by versatile pre-trained foundation models, can create content.

The proven value of foundation models, including large language models (LLMs), positions generative AI as a strategic opportunity for telecoms. Quick wins can demonstrate its value, secure leadership commitment, and motivate teams to expand AI usage.

Promising Telecoms Use Cases

Early engagements with telecoms and AWS ecosystem partners have demonstrated how generative AI can reduce revenue leakage and customer churn, improve cross-selling attach rates, decrease customer support resolution times, and streamline network operations while reducing troubleshooting time.

Generative AI enables staff to focus on higher-value tasks by automating repetitive tasks such as information searching and discovery, planning and summarization, and report and content generation. By improving job satisfaction, generative AI services support employee retention.

The Role of Cloud in Telecom and AI

The cloud has revolutionized telecoms. First, transitioning from hardware-based network appliances to software applications on cloud infrastructure offers operational simplicity and reduced ownership costs. Second, operators can monetize assets and kickstart new business opportunities by integrating cloud at the network edge for enterprise applications and creating private networks for industries.

Third, cloud provides the computing, storage, and services to transform valuable telecom data into actionable insights. AI layered on top of a well-managed and modern data platform, can help operators develop new services, improve operations, and enhance customer experiences.

Paths to Leveraging Generative AI

Operators will consume AI through productivity applications and enterprise software, but the greatest potential value comes from combining private telecom data with generative AI. To achieve this, operators are pursuing multiple approaches:

- **Pre-Trained Models:** Incorporating private data in prompt contexts for pre-trained foundation models is the quickest way to achieve results without extensive customization.

- **Fine-tuning Models:** Adapting pre-trained models to telecom use cases by fine-tuning with private data offers a balance between customization and speed, potentially yielding higher accuracy and performance.

- **Custom Model Training:** Training custom foundation models from scratch using large telecom datasets can provide proprietary capabilities and maximize competitive advantage, but it requires significant time and investment.

We recommend starting with pre-trained models and fine-tuning before embarking on building models. Operators should evolve their approach based on use case requirements, team capabilities, and the maturity of their data modernization efforts.

Practicing Responsible AI

Generative AI presents challenges around security, privacy, and ethics. Telecommunication operators are regulated and choosing a trusted cloud partner committed to comprehensive security across physical infrastructure, hardware, software, people, and processes is crucial. Operators should also establish system safeguards, encrypt data, and restrict model access.

To prevent privacy violations, operators must implement strict controls and anonymize customer data for training or fine-tuning. They should also monitor outputs to prevent the leak of sensitive information.

Additionally, deploying tools to ensure accuracy, promote fairness, and reduce biases is essential. Any potentially inaccurate, problematic, or suspicious outputs should be flagged for human review.

Furthermore, operators should support transparency by documenting their methodology, training data sources, and incorporating explainability. With proper policies, processes, and tools, operators can innovate safely with generative AI.

Why AWS for Generative AI

AWS offers a comprehensive solution for generative AI, combining carrier-grade security, foundation model choice, data-centric approach, and high-performance, cost-effective infrastructure:

- **Secure, Private, and Easy to Use:** AWS enables operators to develop and scale generative AI applications securely and privately. Operators can utilize leading commercial and open-source foundation models, customize them, and benefit from AWS's advanced security and access control.

- **High Performance, Low-Cost Infrastructure:** AWS provides an optimal environment for training and running ML applications, featuring high-performance GPU-based Amazon EC2 P5 instances and purpose-built accelerators like AWS Trainium and AWS Inferentia. Telecom operators with sustainability initiatives will benefit from the power efficiency of AWS infrastructure.

- **Transformative AI-Powered Applications:** AWS offers generative AI-powered applications, such as Amazon CodeWhisperer and QuickSight, that enhance developer and managerial productivity. These applications offer customization, privacy, and seamless integration with private data.

- **Data-Driven Customization:** Amazon AI services like CodeWhisperer support customization, adapting to internal coding styles and telecom-specific APIs and libraries. AWS's extensive AI service portfolio includes Amazon SageMaker, a fully managed ML service for fine-tuning, building, and deploying foundation models. Meanwhile, Amazon Bedrock facilitates the development and scaling of generative AI applications with a choice of models from AI21 Labs, Anthropic, Cohere, Meta, and Stability AI. Accelerating Telecom Generative AI Adoption through Cloud.

Beyond cloud services, telecom operators must recognize the importance of upskilling and retraining employees in cloud and AI. AWS has established free training programs in cloud skills for 29 million individuals. Our Innovation Center, staffed by AI scientists, solution architects, and other experts, works closely with telecom developers and customers to facilitate AI adoption.

Accelerating Telecom Generative AI Adoption through Cloud

Recognizing and adapting to technological shifts is vital, especially in the traditionally conservative yet societally impactful telecom industry. Leaders must understand the substantial change generative AI represents and respond strategically. This involves upskilling staff, recognizing the potential of generative AI, and understanding the capabilities provided by cloud partners like AWS.

Strategic partnerships with foundation model developers and cloud companies like AWS and industry bodies that foster collaboration, such as TM Forum, are vital. The collaboration can accelerate the identification of practical use cases and speed up organization onboarding.

Lastly, adopting a new mindset is necessary. As with cloud adoption, this must permeate the entire organization, from executives to engineers. This workforce transformation is as crucial as the technological shift in ensuring the telecom sector remains at the forefront of innovation and societal progress.

AWS is working step-by-step with customers to help them navigate this transformative journey through reskilling, and building of bespoke solutions that harness the power of generative AI.

To get started, visit our generative AI innovation center and telecom generative AI solutions page https://aws.amazon.com/telecom/generative-ai/