AWS Official Blog

AWS Weekly Roundup: Amazon Q, Amazon QuickSight, AWS CodeArtifact, Amazon Bedrock, and more (May 6, 2024)
April has been packed with new releases! Last week continued that trend with many new releases supporting a variety of domains such as security, analytics, devops, and many more, as well as more exciting new capabilities within generative AI. If you missed the AWS Summit London 2024, you can now watch the sessions on demand, […]

Build RAG applications with MongoDB Atlas, now available in Knowledge Bases for Amazon Bedrock
Foundational models (FMs) are trained on large volumes of data and use billions of parameters. However, in order to answer customers' questions related to domain-specific private data, they need to reference an authoritative knowledge base outside of the model's training data sources. This is commonly achieved using a technique known as Retrieval Augmented Generation (RAG). […]

Stop the CNAME chain struggle: Simplified management with Route 53 Resolver DNS Firewall
Updated 2 May 2024: I removed the reference to Route53 Alias that was incorrectly referred as a chain Starting today, you can configure your DNS Firewall to automatically trust all domains in a resolution chain (such as aCNAMEor DNAMEchain). Let's walk through this in nontechnical terms for those unfamiliar with DNS. Why use DNS Firewall? […]

Add your Ruby gems to AWS CodeArtifact
Ruby developers can now use AWS CodeArtifact to securely store and retrieve their gems. CodeArtifact integrates with standard developer tools like gem and bundler. Applications often use numerous packages to speed up development by providing reusable code for common tasks like network access, cryptography, or data manipulation. Developers also embed SDKs–such as the AWS SDKs–to […]

Amazon Titan Text Embeddings V2 now available in Amazon Bedrock, optimized for improving RAG
Amazon Titan foundation models (FMs) offer a comprehensive suite of pre-trained image, multimodal, and text models accessible through a fully managed API. They're designed for a range of applications while adhering to responsible AI practices.

Amazon Q Business, now generally available, helps boost workforce productivity with generative AI
Amazon Q Business is now available with many new features, including custom plugins, and a preview of Amazon Q Apps, generative AI–powered customized and sharable applications using natural language in a single step for your organization.

Amazon Q Developer, now generally available, includes previews of new capabilities to reimagine developer experience
This generative AI–powered assistant helps developers build applications on AWS, research best practices, perform troubleshooting, and resolve errors.

Run scalable, enterprise-grade generative AI workloads with Cohere Command R & R+, now available in Amazon Bedrock
In November 2023, we made two new Cohere models available in Amazon Bedrock (Cohere Command Light and Cohere Embed English). Today, we're announcing the addition of two more Cohere models in Amazon Bedrock; Cohere Command R and Command R+. Organizations need generative artificial intelligence (generative AI) models to securely interact with information stored in their […]

AWS Weekly Roundup: Amazon Bedrock, AWS CodeBuild, Amazon CodeCatalyst, and more (April 29, 2024)
This was a busy week for Amazon Bedrock with many new features! Using GitHub Actions with AWS CodeBuild is much easier. Also, Amazon Q in Amazon CodeCatalyst can now manage more complex issues. I was amazed to meet so many new and old friends at the AWS Summit London. To give you a quick glimpse, […]

Meta Llama 3 models are now available in Amazon Bedrock
The instruction-tuned large language models have been pre-trained on over 15 trillion tokens of data—a training dataset seven times larger than that used for Llama 2 models.

Guardrails for Amazon Bedrock now available with new safety filters and privacy controls
With Guardrails for Amazon Bedrock, you can implement safeguards in your generative artificial intelligence applications that are customized to your use cases and responsible AI policies.

Agents for Amazon Bedrock: Introducing a simplified creation and configuration experience
With Agents for Amazon Bedrock, applications can use generative artificial intelligence (generative AI) to run tasks across multiple systems and data sources. Starting today, these new capabilities streamline the creation and management of agents: Quick agent creation – You can now quickly create an agent and optionally add instructions and action groups later, providing flexibility […]

Amazon Bedrock model evaluation is now generally available
This new capability helps you to incorporate generative AI into your application by letting you select the foundation model that gives you the best results for your particular use case.

Import custom models in Amazon Bedrock (preview)
With Amazon Bedrock, you have access to a choice of high-performing foundation models (FMs) from leading artificial intelligence (AI) companies that make it easier to build and scale generative AI applications. Some of these models provide publicly available weights that can be fine-tuned and customized for specific use cases. However, deploying customized FMs in a […]

Amazon Titan Image Generator and watermark detection API are now available in Amazon Bedrock
All images generated by Titan Image Generator contain an invisible watermark to help reduce the spread of misinformation by providing a mechanism to identify AI-generated images.

Unify DNS management using Amazon Route 53 Profiles with multiple VPCs and AWS accounts
If you are managing lots of accounts and Amazon Virtual Private Cloud (Amazon VPC) resources, sharing and then associating many DNS resources to each VPC can present a significant burden. You often hit limits around sharing and association, and you may have gone as far as building your own orchestration layers to propagate DNS configuration […]

AWS Weekly Roundup: Anthropic's Claude 3 Opus in Amazon Bedrock, Meta Llama 3 in Amazon SageMaker JumpStart, and more (April 22, 2024)
AWS Summits continue to rock the world, with events taking place in various locations around the globe. AWS Summit London (April 24) is the last one in April, and there are nine more in May, including AWS Summit Berlin (May 15–16), AWS Summit Los Angeles (May 22), and AWS Summit Dubai (May 29). Join us […]

Amazon CloudWatch Internet Weather Map – View and analyze internet health
The Internet has a plethora of moving parts: routers, switches, hubs, terrestrial and submarine cables, and connectors on the hardware side, and complex protocol stacks and configurations on the software side. When something goes wrong that slows or disrupts the Internet in a way that affects your customers, you want to be able to localize […]

Anthropic's Claude 3 Opus model is now available on Amazon Bedrock
We are living in the generative artificial intelligence (AI) era; a time of rapid innovation. When Anthropic announced its Claude 3 foundation models (FMs) on March 4, we made Claude 3 Sonnet, a model balanced between skills and speed, available on Amazon Bedrock the same day. On March 13, we launched the Claude 3 Haiku […]

Congratulations to the PartyRock generative AI hackathon winners
The PartyRock Generative AI Hackathon wrapped up earlier this month. Entrants were asked to use PartyRock to build a functional app based on one of four challenge categories, with the option to remix an existing app as well. The hackathon attracted 7,650 registrants who submitted over 1,200 projects, and published over 250 project blog posts […]