Students Guide to Generative AI

Du'An S. Lightfoot

Sr. Developer Advocate
AWS

Agenda

- What is generative AI?
- Prompting and prompt engineering
- Customizing foundation models
- Generative AI in Cyber Security
- How to get started

What is generative AI?

Working with generative AI

- Text, image, other media, and multi-modal models
- Summarization, analogies, translation and localization, personalization, with long memory and conversational capabilities

A study found office workers using AI assistant Claude were 35.7% more productive, produced higher quality work and made more confident decisi to Claude's data analysis and ins Researchers concluded integrati generative AI like Claude could g boost workplace efficiency and

Un estudio descubrió que los trabajadores de oficina que utilizaban el asistente de IA Claude eran un 35,7% más productivos, producían trabajo de mayor calidad y tomaban decisiones más seguras gracias al análisis de datos y las perspectivas de Claude. Los investigadores concluyeron que integrar la IA generativa como Claude podría aumentar enormemente la eficiencia y la toma de decisiones en el gar de trabajo

Image Examples

IMAGE GENERATION, TRANSFORMATION, UPSCALING

Generated by Stable Diffusion 2.0

Image transformation

4x

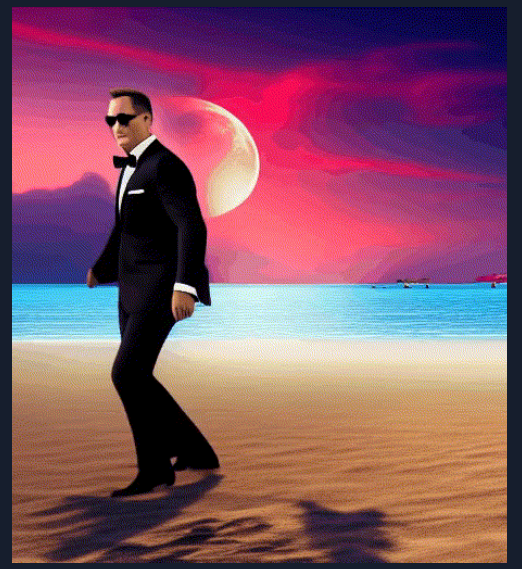
Upscaling

Generating new video content

Open Source project example

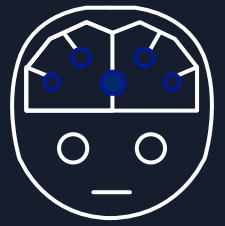
ControlVideo (May 2023)

Prompt: "A sleek boat glides effortlessly through the shimmering river, Van Gogh style"



Prompt: "James Bond moonwalk on the beach, animation style"

Where does Generative AI fit?



Artificial intelligence (AI)

Any technique that allows computers to mimic human intelligence using logic, if-then statements, and machine learning

Machine learning (ML)

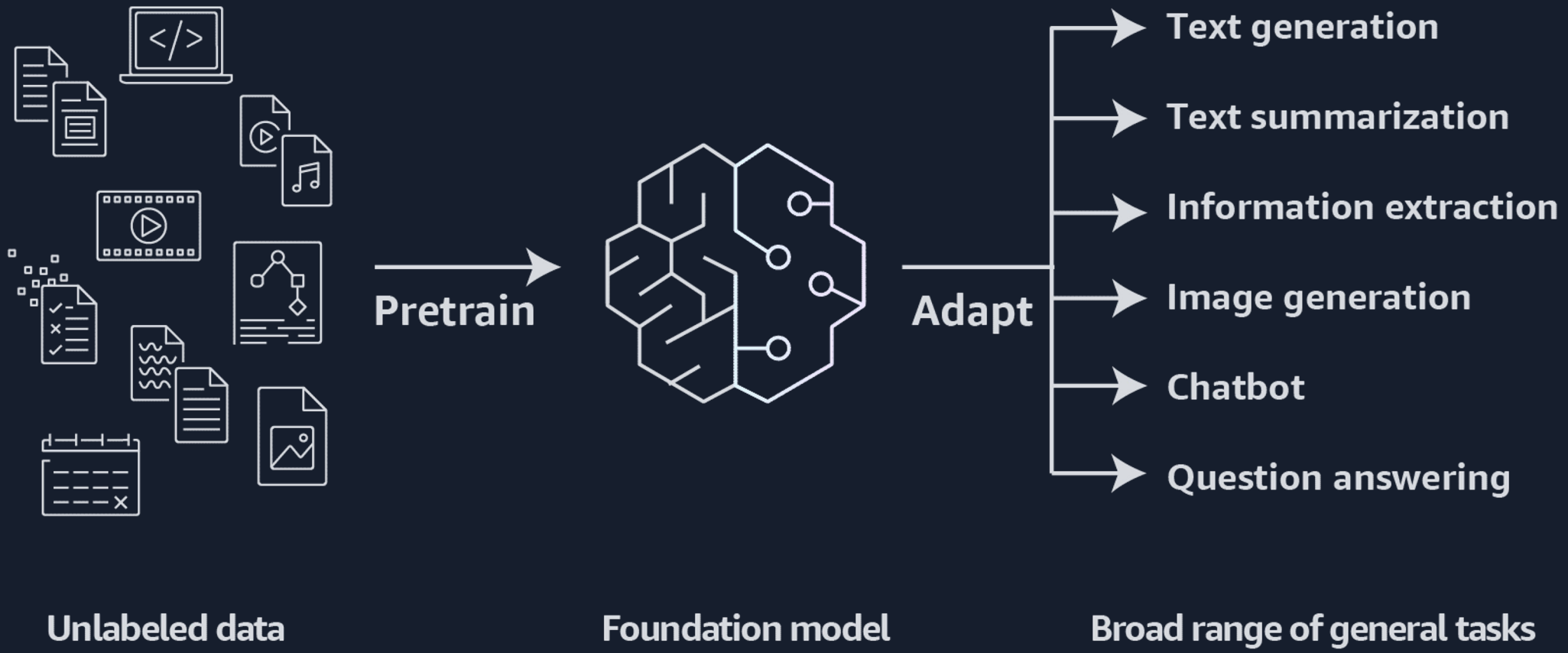
A subset of AI that uses machines to search for patterns in data to build logic models automatically

Deep learning (DL)

A subset of ML composed of deeply multi-layered neural networks that perform tasks like speech and image recognition

Generative AI

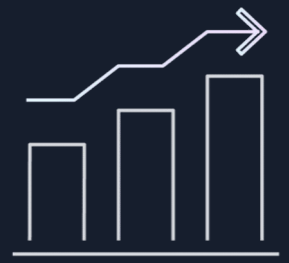
How does a foundation model function?



LLM use cases

Improves customer experience

- Chatbots
- Call analytics
- Agent assist



Boosts employee productivity

- Conversational assist
- Code generation
- Automated report generation

Enhances creativity and content creation

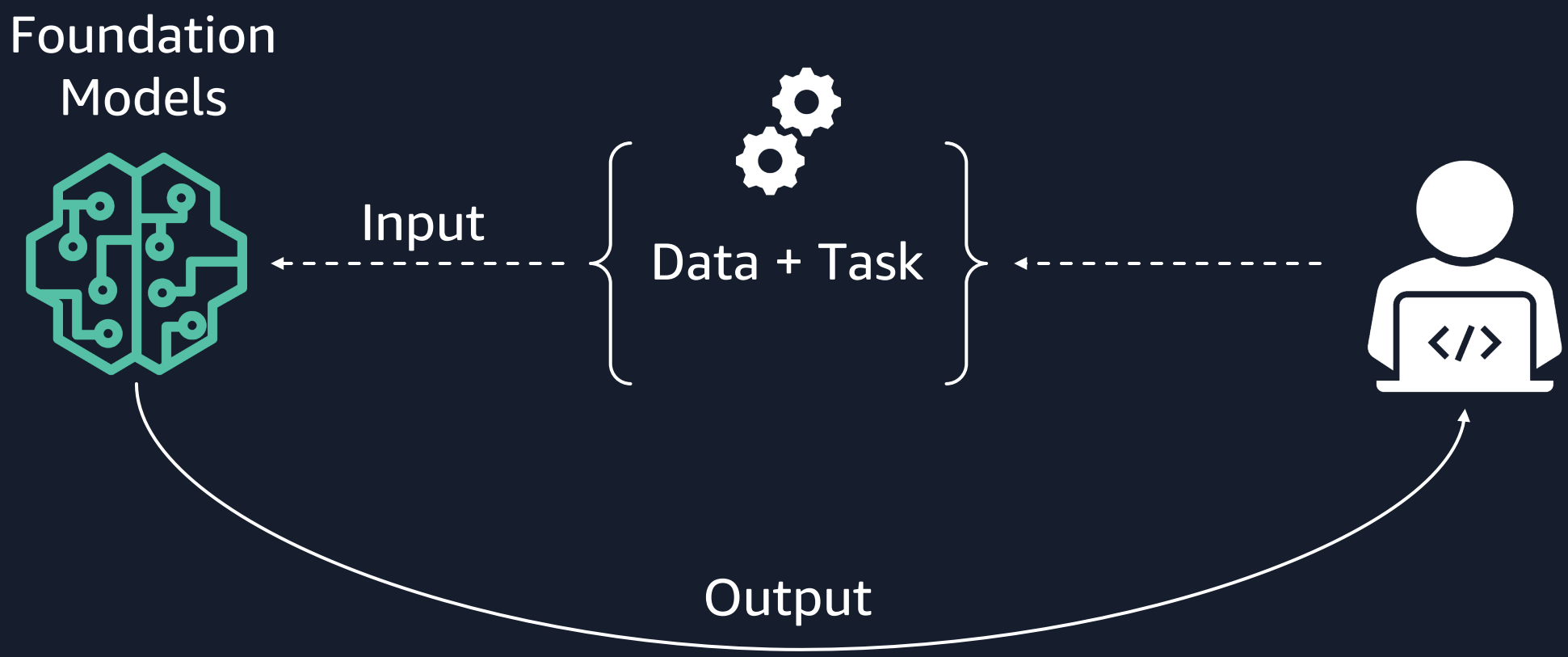
- Marketing
- Sales
- Product development
- Media and entertainment
- News generation

Accelerates process optimization

- Document processing
- Fraud detection
- Supply-chain optimization

Prompt engineering

Prompt engineering, new way of using ML!



Elements of the prompt example

Instructions and output indicator

Context

Input data

Prompt	Output
Write a summary of a service review using two sentences. Store: Online Service: Shipping Review: Amazon Prime Student is a great option for students looking to save money. Not paying for shipping is the biggest save in my opinion. As a working mom of three who is also a student, it saves me tons of time with free 2-day shipping, and I get things I need quickly and sometimes as early as the next day, while enjoying all the free streaming services and books that a regular Prime membership has to offer for half the price. Amazon Prime Student is only available for college students, and it offers so many things to help make college life easier. This is why Amazon Prime is the no-brainer that I use to order my school supplies, my clothes, and even to watch movies in between classes. I think Amazon Prime Student is a great investment for all college students. Summary:	Amazon Prime Student is a fantastic option for college students, offering free 2-day shipping, streaming services, books, and other benefits for half the price of a regular Prime membership. It saves time and money, making college life easier.

Best practices for designing effective prompts

Be clear and concise

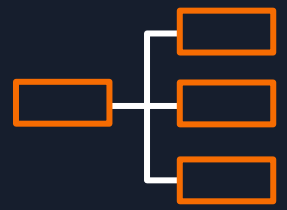
Include context

Use directives

Include output

Start with a question

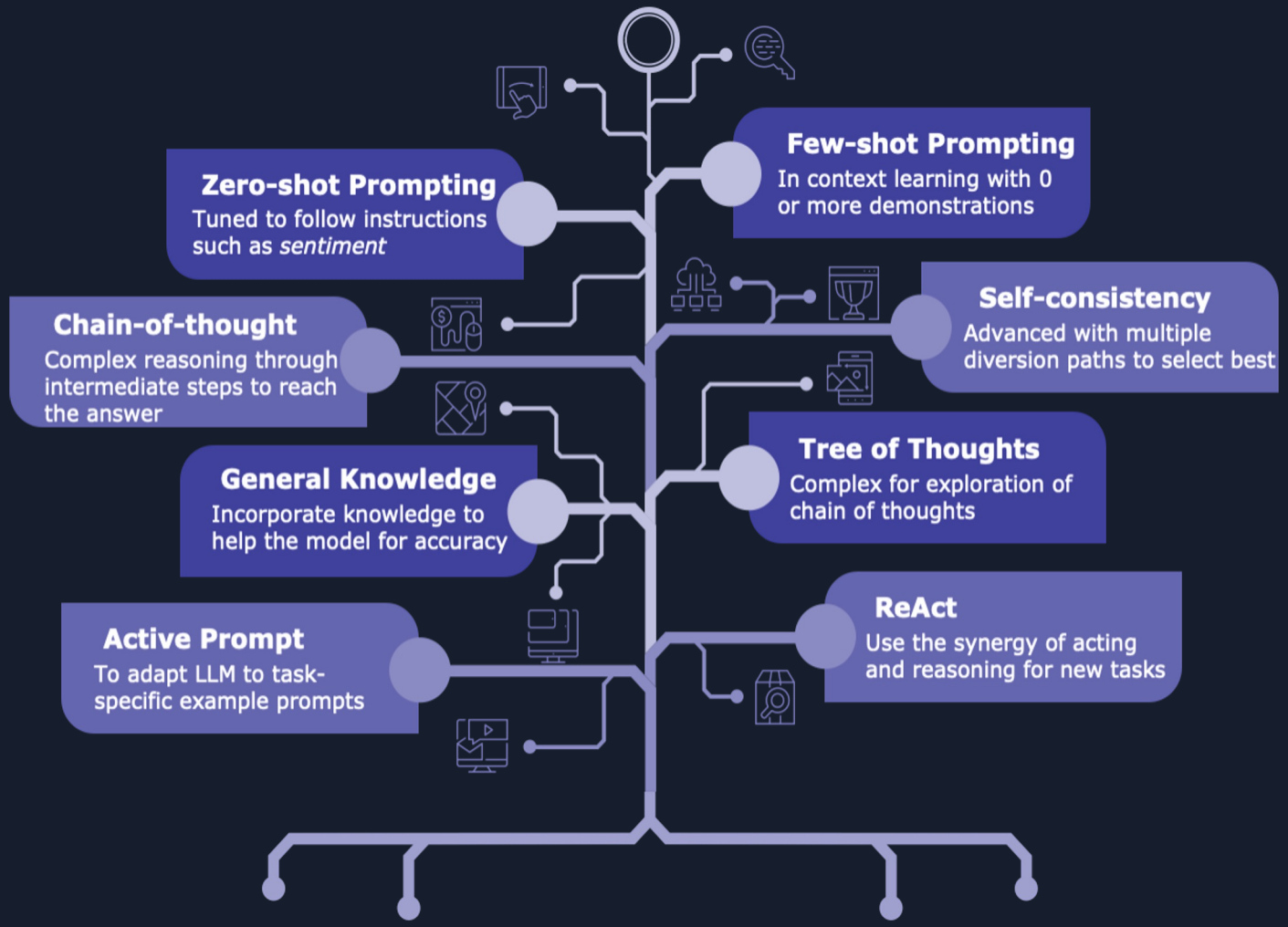
Provide example responses



Break up complex tasks

Experiment and be creative

Prompting Techniques



Prompt template

Role

Task

Background and objective

Variations

Guidelines

Examples (3-5)

Latest information

Chain of thought

Prompt Engineering

> You are a personal assistant. You are friendly, polite and casual. You help with…

> You are a classifying agent that filters user inputs into categories. Your job is to sort these inputs before they are passed along to our function calling agent. The purpose of our function calling agent is to call functions in order to answer user's questions.

Prompt Engineering

“
```
The user input is between the
<question></question> XML tags:
<question>
What is the weather in Oslo right now?
</question>
```
”

Which prompt follows prompting best practices?

INCLUDE CONTEXT IF NEEDED

Prompt 1
Summarize this article: [insert article text]

or

Prompt 2
Provide a summary of this article to be used in a blog post: [insert article text]

Which prompt follows prompting best practices?

INCLUDE CONTEXT IF NEEDED

Prompt 1
Summarize this article: [insert article text]

or

Prompt 2
Provide a summary of this article to be used in a blog post: [insert article text]

Which prompt follows prompting best practices?

PROVIDE AN EXAMPLE RESPONSE

Prompt 1
Determine the sentiment of this social media post: “[insert post]”

or

Prompt 2
Determine the sentiment of the following social media post using these examples: post: “great pen” // Positive post: “I hate when my phone battery dies” // Negative “[insert post]” //

Which prompt follows prompting best practices?

PROVIDE AN EXAMPLE RESPONSE

Prompt 1
Determine the sentiment of this social media post: “[insert post]”

or

Prompt 2
Determine the sentiment of the following social media post using these examples: post: “great pen” // Positive post: “I hate when my phone battery dies” // Negative “[insert post]” //

Challenges with LLMs

Lack of Domain Knowledge

Temporal Unawareness

Training Data Cut-off Date

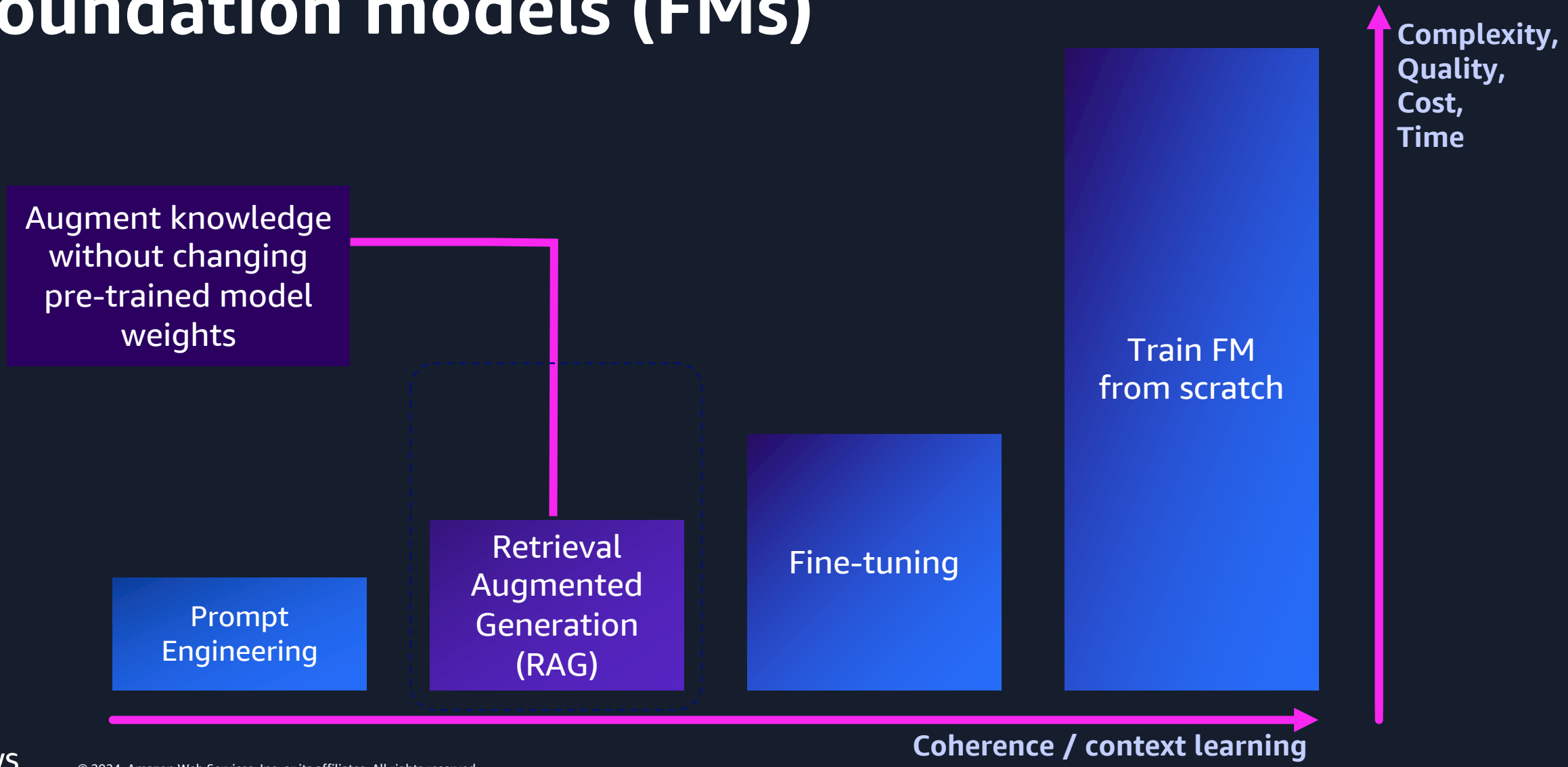
Model hallucination

Lack of Proprietary Knowledge

Privacy and Security

Customizing foundation models

Common approaches for customizing foundation models (FMs)



What is Retrieval Augmented Generation?

Retrieval

Fetches the relevant content from the external knowledge base or data sources based on a user query

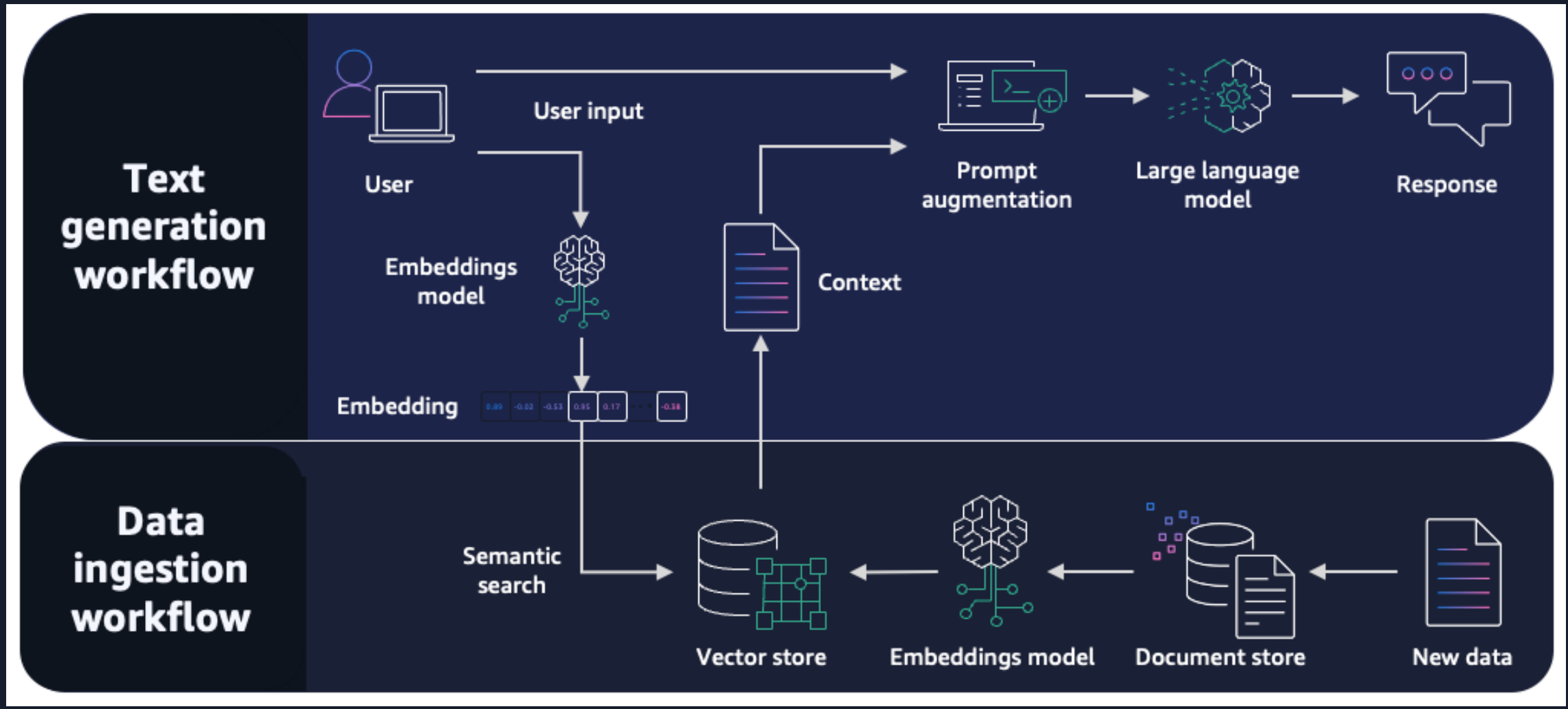
Augmentation

Adding the retrieved relevant context to the user prompt, which goes as an input to the foundation model

Generation

Response from the foundation model based on the augmented prompt.

How RAG works



Generative AI Stack

APPLICATIONS THAT LEVERAGE LLMs AND OTHER FMs

Amazon Q Business | Amazon Q Developer | Amazon Q in QuickSight | Amazon Q in Connect

TOOLS TO BUILD WITH LLMs AND OTHER FMs

Amazon Bedrock

Guardrails | Agents | Studio | Customization Capabilities | Custom Model Import

INFRASTRUCTURE FOR FM TRAINING AND INFERENCE

GPUs | Trainium | Inferentia | SageMaker

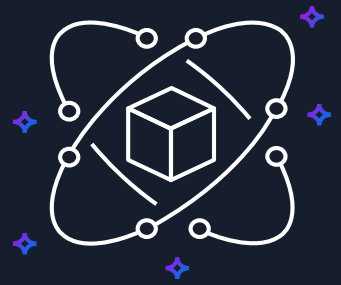
UltraClusters | EFA | EC2 Capacity Blocks | Nitro | Neuron

Generative AI in cyber security

Three Ways to think about generative AI + security

Security of generative AI

How do I secure my business applications that leverage generative AI?



Generative AI for security

How can I use generative AI to minimize vulnerabilities, threats, and risks?

Security from generative AI-powered threats

How can I protect against threat actors using generative AI?

Potential sources of AI-powered threat

A NON-EXHAUSTIVE LIST

Denial of Service

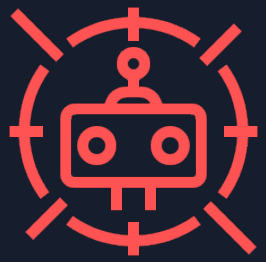
- SYN Floods
- Reflection Attacks
- Web Request Floods

App Vulnerabilities

- SQL Injection
- Cross-site Scripting (XSS)
- OWASP Top 10
- Common Vulnerabilities and Exposures (CVE)
- Privilege Escalation
- Data Exfiltration

Good Bots

- Search Engine Crawling
- Website Health Monitoring
- Vulnerability Scanning



Bad Bots

- Comment Spam
- SEO Spam
- Fraud
- Account Takeover

The OWASP® Top for 10 Large Language Models (LLMs)

LLM01

Prompt Injection

This manipulates a large language model (LLM) through crafty inputs, causing unintended actions by the LLM. Direct injections overwrite system prompts, while indirect ones manipulate inputs from external sources.

LLM02

Insecure Output Handling

This vulnerability occurs when an LLM output is accepted without scrutiny, exposing backend systems. Misuse may lead to severe consequences like XSS, CSRF, SSRF, privilege escalation, or remote code execution.

LLM03

Training Data Poisoning

This occurs when LLM training data is tampered, introducing vulnerabilities or biases that compromise security, effectiveness, or ethical behavior.

LLM04

Model Denial of Service

Attackers cause resource-heavy operations on LLMs, leading to service degradation or high costs. The vulnerability is magnified due to the resource-intensive nature of LLMs and unpredictability of user inputs.

LLM05

Supply Chain Vulnerabilities

LLM application lifecycle can be compromised by vulnerable components or services, leading to security attacks. Using third-party datasets, pre- trained models, and plugins can add vulnerabilities.

LLM06

Sensitive Information Disclosure

LLMs may inadvertently reveal confidential data in its responses, leading to unauthorized data access, privacy violations, and security breaches. It's crucial to implement data sanitization and strict user policies to mitigate this.

LLM07

Insecure Plugin Design

LLM plugins can have insecure inputs and insufficient access control. This lack of application control makes them easier to exploit and can result in consequences like remote code execution

LLM08

Excessive Agency

LLM-based systems may undertake actions leading to unintended consequences. The issue arises from excessive functionality, permissions, or autonomy granted to the LLM-based systems.

LLM09

Overreliance

Systems or people overly depending on LLMs without oversight may face misinformation, miscommunication, legal issues, and security vulnerabilities due to incorrect or inappropriate content generated by LLMs.

LLM10

Model Theft

This involves unauthorized access, copying, or exfiltration of proprietary LLM models. The impact includes economic losses, compromised competitive advantage, and potential access to sensitive information.

Source: https://owasp.org/www-project-top-10-for-large-language-model-applications/

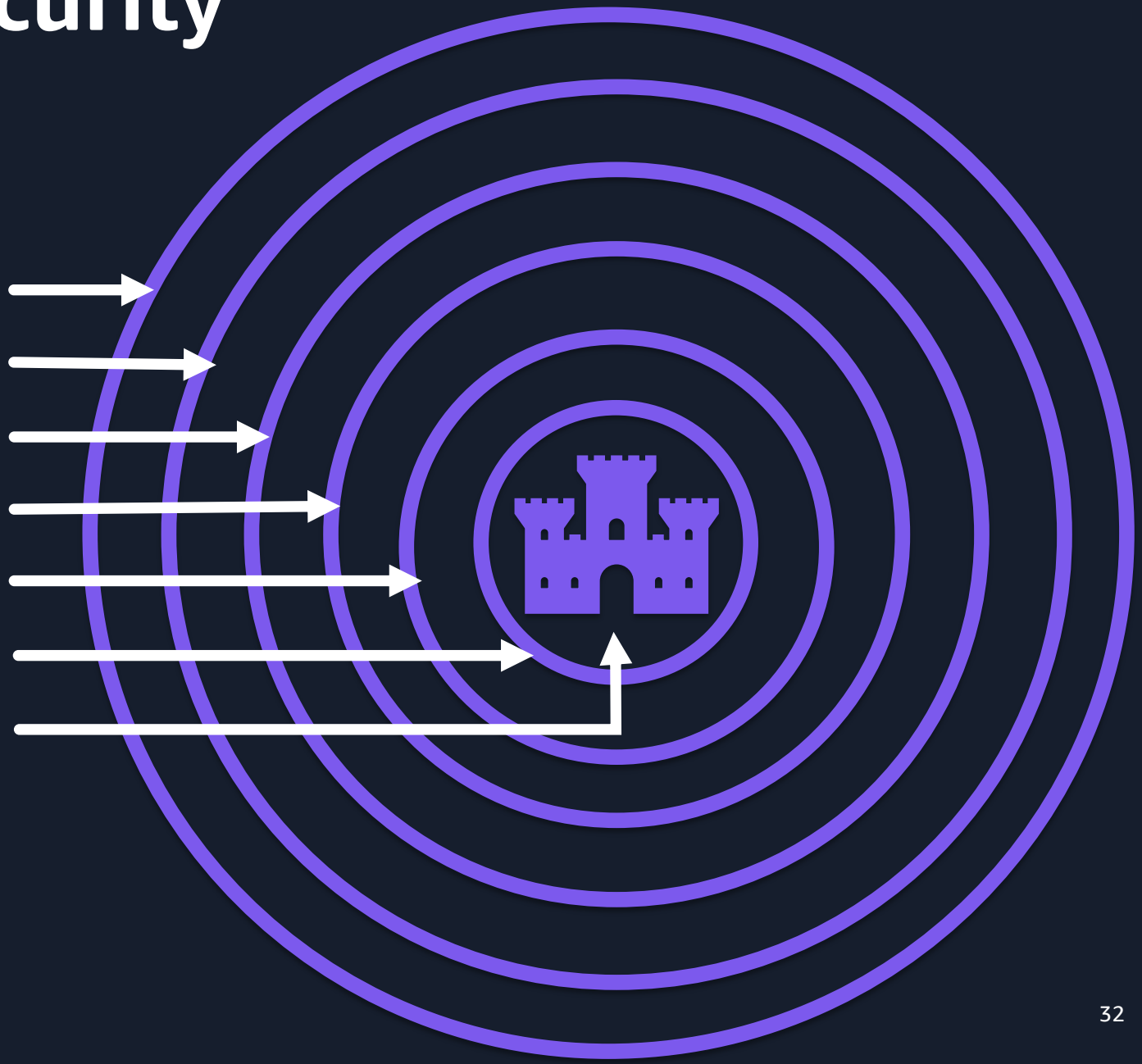
There's no silver bullet solution with cyber security, a layered defense is the only viable defense.

James Scott

Institute for Critical Infrastructure Technology

Defense-in-depth security

- Policies, Procedures & Awareness
- Network & Edge Protection
- Identity & Access Management
- Threat Detection & Incident Response
- Infrastructure Protection
- Application Protection
- Data Protection



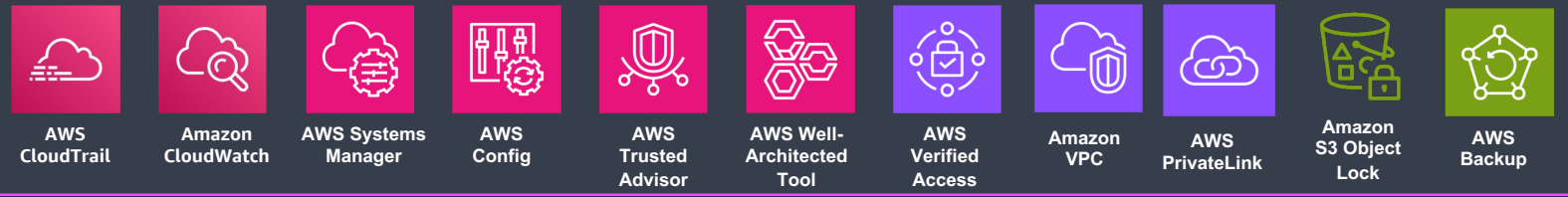
AWS generative AI and security integrated together

FOUNDATIONAL AWS SECURITY + ADDITIONAL SECURITY FEATURES OF GENERATIVE AI SERVICES

AWS Generative AI Services

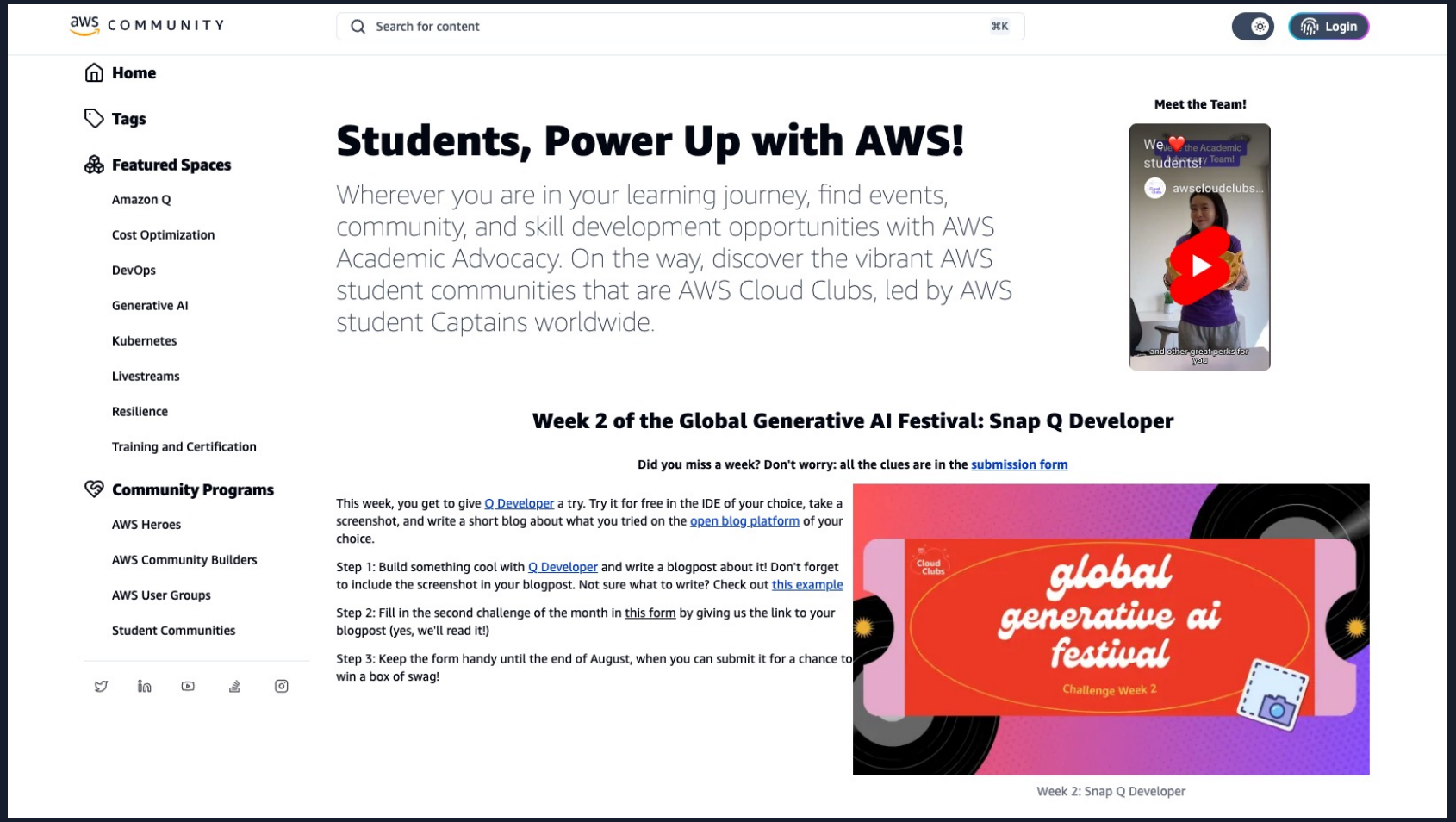
AWS Security, Identity & Compliance Services

AWS Cloud Ops, Networking, and Storage



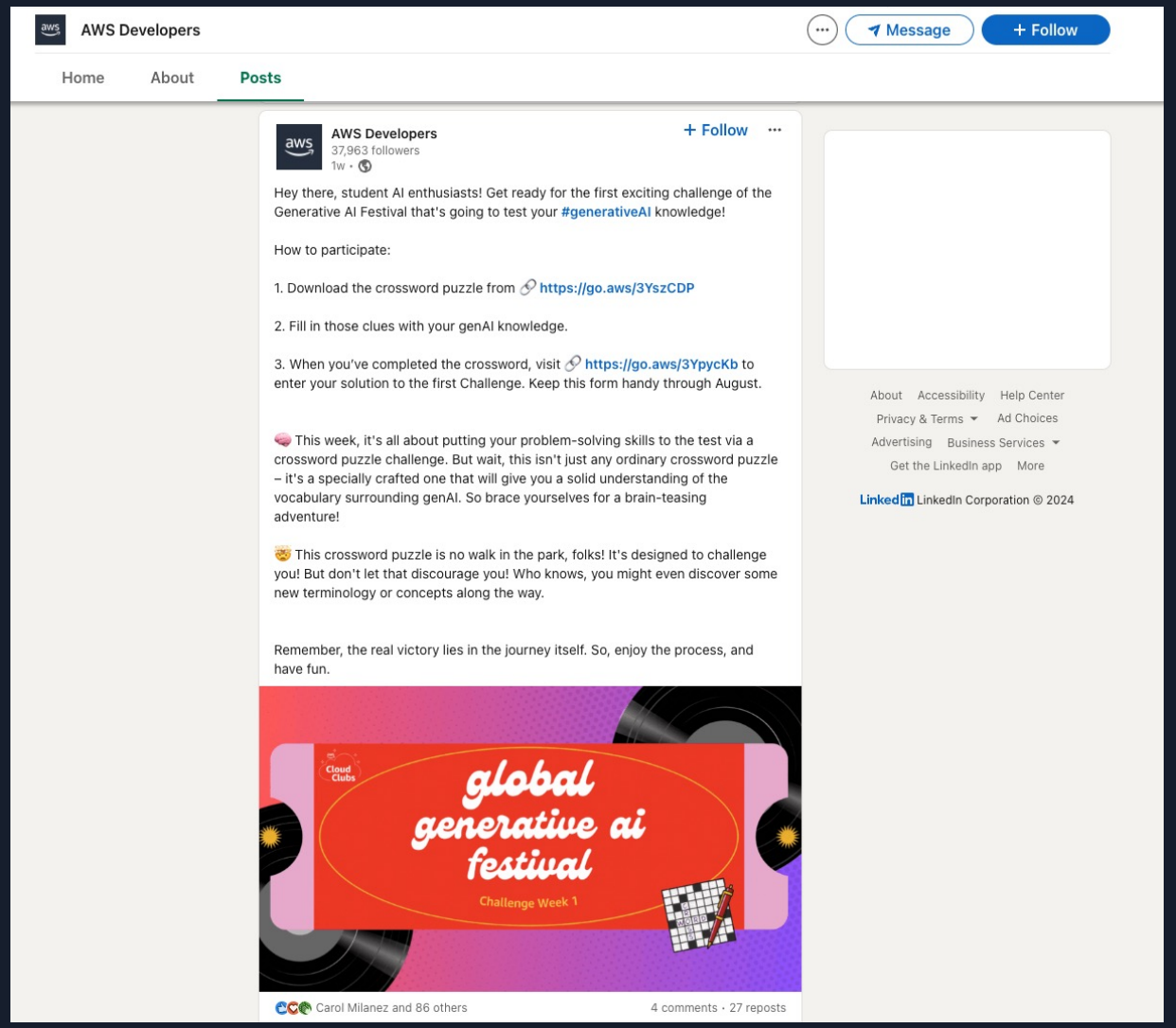
How to get started

AWS Cloud Clubs – Generative AI Festival



https://community.aws/students

Follow AWS Developers on LinkedIn



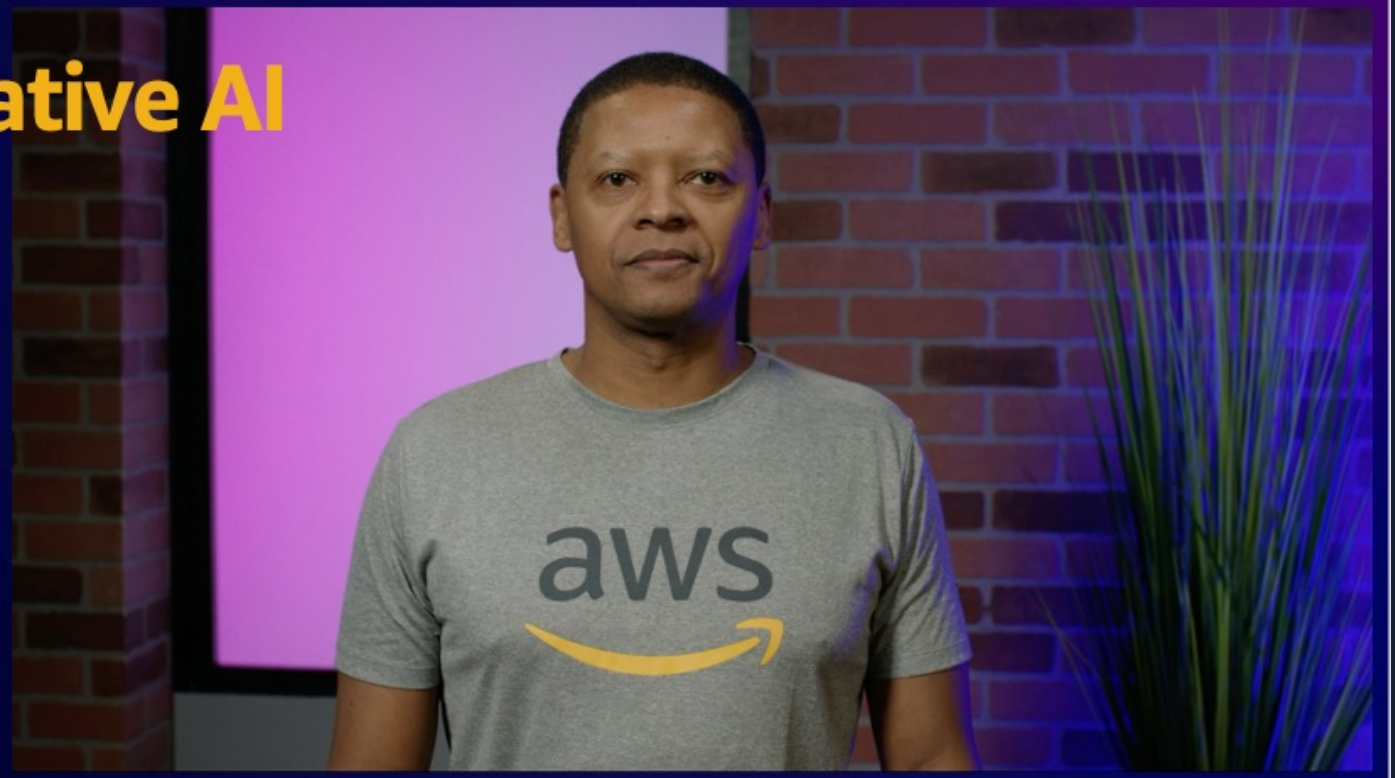
https://www.linkedin.com/showcase/aws-developers

Fundamentals of Generative AI for Beginners

Enroll for free!

coursera

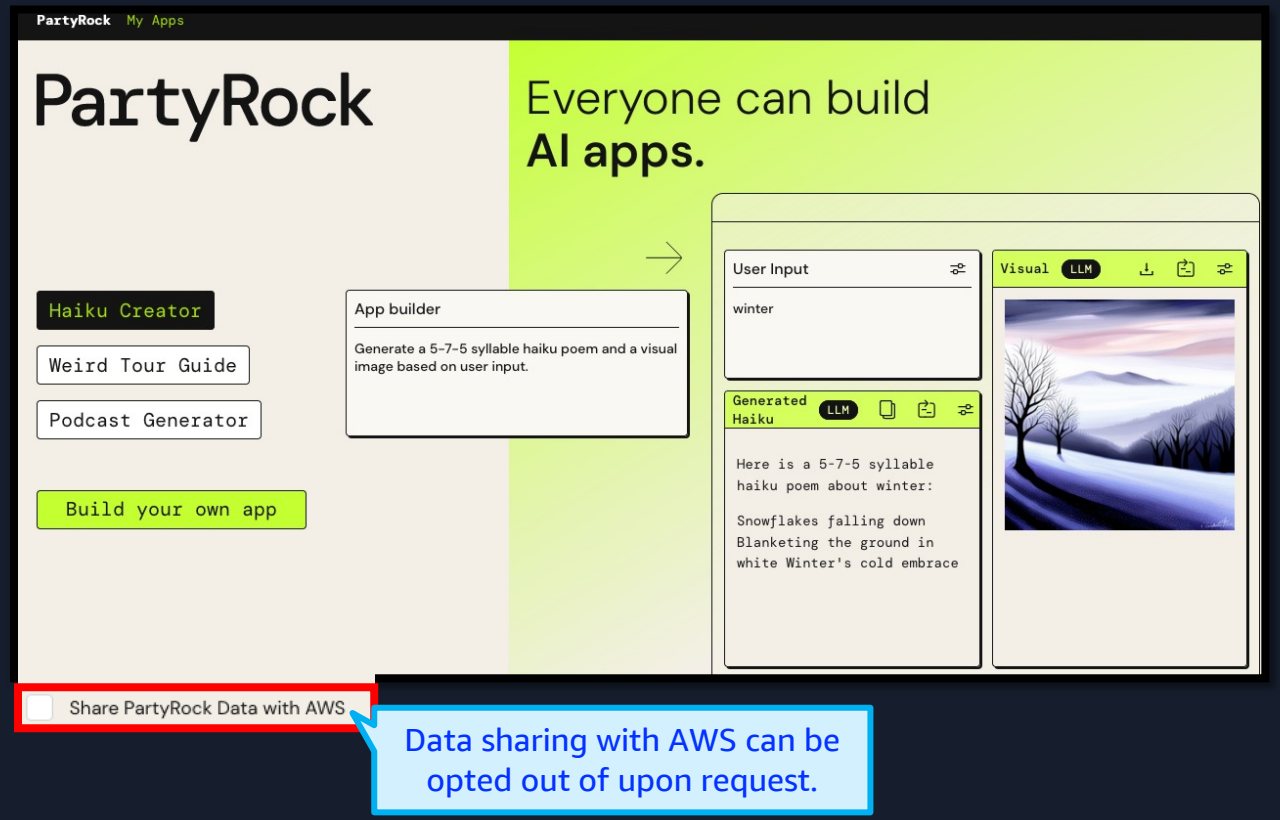
https://s12d.com/gen-ai-for-beginners



PartyRock – Try out, create, and share generative AI apps

A free service that allows you to create generative AI apps powered by Amazon Bedrock without coding (X No AWS account required, free usage up to a certain limit)

https://partyrock.aws/



- A "playground" where you can just type a prompt to generate AI applications and share them with others.
- No account needed. You can instantly register as a user and start using the service just by having the following ID.
  - Google
  - Apple ID
  - Amazon
- You can combine and use the following with widgets:
  - User input
  - Static text
  - Chatbot (by Foundation Model)
  - Text generation (by Foundation Model)
  - Image generation (by Foundation Model)

PartyRock – The “widget” feature of generative AI apps

The application's UI allows placing 5 types of widgets. Based on user input, static text, etc., it is possible to invoke the underlying model.

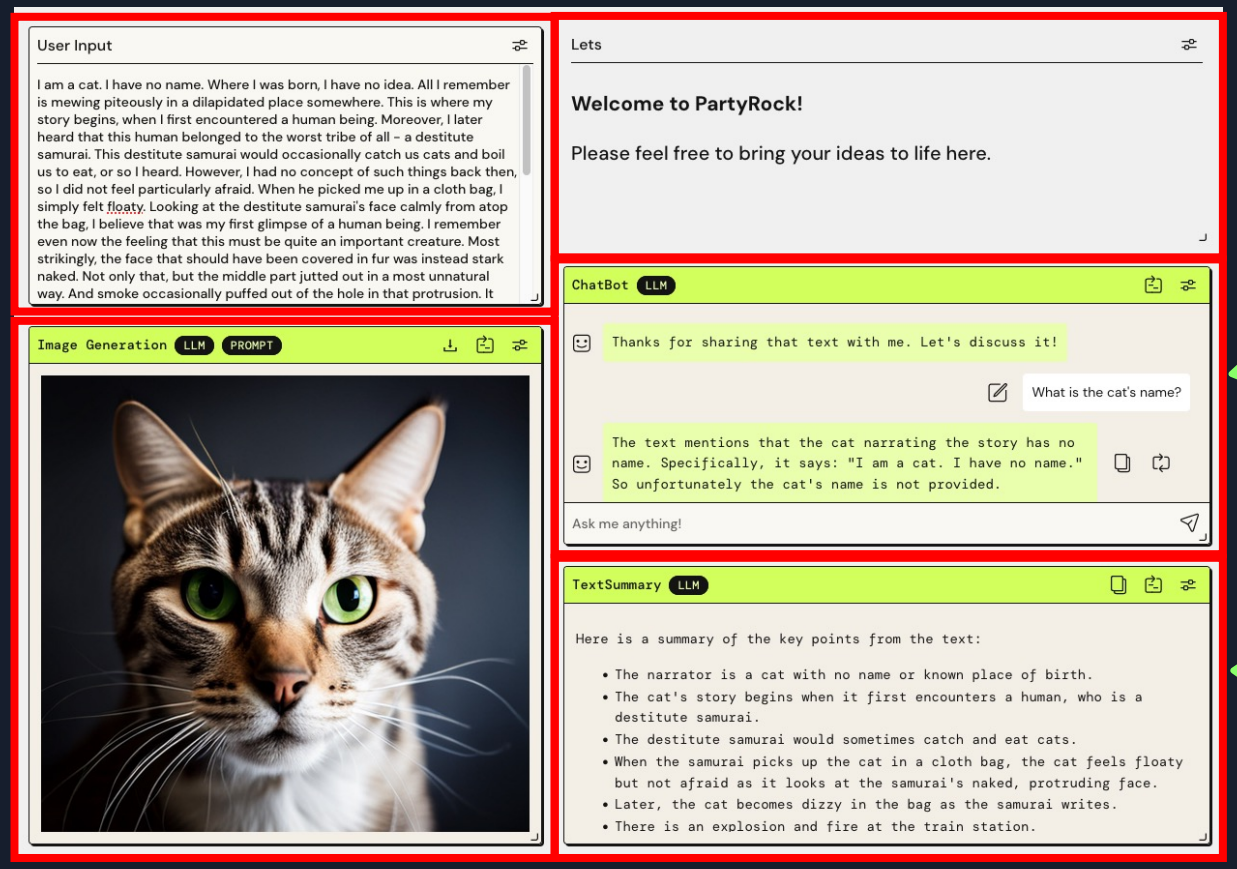
1. User Input

2. Image Generation

3. Static Text

4. Chatbot

5. Text Generation



Amazon Q Developer

Reimagines the experience across the entire software delivery lifecycle

Helps developers and IT professionals build and manage secure, scalable, and highly available applications

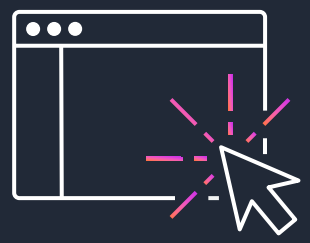
Helps you write, debug, test, optimize, and upgrade your code faster

Converses with you to explore new AWS capabilities, learn unfamiliar technologies, and architect solutions

Amazon Q is built with security and privacy in mind from the start, making it easier for organizations to use generative AI safely.

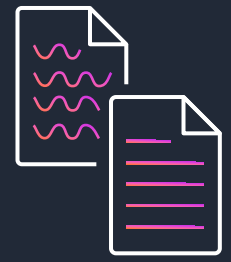
https://aws.amazon.com/q/developer/

Amazon Q is available where you do your work



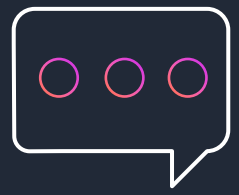
AWS Consoles

IDEs



AWS Documentation

AWS Console Mobile App



Slack and Teams (via AWS Chatbot)

Amazon CodeCatalyst

Thank you!

Du'An Lightfoot

Socials: @labeveryday

Email: duanlig@amazon.com

https://d6hwwmxwo18jz.cloudfront.net/