aws

GENERATIVE AI FOR PUBLIC SECTOR
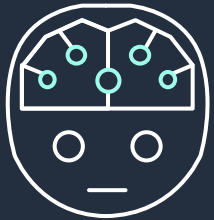
# Miami Learning Day

Erick Cedeno

Sr. Solutions Architect
AWS – State & Local Government

# Agenda

- Artificial Intelligence Definitions

- Public Sector Landscape and Challenges

- Use-case Examples

- AWS's Approach to Gen AI

    - Services and Capabilities

    - Responsible AI

- Getting Started with Gen AI

# AI/ML/Gen AI hierarchy

## 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

**Powered by large models that are pretrained on vast corpora of data and commonly referred to as foundation models (FMs)**

What is Generative AI?

AI that can produce original content close enough to human generated content for real-world tasks

Powered by foundation models pre-trained on large sets of data with several hundred billion parameters

Tasks can be customized for specific domains with minimal fine-tuning

Applicable to many use cases like text summarization, question answering, digital art creation, code generation, etc.

Reduces time and cost to develop ML models and innovate faster

# Challenges from State and Local Government Customers

**Demand for government services is rising** while resources and capacity to deliver them **aren't keeping pace**

Citizens increasingly expect government to **provide modern digital experiences** for conducting online transactions

**Aging infrastructure** for data capture, storage, and management **creates friction** for leveraging data for analytics and AI use-cases

**Complex security, privacy, and compliance requirements** create barriers to change and block adoption (Responsible AI)

**Risk averse culture** slows down innovation

aws

# USE CASE EXAMPLES

## Chatbots and Virtual Assistants

Automate responses for citizen and public employee service queries through generative AI-powered chatbots, voice bots, and virtual assistants.

## Agent assist

Enhance agent performance to improve first contact resolution, augment tasks such as knowledge search, call summarization, and problem-solving.

## Conversation analytics

Provide insights from recorded conversations, forms, or surveys to better understand customer needs, identify call drivers, and detect emerging trends.

## Report generation

Unlock powerful insights from your data and enable every citizen and public employee to make faster and better, data-driven decisions

# USE CASE EXAMPLES

## Content creation

Empower public employees to create content faster and smarter across departments—RFx's, emails, reports.

## Personalization

Deliver better personalized experiences and increase citizen engagement with individually curated offerings and communications.

## Search, summarization and analysis

Summarize data from various sources into actionable insights and perform comparative analysis on your enterprise content.
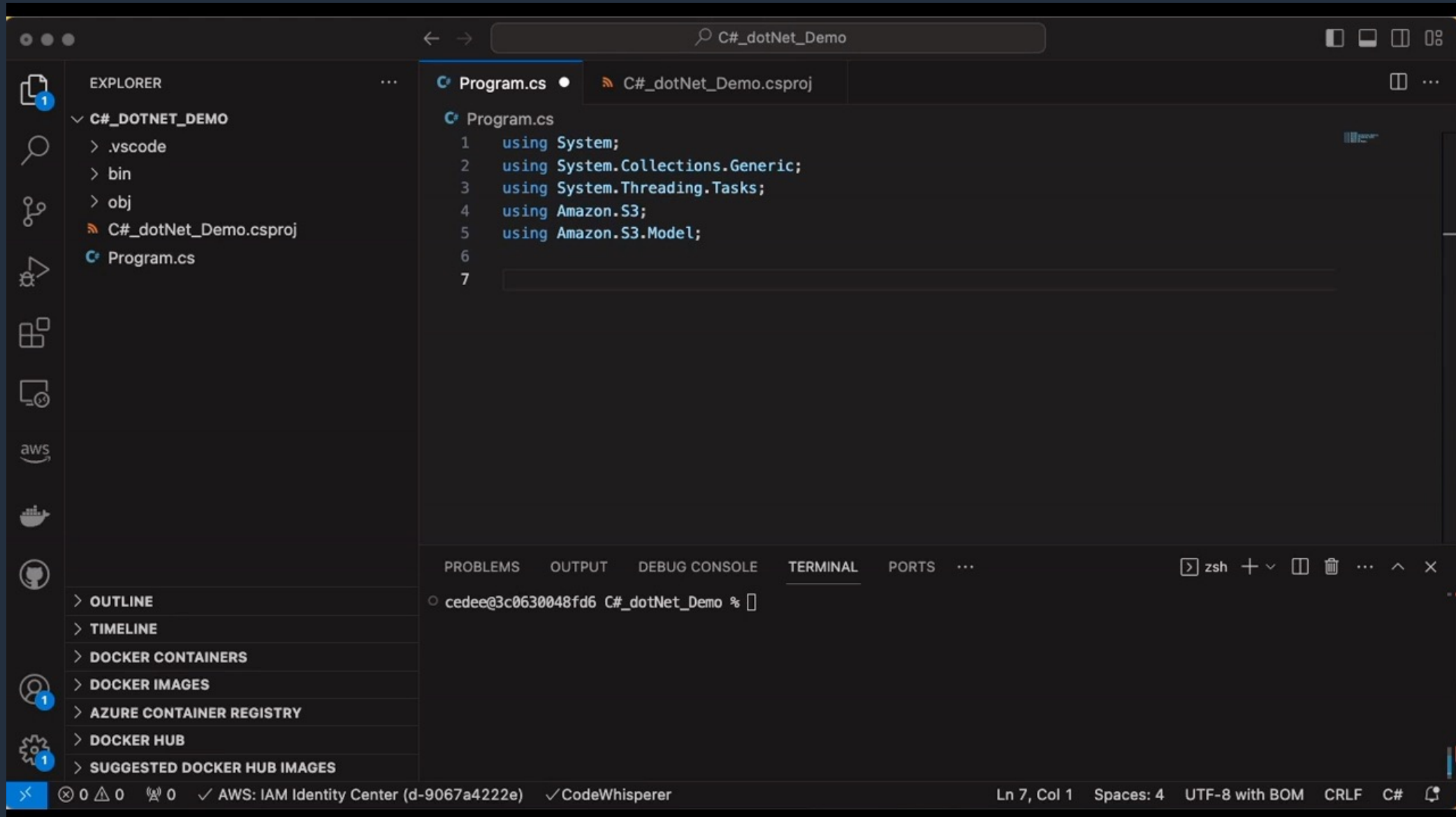
## Code generation

Accelerate application development with automatic code recommendations based on the code and comments in your IDE.

Amazon Q Developer – Demo

# AWS approach to Gen AI

## Amazon
## **Bedrock**

The easiest way to build and
scale generative AI applications
with foundation models (FMs)

Choice of leading FMs through a single API

Model customization

Retrieval Augmented Generation (RAG)

Agents that execute multistep tasks

Security, privacy, and safety

# **Amazon Bedrock**
# Helps keep your data secure and private

None of the customer's data is used to train the underlying model

All data is encrypted in transit and at rest; data used for customization is securely transferred through customer's VPC

Data remains in the Region where the API is processed

Support for GDPR, SOC, ISO, CSA compliance, and HIPAA eligibility

# Guardrails for Amazon Bedrock

Implement safeguards customized to your application requirements and responsible AI policies

Apply guardrails to multiple foundation models and Agents for Amazon Bedrock

Configure harmful content filtering based on your responsible AI policies

Define and disallow denied topics with short natural language descriptions

Redact or block sensitive information such as PIIs, and custom Regex

# How it works



**Amazon Bedrock**

The easiest way to build and scale generative AI applications with FMs

**Choose from leading FMs**

Access a range of top FMs using a single API

**Customize with your data***

Privately fine-tune FMs with your data

**Connect data sources**

Connect FMs with your data sources for RAG

**Enable task execution**

Configure agents to dynamically invoke APIs

* Your data is not used for service improvements and is not shared with third-party model providers

# Amazon Bedrock supports leading foundation models

**amazon**

### Amazon Titan

Text summarization, generation, classification, open-ended Q&A, information extraction, embeddings, and search

**AI21labs**

### Jurassic-2

Multilingual LLMs for text generation in Spanish, French, German, Portuguese, Italian, and Dutch

**ANTHROP\C**

### Claude 2

LLM for conversations, question answering, and workflow automation based on research into training honest and responsible AI systems

**cohere**

### Command

Text-generation model for business applications and embeddings model for search, clustering, or classification in 100+ languages

**Meta**

### Llama 2

Fine-tuned models ideal for dialogue use cases and language tasks

**stability.ai**

### Stable Diffusion

Generation of unique, realistic, high-quality images, art, logos, and designs

*\* Mistral AI's Mistral Large LLM also added in April 2024*

# **Amazon** Q

Your generative AI assistant designed for work that can be tailored to your business, data, code, and operations

Provides interactive answers, solves problems, generates content, and takes action

Understands your organizations information, code, and systems

Personalizes interactions based on your role and permissions

Built to be secure and private

Generative AI
Application

Data
Foundation

**STORAGE**

**GOVERNANCE
& COMPLIANCE**

**DATABASES,
ANALYTICS,
& DATA LAKES**

**DATA
INTEGRATION**

# Responsible AI – what could go wrong?

| Inaccuracies | Bias | Copyright and IP | Security and privacy |
|---|---|---|---|
| Answers that are factually incorrect, irrelevant, or nonsensical, because of limitations in their training data and architecture | Answers that display discriminatory behaviour resulting in prejudiced or unequal treatment of a particular group or groups | The rights of content creators from whom training data is collected remains uncertain and is currently being challenged | Some model providers use and store data for training purposes. Entire end-to-end data pipelines require security and data privacy controls. |
| *"The world record for crossing the English channel on foot is 15 hours"* | *"Generate a picture of a person cleaning" returns overwhelmingly women* | *Artists suing creators of foundation models alleging the improper use of its photos* | *Engineers accidentally releasing source code by putting into a foundation model for debugging* |

# Responsible AI dimensions

## FAIRNESS

Considering impacts
on different groups
of stakeholders

## EXPLAINABILITY

Understanding
and evaluating
system outputs

## CONTROLLABILITY

Having mechanisms
to monitor and steer
AI system behavior

## SAFETY

Preventing harmful system
output and misuse

## PRIVACY & SECURITY

Appropriately obtaining,
using and protecting
data and models

## GOVERNANCE

Incorporating best practices
into the AI supply chain,
including providers
and deployers

## TRANSPARENCY

Enabling stakeholders to
make informed choices
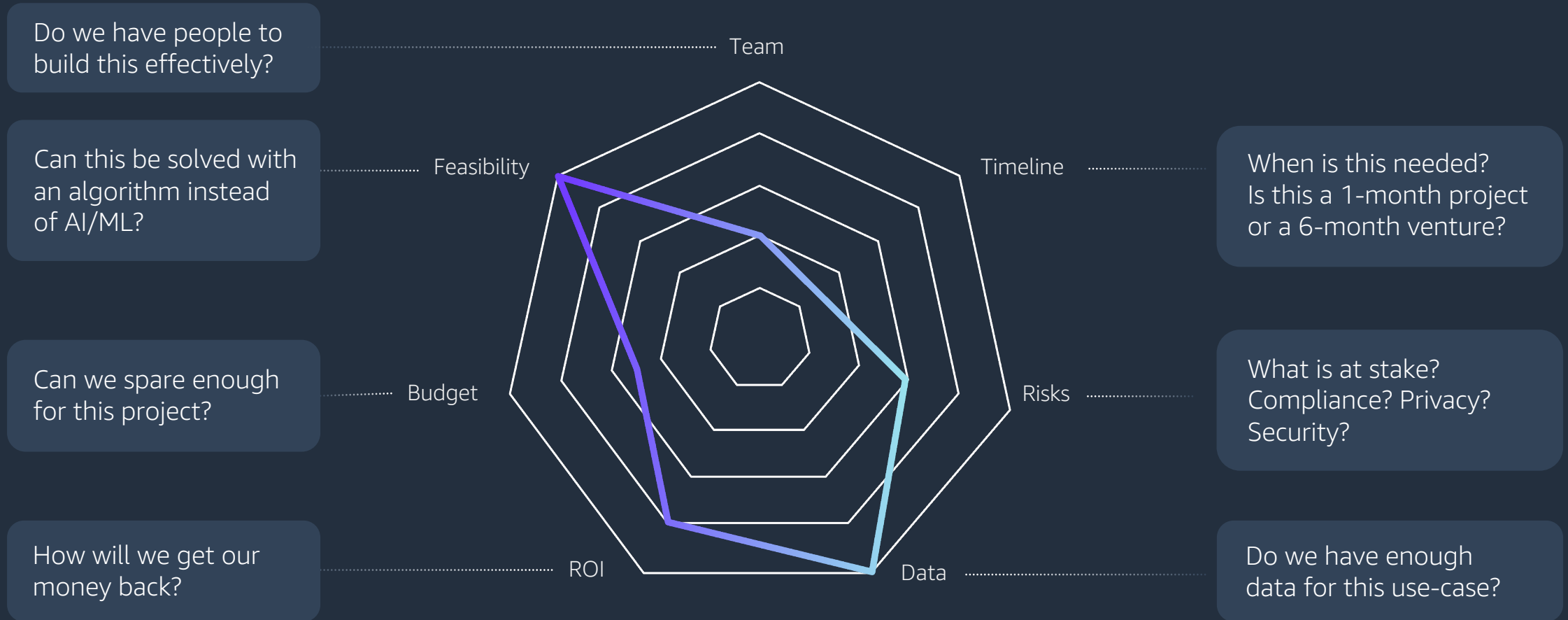about their engagement
with an AI system

## VERACITY &
ROBUSTNESS

Achieving correct
system outputs,
even with unexpected
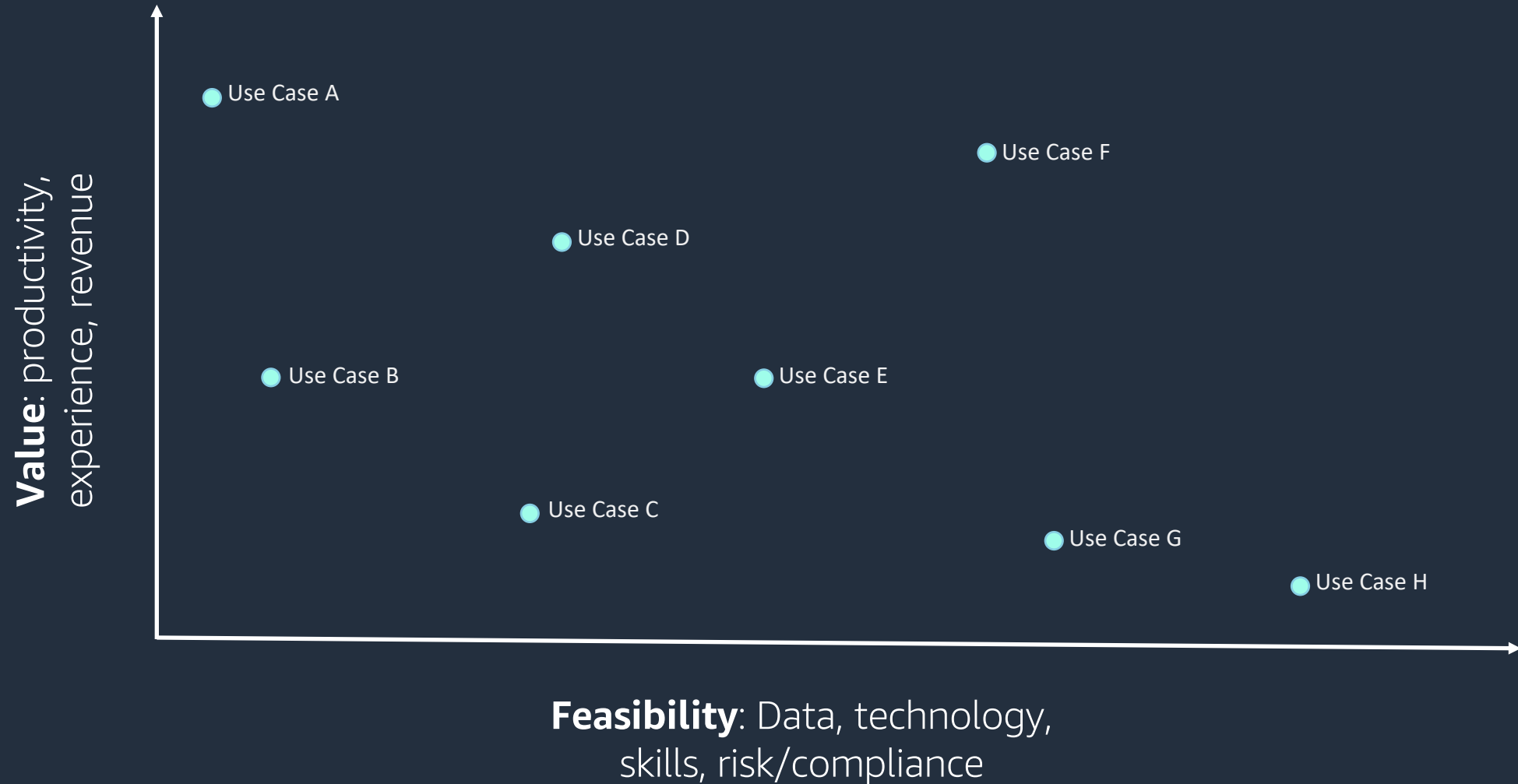or adversarial inputs

# **Getting Started** with generative AI

**01**

Select the
right use case

# **Assessing the fit** of generative AI use cases



Do we have people to build this effectively?

Can this be solved with an algorithm instead of AI/ML?

Can we spare enough for this project?

How will we get our money back?

Team

Timeline — When is this needed? Is this a 1-month project or a 6-month venture?

Feasibility

Budget

ROI

Data — Do we have enough data for this use-case?

Risks — What is at stake? Compliance? Privacy? Security?

# **Selecting** the right use cases



**Value**: productivity, experience, revenue (y-axis)

- Use Case A
- Use Case F
- Use Case D
- Use Case B
- Use Case E
- Use Case C
- Use Case G
- Use Case H

**Feasibility**: Data, technology, skills, risk/compliance (x-axis)

# **Getting Started** with generative AI

## 01

Select the
right use case

## 02

Empower your teams to
innovate with tools and
trainings

# Comprehensive, customized training
## for every stage of your cloud journey

### Skills assessment and personalized training plans

Our **Learning Needs Analysis** helps you understand where your gaps are and design a plan to target the AI and ML skills that are most important to your business goals.

### Blended training designed by AWS experts

Digital courses, instructor-led classes, and game-based learning meet teams' diverse training needs.

**AWS Skill Builder**: 100+ learning resources on AI and ML, including generative AI. Interactive challenges in a secure sandbox environment.

### Industry-recognized credentials

Validate expertise with certifications like **AWS Certified Machine Learning – Specialty.** Retain top talent by funding exam fees.

# **Getting Started** with generative AI

## 01
Select the
right use case

## 02
Empower your teams to
innovate with tools and
trainings

## 03
Get started on
your top use cases

# Find the right **AWS Partners** to fit your business needs

## AWS Marketplace

Explore and find partner practices and solutions to help bring generative AI into your business.

**Global Systems Integrators**
**Systems Integrators**
**Independent Software Vendors**
**Business Consulting & Advisory**

## Hundreds of Use Cases

AWS Partners have released offerings and solutions for Generative AI use cases across all major industries and cross-industry business functions.

## Generative AI Competency Partners

Boost productivity, build differentiated experiences, and innovate faster with AWS partners with proven success deploying generative AI on AWS

**Consulting Services**
**Generative AI Applications**
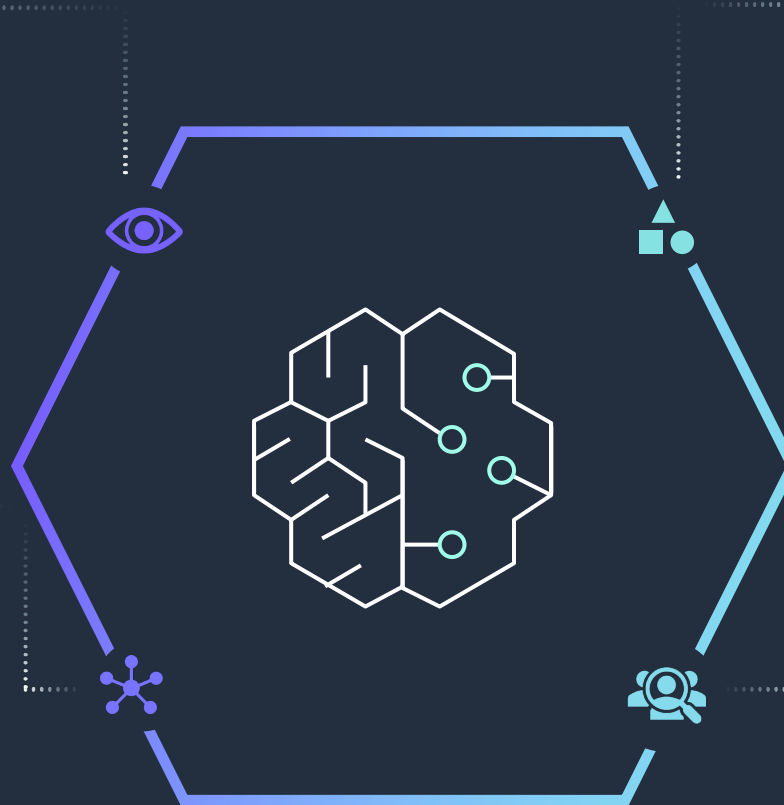**Foundation Models & App Development**
**Infrastructure & Data**

# **Strategic recommendations** on generative AI

## EYES WIDE OPEN

- Culture of continuous experimentation

- Prevent early dependencies

## NO ONE SIZE FITS ALL

- Generative AI as an enhancement – not a replacement.

- Requirements in enterprises differ based on financial resources, security needs, governance and skills.

- Evaluate buy-vs-build

## FLEXIBILITY IS KEY

- Innovation requires flexibility free from technical or contractual lock-ins.

- Infrastructure supporting third-party Generative AI integration

- A breadth of services ensures long-term flexibility and business value

## LISTEN AND ENABLE

- Work backward from your constituents

- Listen to your domain experts

- Enable employees with the right set of tooling

aws

**Please Take the Survey:**

# Thank you!

Erick Cedeno
cedee@amazon.com

AI/ML Track
Generative AI for Public Sector