GenAI is Eating the World

Exploring AWS GenAI Technologies for
SAP Solutions

David Nye

Solutions Architect, SAP
Amazon Web Services

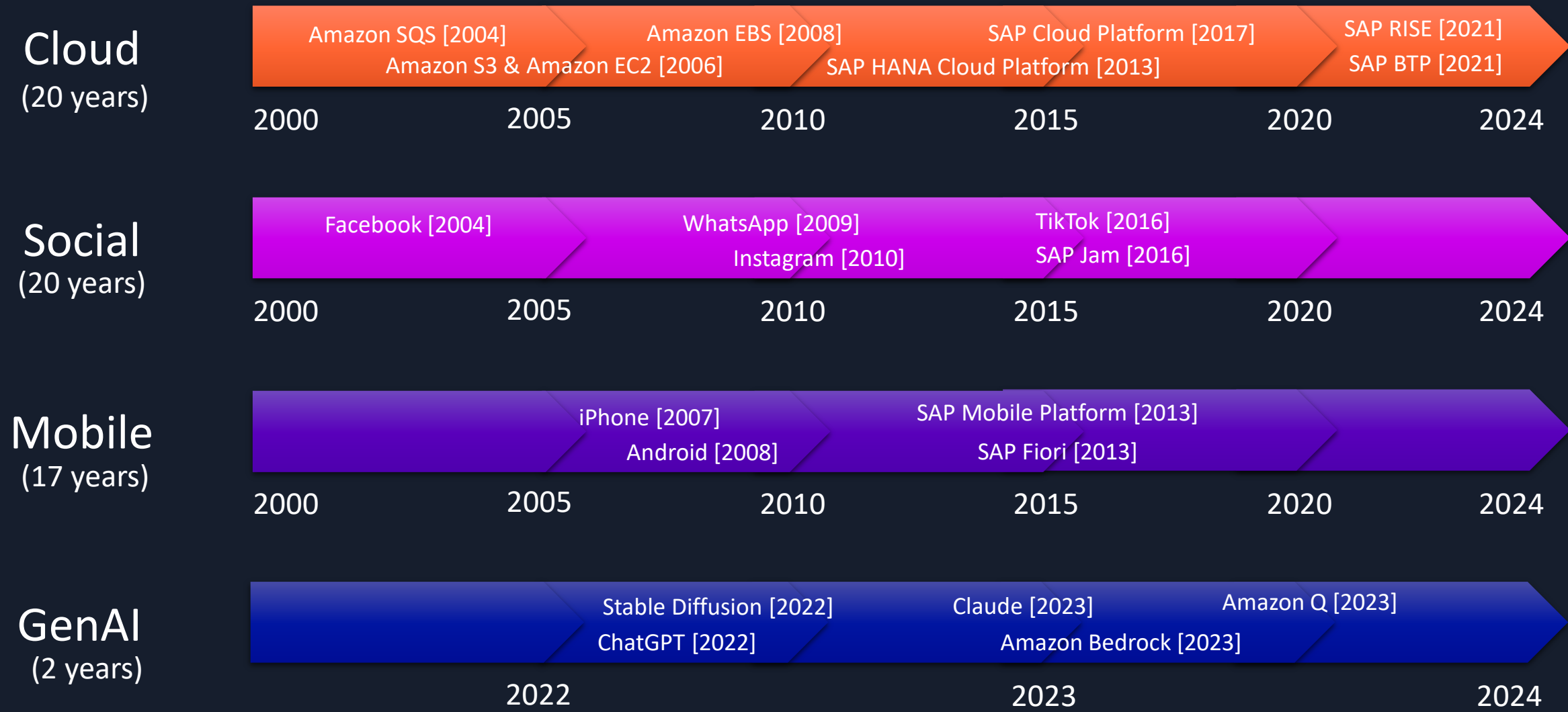# Software is eating the world

Marc Andreessen [2011]

a16z

By JD Lasica - https://www.flickr.com/photos/jdlasica/10082059294/,
CC BY 2.0, https://commons.wikimedia.org/w/index.php?curid=145398495

GenAI is Eating the World

Cloud
(20 years)

| Amazon SQS [2004] | Amazon EBS [2008] | SAP Cloud Platform [2017] | SAP RISE [2021] |
| Amazon S3 & Amazon EC2 [2006] | SAP HANA Cloud Platform [2013] | | SAP BTP [2021] |

2000   2005   2010   2015   2020   2024

Social
(20 years)

| Facebook [2004] | WhatsApp [2009] | TikTok [2016] |
| | Instagram [2010] | SAP Jam [2016] |

2000   2005   2010   2015   2020   2024

Mobile
(17 years)

| iPhone [2007] | SAP Mobile Platform [2013] |
| Android [2008] | SAP Fiori [2013] |

2000   2005   2010   2015   2020   2024

GenAI
(2 years)

| Stable Diffusion [2022] | Claude [2023] | Amazon Q [2023] |
| ChatGPT [2022] | Amazon Bedrock [2023] | |

2022   2023   2024

Your Presenter

DAVID NYE

SOLUTIONS ARCHITECT

AMAZON WEB SERVICES

TECHNOLOGY | SECURITY | INTEGRATION

~20 YEARS WORKING WITH SAP TECHNOLOGIES

COMPUTER SCIENCE MAJOR

MOUNTAIN BIKER

NYDAVID@AMAZON.COM

HTTPS://WWW.LINKEDIN.COM/IN/NYEDAVID/

Exploring AWS GenAI Technologies for SAP Solutions

01 What is GenAI

02 AWS GenAI Services

03 SAP & GenAI

04 Getting Started

What is GenAI?

What is Generative AI?

AI that can
generate content
close enough to human created content
for real-world tasks

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

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

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

New Volvo car concept design by midjourney
Credit: @sugardesign_1 Instagram
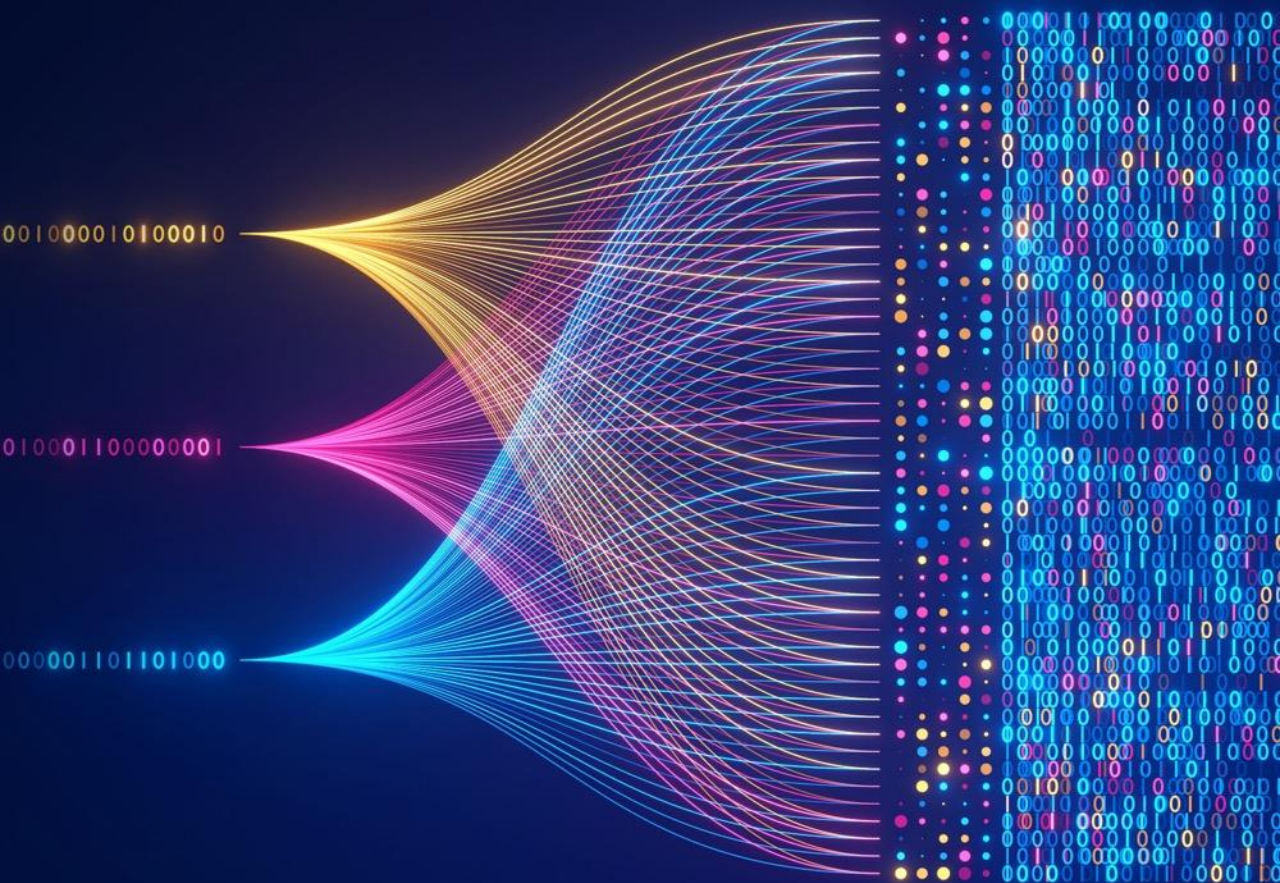
Generative AI could add the equivalent of $2.6 trillion to $4.4 trillion in annual economic benefits to the global Economy

"The economic potential of generative AI: The next productivity frontier," McKinsey, June 14, 2023

Generative AI is powered by Foundation Models

Parametric Knowledge

$$\text{Inference} = w_0\,P_0 + w_1\,P_1 + w_2\,P_2 + \ldots$$

Public domain knowledge

Training

Vocabulary

Parameters

Weights

Foundational Model
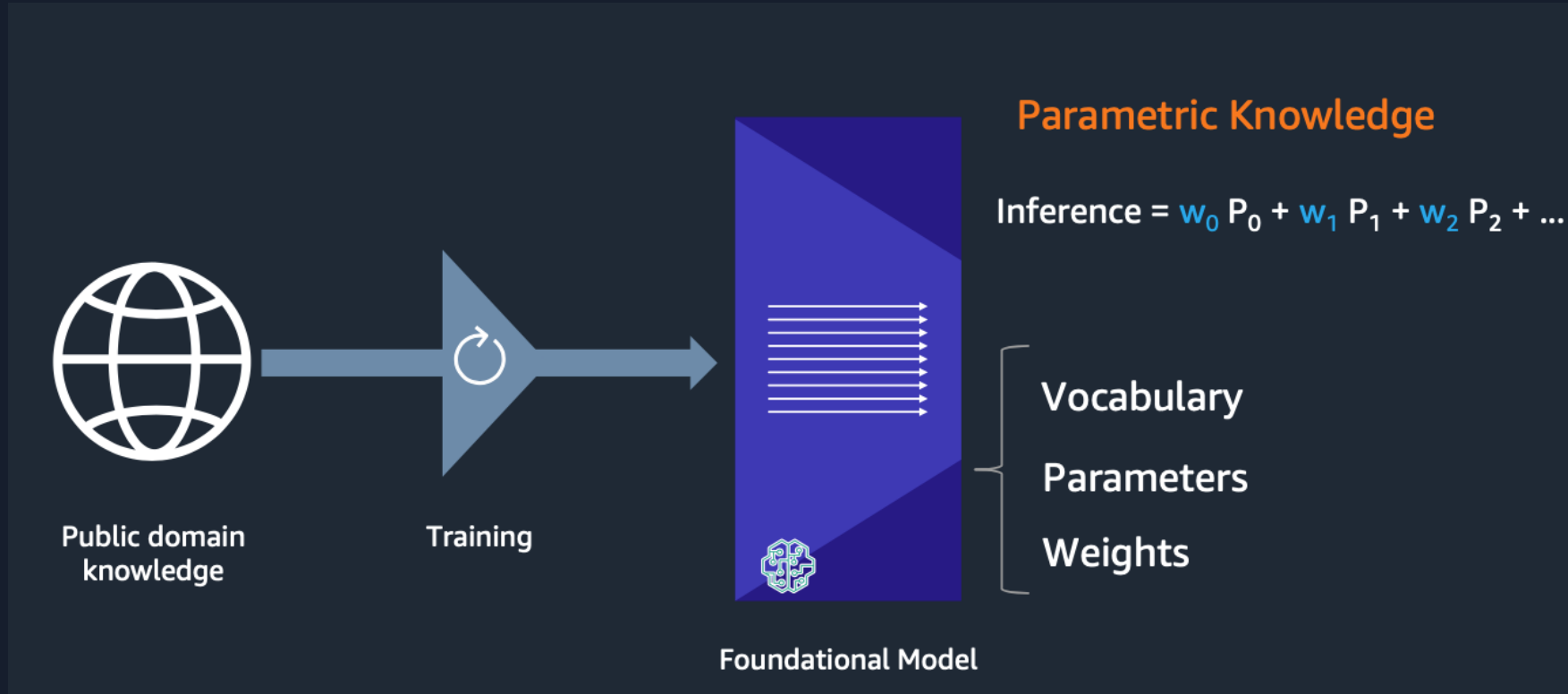
- Pretrained on vast amounts of unstructured data

- Contain large number of parameters that make them capable of learning complex concepts

- Can be applied in a wide range of contexts (based on Neural Networks)

- Customize FMs using your data for domain specific tasks

Types of foundation models

Input		FM		Output
	▶		▶	
"Summarize the articles on impact of walking on heart health"		**Text-to-text** Generate text from natural language prompts		"Ten thousand steps per day is optimum for maintaining a healthy heart"
"What is the top-5 selling products by total sales last month ?"		**Text-to-SQL** Generate SQL Query of text		"select product, sum(amount) from sales group by product order by sum(amount) desc limit 5
"a photo of an astronaut riding a horse on mars"		**Multimodal** Create and edit images using natural language prompts		

Challenges with Foundation Models

1 Hallucinations | Model response with Incorrect Information confidently

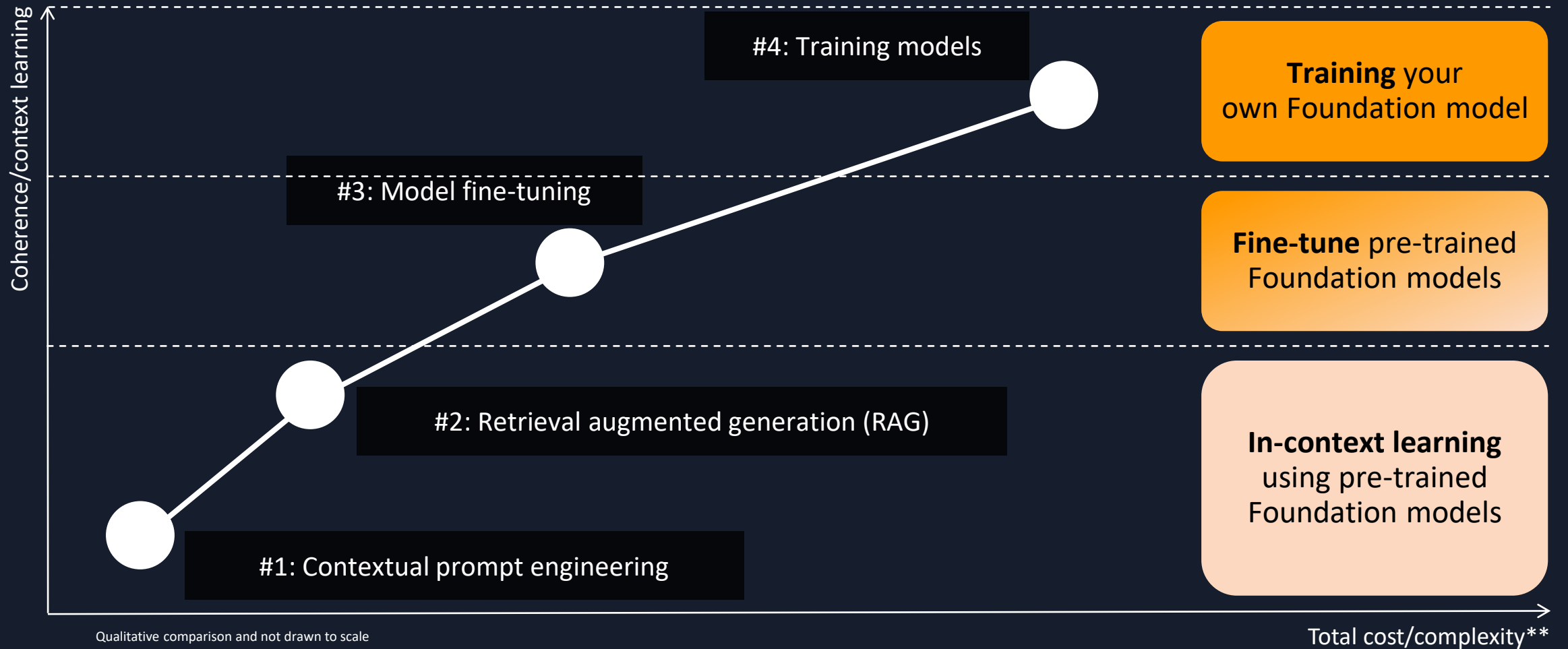**2** Knowledge Cut off | Training data is dated, knowledge in model is not up-to-date

3 Imprecise response | Foundation Models are built on public knowledge and may not align response with enterprise context. Model needs to be Grounded

Generative AI Application Lifecycle

| Use Case | → | Evaluate / Select Models | → | Improve performance:
1. Prompt Engineering
2. Retrieval Augmented Generation (RAG)
3. Fine Tuning
4. Training Models (Scratch Build) | → | Evaluate Results | → | Deploy, Monitor |
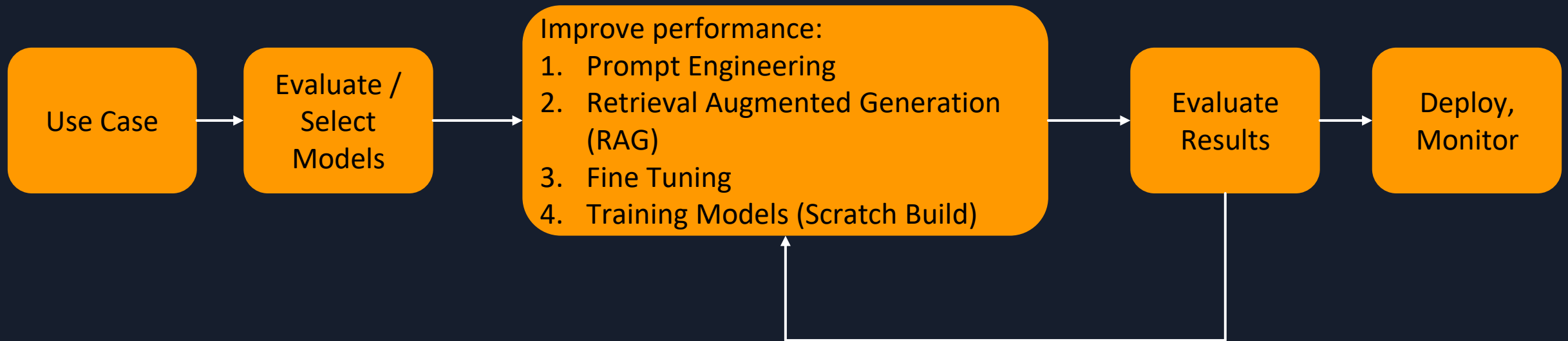
Generative AI model implementation patterns

Coherence/context learning

#4: Training models

#3: Model fine-tuning

#2: Retrieval augmented generation (RAG)

#1: Contextual prompt engineering

Qualitative comparison and not drawn to scale

Total cost/complexity**

Training your own Foundation model

Fine-tune pre-trained Foundation models

In-context learning using pre-trained Foundation models

** Total cost/complexity of model preparation, inference, Data management

Generative AI Innovations at Amazon
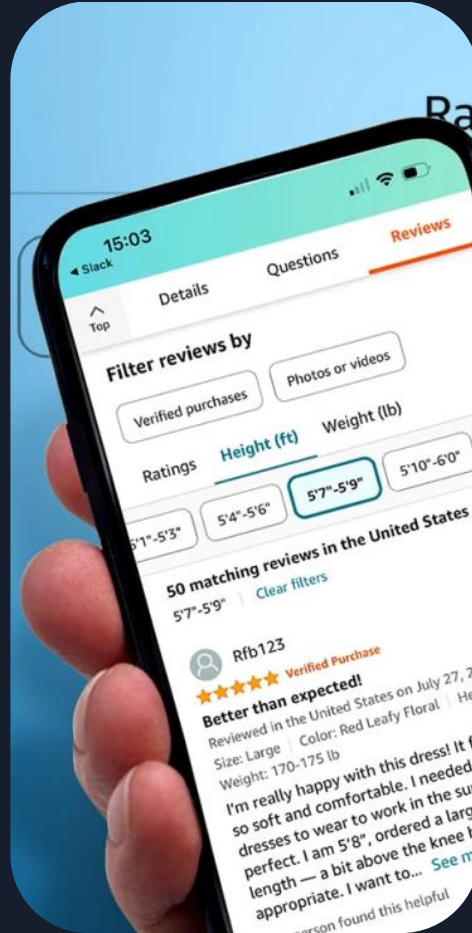
Customer reviews
on Amazon.com
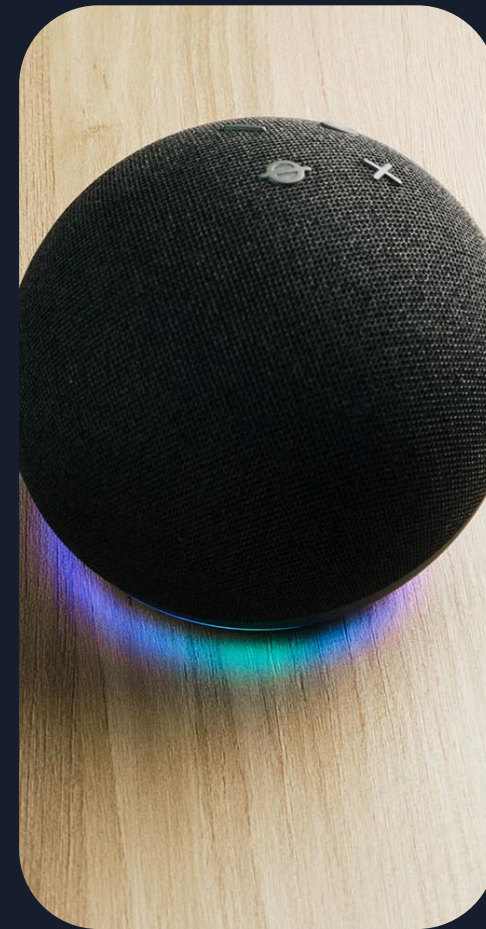
Alexa
Billions of interactions each week
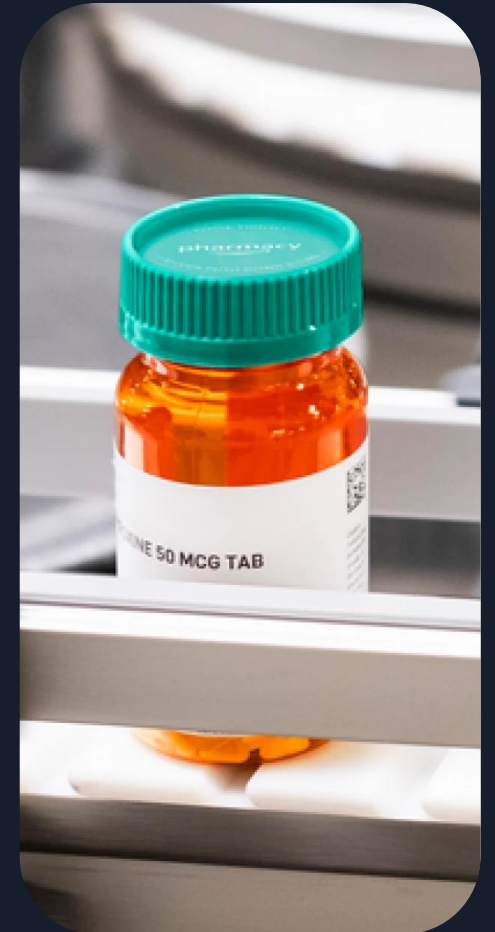
Amazon Pharmacy
faster prescriptions and more helpful support

14

AWS GenAI Services

15

Generative AI Stack
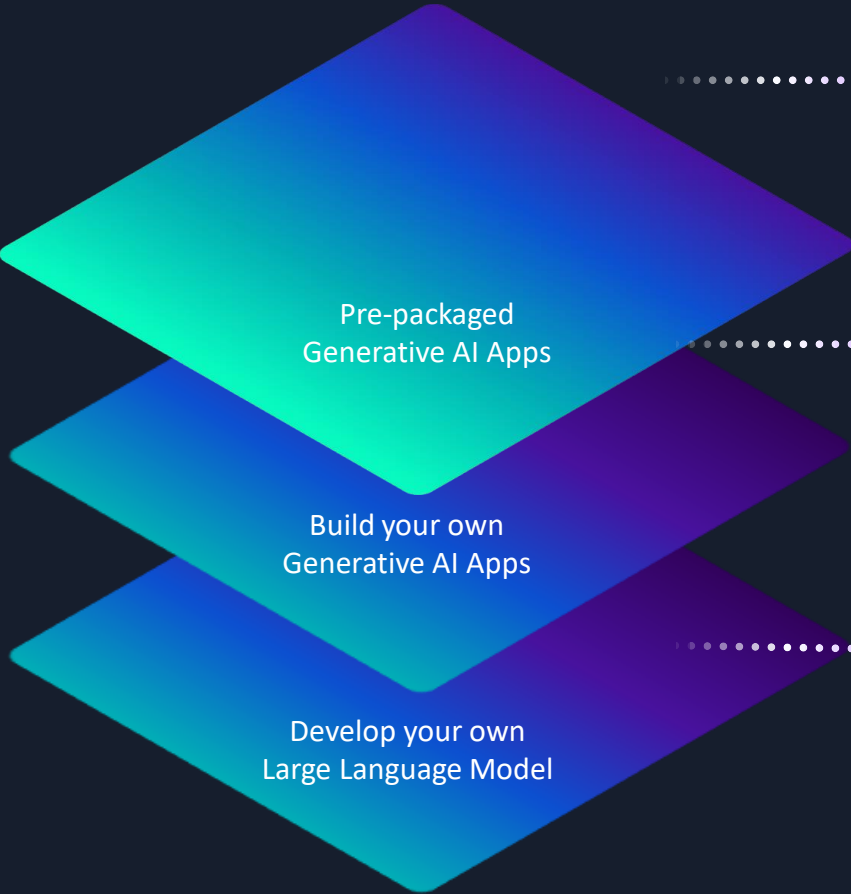
APPLICATIONS THAT LEVERAGE FMs

Amazon Q

Amazon Q in Amazon QuickSight

Amazon Q in Amazon Connect
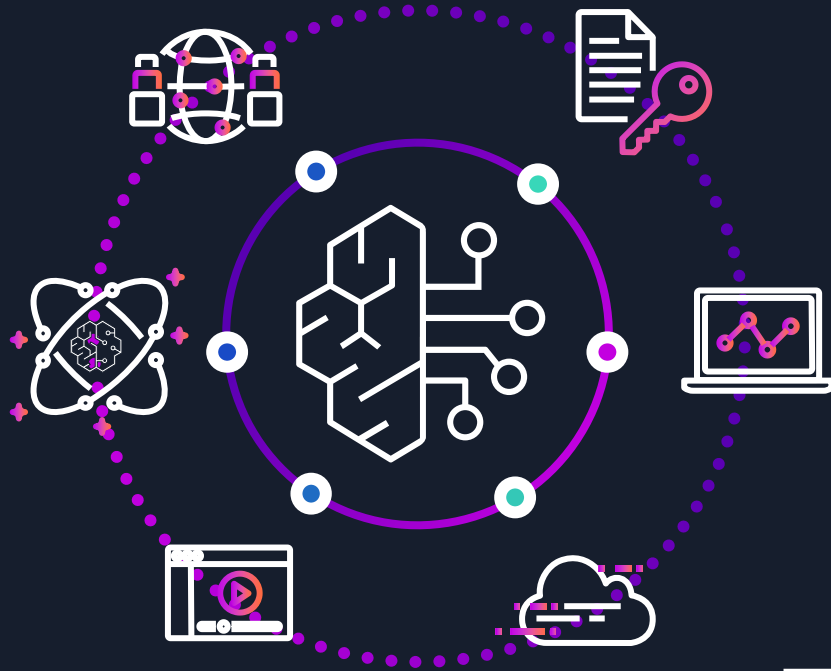
Amazon CodeWhisperer

TOOLS TO BUILD WITH FMS AND LLMS

Amazon Bedrock

Guardrails   Agents   Customization capabilities

INFRASTRUCTURE FOR FM TRAINING & INFERENCE

GPUs   Trainium   Inferentia   SageMaker

UltraClusters   EFA   EC2 Capacity Blocks   Nitro   Neuron

Pre-packaged Generative AI Apps

Build your own Generative AI Apps

Develop your own Large Language Model

Amazon Bedrock

THE EASIEST WAY TO BUILD AND SCALE GEN AI APPLICATIONS

- 21+ best in class FMs from 7 providers
- Privately customize models with your data
- Native support for Retrieval Augmented Generation (RAG)
- Agents that execute multistep tasks
- Model Evaluation
- Guardrails (preview)
- Security, privacy, and safety

Amazon Bedrock

AI21labs   ANTHROP\C   cohere   Meta   stability.ai   amazon

Jurassic	Claude	Command + Embed	Llama 3	Stable Diffusion	Amazon Titan
Contextual answers, summarization, paraphrasing	Summarization, complex reasoning, writing, coding	Text generation, search, classification	Q&A and reading comprehension	High-quality images and art	Text summarization, generation, Q&A, search

Amazon Q is your business expert

BOOST YOUR WORKFORCE PRODUCTIVITY WITH GENERATIVE AI

> Delivers quick, accurate, and relevant answers to your business questions, securely and privately and document repositories.

> Provides responses with references and citations for easy fact-checking

> Respects existing access control based on user permissions

> Connects to over 40 popular enterprise applications and document repositories

> Enables administrators to easily apply guardrails to customize and control responses

New Conversation ✕

Draft a professional email
November 15, 2023

Amazon Q

👋 Hello,
I am your AI assistant. Enter a prompt or choose a suggestion to start a conversation. I'll respond using data from within your organization.

Draft a professional email	Summarize a report
Help plan a workshop	Analyze market trends
Write a blog post	Brainstorm taglines

Enter a prompt

Enable Streaming

Amazon Q uses generative AI, it may give inaccurate or inappropriate responses. For more information: AWS Responsible AI Policy

ⓘ Info

Key Features

Trusted answers generated from enterprise data

Citations and source attribution

Conversation history and context

Upload files and analyze content

Execute Actions across multiple Enterprise Apps

Top use cases

Streamline search experience

Connect internal wikis, SharePoint sites, Confluence, Salesforce, and 40+ other applications for a unified, conversational search experience.

Accelerate content creation

Empower employees to create content faster and smarter across departments—from marketing to sales to engineering.

Generate summaries

Quickly understand the essence of documents by creating summaries of uploaded files or existing enterprise documents.

Extract key insights

Perform comparative analysis on your documents and unlock powerful insights so that your employees can make faster, better data-driven decisions.

Amazon Q in QuickSight

YOUR GENERATIVE BI ASSISTANT THAT MAKES IT EASY TO BUILD AND CONSUME INSIGHTS

> Enables business analysts/users' to quickly build and refine compelling visuals, summarize insights, and answer data questions.

> Accelerate analysis by easily creating calculations without looking up or learning specific syntax

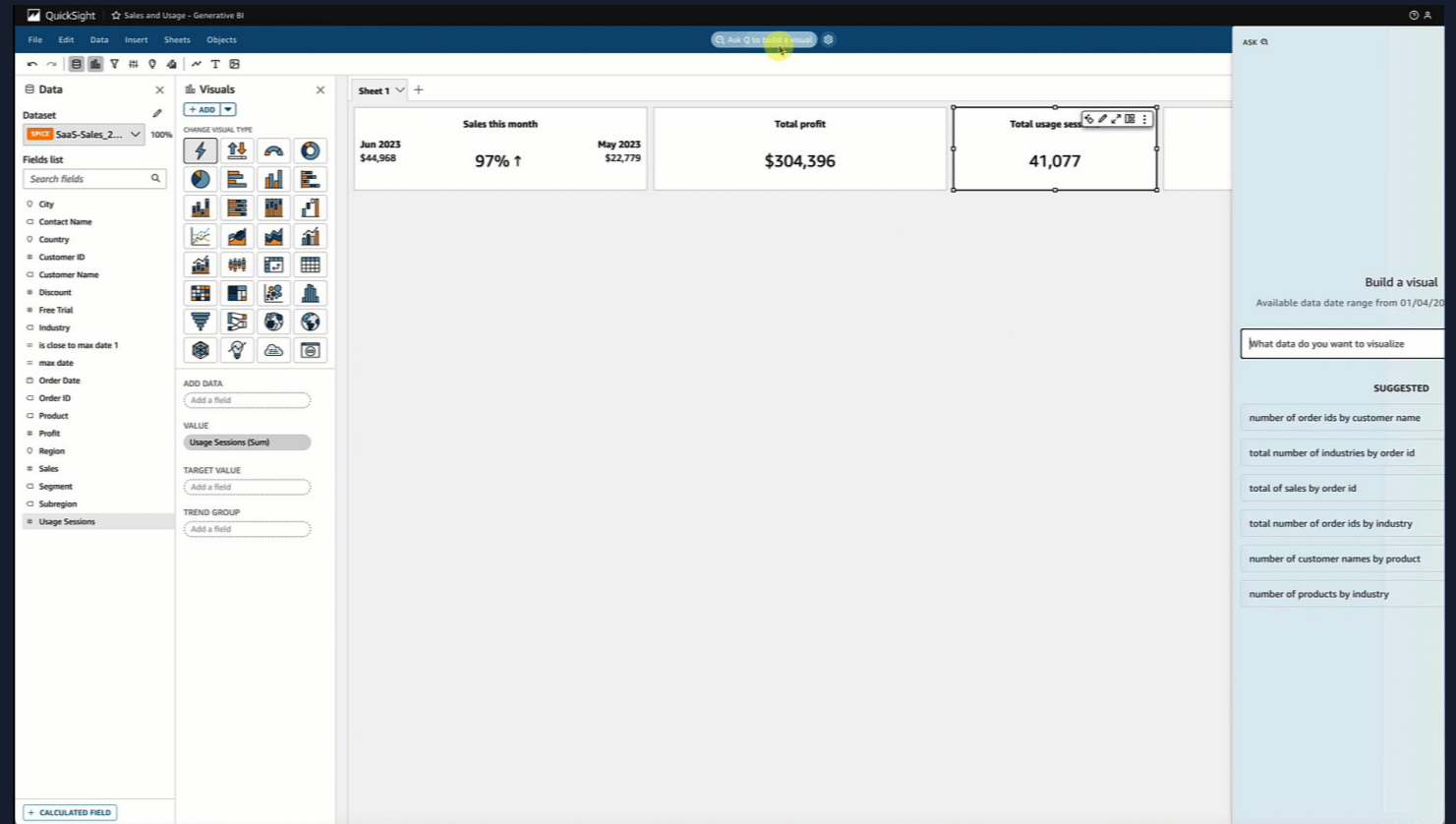> Interpret and share insights with visually compelling narratives with simple UI-based workflows
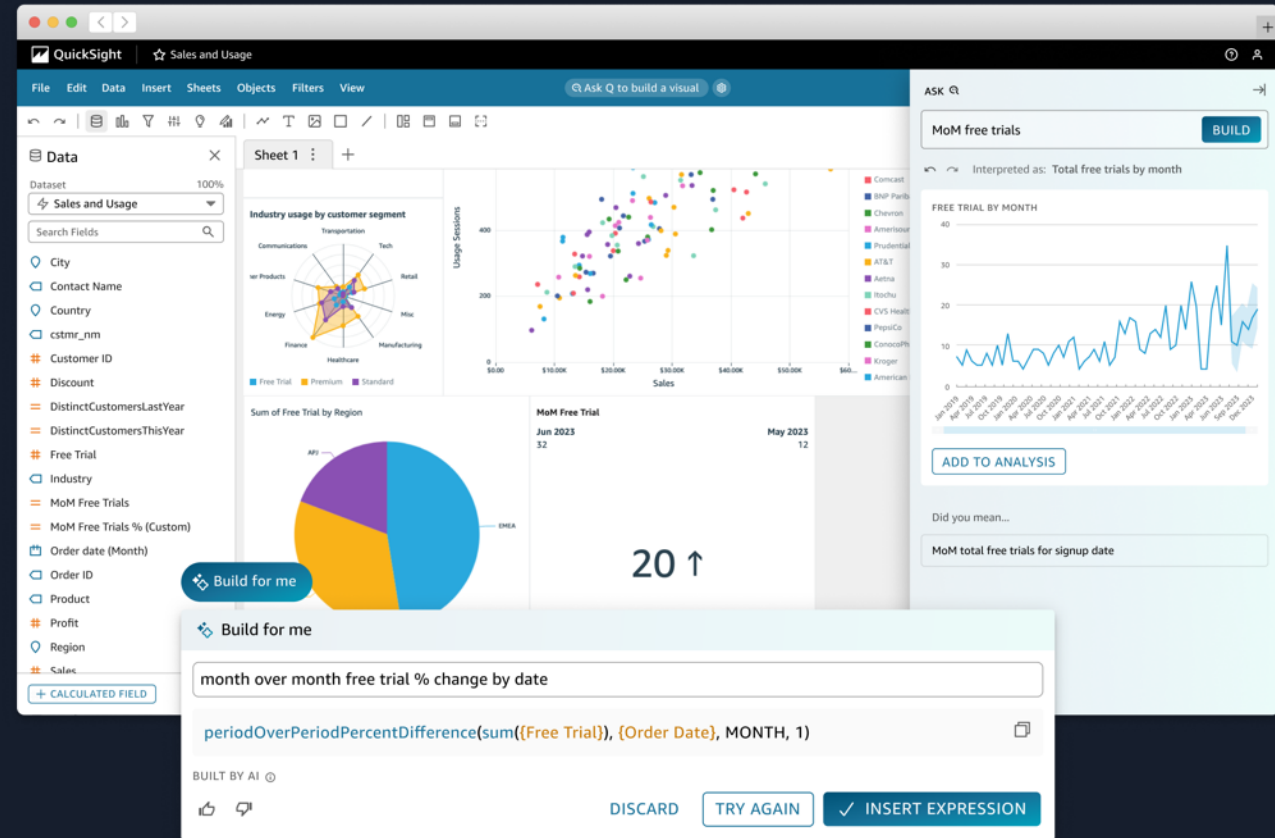
Author visuals using natural language in seconds

A NEW DASHBOARD BUILDING EXPERIENCE POWERED BY GENERATIVE BI

Natural language visual creation

Use vague or precise language to generate and add visuals to a dashboard

Preview

Generative BI capabilities in Amazon QuickSight

ACCELERATE THE PATH TO DATA-DRIVEN ACTION FOR ALL USERS

Accelerate authoring with natural language

- Generate visuals in seconds using precise or vague language
- Fine-tune visuals with generative editor
- Create new calculations without syntax lookups

Generate visually compelling stories

- Contextualize insights to drive actions
- Automatically generate data stories with natural language prompts
- Create refined narratives using AI, themes, or manual edits
- Securely share narratives that stay up-to-date

Enterprise-ready: Governed & secure

- Govern with QuickSight security rules and policies
- All data stays within customer's AWS account
- No customer data is used to train the LLMs that underlie the Generative BI capabilities
- All calculations occur in QuickSight analytical engine, driving accuracy

Powered by Amazon Bedrock and built on QuickSight ML foundation

SAP & GenAI

SAP Concur enhanced with Amazon Bedrock

SAP AI with Amazon Bedrock Reference Architecture
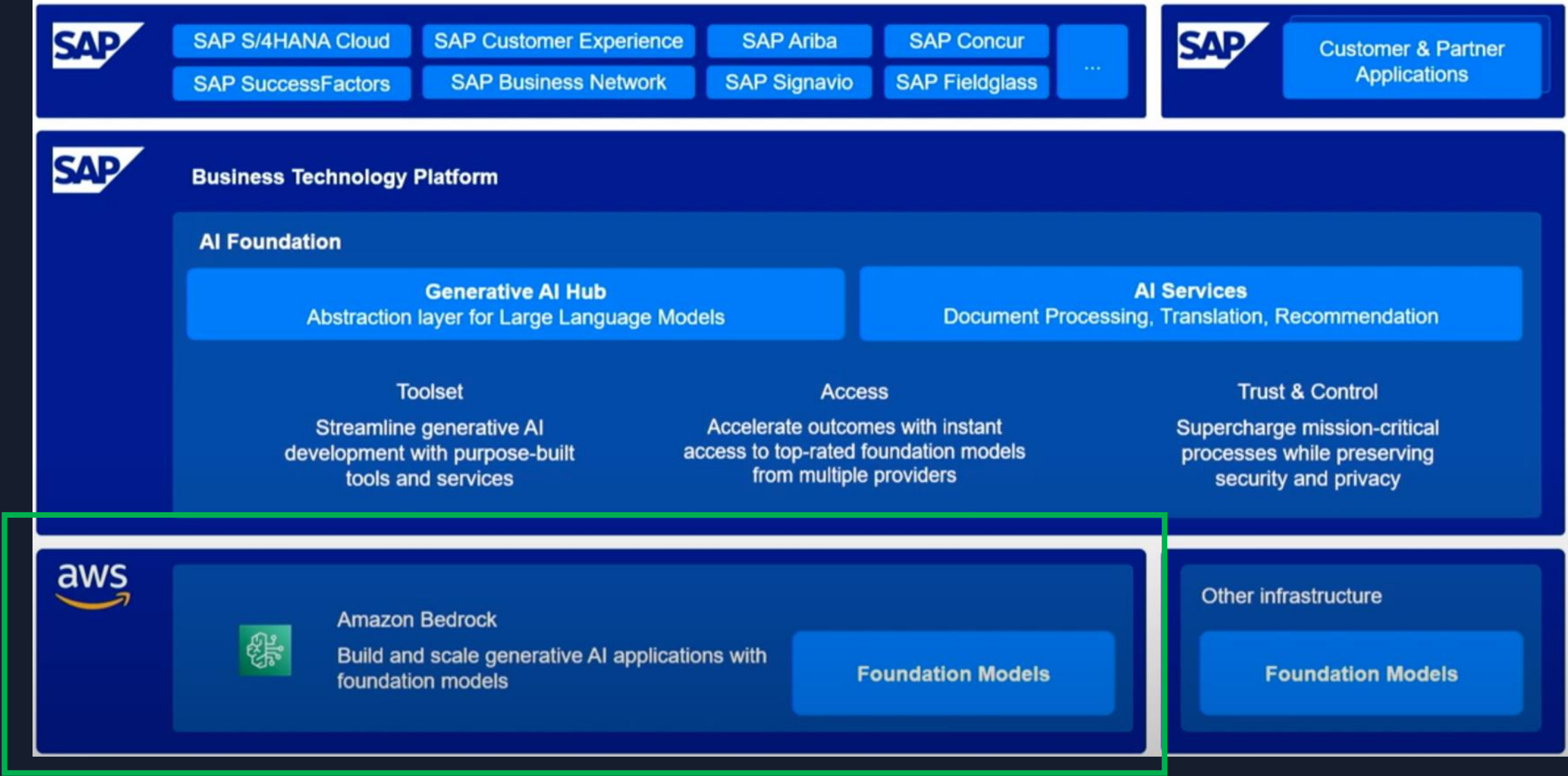
Accelerate SAP automation and innovation with generative AI from AWS

Sales Inventory Analyst

Gather real-time insights in natural language prompts with **Amazon Bedrock and SAP Datasphere**

Business User (Cross-Industry)

Improve user productivity with Conversational AI using **Amazon Q Business**

SAP Developer: Integrate & Extend Core

Accelerate development using SAP BTP and AWS services with **Amazon Bedrock and Amazon Q Developer**

SAP System Administrator

Complete development tasks faster using **Amazon Q Developer**
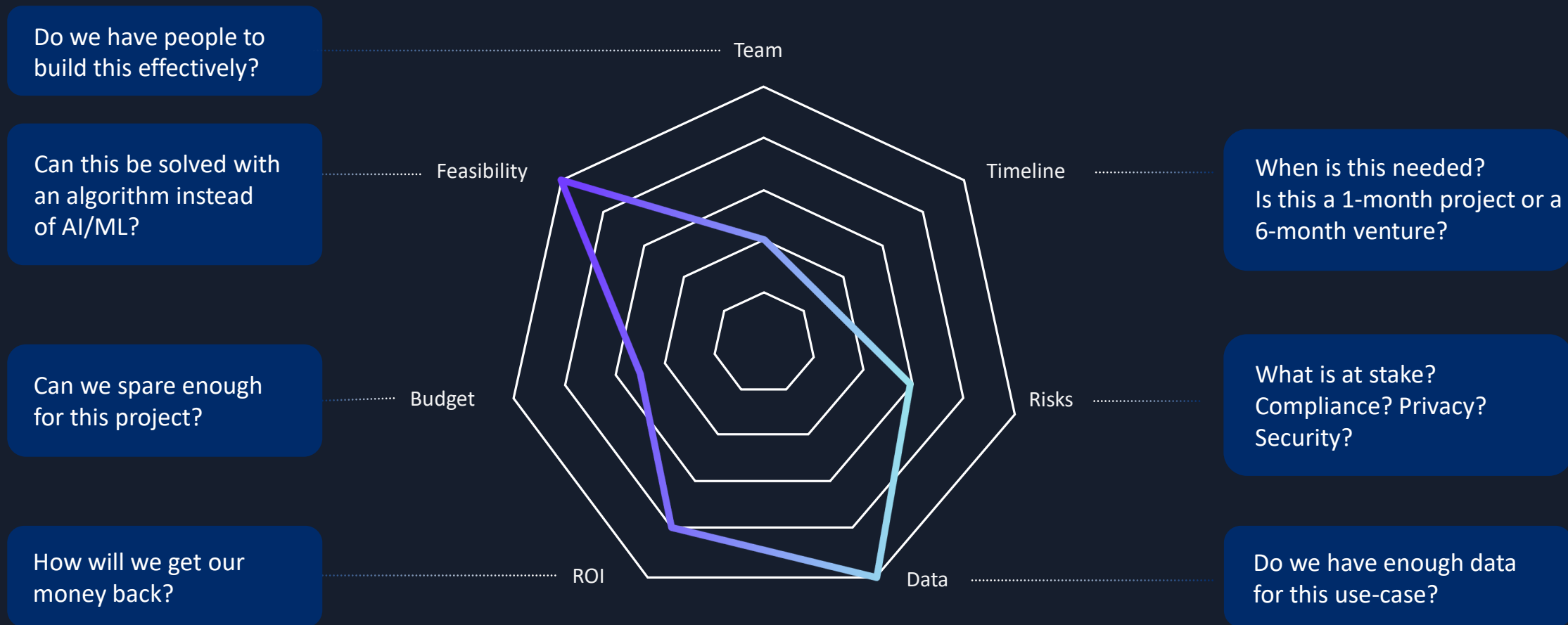
Order-to-Cash Analyst

Improve decision making and analyze SAP data using generative BI capabilities with **Amazon Q in QuickSight**

Sales Account Manager

Gain customer insights by summarizing and analyzing documents with **Amazon Bedrock SAP Fiori and SAP BTP**

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?

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

What is at stake? Compliance? Privacy? Security?

Do we have enough data for this use-case?

Team

Timeline

Risks

Data

ROI

Budget

Feasibility

GenAI Use cases – UX strategy

Build Sidecar application

SAP Fiori /
SAP Build
Work Zone

AWS Amplify

Connect to Amazon Bedrock
via APIs

Embed Amazon Q for Business
Into Fiori LaunchPad

Customise SAP ERP

AWS SDK for SAP
ABAP

Connect to Amazon Bedrock
with 5 lines of code

Extend with Business Applications

Amazon QuickSight

Amazon
Connect

Amazon
Supply Chain

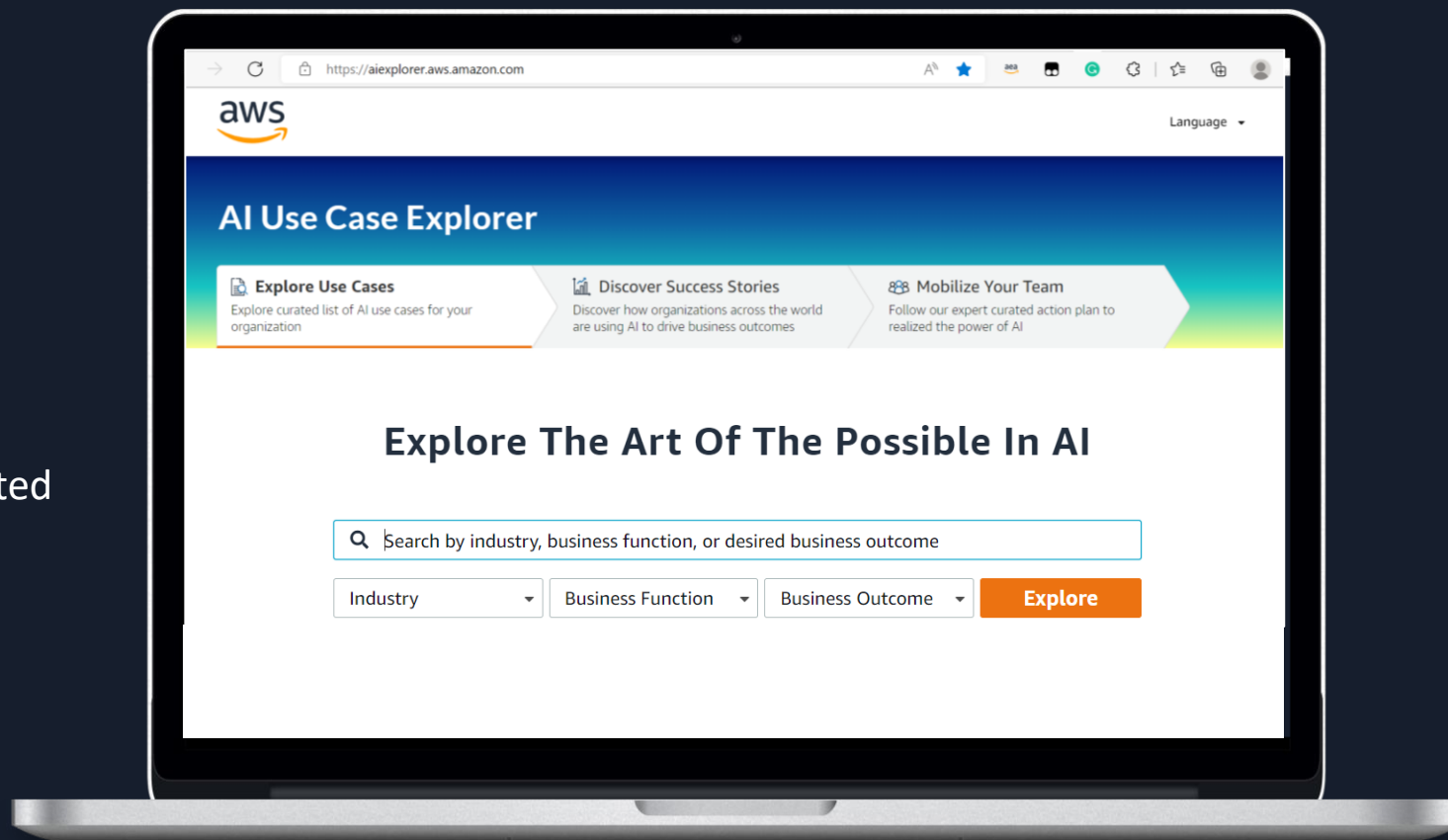Amazon Q embedded into
business applications

Getting Started

AI Use Case Explorer

Easily find the most relevant AI use cases with related content and guidance to make them real

aiexplorer.aws.amazon.com

Experiment

Experiment with a Proof-of-Concept!

Engage us for SAP on AWS innovation topics & workshops

Thank you!

David Nye

nydavid@amazon.com

https://www.linkedin.com/in/nyedavid/