



Generative AI with AWS

Brian Helms
Solutions Architect, Partner & Channel Development
brihelms@amazon.com

Agenda

- Innovation opportunities
- Data is your differentiator
- Generative AI with AWS
- What's next?

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 has potential to create significant business value



NEW EXPERIENCES

Create new innovative and engaging ways of interacting with your customers and employees



PRODUCTIVITY

Radically improve productivity across all lines of business



INSIGHTS

Extract insights and clear answers from all your corporate information, enabling faster and better decisions



CREATIVITY

Create new content and ideas, including conversations, stories, images, videos, and music

Generative AI use cases across industries

ENHANCE CONSTITUENT EXPERIENCES

CHATBOTS
VIRTUAL ASSISTANTS
CONVERSATION
ANALYTICS
PERSONALIZATION

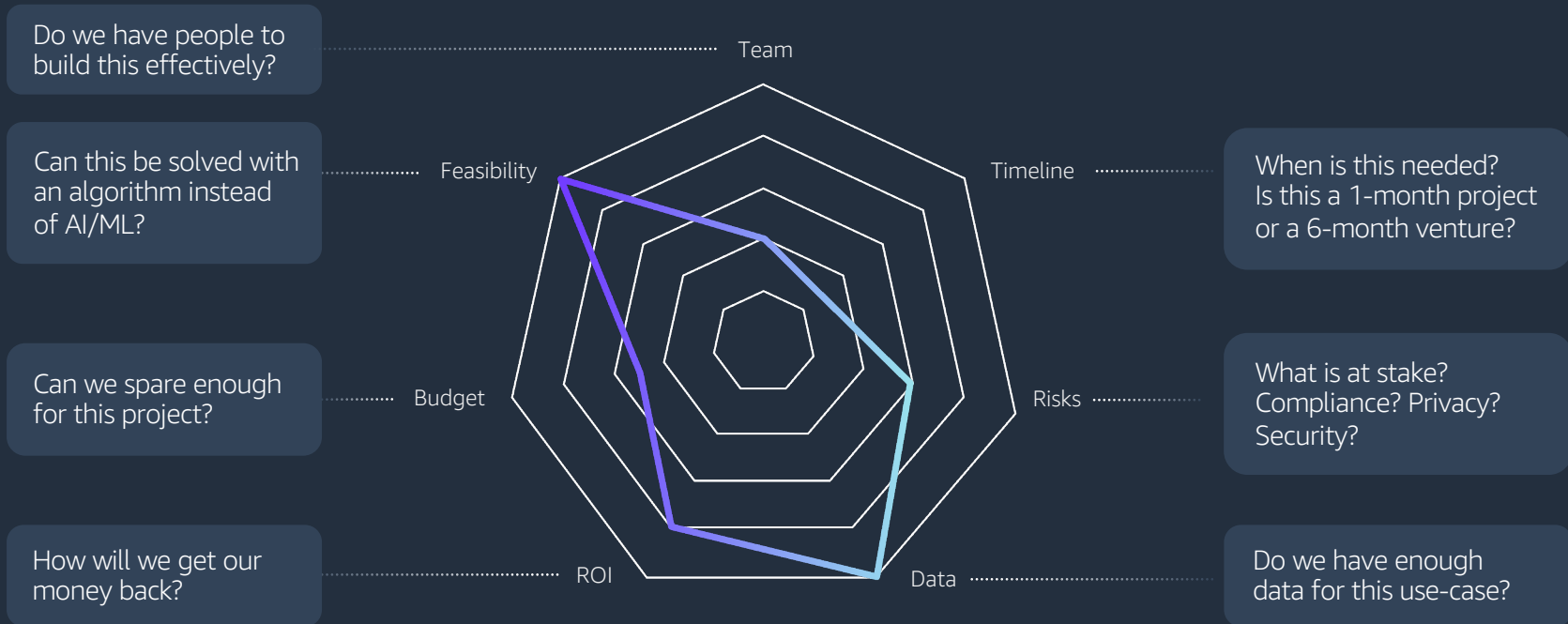
BOOST EMPLOYEE PRODUCTIVITY & CREATIVITY

CONVERSATIONAL
SEARCH
SUMMARIZATION
CONTENT CREATION
CODE GENERATION
DATA TO INSIGHTS

OPTIMIZE BUSINESS PROCESSES

DOCUMENT PROCESSING
DATA AUGMENTATION
FRAUD DETECTION
PROCESS OPTIMIZATION

Assessing the fit of generative AI use cases





Generative AI
Application

Data Foundation

STORAGE

GOVERNANCE
& COMPLIANCE

DATABASES,
ANALYTICS,
& DATA LAKES

DATA
INTEGRATION

Your data is the **differentiator**



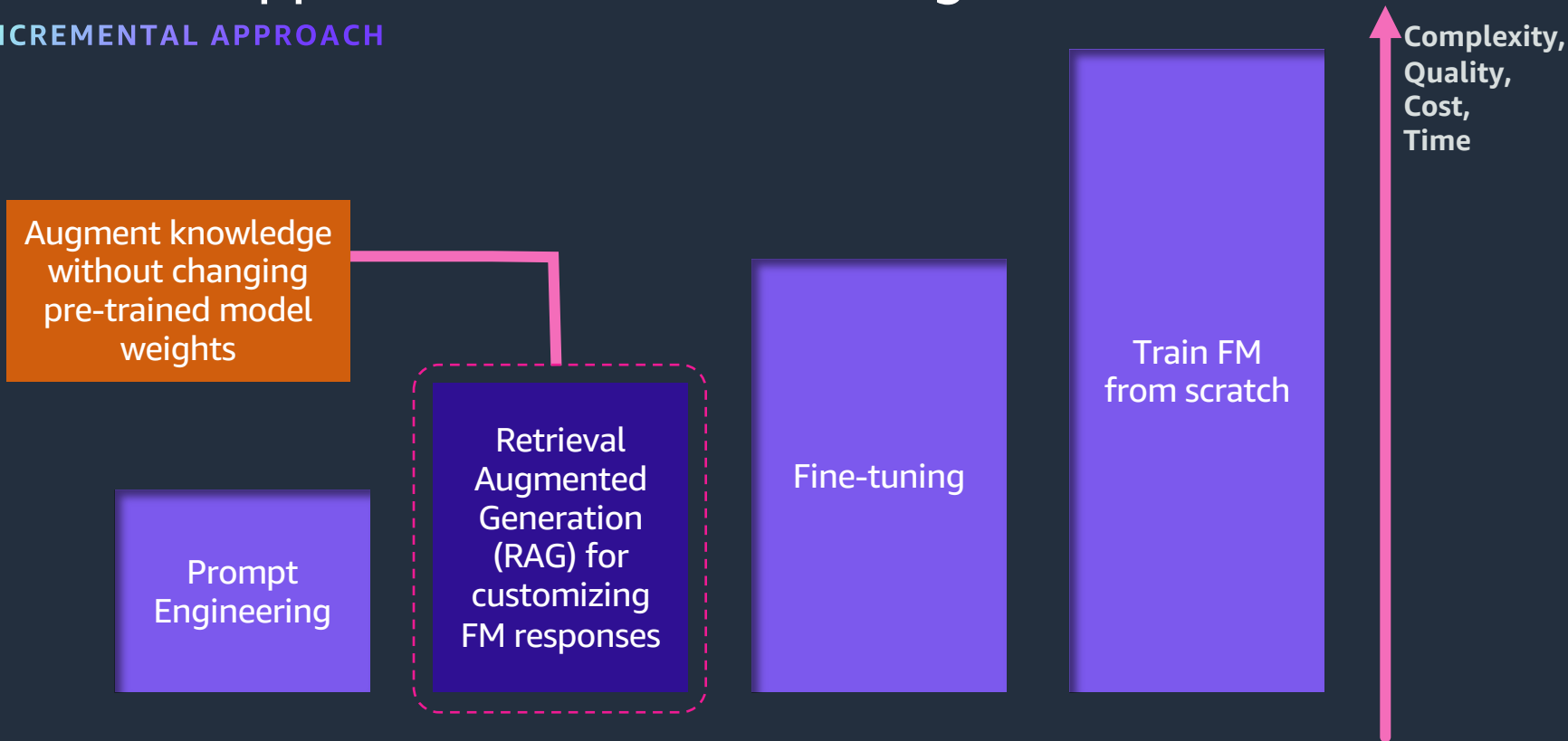
Generic
generative AI



Generative AI that
knows your business
and your customers

Common approaches for customizing FMs

INCREMENTAL APPROACH



Security considerations for generative AI

COMPLIANCE & GOVERNANCE

The policies, procedures, and reporting needed to empower the business while minimizing risk

Create generative AI usage guidelines

Establish process for output validation

Develop monitoring & reporting processes

LEGAL & PRIVACY

The specific regulatory, legal, and privacy requirements for using or creating generative AI solutions.

Retain control of your data

Encrypt data in transit and at rest

Support regulatory standards

CONTROLS

The implementation of security controls that are used to mitigate risk.

Human-in-the-loop

Explainability & audibility

Testing strategy

Identity and access management

RISK MANAGEMENT

Identification of potential threats to generative AI solutions and recommended mitigations.

Threat modeling

Third-party risk assessments

Ownership of data, including prompts and responses

RESILIENCE

How to architect generative AI solutions to maintain availability and meet business SLAs.

Data management strategy

Availability

High Availability and Disaster Recovery strategy

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 PII, and custom Regex.

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

What does **the future**
hold for generative AI?

Agents

Multi-modal

Multiple models

AI Policies & Standards

Get moving with generative AI



Move quickly



Customize



Scale



Thank you!

Brian Helms

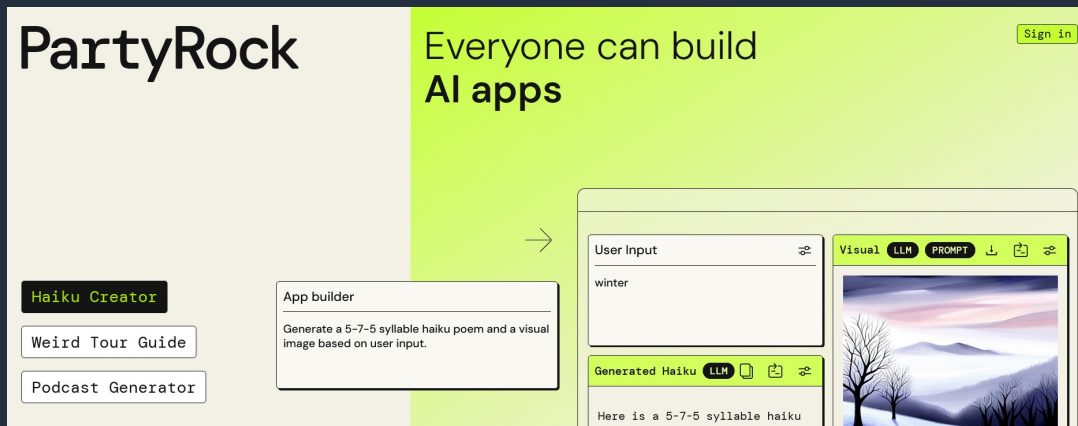
Solutions Architect | Partner & Channel Development

brihelms@amazon.com



Amazon Bedrock Playground PartyRock

- AWS account not required.
- Try out featured apps built with PartyRock.
- Learn generative AI fundamentals. No coding required.
- Access to powerful foundation models from Amazon Bedrock.
- Simple to share, easy to get inspired.



Scan to get started

Public Sector eBook

Get Started with Generative AI

Generative AI can open doors to endless possibilities increasing creativity, productivity, and progress within public sector organizations. With AWS, you can build and scale generative AI applications with security, privacy, and responsible AI built in from day one.

- -Jeff Kratz, AWS Vice President for Worldwide Public Sector Industry Sale



Scan to download eBook