

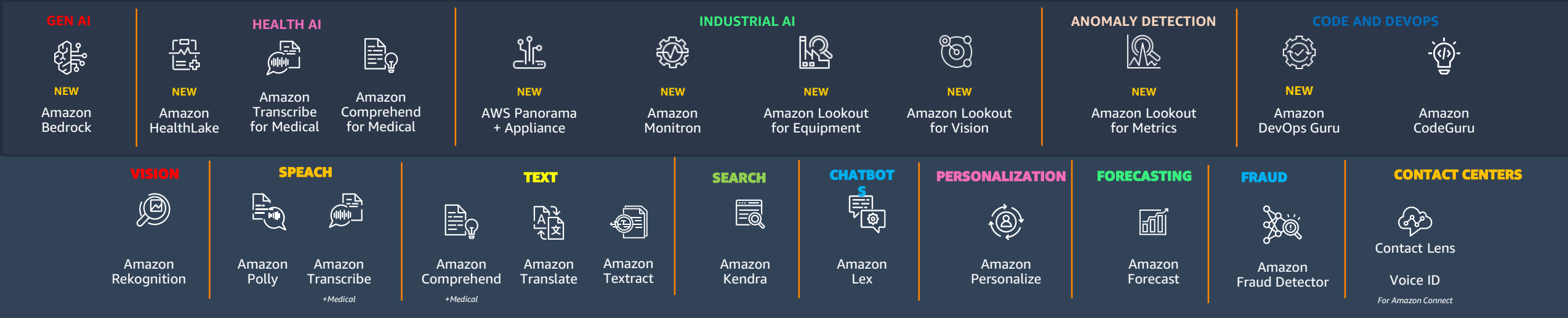


AI and Machine Learning : *Accelerating Innovation and Augmenting Capacity*

Mary Strain
Artificial Intelligence and Machine
Learning Services, EDU
AWS

The AWS AI/ML Stack

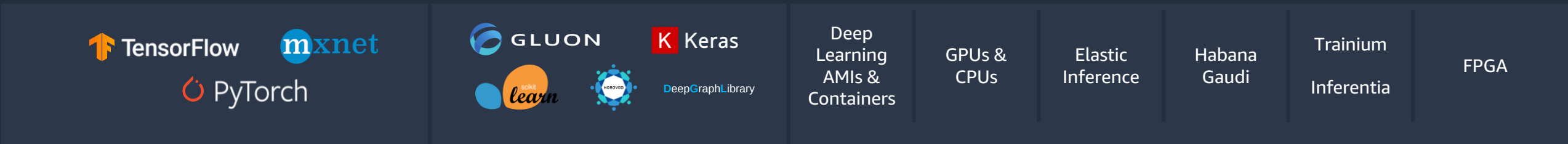
AI SERVICES



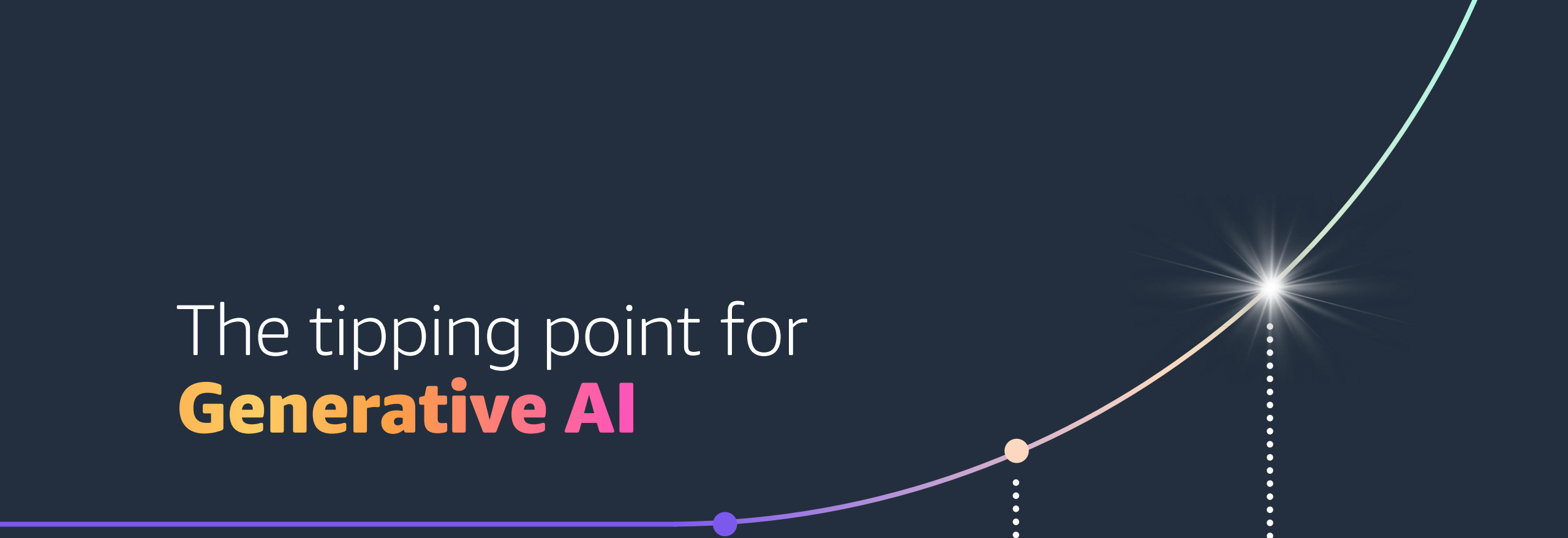
ML SERVICES



FRAMEWORKS & INFRASTRUCTURE



The tipping point for **Generative AI**



MASSIVE PROLIFERATION
OF DATA

AVAILABILITY OF
SCALABLE COMPUTE
CAPACITY

MACHINE LEARNING
INNOVATION

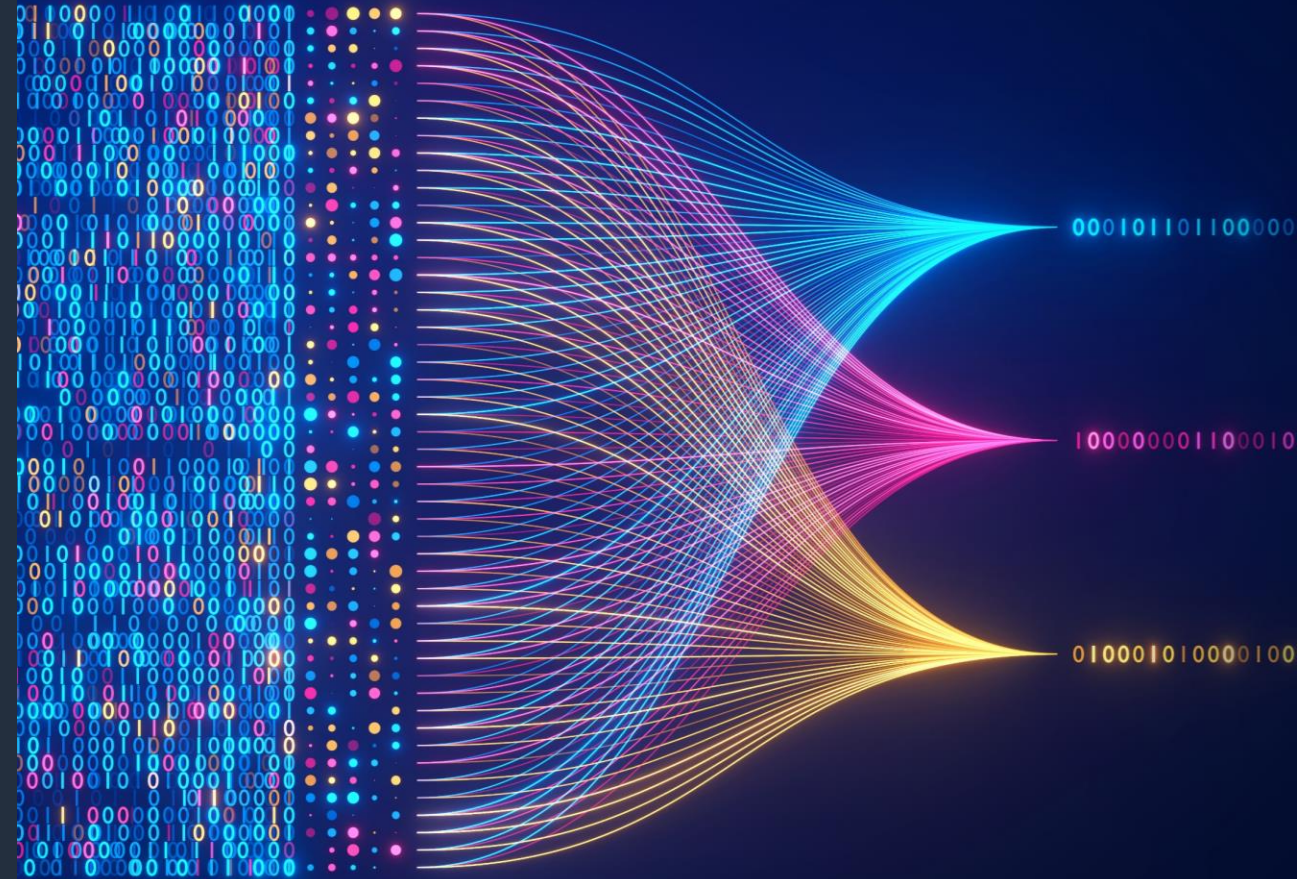
Generative AI is powered by foundation models

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

Customize FMs using your data for domain specific tasks



What could go wrong?



Hallucinations

Answers that are factually incorrect, irrelevant or nonsensical, because of limitations in their training data and architecture

"The world record for crossing the English channel on foot is 15 hours"



Bias

Answers that display discriminatory behaviour resulting in prejudiced or unequal treatment of a particular group or groups

"Generate a picture of a person cleaning" returns overwhelmingly women



Copyright & IP

The rights of content creators from whom training data is collected remains uncertain and is currently being challenged

Getty suing creators of Stable Diffusion alleging the improper use of its photos



Security & privacy

Some model providers use and store data for training purposes. Entire end-to-end data pipelines require security and data privacy controls.

Samsung engineers accidentally releasing source code by putting into ChatGPT for debugging

Emerging Regulatory Environment

K-12 GENERATIVE AI READINESS CHECKLIST



LEAD PARTNERS



COSN
LEADING EDUCATION INNOVATION

ARTIFICIAL INTELLIGENCE

Pennsylvania, Virginia Governors Issue AI Executive Orders

HILL

News

Policy

Business

Health

Opinion

Events

Jobs

Hill
TV



ADMINISTRATION

Biden issues sweeping executive order on AI

CLASSROOM TEC

The 93 Questions Schools Should Ask About AI



By [Alyson Klein](#) — October 17, 2023 ⌚ 2 min read

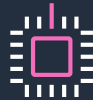
How AWS is Approaching Generative AI



Flexibility



Secure
customization



The most cost-
effective
infrastructure



The easiest way
to build with FMs



Generative AI-
powered solutions

Amazon Bedrock

Choice of foundation models

AI21labs

JURASSIC-2

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

cohere

COMMAND + EMBED

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

stability.ai

STABLE DIFFUSION XL 1.0

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

ANTHROPIC

CLAUDE 2

LLM for thoughtful dialogue, content creation, complex reasoning, creativity, and coding, based on Constitutional AI and harmlessness training

amazon

AMAZON TITAN

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



Build with publicly available foundation models

AVAILABLE ON SAGEMAKER JUMPSTART

AI21labs

Models
Jurassic-2 Ultra, Mid
Contextual answers
Summarize
Paraphrase
Grammatical error
correction

Tasks
Text generation
Long-form
generation
Summarization
Paraphrasing
Chat
Information
extraction

Meta AI

Models
Llama 2 7B, 13B, 70B

Tasks
Question answering
Chat
Summarization
Paraphrasing
Sentiment analysis
Text generation

cohere

Models
Cohere
Command XL

Tasks
Text generation
Information
extraction
Question answering
Summarization

Hugging Face

Models
Falcon-7B, 40B
Open LLaMA
RedPajama
MPT-7B
BloomZ 176B
Flan T-5 models (8 variants)
DistilGPT2
GPT NeoXT
Bloom models
(3 variants)

Tasks
Machine translation
Question answering
Summarization

stability.ai

Models
Stable Diffusion XL 1.0
2.1 base
Upscaling
Inpainting

Tasks
Generate photo-realistic
images from text input
Improve quality of
generated images

Features
Fine-tuning on Stable
Diffusion 2.1 base
model

Lighton

Models
Lyra-Fr
10B, Mini

Tasks
Text generation
Keyword extraction
Information extraction
Question answering
Summarization
Sentiment analysis
Classification

databricks

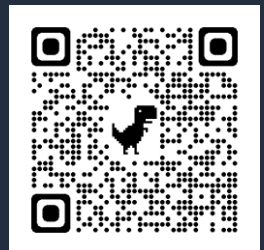
Models
Dolly

Tasks
Question answering
Chat
Summarization
Paraphrasing
Sentiment analysis
Text generation

alexa

Models
AlexaTM 20B

Tasks
Machine translation
Question answering
Summarization
Annotation
Data generation



Healthcare & Life Sciences

Ambient digital scribe

Medical imaging

Drug discovery

Enhance clinical trials

Research reporting

Industrial & Manufacturing

Product design

Operational efficiency

Maintenance Assistants

Supply chain optimization

Equipment diagnostics

Financial Services

Portfolio management

Financial documentation

Intelligent advisory

Fraud detection

Compliance assistant

Retail

Pricing optimization

Virtual try-ons review

Marketing Optimization

Product descriptions

Pers. Recommendations

Media & Entertainment

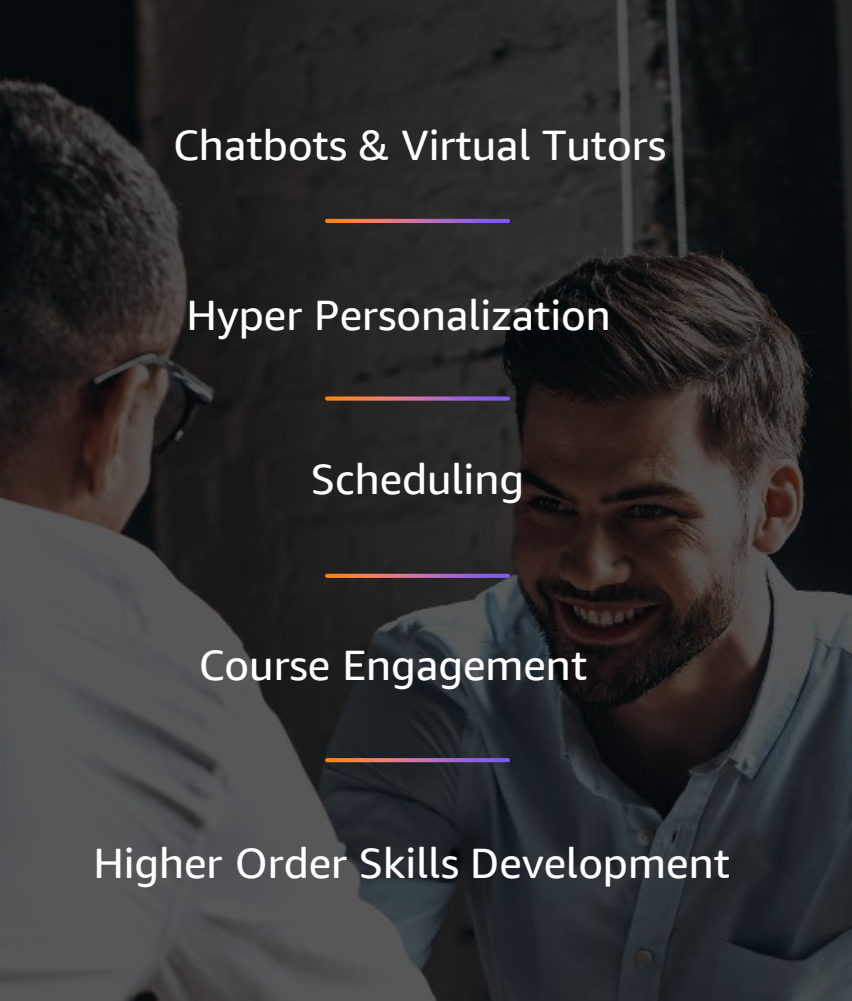
HQ content at scale

Enrich broadcast content

Automated content tagging

Optimize subscriber exper.

Automated highlights gen.



Chatbots & Virtual Tutors

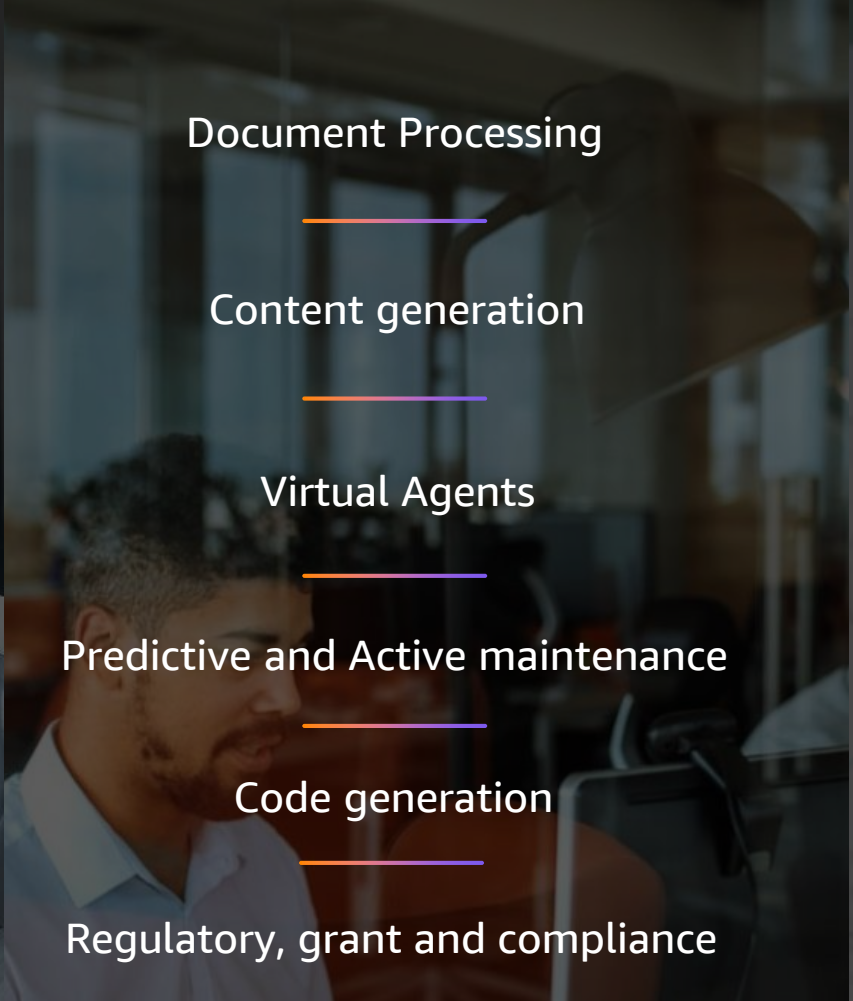
Hyper Personalization

Scheduling

Course Engagement

Higher Order Skills Development

**Teaching and
Learning**



Document Processing

Content generation

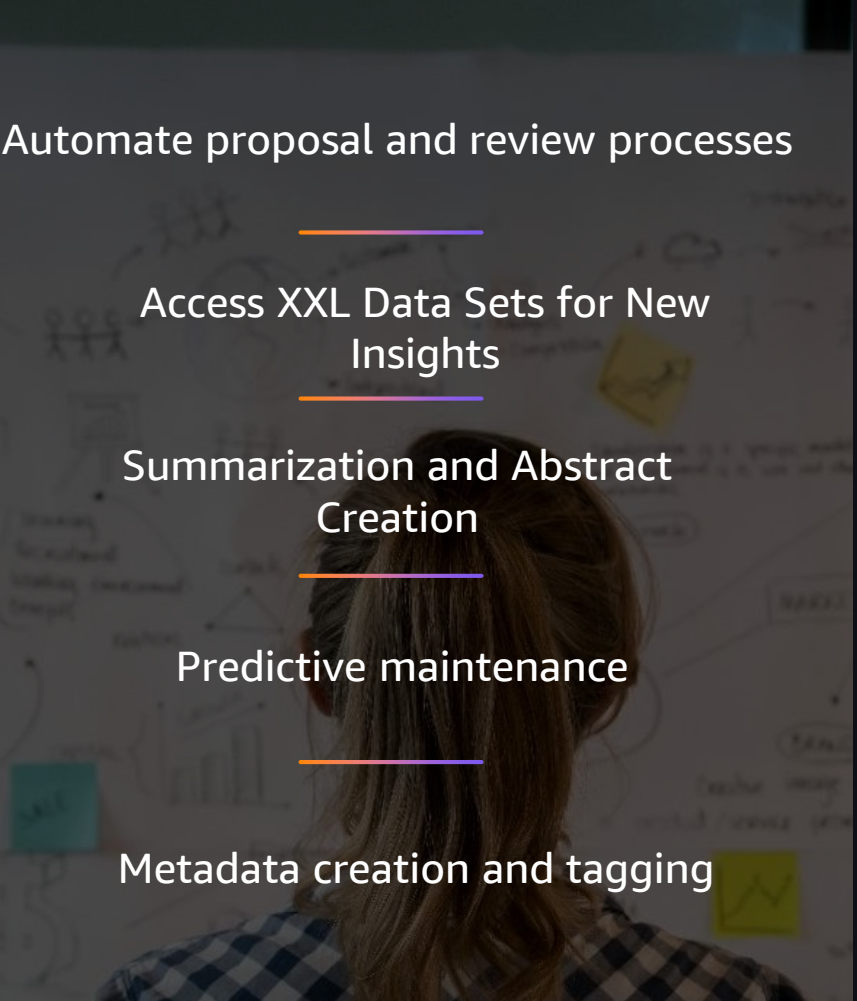
Virtual Agents

Predictive and Active maintenance

Code generation

Regulatory, grant and compliance

Administration



Automate proposal and review processes

Access XXL Data Sets for New
Insights

Summarization and Abstract
Creation

Predictive maintenance

Metadata creation and tagging

**Libraries ,
Research, Athletics**

Use Cases – AI/ML Powered Scientific Discovery

- Neural solvers of differential equations
- Hydropower station location planning
- High-throughput virtual screening
- Navigation in the hypothesis space
- Super-resolution 3D live-cell/Molecular imaging
- Planning chemical synthesis pathway
- Weather forecasting
- Synthetic EHR generation
- Rate event selection in particle collisions
- Language modeling for biomedical sequences
- Battery design optimization
- Magnetic control of nuclear fusion reactors

Pillars for the responsible use of AI

Value alignment

Systems should be designed and used in ways that align with company mission, social norms, and legal compliance

Inclusion

Inclusion of unique skills, experiences, perspectives, and cultural backgrounds

Training & education

Appropriate knowledge sharing and education to understand purpose, use, and impact

Accountability

Structured maintaining human involvement and responsibility for design, development, decision processes, and outcomes

Privacy & security

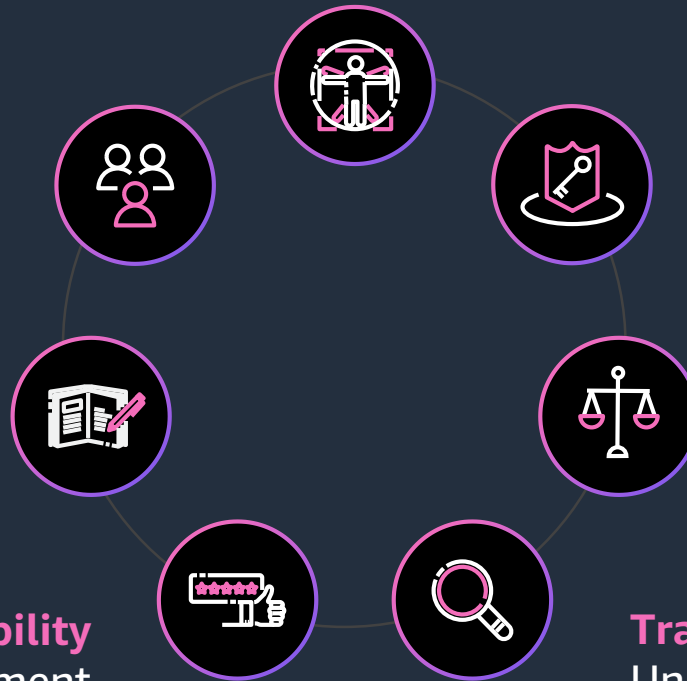
Protects the quality and integrity of data used, its relevance, access, and processing

Fairness

Systems must be designed to minimize bias and promote inclusive representation

Transparency & explainability

Understanding how data is used, how decisions and outcomes are made in a human understandable way



Gen AI Augmented Research Assistants

- Conduct literature reviews
- Prepare materials for submission to granting agencies and foundations
- Prepare materials for UAB Committee review
- Generate recruitment and interview questions
- Transcribe and summarize interviews, project meetings, seminars
- Summarize project results, generate reports, presentations for funding agencies
- Manuscript and proposal writing





Thank you!

Mary Strain

straimar@amazon.com