# Enabling GenAI in the Enterprise

August 23, 2024

Cindi, Brad, Heiko

# Who we are

Cindi Reynolds
Principal Cyber Security Engineer

Heiko Zuerker
Principal Architect

Brad Kostreva
Assistant Director

Revenue $36 billion

FORTUNE 500® rank: #110

10,300+ Financial Professionals

8,200+ Employees

Headquartered in Milwaukee, Wisconsin

Downtown NYC Corporate Office

For more than 165 years, Northwestern Mutual has been helping families and businesses achieve financial security.

B

# Disclaimer



This presentation is not from the "shiny side" side of Generative AI.

We're going to look under the hood and hear from the teams that made it happen in a regulated environment.

**H**

# Knock Knock...Who's There?

Go away, you're just a fad.

It's GenAI.

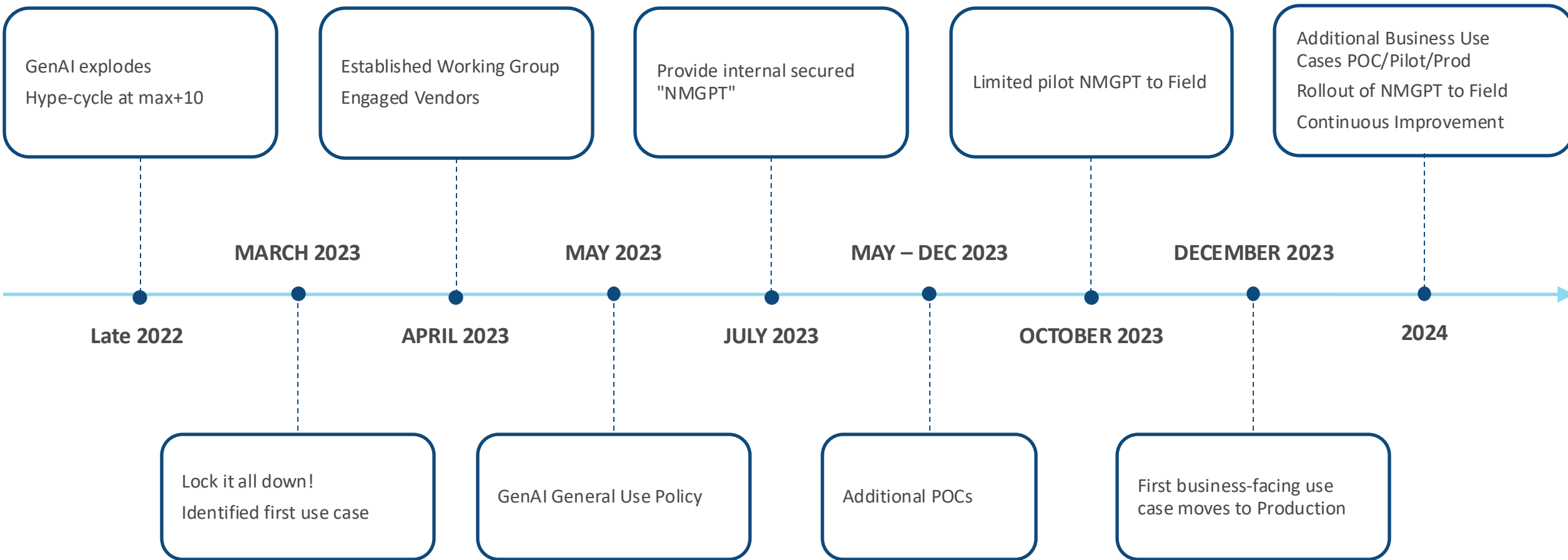Ah great... Now what?

Oh no, this is happening!

I'm glad you asked...

H

# Fact!

This GenAI thing is happening,
if you want it or not.

**Find your way to "YES"!**

# Our Journey – 10,000 ft. View

**GenAI explodes**
Hype-cycle at max+10

**Established Working Group**
Engaged Vendors

**Provide internal secured "NMGPT"**

**Limited pilot NMGPT to Field**

**Additional Business Use Cases POC/Pilot/Prod**
Rollout of NMGPT to Field
Continuous Improvement

**MARCH 2023**

**MAY 2023**

**MAY – DEC 2023**

**DECEMBER 2023**

**Late 2022**

**APRIL 2023**

**JULY 2023**

**OCTOBER 2023**

**2024**

**Lock it all down!**
Identified first use case

**GenAI General Use Policy**

**Additional POCs**

**First business-facing use case moves to Production**

# Fact!

GenAI does **<u>not</u>** fit

into your existing processes.

H

# We had to embrace uncertainty!

- Technology continues to evolve fast

- Rules, laws, and regulations continue to evolve

- Initially use more guidelines and less standards

- There's no correct answer

- Use best judgement & course correct as-needed

- Manage risk

**H**

# Where do you start?

1. Identify key stakeholders

2. Assemble a "2-pizza" core team

   • Senior level with a yes attitude to navigate organizational challenges

3. Take a chance and just get it done

   • Do not boil the ocean → what's good enough?

4. Create clear internal (Gen)-AI usage policies

5. Create a playbook – repeatable patterns, process, and procedures

c

# Speed, Agility, Knowledge...and Picking Your Kickball Team

**GEN AI WORKING GROUP WORKSTREAMS:**

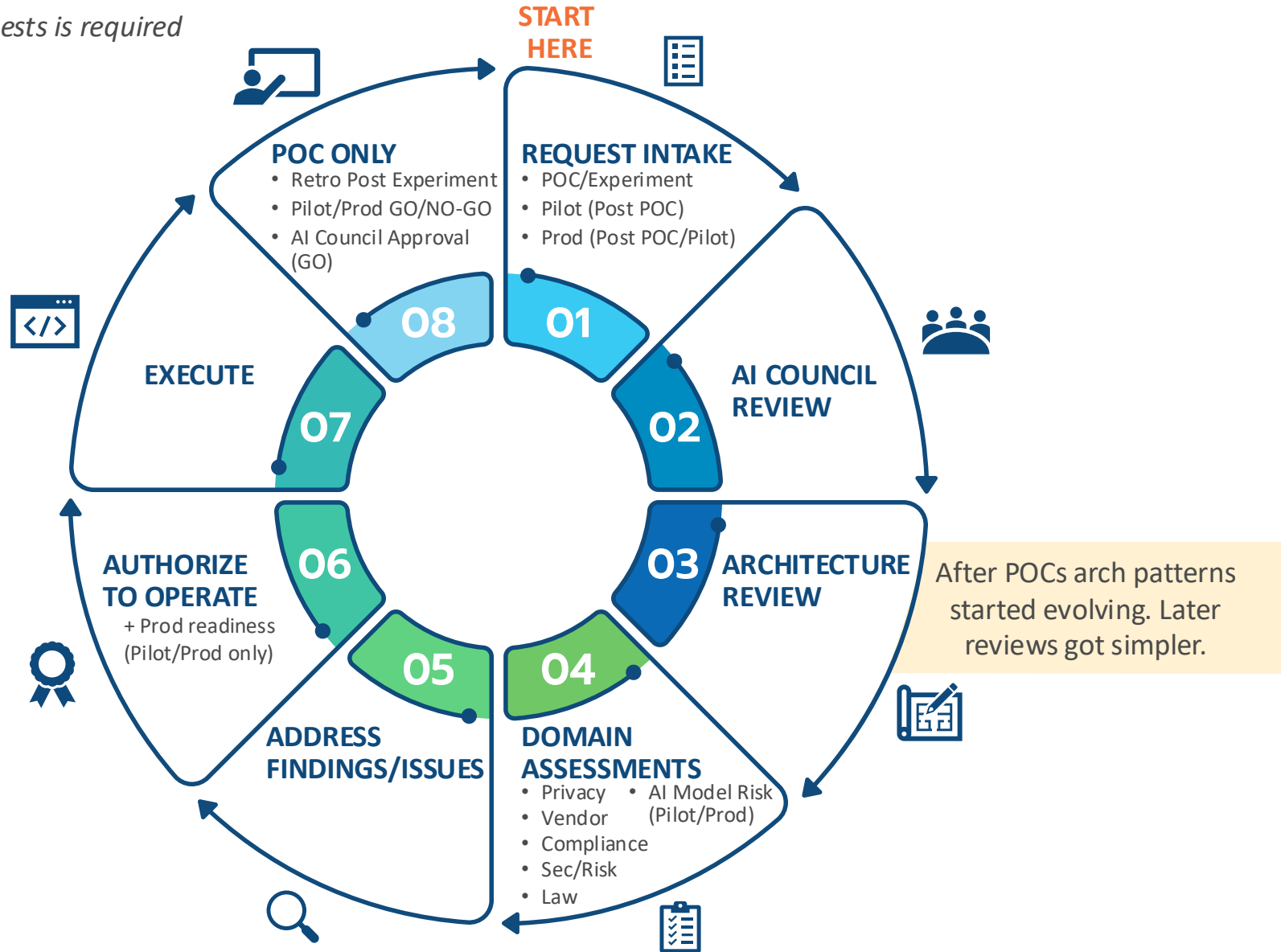| Architecture Patterns & Assessment Suitability | Cost Management Requirements & Guidelines | Data Access & Management | Engineering Enablement | Experiment Process Flow |
|---|---|---|---|---|
| General Access API Enablement | Law/Privacy/ Compliance Requirements | Logging Strategy & Requirements | Model Governance | Security Guardrails |

**GOAL:**
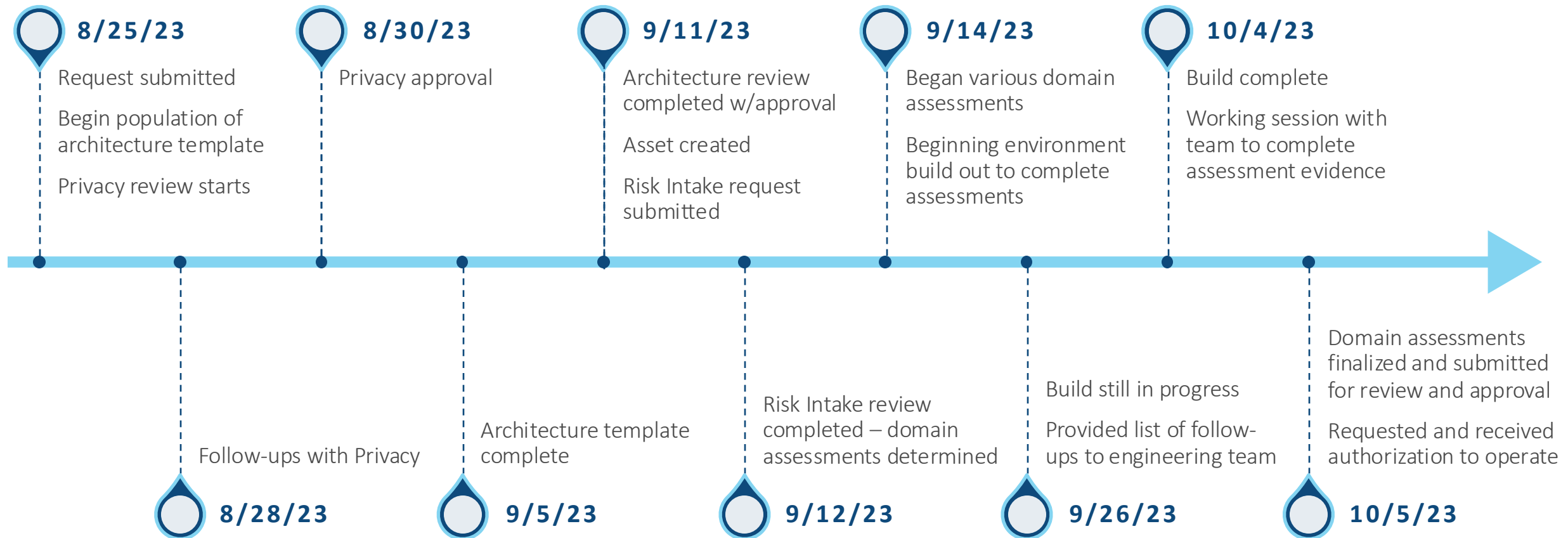Create a standardized, efficient, safe (secure), and scalable process that will be adopted across the organization.

C

# High Level Process: POC/Pilot/Prod

*Lifecycle management for all requests is required*



**START HERE**

**REQUEST INTAKE**
- POC/Experiment
- Pilot (Post POC)
- Prod (Post POC/Pilot)

**O1**

**O2**

**AI COUNCIL REVIEW**

**O3**

**ARCHITECTURE REVIEW**

After POCs arch patterns started evolving. Later reviews got simpler.

**O4**

**DOMAIN ASSESSMENTS**
- Privacy
- Vendor
- Compliance
- Sec/Risk
- Law
- AI Model Risk (Pilot/Prod)

**O5**

**ADDRESS FINDINGS/ISSUES**

**O6**

**AUTHORIZE TO OPERATE**
+ Prod readiness (Pilot/Prod only)

**O7**

**EXECUTE**

**O8**

**POC ONLY**
- Retro Post Experiment
- Pilot/Prod GO/NO-GO
- AI Council Approval (GO)

# Test Driving The Process

## REAL EXPERIMENT/POC:

**8/25/23**

Request submitted

Begin population of architecture template

Privacy review starts

**8/30/23**

Privacy approval

**9/11/23**

Architecture review completed w/approval

Asset created

Risk Intake request submitted

**9/14/23**

Began various domain assessments

Beginning environment build out to complete assessments

**10/4/23**

Build complete

Working session with team to complete assessment evidence

Follow-ups with Privacy

**8/28/23**

Architecture template complete

**9/5/23**

Risk Intake review completed – domain assessments determined

**9/12/23**

Build still in progress

Provided list of follow-ups to engineering team

**9/26/23**

Domain assessments finalized and submitted for review and approval

Requested and received authorization to operate

**10/5/23**

**30 BUSINESS DAYS**

# Introducing...The Pizza Tracker

👍 Self-service

Always up to date

Visibility to process step status

Links to use case details, asset inventory & backlog

👎 Manually updated

| Use Case | Request Type | AI Council | Architecture | Privacy | Risk | Compliance | AI Model Risk | ATO | Deployment |
|----------|--------------|------------|--------------|---------|------|------------|---------------|-----|------------|
| Use Case 1 | POC | APPROVED | APPROVED | APPROVED | APPROVED | APPROVED | N/A | COMPLETE | COMPLETE |
| Use Case 1 | PILOT | PENDING | | | | | | | |
| Use Case 2 | POC | PENDING | | | | | | | |
| Use Case 3 | POC | APPROVED | WIP | APPROVED | | | | | |

C

# Fact!

You <u>will</u> experiment with production data.

B

# The data challenge

- Synthetic data great for initial & basic development

- Accuracy testing in non-prod requires production data

- Provide non-prod environments with production controls

- Data privacy, compliance, and security is our top priority
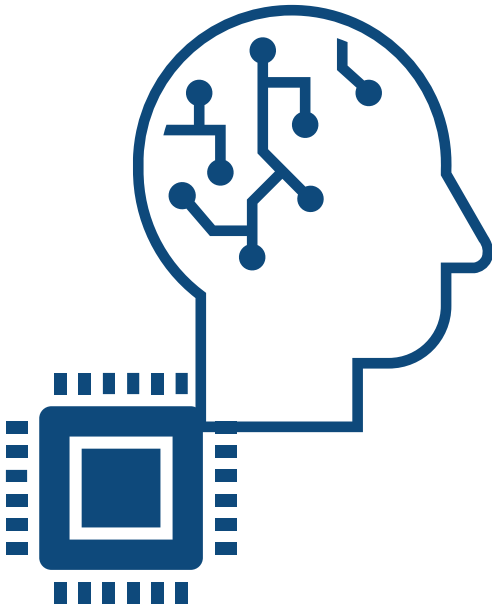
- Get those teams involved from the start

B

# More data challenges

- Use RAG (retrieval augmented generation)

- Avoid training or fine-tuning models, due to high cost and time

- Only use well curated data sources

- Tag & label your data

- Feedback mechanism critical to improve

- Review data access

- Start with non-sensitive data sources - vetted for acceptable use

**B**

# GenAI Infrastructure

## COMMON COMPONENTS

- AI Endpoint

- Search Service

- Vector Database

- Index / File Storage for Data Pipeline(s)

- Document Recognition & OCR

## LOGGING PROXY

- Authentication

- Attestation

- Incident Response

*Fraud / Abuse Liability to Vendor

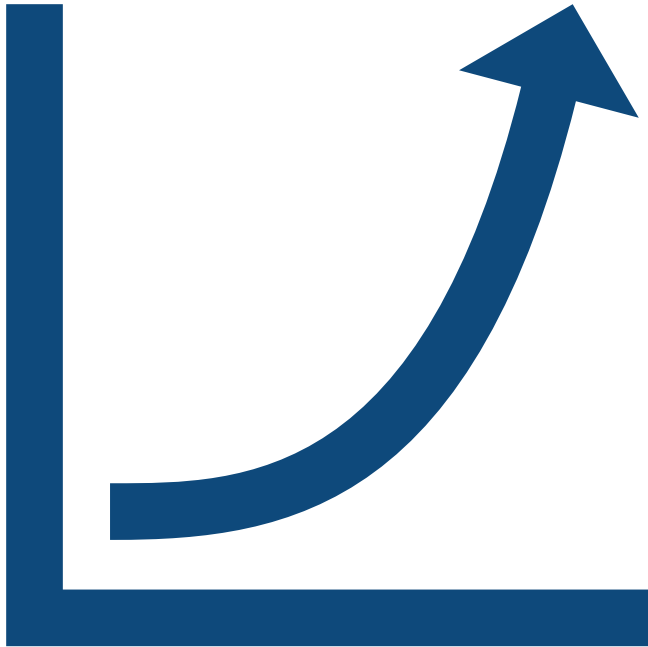**B**

# Encourage safe experimentation

- "GenPop GenAI API Endpoint"

- Low bar to entry through basic training & attestation

- Encourage use through "hackathons" / crowdsourced ideas

- Clear rules! No production data! No business use!
  - Technical controls to prevent access from application platforms
  - Ephemeral only to prove out concepts and build business cases

- EXTREMELY valuable
  - Business and personal development

# It worked!

- GenAI huge success!

- 10,000+ users of "NMGPT"

- Lots of new proposals every month

- Commoditize process & infrastructure

- Continuously improving

**B**

# Lessons learned

1. Can't avoid GenAI, it will soon be everywhere.

2. Provide a safe and accessible space to learn and experiment.

3. Don't sacrifice security/privacy/compliance for speed.

4. Don't get paralyzed by ambiguity and uncertainty. Can't be rigid, must be flexible and adjust on the fly.

5. Don't chase technologies, pick the right partners. Adopt native technologies!

6. It's a team sport! Involve all stakeholders from the start.

H