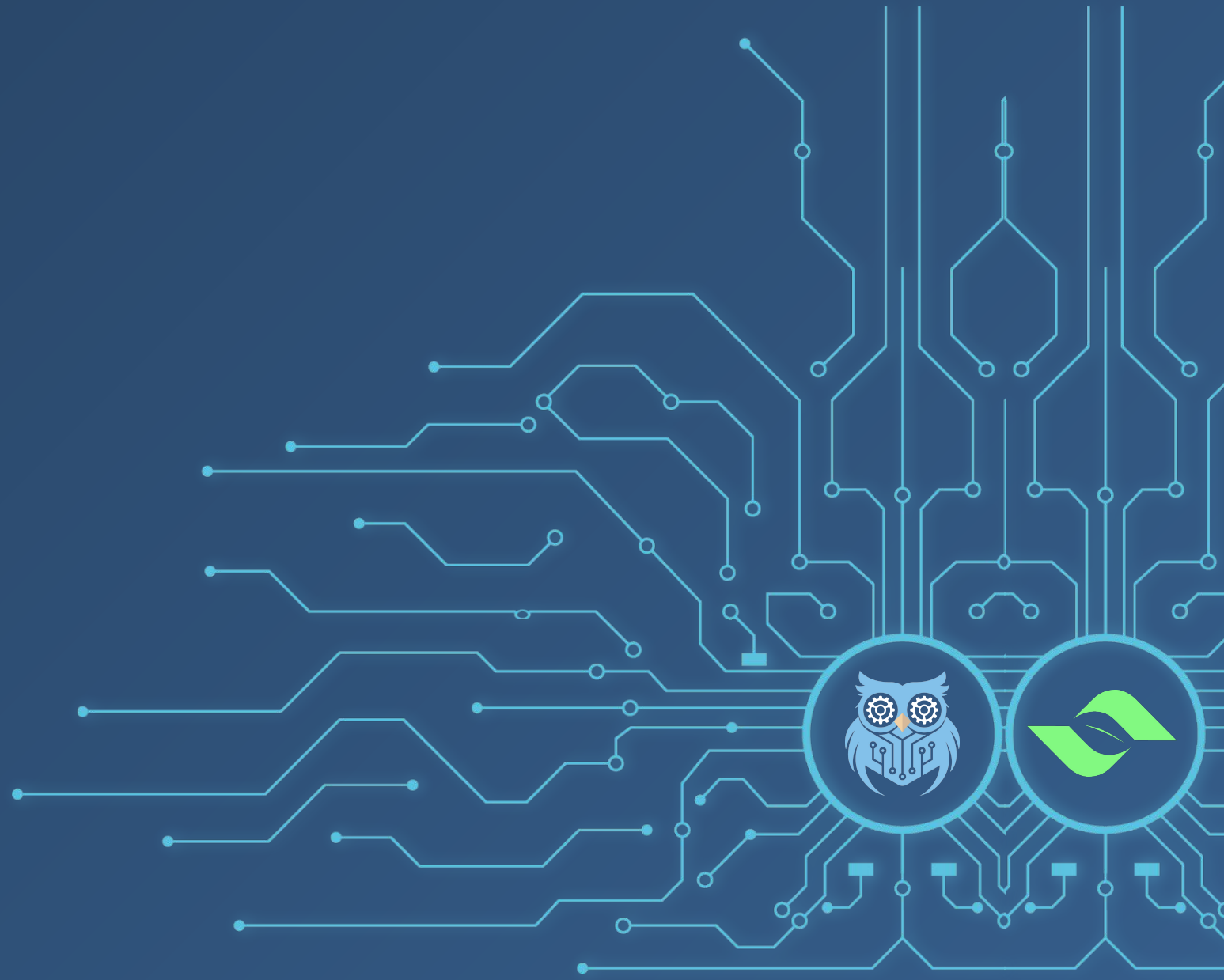
Federal AI Playbook:

From Data Chaos to Intelligence

Charting the Course for Federal AI Excellence



Where are we now?

- → AI Adoption Trends
- → Key Drivers
  - Efficiency
  - National Security
  - Citizen Services
- → Challenges
  - Legacy Systems
  - Data Silos
  - Regulatory Demands
- → Real world example: **Department of Commerce**
  - Internal & External
  - Data Source of Record
  - Multi-Stakeholder

75%

Of all Federal AI programs report facing substantial barriers.[1]

+147%

Increase in public AI applications from 2023 to 2024.[2]

ADMINISTRATIVE CHALLENGES

USDS (DOGE)

→ **Primary Goals of DOGE:**
- Modernize Federal Technology / Software
- Reduce Federal Spending
- Automate Governmental Tasks

→ **Challenges Presented**
- Workforce Impacts (personnel & chaos)
- Data Security & Privacy
- Transparency and Oversight

→ **Opportunities in Leveraging AI for Efficiency**
- Enhanced Decision-Making (context)
- Improved Public Services (speed)
- Cost Savings

Life after DOGE...

→ **Be proactive** - scrutinize yourself before someone else does.

→ Make your **personnel superhuman** through automation and AI augmentation (training & tools).

→ Create "**efficiency reports**" to showcase improvements.

→ Balance **technical improvements** with **functional outcomes and data security**.

→ Get **data-driven** about everything.

THE DATA-FIRST IMPERATIVE

AI Starts with Data

- → Garbage in, garbage out
- → Data Quality is Foundational (and multi-faceted)
- → Common Federal Data Challenges
  - Duplication
  - Inconsistency
  - Accessibility
- → Example: Census Bureau's approach to structuring large-scale datasets.
  - Enterprise Data Lake (EDL)
  - Data-centric (vs survey-centric)
  - TIGER for mapping geospatial
  - Application of Big Data Techniques

42%

Data Leaders cite data quality as the main obstacle to Gen AI.[3]

CPE Question #1

Which of the following best describes the biggest challenge your agency faces in adopting AI for modernization?

Follow-up: How has your agency approached addressing this challenge?

BREAKING DOWN SILOS

Collaboration Over Competition

- **The cost of data silos:**
  - Inefficiency and Increased Cost
  - Reduced Data Quality and Consistency
  - Missed Insights & Opportunities
- **Strategies to enable cross-agency collaboration**
  - Shared Data Governance Framework
  - Interoperable Technology Standards
  - Prioritize Multi-Agency Data Sharing Agreements

DoD's Joint AI Center (JAIC)

- Established to accelerate the adoption of AI across the DoD, the JAIC attempted to centralized AI efforts that were previously fragmented across various branches and agencies.
- Structured to unify AI projects under a single umbrella, to fosters cross-department collaboration and enable shared learning and resource optimization.
- Due to a number of challenges, the JAIC struggled to gain traction across the DoD:
  - Organizational Structure and Authority
  - Lack of cross-agency Governance and Security Strategy
  - Data Management and Accessibility

AI GOVERNANCE AND ETHICS

Building Trust in AI

→ **Overview: AI Ethics Guidelines**

- NIST AI Rist Mgmt Framework
- EO 13960
- DoD Ethical Principles for AI
- AI Bill of Rights Blueprint
- FTC Guidance on AI & Algorithms
- National AI Initiative Act of 2020
- GSA AI Policies

True Goals:

1. Explainability
2. Accountability
3. Transparency

What's Missing?

Cracking open the black box.

GSA's Implementation of Ethical AI Frameworks.

By incorporating principles of accountability, the GSA ensures that AI projects are aligned with ethical guidelines and that responsible parties are clearly identified throughout the lifecycle of AI deployment. Transparency is further promoted through public-facing documentation and pilot projects, allowing for broader stakeholder engagement and oversight.

Stages to AI Success

Phase 1: Data audit and governance foundation.

- → Clear understanding of data assets (quality, structure, relevance)
- → Robust governance for consistency & compliance
- → Data Catalog to trace decision-chain and transparency

Phase 2: Building scalable, interoperable systems.

- → Scaling should be both horizontal (increasing data volume) and vertical (advancements in tech)
- → Prioritize interoperability (ODNI)
- → Leverage cloud-based arch and standardized APIs for future-insulation (no future-proof)

Phase 3: Incremental AI deployment & measurement.

- → Begin with pilots to validate models & approaches
- → Gradually introduce “human-in-the-loop” systems to ensure accountability & trust
- → Progress towards ful AI
- → Continuously measure performance and outcomes

From Theory to Action

Case Study 1:

IRS's journey to modernize fraud detection with AI.

ROI: 20-fold increase in data processing speeds, enhancing detection and prevention of tax fraud.

Case Study 2:

NASA's AI-driven satellite imagery analysis.

Outcome: creation of the HLS Geospatial Foundation Model, an open-source AI tool.

Lessons Learned:

- ➔ POC Approach to AI Development
- ➔ Show Impact Quickly and Often
- ➔ Maintain Control on Requirements and Outcomes

CPE Question #2

True or False: "Interagency data sharing has improved significantly over the past five years, primarily due to advancements in secure data-sharing technologies.

Follow-up: What technologies, policies, or cultural shifts have been most impactful in facilitating or impeding interagency data sharing?

MANAGING TODAY WHILE BUILDING FOR TOMORROW

Balancing Legacy Systems & Innovation

- → Strategies for integrating AI in legacy systems
  - Metadata extraction / APIs
  - Managed migration
- → The role of hybrid architectures and cloud migration
  - “Front Office” vs “Infrastructure”
- → The “why” vs The “how”
  - Why: AI can offer rapid digestion & contextualization of aging resources
  - How: It Depends

THE ROLE OF PARTNERSHIPS

Leveraging the Ecosystem

- → The power of public-private collaboration
- → Examples: Cyberhill's work with Federal Agencies to enhance secure commercial software deployments
- → Industry advancements in data cataloging and sharing.
- → Build v Buy

+7M

Low-income homes have improved energy efficiency w/ DOE public-private partnerships

Charting Your Path to Success

1. Start with Data Readiness

Conduct a comprehensive audit to understand your data's quality, structure, and relevance. Implement governance practices to ensure data accuracy, transparency, and compliance from the outset.

2. Approachable & Actionable

Identify low-risk, high-impact AI opportunities to pilot. Focus on small, achievable goals to build confidence and refine processes incrementally.

3. Align with Federal AI Guidance

Leverage existing federal frameworks and initiatives, such as the NIST AI RMF, to ensure compliance and strategic alignment. This ensures your efforts are scalable and interoperable across agencies.

4. Stepped Modernization

Integrate AI into legacy systems using hybrid architectures and scalable solutions. Avoid costly overhauls by building on existing infrastructure and evolving incrementally.

CPE Question #3

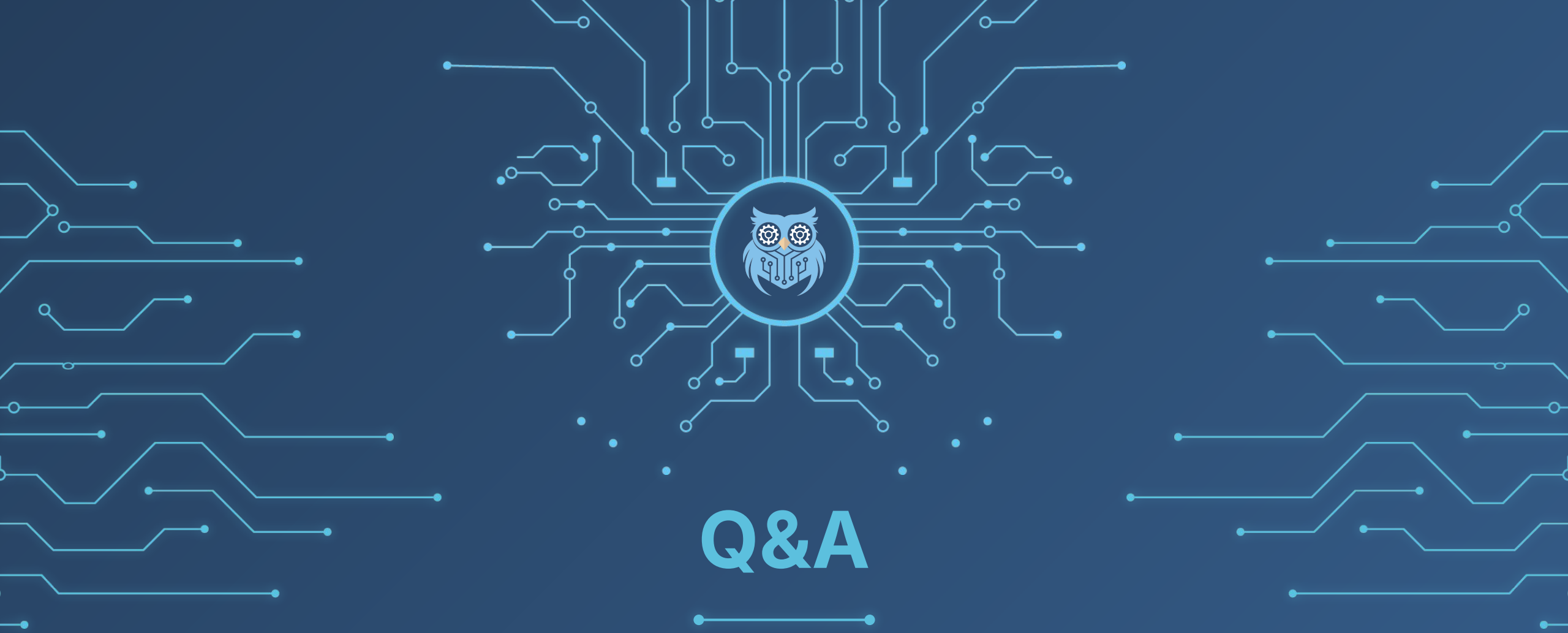
How does your agency currently prioritize the integration of AI tools for modernization?

Follow-up: What criteria or metrics are most important when deciding where to apply AI within your agency?

The critical role of data interoperability, access, and trustworthiness

data interoperability, access, and trustworthiness are critical for AI, for doing command and control across all domains, and for being a modern Defense Department.

Deputy Secretary of Defense
Kathleen Hicks



Q&A

Patrick McGarry

GM Federal, data.world

patrick@data.world

Jake McAndrew

Partner, Cyberhill

jake@cyberhillpartners.com