

Metadata Matters: The Power of AI

Tyler Morris, Senior Director, Public Sector Strategy & Growth

September 6, 2023

Agenda

- Introduction to Artificial Intelligence
- Value and Benefits of AI & Metadata
- Applications in Records Management
- Government Use Cases





Data Expansion



Daily...



306 billion emails and
500 million Tweets are made

By 2025...



463 exabytes of data
will be generated



200+ zettabytes of data
will be in cloud storage



The U.S. government is a major contributor to the information surge. Government data generation is on the rise.



Mobile devices
& applications



Smart sensors
& devices



Cloud computing
solutions



Citizen-facing
portals





Growth of Artificial Intelligence



\$1B

Of funding committed by US government to AI in 2020

70%

Of leaders support AI-driven government projects

80%

Of government agencies are at the initial or developing stages of reaching digital maturity

75%

Of governments will have at least three enterprise-wide hyper-automation initiatives by 2024



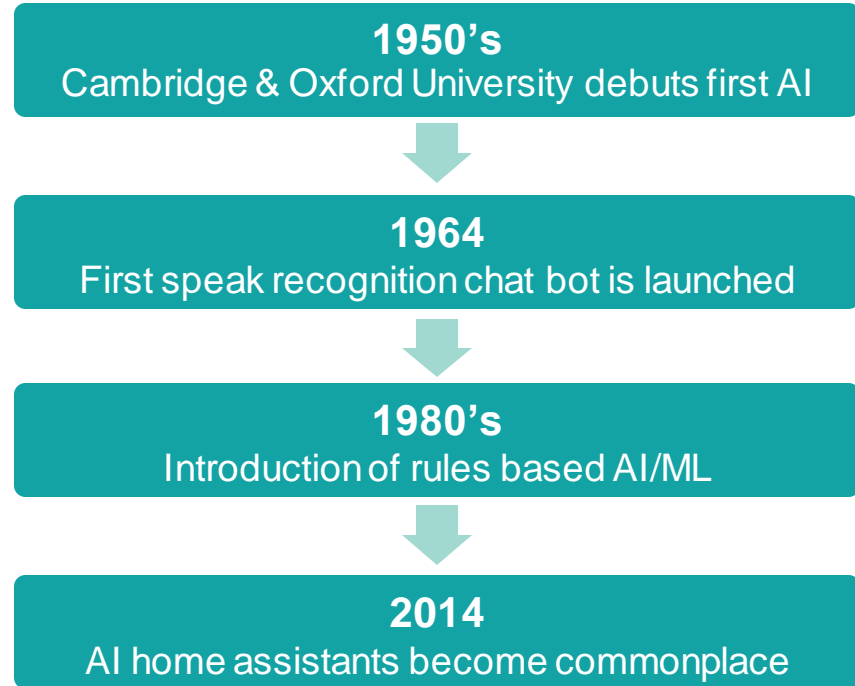


Introduction to AI/ML

“The term ‘artificial intelligence’ means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments.”

– *The National Artificial Intelligence Act of 2020*

Sources: [Timeline of AI](#)





Artificial Intelligence

The ability of a machine or computer system to perform tasks that would normally require human intelligence.

- Learning
- Solving problems
- Recognizing patterns
- Predicting outcomes
- Making decisions



Machine Learning

A subset of AI, ML learns from data without being explicitly programmed.

- Advanced algorithms make predictions and learn from outcomes
- No human intervention required
- ML improves AI over time



“Data about Data”

Metadata tells us what content a record possesses and its characteristics.

Metadata classifies records according to:

Type	Creation date
Author	Confidentiality levels
Subject matter	Data handling rules

Types of Metadata Creation



Manually input



Automated by intelligent document processing

The Importance of Metadata Creation for Digital Records

Components of Good Metadata

Provenance – who has ownership of the record?

Technical Requirements – what is needed to interact with the record?

Authenticity – is this record what it says it is?

Details – what key fields are in this record?



Only 24% of federal agencies say their electronic records contain appropriate metadata

[NARA 2021 Annual Report](#)

Metadata and its Role in Government

Metadata is essential for the public sector for various reasons:

Organization



Locating electronic records, data, and information

Accessibility



Complying with legal holds requiring preservation of data or metadata

Preservation



Identifying records that have fulfilled their retention period

Interoperability



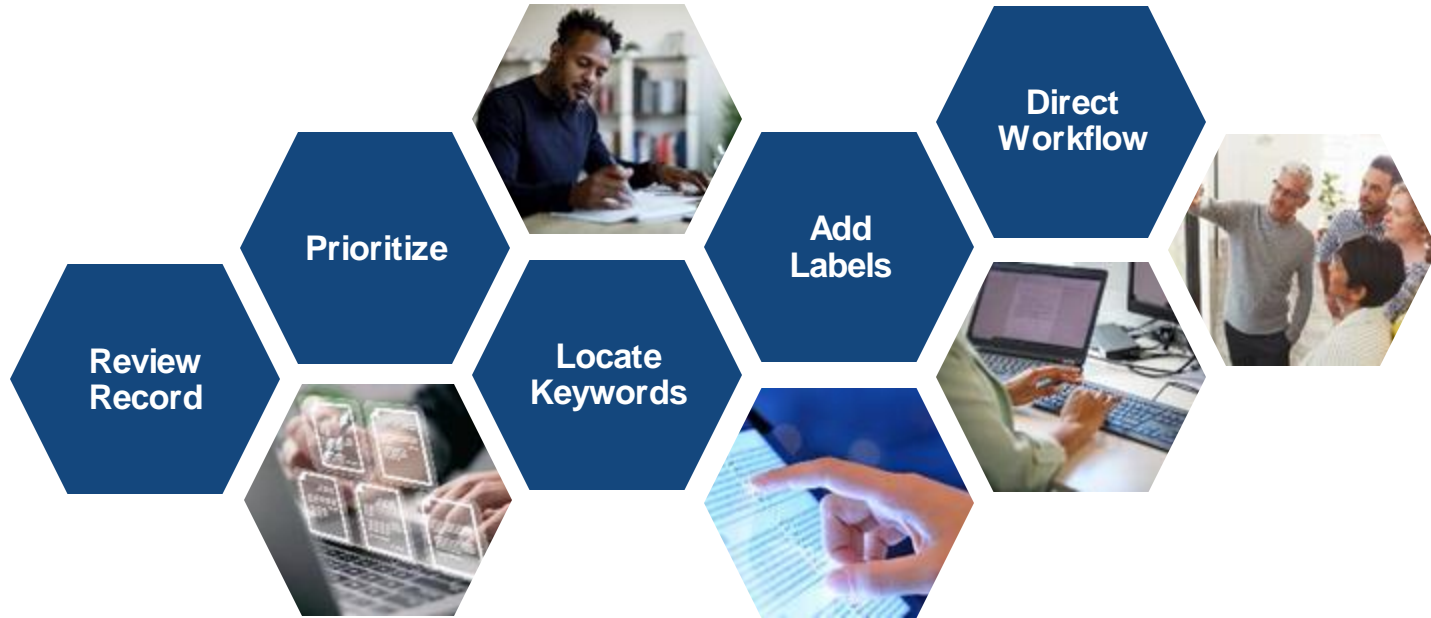
Implementing and using data analytics



Ensuring the data required for downstream application is digitized

Sources: [IRM Whitepaper](#), [IRM Executive Summary](#)

Metadata without AI





Benefits of AI



Improved efficiency



Enhanced search and retrieval



Improved security



Optimized decision-making



Significant Potential for Government Agencies

High Investment (tasks speed up by 200%)

1.2B hours

\$41.1B

Low Investment (tasks speed up by 20%)

96.7M hours

\$3.3B

Source: [Deloitte report](#)

AI/ML for Records Management



Supervised Learning Algorithms **Trainable Pattern Recognition**

Learns to make predictions from labeled data

Useful for interpreting patterns and classifying new, unlabeled records



Unsupervised Learning Algorithms **Cluster and Classify**

Used for rapidly scouring unlabeled data to identify patterns

Ideal for categorization of records rather than interpretation of records

Practical Applications



Document Classification

AI/ML can classify documents into categories predefined by the records manager

Tools Used: OCR, RPA



Document Categorization

AI/ML can group records into clusters based on characteristics defined by the records manager

Tools Used: OCR, RPA, Clustering



Information Retrieval and Search

AI/ML can find records and recommend related records in an instant. Vast results can be summarized for brevity.

Tools Used: NLP, Predictive Analytics, ML-Powered Search

Practical Applications



Redaction and Data Privacy

Detect sensitive PII or confidential information before a record is shared; and provide access security

Tools Used: OCR, RPA, Predictive Analytics, Blockchain



Information Governance

Record analysis can draw unseen insight, helping the records manager make a more informed decision

Tools Used: RPA, ML-Powered Search, Predictive Analytics



Government Use Cases

A modern, multi-story office building with a glass and metal facade. The building is partially obscured by a stylized blue mountain range at the bottom. The Iron Mountain logo, consisting of a blue triangle and the text 'IRON MOUNTAIN', is prominently displayed on the right side of the building.

IRON
MOUNTAIN

Practical Uses of AI Within Government

De-Duplication Model

- Conduct entity resolution (of records) to create a complete picture of a person
- AI Tools: **ML** to identify which records belong to the same unique individual
- Outcome: USCIS can create a full immigration history for an individual without time-consuming research across disparate systems

A-Number Management Model

- Identify which records within the USCIS identity database best match search criteria
- AI Tools: **ML** to ensure that search results presented to authorized external partners for integrations and servicing have a high degree of confidence
- Outcome: Critical in USCIS entity resolution and person surfacing



U.S. Department of Homeland Security

Person-Centric Identity

Services (PCIS) is to be the authoritative source of trusted information that provides visibility into an individual's comprehensive immigration history and status

Source: <https://www.ai.gov/ai-use-case-inventories/>

Practical Uses of AI Within Government

Natural Language Processing and Similarity Measures Application

- Conducts advanced information retrieval and estimate similarities between records
- AI Tools: NLP and similarity algorithms
- Outcomes: Identify opportunities for resource prioritization and DOE predictions (i.e., environmental reports, COVID-19)

Data Analytics and Machine Learning (DAMaL)

- Automate records classification, provide trends insights and indications of importance and risk
- AI Tools: NLP, classification algorithms, unsupervised ML learning
- Outcomes: Identify important records; identify topics to explore potential safety issues; obtain insights from operating experience; improve decision-making



U.S. Department of Energy

The Office of Environment, Health, Safety & Security (EHSS) is the DOE's central organization responsible for health, safety, environment, and security

Source: <https://www.ai.gov/ai-use-case-inventories/>



Infused with the *Power of AI*, data becomes the ignition for an unstoppable engine, propelling innovation, efficiency, and transformative success in fulfilling government missions.



Thank you!



Visit us at ironmountain.com/publicsector



Follow us on LinkedIn
@IronMountainGovernmentSolutions



IRON
MOUNTAIN®

GOVERNMENT SOLUTIONS