AI Pilot Results

US Department of State

June 24, 2024

Bureau of Administration

Information Programs & Services

Who We Are

A/GIS/IPS is the principal advisor on all matters concerning the management of information as a critical resource specifically relating to records life-cycle management, public and need-to-know access to information.

Our Mission

Develop, implement and evaluate information and records-related programs on behalf of the Secretary of State:

-FOIA and Other Document Production Activities
-Classification and Declassification Management
-Federal Records Management
-Ralph J. Bunche Library and Diplomatic Research Services

Who We Support

Employees across every bureau and in over 200 missions, from working-level to the Secretary of State.

The U.S. Interagency and Global Partners.

Research Communities and the Public.

AI For Declassification

The Challenge

Executive Order 13526 requires that classified records of permanent historical value are automatically declassified after 25 years, unless a review determines an exemption.

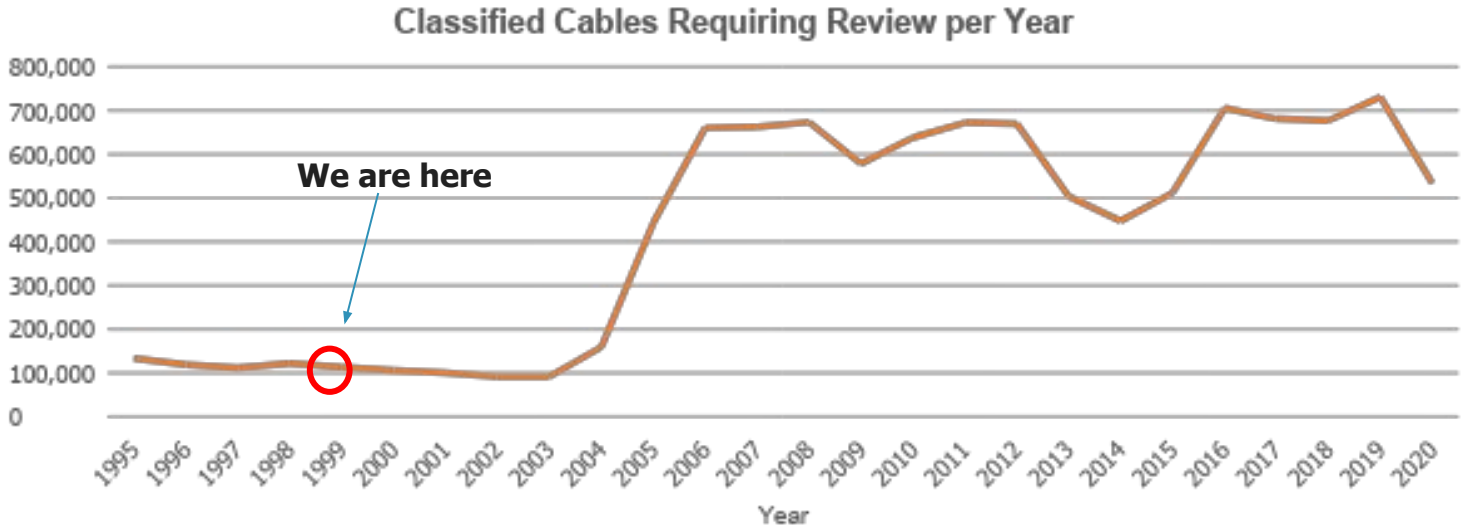
BUSINESS CHALLENGE

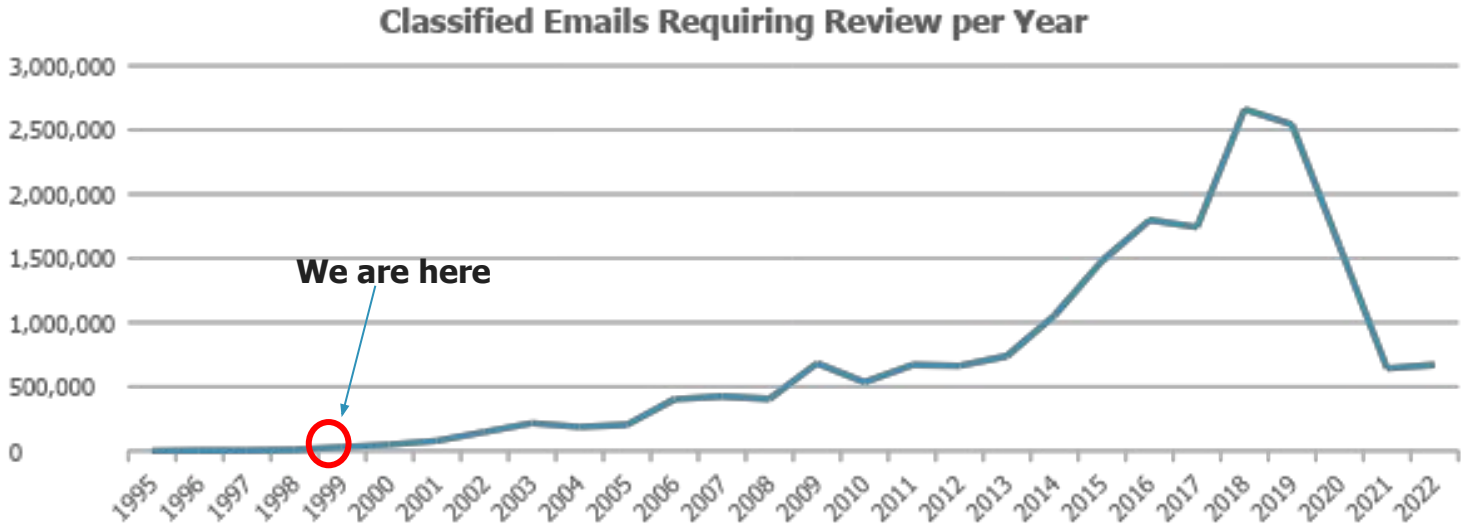
-**Cables** are the **authoritative** reporting by U.S. diplomatic and consular posts overseas and are some of the most historically significant records produced by State.

-The volume of cables requiring review is increasing, rendering **manual/human review unsustainable.**

-Similar problems will soon exist **for classified emails** and other **electronic record types**.

-Transfer of records from State to NARA takes additional time which delays **public access**.





Note: Axes are ***not*** same scale!

The Approach

Started small with a 3-month, limited scope pilot

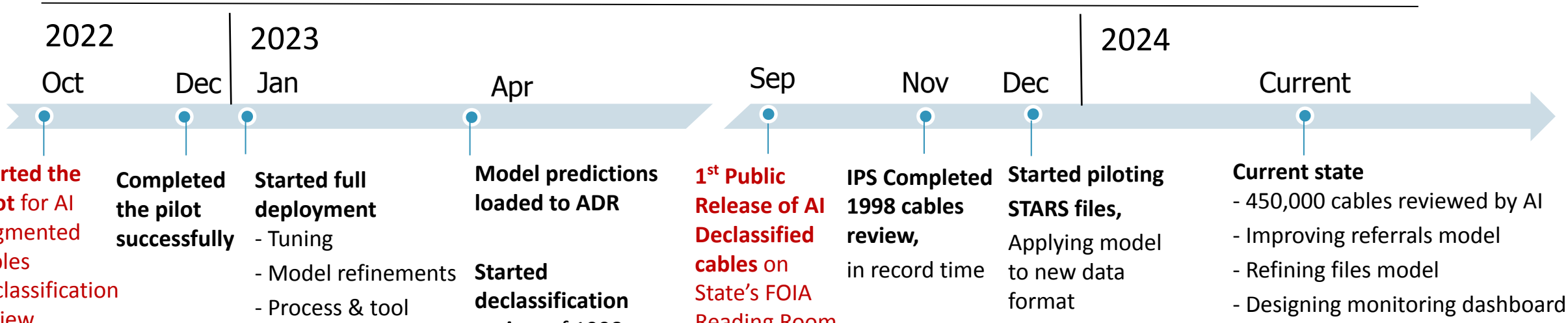
- Chose one electronic record type: cables
  - Cables are uniformly structured, readily available in eRecords
- Used 1995-1997 cables, already human reviewed (labeled data)

Used past decisions by human reviewers

- Trained ML models on past decisions by human reviewers (whether to "declassify" or "exempt from declassification")

Retained human review, not 100% automation (by design)

- Humans will always be in the loop to:
  - Review/label training data as necessary
  - Perform Quality Control (QC) checks
  - Review cables the model is unsure of
  - Pick up on topic drift over time



Results

Expanded public access to records of permanent historical value through unifying automatic declassification review with proactive disclosure under the FOIA.

Reduced volume for humans to review (by 80%) freeing them up to focus on other initiatives.

Reduced annual declassification review cycles.

Adaptable process to handle upcoming explosive growth in classified permanent records volumes.

More consistent declassification decisions.

Enhanced policies and training.

Approach is adaptable to other record types and potentially for other purposes.

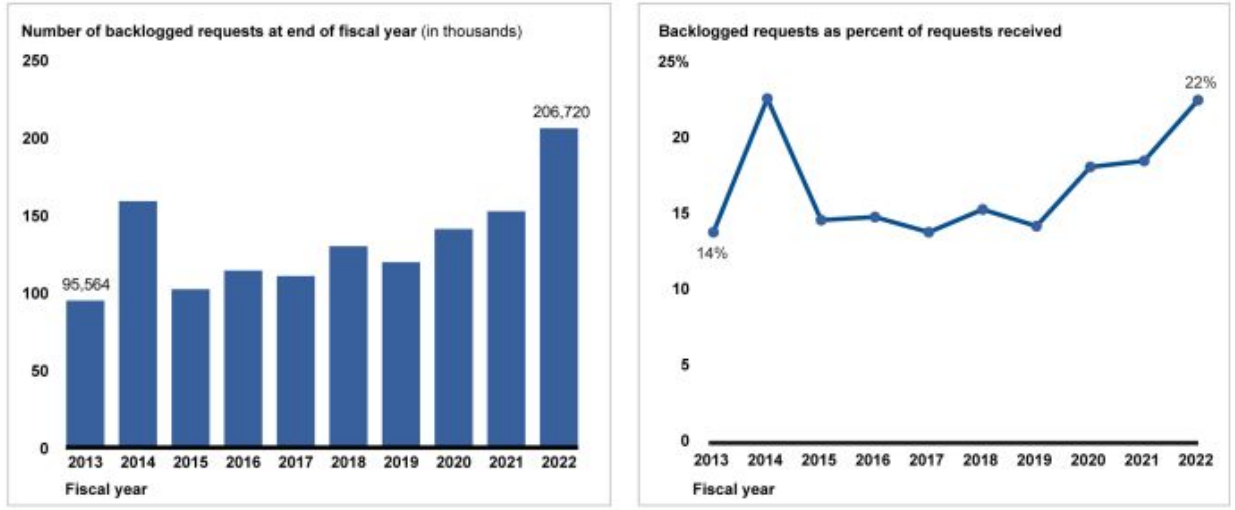
Click HERE for access to the proactively disclosed cables or visit FOIA.State.gov and search case number: S-2023-00002.
Additional cables will be released monthly.

AI For FOIA

The Challenge

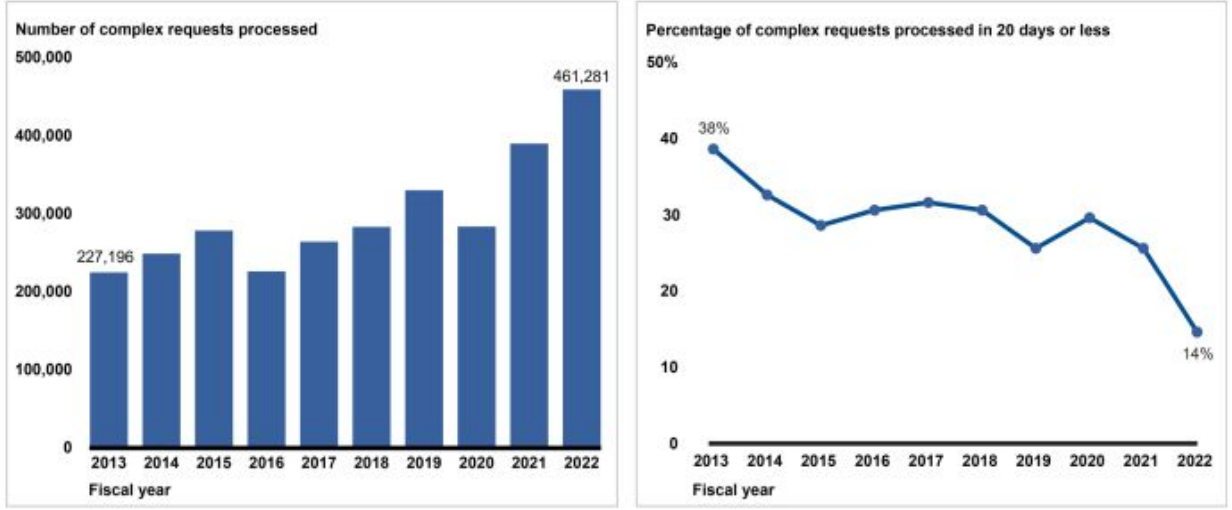
FOIA Business Challenge	State Department FOIA Snapshot
Government-wide request backlog has risen over the past decade. Requests are increasing in complexity. Silos between tools, data and teams lead to inefficiencies, delays, and duplication of efforts.	$65 million in annual processing costs 267 *Full-time* FOIA staff >20,000 pending FOIA cases >15,000 new requests received annually 190 days on average to process requests

Government-wide Backlogged Requests FY2013-2022



Source: GAO analysis of agency-reported Freedom of Information Act (FOIA) data on FOIA.gov. | GAO-24-106535

Government-wide Complex Requests FY2013-2022



Source: GAO analysis of agency-reported Freedom of Information Act (FOIA) data on FOIA.gov. | GAO-24-106535

The Approach

Started small with a 2-phase limited scope pilot

Phase 1 - Assessing the feasibility and identifying AI/ML use cases in the existing FOIA processes and infrastructure.

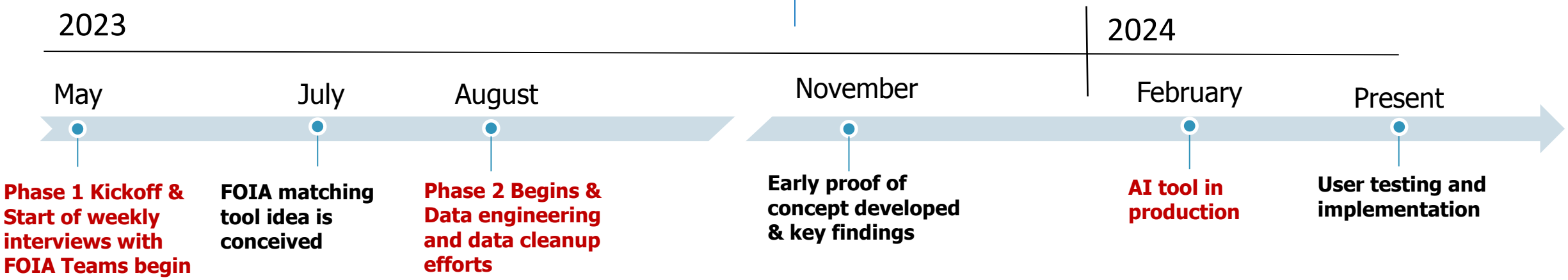
Phase 2 - Piloting the desired models – experiments, assessment and pilot of Proof Of Concept (POC) AI/ML models, based on results of Phase 1 recommendations.

Project Characteristics

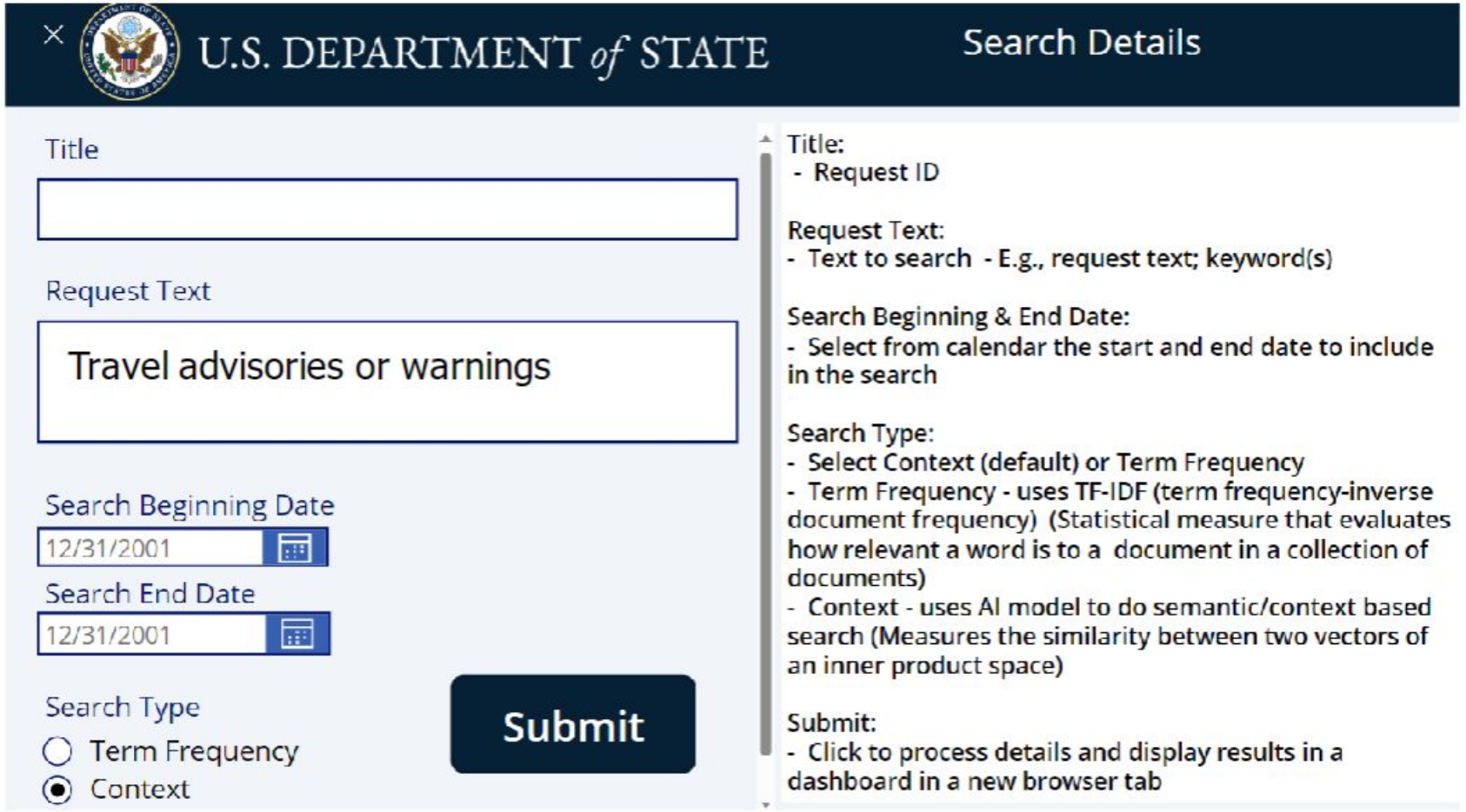
Data Preparation - FOIA data at State resides on 3 different systems.

Infrastructure - Must function on the UNCLASSIFIED network.

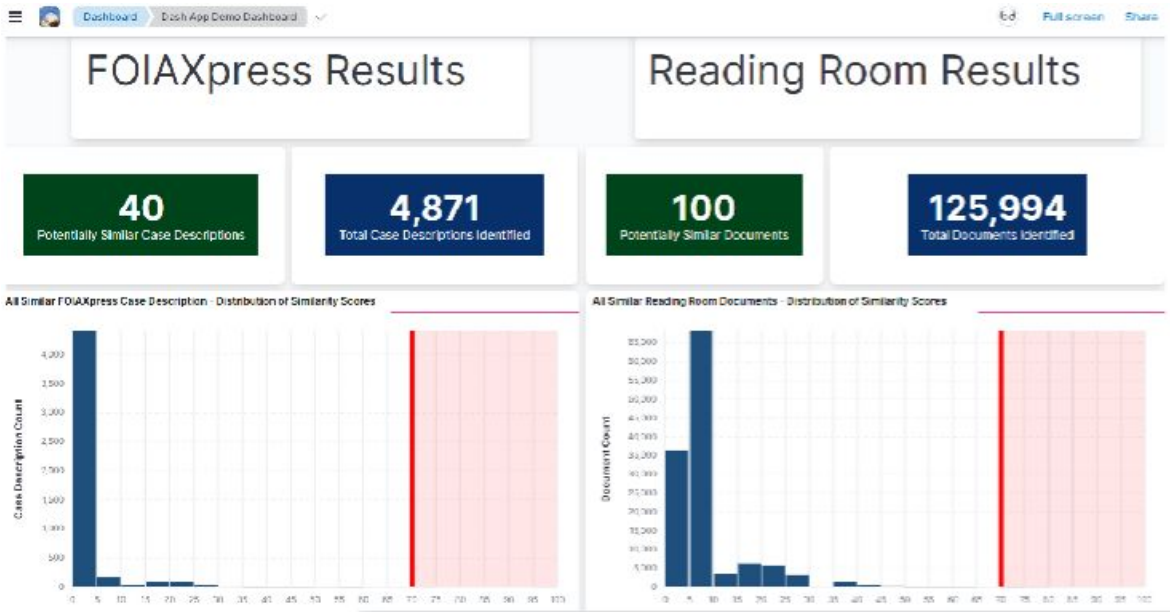
Advanced Analytics – Must be capable of understanding natural language and perform a wide-range of tasks.

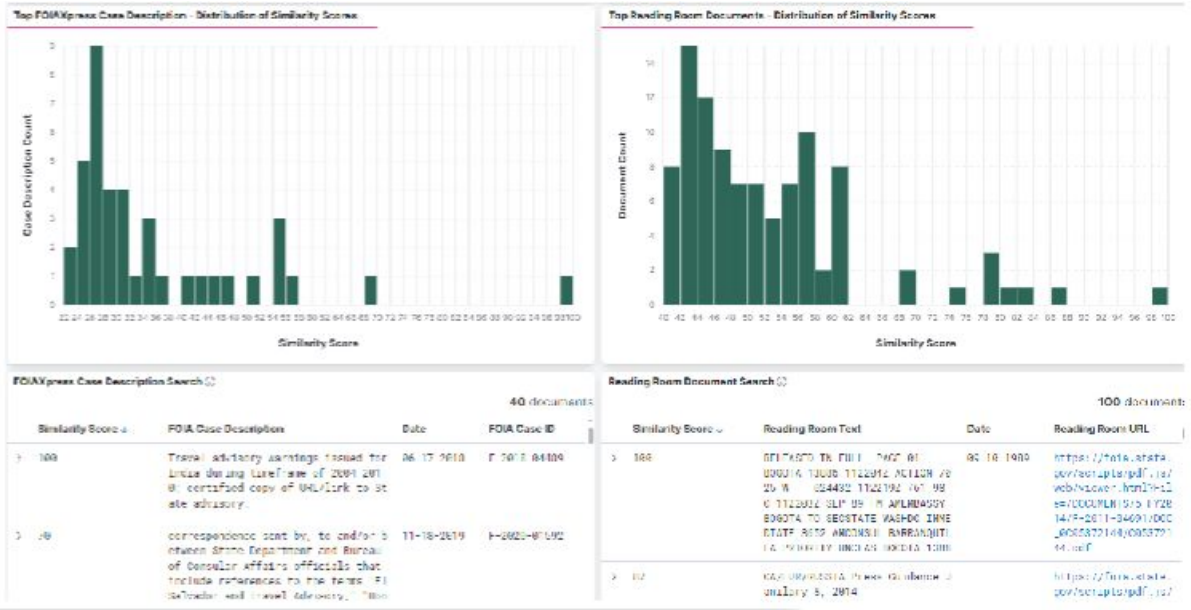


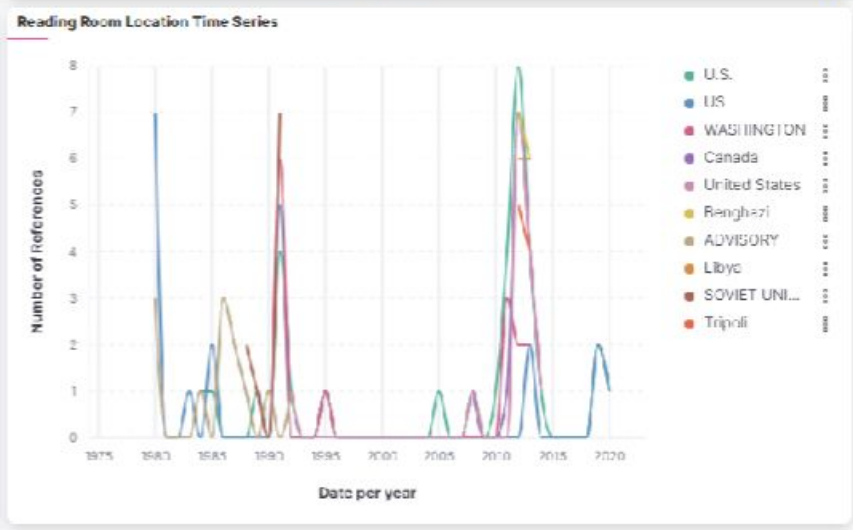
Results– 360 FOIA Matching Tool



Results– 360 FOIA Matching Tool







Lessons Learned

Lessons Learned

Start Small

- Focus on one fundamental challenge.
- Define clear objectives.
- Set technical and business metrics.
- Fail fast and learn.

Assemble the Right Team

- Seek champions within and outside of your organization.
- Secure executive sponsorship.
- Be prepared for a learning curve.

Plan Ahead

- AI data readiness is critical.
- Plan for long term impacts, scalability and maintenance.
- Consider impacts of model retraining, tuning and human quality control to combat data drift.
- Be prepared to explain how the AI arrived at a result.

Thank you

Timothy Kootz
KootzTJ@State.gov
U.S. Department of State