



# **Automating Declassification and FOIA Requests: When AI Leads to Increased Government Transparency**

**US Department of State**  
**A/GIS, M/SS/CFA, IRM/MSO**

**March 7, 2024**



# Declassification Review Program at A/GIS

**Eric Stein**, Deputy Assistant Secretary | Global Information Services (A/GIS)

## Information Programs and Services (A/GIS/IPS)

**Director: Timothy J. Kootz**

- **Administers the Department's records management, privacy, classification, declassification, and public access programs.**
- Provides a range of services to both members of the Department and the public, in accordance with the Freedom of Information Act, the Privacy Act of 1974, the E-Government Act of 2002, Executive Order 13526 (Classified National Security Information) and related program legislation.

## Systematic Review Program (A/GIS/IPS/SRP)

**Division Chief: Jeffrey Charleston**

- **Manages the Department's information classification and declassification programs** under Executive Order 13526 and associated regulations, statutes, and international agreements.

## Key Partners



- **Center for Analytics (M/SS/CfA):** All things data, analytics and AI.
- **Messaging System Office (IRM/OPS/MSO):** Infrastructure and systems for data/records.
- **Other Federal Agencies:** Collaboration and coordination.



# M | Strategy & Solutions

## Center for Analytics



*D-MR and Chief Data and AI Officer Dr. Matthew Gravis  
sign the first-ever Enterprise Data Strategy (EDS)*

## Who We Are

M/SS/CfA is the Department of State's **enterprise data management and analytics capability**.

Led by the Chief Data and AI Officer, we **transform data into bold insights** to help make better management and foreign policy decisions.

## Our Mission

**Expand data access and grow analytic expertise across the Department globally**, enabled through our data and technology platform called Data.State.

## Who We Support

We empower employees across every bureau and in over 200 missions, from working-level to the Secretary.

“ We also want to go much further in using technology, innovation, and **data** to solve foreign policy challenges. We unveiled the State Department's **first-ever enterprise data strategy** last month... The Department has vast and diverse data sets, but we haven't done a good enough job making data available to you in a timely and useful way, to help you make mission or management decisions more effectively. **We're changing that.** ”

- Secretary of State Antony Blinken



# Enterprise AI Strategy



## EAIS VISION

The Department of State will responsibly and securely harness the full capabilities of trustworthy artificial intelligence to advance United States diplomacy and shape the future of statecraft.

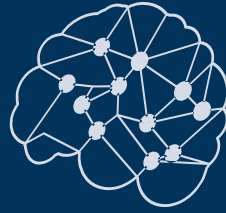
## EAIS GOALS & OBJECTIVES

1. Leverage Secure AI Infrastructure
2. Foster a Culture that Embraces AI Technology
3. Ensure AI is Applied Responsibly
4. Innovate

## EAIS DEVELOPMENT

Guided by the Office of Management Strategy & Solutions' Center for Analytics, the Fiscal Year 2024 through Fiscal Year 2025 EAIS is the product of the State Department's AI leaders and policy experts from over 25 bureaus and offices across the enterprise

# Value of Responsible AI



Decision Advantage and Operational Efficiency

## What Have We Done?

- ✓ Hired a Responsible AI Official (RAIO)
- ✓ Launched an AI Steering Committee
- ✓ Published FAM AI Policy

## Overarching Federal Guidance

AI Executive Order

National Security Memorandum (*pending*)

NIST AI Risk Framework

OMB Memorandum

## Enterprise AI Strategy: Empowering Diplomacy Through Responsible AI

### Vision

The Department of State will responsibly and securely harness the full capabilities of trustworthy AI

### Goals

- 1 Leverage AI Infrastructure
- 2 Foster an AI Culture
- 3 Ensure AI is Applied Responsibly
- 4 Innovate



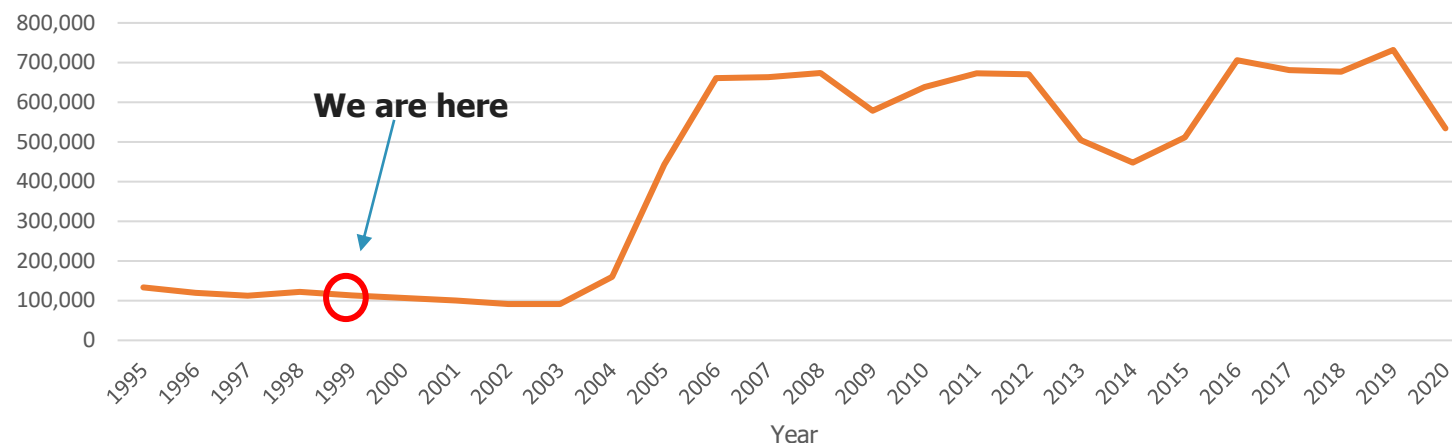
# AI/ML Project Approach and Model Details



# Mandated 25-Year Declassification Review

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

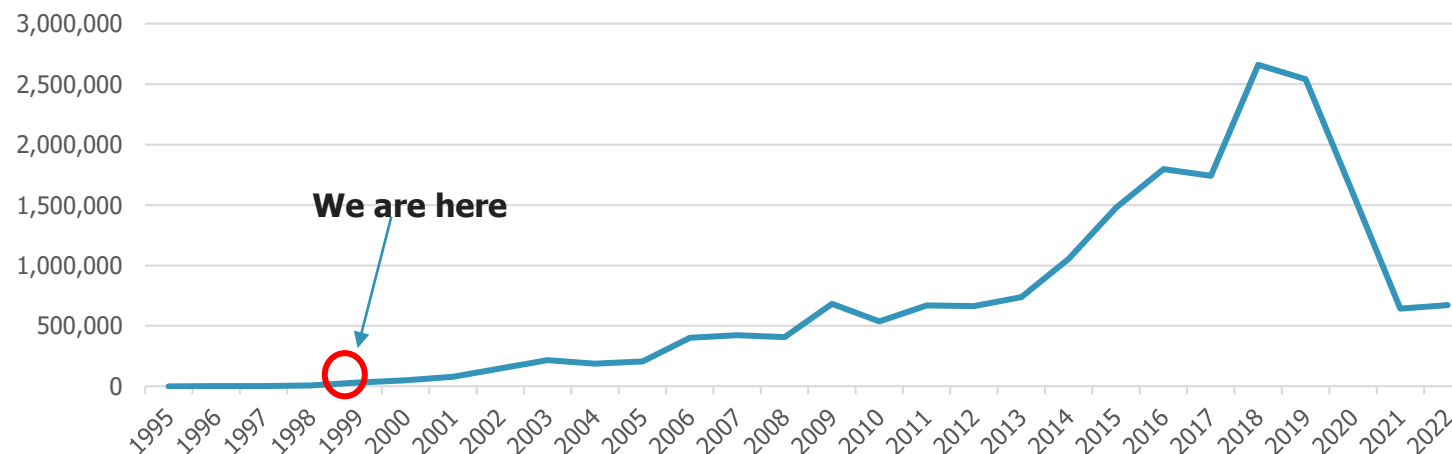
Classified Cables Requiring Review per Year



## BUSINESS PROBLEM

- A/GIS already spends significant labor on the 25-year ADR of classified cables
- Inability to review all cables by a year's end poses a **national security risk** to the Department
- Volume of cables for review is increasing, rendering **manual review unsustainable**
- Similar problems exist for classified emails and other electronic record types

Classified Emails Requiring Review per Year







# Project and Solution Approach

## Started small with a 3-month, limited scope pilot

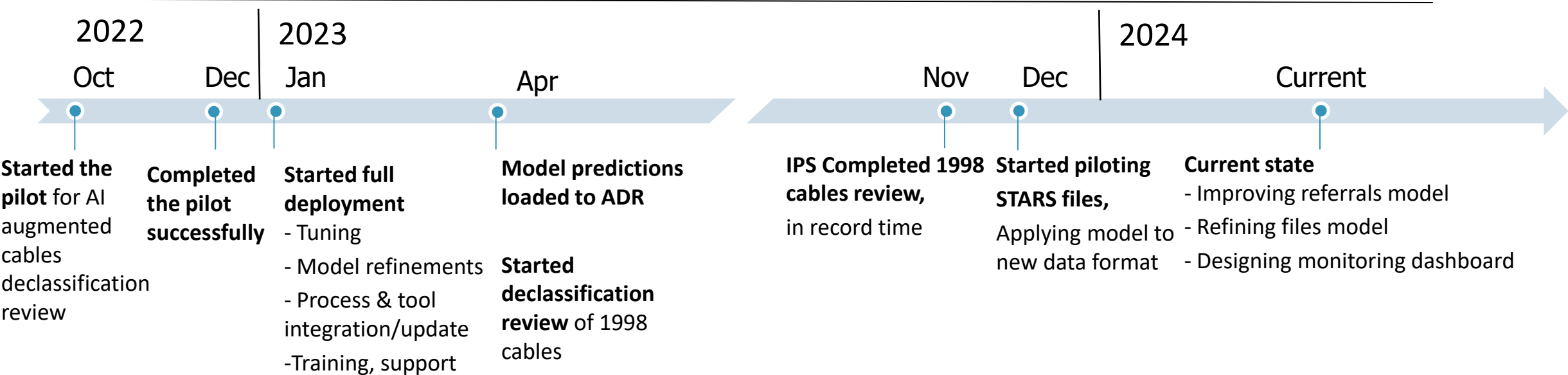
- Chose one electronic record type: cables
  - Cables are uniformly structured, readily available in eRecords
- Used 1995-1997 cables, already human reviewed (labeled data)

## 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







# Declass Model Overview

## Supervised ML

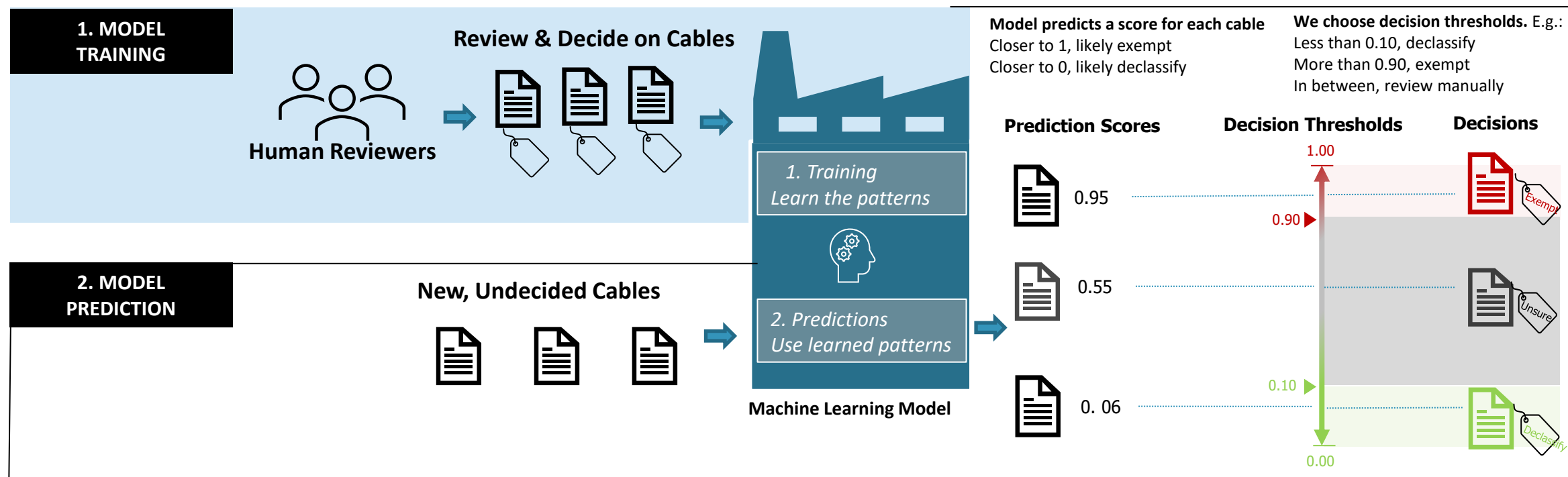
- Used three years of “labeled” cables as training data
- Trained the ML model to learn patterns in the training data
- On new cables, model predicts scores representing declassification or exemption

Supervised = We train it

Label = Human decision to declassify or exempt

Feature = word, phrase, metadata

## AI/ML Model for Cable Declassification: Training and Predictions





# Integration into the Current Manual Review Process

## Model classifies cables into three groups:

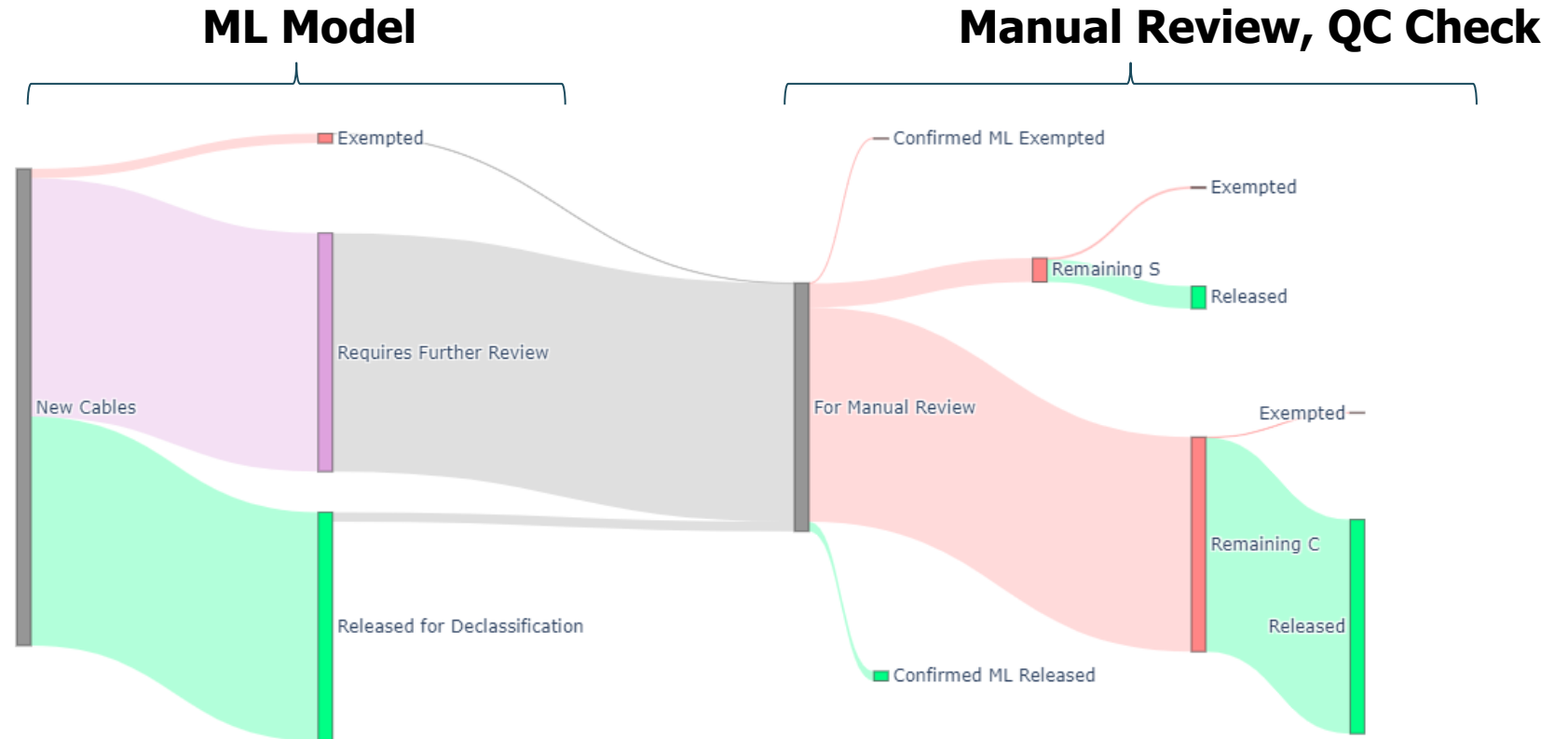
- Confidently Exempt
- Confidently Declassify
- Requires manual review

## Reviewers:

- QC a random sample of Exempt and Declassify groups of cables
- Manually review cables in the "Manual Review" group

## Iterate (retrain, re-predict) based on human QC feedback loop to:

- Lower error rate and deliver more confident decisions
- Reduce volume of cables requiring Manual Review



*Slide mentions only Declassify/Exempt model, but Exemption Reasons and Referral Agency models work similarly*



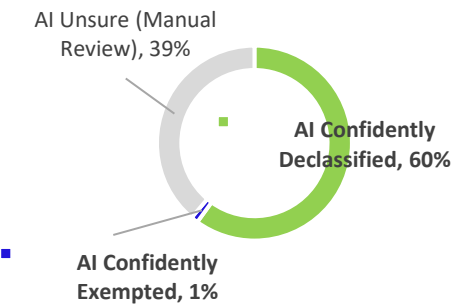
# Deployment - AI Augmented Declassification Review of 1998 Cables

Total cables due for review: 121,536

## 1. INITIAL MODEL PREDICTIONS

- AI model confidently predicted declassification or exemptions for 59% cables.
- 41% were left for manual review.
- A 2.5% random sample of predicted cables were flagged for human Quality Control (QC) check.

### Predictions Breakdown



## 2. FIRST HUMAN QC CHECK

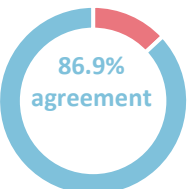
- First QC check by human reviewers showed that the observed accuracy of AI predictions were within acceptable limit and better than projected.

### AI-Declassified Cables Human QC Check Result



Expected agreement: 99.3%  
QC cables count: 1,587

### AI-Exempted Cable Human QC Check Result

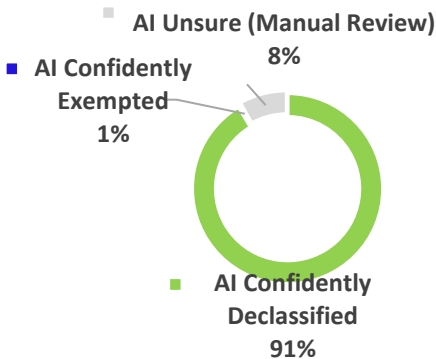


Expected agreement: 81.4%  
QC cables count: 1,427

## 3. RETRAINING, RE-PREDICTIONS

- After retraining the model, it predicted more confidently.
- It confidently predicted on 92% of the remaining (unreviewed) cables
- Vastly reduced number of cables requiring manual review

### Re-predictions Breakdown



## 4. SECOND HUMAN QC CHECK

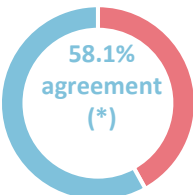
- 1998 cables required human review on only 20% of cables.
- Model predictions were used for the rest, with the overall error rate expected to be approximately 1%.
- DOS completed annual declassification review of cables in record time and within the year.

### AI-Declassified Cables Second QC Check Result



Expected agreement: 99.0%  
QC cables count: 2,037

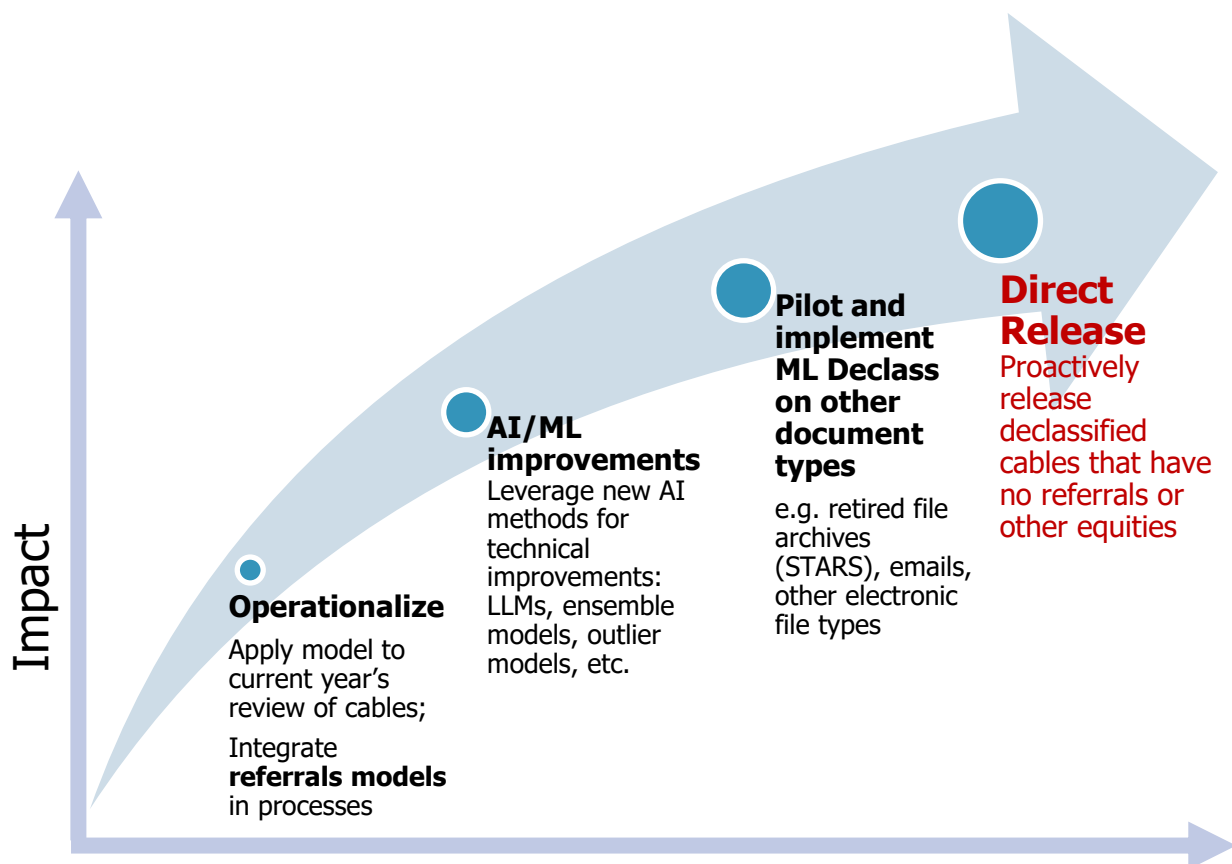
### AI-Exempted Cables Second QC Check Result



Expected agreement: 85.0%  
QC cables count: 405  
(\* ) Low performance resolved after reinforcing declassification policy with the reviewers



# ML for Cables Declassification: Current and Upcoming Goals



U.S. DEPARTMENT of STATE

## Current and Upcoming Goals



**Agency Collaboration:** The Department would like to begin collaborating on improving referrals.



**FOIA:** We have begun a similar FOIA AI project to accelerate responses to FOIA requests using AI/NLP to reduce duplication of effort across cases and silos.



**Emails:** Millions of emails requiring review in the coming years. Have a scalable process and tool in place



**Special Projects:** Ad-hoc requirements for rapid decisions about highly visible issues.



**Operationalize Direct Release:** Establish process to publish *declassified cables with no PII or other equities* to DOS' FOIA Reading Room

- Final checkgate review; remove classification markings, etc.
- Release incrementally in manageably sized tranches.



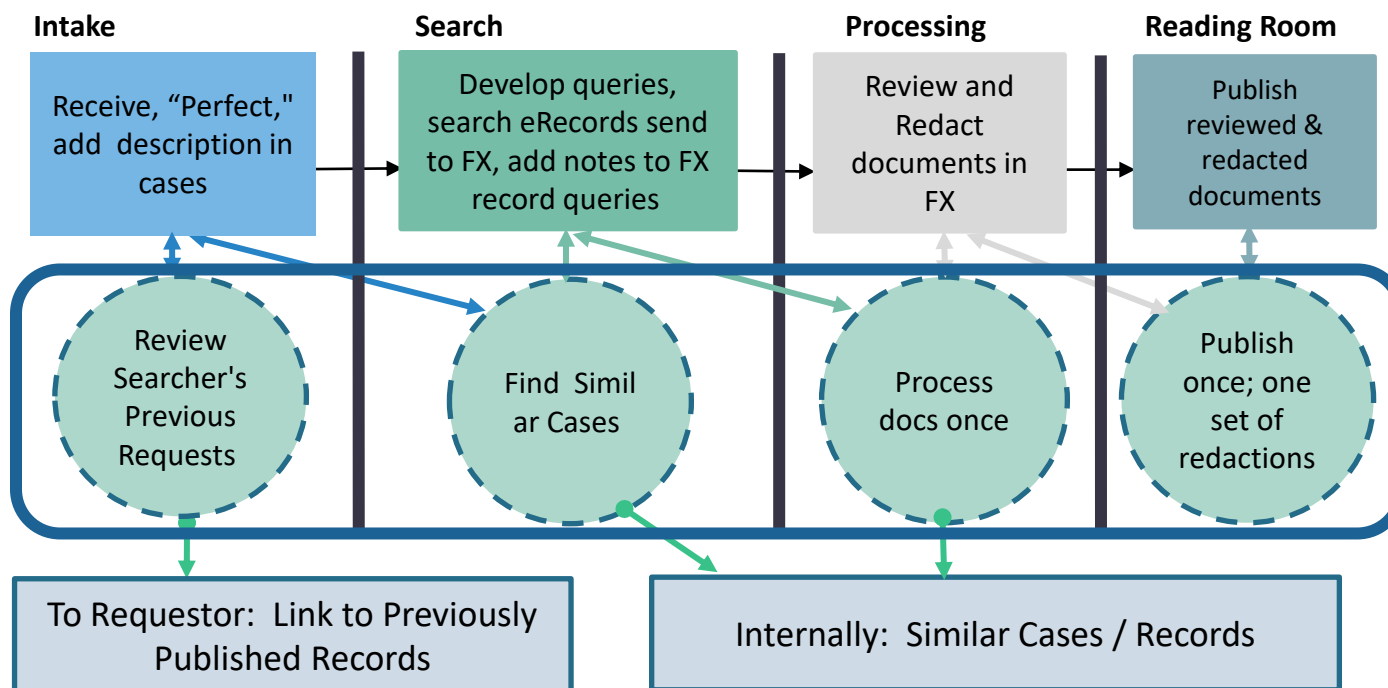
# FOIA AI: Background

## FOIA Business Challenge

- Address increase in costs of FOIA processing and increasing backlog.
- Efficiently respond to increasing FOIA requests and reduce the burden that FOIA places on bureaus.
- Remove siloes between tools and teams that lead to inefficiencies, delays, and duplication of efforts.

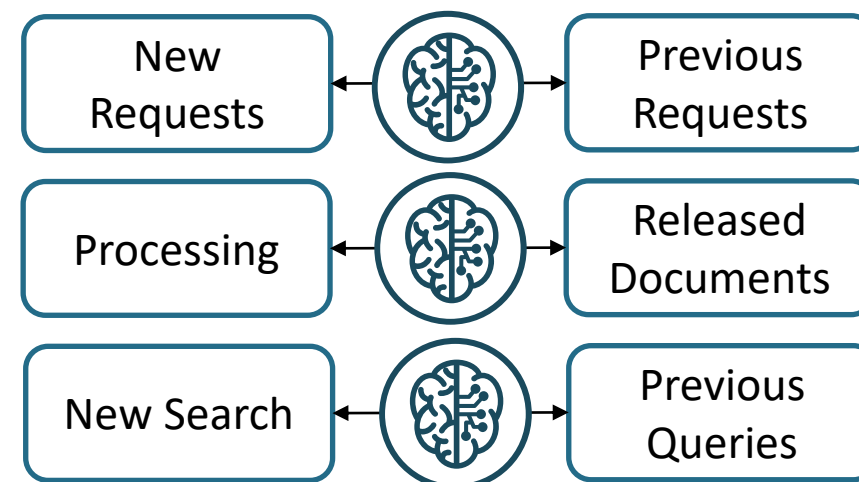
<b>\$65 million</b>	FOIA Annual Costs
<b>~20,000</b>	Pending Cases
<b>~15,000</b>	New Requests
<b>267 personnel</b>	FTE

## Our Mandate: Areas of Focus for Leveraging AI to Improve Customer Experience and Review Process Efficiency



## AI Solution

### Matching AI Solution for Various FOIA Steps





# Unique Challenges with AI/ML and Solutions



## Quality Concerns

- Machine learning model may not perform as well as expected
- Model may have greater error rate than expected



- Engage process; quality control check by human reviewers of a random sample of ML results



## Data Drift

- Ever-changing political & historical themes deteriorate model performance
- Left unchecked, the model will remain rigid & lose predictive power



- Frequent model retraining, tuning and human quality control to combat historical data drift



## Imbalanced Data

- Few exemptions (<5%) compared to declassifications
- Most cables don't get referred to given agency
- Model doesn't have enough examples of minority classes to train on



- Over/under sampling, outlier models to identify outliers



## Model Explainability

- Model outputs decisions & confidence with little explanation
- Model makes pass/fail decision on entire cables, not subsections



- More detailed feature importance plus analysis of cable sections improve buy-in & transparency



# Process Improvements: Lessons Learned

## Data Management



**Data quality and accessibility are critical to any AI effort;** the Department of State eRecords platform empowered the team to train, test, and deploy with high quality data

**Continuously improve** tools and processes with AI/ML features

## Start Small



**Start small, with a pilot.** The experimental approach with well defined **performance metrics** helped the team measure improvement with each iteration.

One **well-suited document type** with manageable volume has helped the team **scale their approach**

## Process Transformation



**Consider** early how AI/ML **transformation** will be incorporated into **process and tools**

**Explore concurrent improvements** to process and tools, even if not related to ML





# Q&A

**Dr. Gio Altamirano Rayo - [altamiranorayog@state.gov](mailto:altamiranorayog@state.gov)**

**Timothy Kootz - [kootztj@state.gov](mailto:kootztj@state.gov)**

**Dr. Sam Stehle – [stehlesk@state.gov](mailto:stehlesk@state.gov)**

**Rabi Tripathi – [tripathir@state.gov](mailto:tripathir@state.gov)**

