Using Machine Learning for Document Reviews and Searches

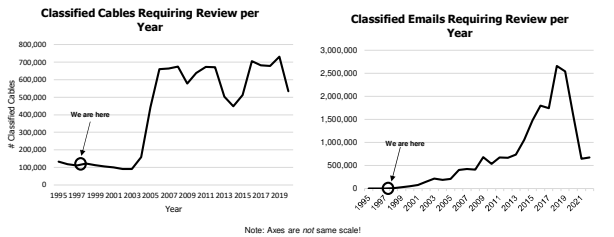
☆☆☆

Eric F. Stein
Global Information Services (A/GIS)
June 29, 2023

1

☆☆☆

Future Growth of Volume of Records to Review



Note: Axes are *not* same scale!

2

☆☆☆

Overview

- Overview of Data and Artificial Intelligence (AI) – Role of Records Management
- 3 State Department Examples, Lessons Learned, Improvements Sought
  - Machine Learning Declassification Review Pilot
  - Freedom of Information Act (FOIA) Customer Experience (CX) Pilot
  - FOIA Search Pilot
- Lessons Learned Themes
- Discussion

3

☆☆☆

Overview of Data and Artificial Intelligence (AI) – Role of Records Management

- State Department Policies are in the Foreign Affairs Manual or FAM available at https://fam.state.gov/.
- The FAM has definitions of key terms like data, artificial intelligence (AI), and record.
- In December 2022, the Department issued its Data Policy - https://fam.state.gov/Volumes/Details/20FAM
- In April 2023, the Department released it AI Policy - https://fam.state.gov/FAM/20FAM/20FAM020101.html
- The Department's Chief Data Officer (CDO) is responsible for Data and AI. The CDO leads the Department's Center for Analytics or CfA in the Office of Management Strategy and Solutions or M/SS.

4

☆☆☆

M/SS Center for Analytics (CfA)

cfa

D-MR signs the first-ever Enterprise Data Strategy (EDS) with CDO Garvias

Who We Are

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

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

CfA composed both the Department's Enterprise Data Strategy and the Department's new AI Policy: 20 FAM 201, AI Policies & Procedures

Who We Support

We empower employees across every bureau and 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 Antony Blinken*

5

☆☆☆

Data and Artificial Intelligence (AI) Defined (20 FAM 101.1-3 DEFINITIONS)

Data - Recorded information, regardless of form or the media on which it is recorded.

20 FAM also defines over 30 other terms associated with data.

Artificial Intelligence - The federal government has defined Artificial Intelligence per section 238(g) of the National Defense Authorization Act for Fiscal Year 2019 (P.L. 115-232): the term "artificial intelligence" includes the following:

- (1) Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.
- (2) An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.
- (3) An artificial system designed to think or act like a human, including cognitive architectures and neural networks.
- (4) A set of techniques, including machine learning, that is designed to approximate a cognitive task.
- (5) An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision making, and acting.

6

☆☆☆

Data and Artificial Intelligence (AI) Defined (From 20 FAM 200)

- **Generative AI**: Generative AI refers to a sub-category of artificial intelligence techniques that generate new data in the form of text, images, or video, using the data they were trained on as a model, and an external input (like a question from a user) as a prompt.
- **AI Use Case:** An AI Use Case is any Department application or use of AI described in either of the following:
  - (1) AI designed, developed, acquired, or used specifically to advance the execution of the Department's missions, enhance decision making, or provide the public with a specified benefit.
  - (2) Both existing and new uses of AI; both standalone AI and AI embedded within other systems or applications; AI developed both by the Department or by third parties on behalf of the Department for the fulfillment of specific Department missions, including relevant data inputs used to train AI and outputs used in support of decision making; and the Department's procurement of AI applications.
- **AI service**: An AI application that consists of an AI system/tool/app provided by a third party (e.g., an outside service provider).
- **Discriminative AI**

7

☆☆☆

Overview of Data and Artificial Intelligence (AI) – Role of Records Management

- Records and Lifecycle Management is the core of everything.
  - **Records**: Includes all recorded information, regardless of form or characteristics, made or received by a Federal agency under Federal law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, operations, or other activities of the United States Government or because of the informational value of data in them. https://fam.state.gov/FAM/05FAM/05FAM0410.html#M415
- Office of Management and Budget (OMB) and National Archives and Records Administration (NARA) Mandates – 2016, 2019
  - M-19-21 Transition to Federal Records
- eRecords Archive
  - Pivotal to three examples about to cover
  - Manages Department's state.gov email and permanent electronic records
    - Receives over 2 million unique records daily; currently has over 3 billion records

8

☆☆☆

3 State Department Examples and Lessons Learned

- Example 1 – Machine Learning Declassification Review Pilot
  - October 2022 to January 2023
- Example 2 – Freedom of Information Act (FOIA) Customer Experience (CX) Pilot
  - Putting the AI in FOIA
  - June 2023 – February 2024
- Example 3 – FOIA Search Pilot
  - Putting the AI in FOIA
  - June 2023 – February 2024

9

☆☆☆

Example 1 – Machine Learning Declassification Review Pilot

- Pilot occurred from October 2022 to January 2023
- Partnership for Public Service Course – https://ourpublicservice.org/course/ai-federal-leadership-program/
  - October 2021 to May 2022

Artificial intelligence has the potential to improve how government works—more so than any other recent technological innovation. From increasing efficiency to finding data insights that enhance the customer experience, AI is an invaluable tool for federal leaders to serve the public and transform their agencies. But to capitalize on these benefits, leaders must understand AI fundamentals and how to use AI effectively.

Through a partnership with Microsoft and Google, we are creating a cohort of senior leaders across government who are prepared to guide their agencies' AI strategy. This program is designed to:

- Educate agency decision-makers on the opportunities around AI.
- Highlight best practices for how to make the case for and develop AI solutions.
- Prepare leaders to incorporate AI technology into their strategies and equip their workforce.
- The Partnership has extensive experience delivering leadership development programs that support federal employees at all levels. **This program is offered to select senior executives and GS-15s at no cost to federal agencies.**

10

☆☆☆

Current Review Process: eRecords Declassification Module

The **eRecords** archive system was built for Dept. records, including **Cables**, emails & files

eRecords used for 25 use-cases in DOS e.g.
- FOIA
- Declassification
- Historian's office

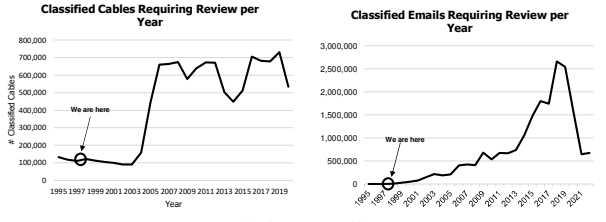
eRecords' declassification module is currently used by reviewers for 25-year declass review of Cables

NOTE: Screenshots are from dev environment, not actual secret data

11

☆☆☆

Future Growth of Volume of Records to Review



Note: Axes are *not* same scale!

12

☆☆☆

Example 1 – Machine Learning Declassification Review Pilot

Project Charter from 2022:

12-1-0 Declassification Review Pilot

Updated: May 13, 2022

Challenge Statement: Electronic records and data sets have grown exponentially over the past two decades complicating mandatory review efforts of this information. Executive Order 13526 on Classified National Security Information establishes an automatic declassification date for classified information after 25 years. Agencies have established declassification review programs for paper and electronic records at the 25 year mark. Current processes and resource levels will not work to review electronic records, including classified emails, created in the early 2000s and beyond, meaning that in 2025 these declassification reviews will not succeed using existing technology and processes.

Project Justification: Pilot the declassification review of electronic records going through the 25 year review process in 2022; this pilot would occur concurrently with the existing process. If successful, the pilot would be implemented to conduct the electronic declassification review in 2023; human resources dedicated to this manual year-long review process now could be redirected to other program areas. The vision is to take 12 months of work, pilot a month long review using machine learning, and achieve declassification proposal results in weeks (zero months) starting in 2023.

Key Personnel/Stakeholders
Sponsor: Department of State – Declassification Program

Team Leader: Declassification Division (Bureau of Administration)
Team Members: Information Resource Management (IRM) - State IT
Stakeholders providing input: Chief Data Officer/Center for Analytics, Office of the Historian (FSI/OH), A, IRM
Stakeholders utilizing input: A, FSI/OH, IRM

13

☆☆☆

Example 1 – Machine Learning Declassification Review Pilot

Project Charter – Continued

SMART Goals

- Test algorithm against the 25 year electronic data that was reviewed in 2021 to determine baseline response from a completed review.
- Run algorithm against the 25 years electronic data currently going through the manual process and compare results with manual review.
- Establish baseline of terms and factors for 25 years declassification review scheduled to occur in 2023.

In Scope: Review of electronic classified (CONFIDENTIAL and SECRET) permanent records that are 25 years old in 2022. Use of the Department's AI and machine learning capabilities in the eRecords archive storing these records, as well as existing data scientist resources at the Department.

Out of Scope: Review of paper classified records (CONFIDENTIAL, SECRET, TOP SECRET) that are 25 years old in 2022.

Key Tasks and Deliverables

Milestone 1:	June 1, 2022 – Hold stakeholder meeting to launch initiative.
Milestone 2:	July 18, 2022 – Convene Team Leader and Members to start work on algorithms.
Milestone 2:	September 1, 2022 – Begin month-long pilot using machine learning capabilities in eRecords database/system.
Milestone 3:	October 30 – Analyze results from pilot and test new algorithms.
End:	Complete pilot by October 31, 2022; determine in November 2022 whether feasible to use machine learning/AI for 2023 electronic declassification review starting in January 2023.

14

☆☆☆

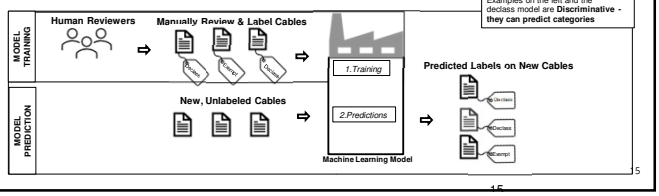
Machine Learning Conceptual Overview

Supervised Machine Learning – Use cables that already have exempt/declass decisions ("labeled" data) to train a ML model. The model learns the patterns and then predicts labels (exemption/declassification) on new, undecided cables.

Successful Applications – Supervised ML models helps solve real-world problems at scale. Examples include:
- Categorizing emails into spam or not spam
- Facial recognition

AI Fun Fact:
ChatGPT and DALL.E are **Generative** models. They can be used to generate new content - text, image etc.

Examples on the left and the declass model are **Discriminative** - **they can predict categories**

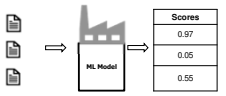


15

☆☆☆

Cable Declassification: Prediction Scores

1. Model outputs a confidence score from 0 to 1



Scores
0.97
0.05
0.55

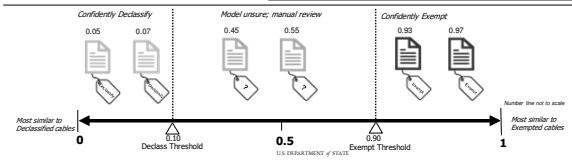
2. We interpret the score by choosing thresholds

Prediction
Exempt
Declassify
Unsure

Thresholds = cutoff scores for decision

Score below Declass Threshold (say 0.05) = predict Declassify
Score above Exempt Threshold (say, 0.90) = predict Exempt

Thresholds adjusted to balance risk (accuracy) and reward (labor saved)



16

☆☆☆

1997 Machine Learning Cable Predictions & Performance

Total Test Set Size = **78,023**
0.01 < no decision < 0.90

1997 Cables	Confidently Declass [48,707]	No Decision	Confidently Exempt [883]
Correct (no conflict)	**48,361**	28,433	**719**
Incorrect (conflict)	**346**		**164**
	Error Rate = **0.71%**		Error Rate = **18.57%**

Risk
Threshold Accuracy = **98.97%**
Baseline* Accuracy = **96.53%**
Error Rate = **1.03%**

"10.3 incorrect cable decisions per 1000"

Reward
Potential Work Saved = **63.56%**
- 49,590 out of 78,023 cable decisions confidently predicted
- 28,433 unpredicted cables require further review

Baseline*: Accuracy refers to accuracy achieved if all cables were to be automatically declassified (3.47% of this set of cables were truly exempt)

17

☆☆☆

1998 Machine Learning Cable Predictions

Total Test Set Size = **121,536**
0.01 < no decision < 0.90

1998 Cables	Confidently Declass [72,891]	No Decision	Confidently Exempt [1,427]
Correct (no conflict)	**72,891**	47,218	**1,427**
Incorrect (conflict)			
	Error Rate = **??%**		Error Rate = **??%**

Risk
Threshold Accuracy = **??%**
Baseline* Accuracy = **??%**
Error Rate = **??%**

"?? incorrect cable decisions per 1000"

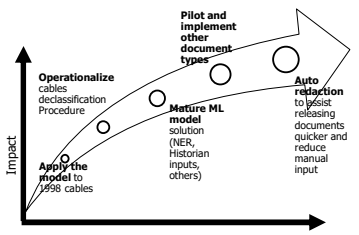
Reward
Potential Work Saved = **61.15%**
- 74,318 out of 121,536 cables confidently predicted
- 47,218 unpredicted cables require further review
- 3,014 out of 76,820 flagged for *quality control* review

Baseline*: Accuracy refers to accuracy achieved if all cables were to be automatically declassified (??% of this set of cables were truly exempt)

18

☆☆☆

ML for Cables Declassification: The Future



NEXT STEPS

Emails: *Millions* of emails will require review in a few years time

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

FOIA: The Department spends immense resources reviewing and redacting FOIA files

19

☆☆☆

Example 1 – Lessons Learned

- Need quality data.
  - Similar data set was available and worked well.
  - Challenges when new data sets/types introduced.
- Partnerships are critical to success – IT Office, Data Office, Program Office.
- Start small – identify a specific project and scope for the project.
- Be open to all results and feedback.
  - Identified ways to approach new and existing challenges
- Patience – Avoid quick judgements or jumping to conclusions (i.e., it works, let's use AI for everything now!).
- Develop quality control checks for results.
- Recurring, sustainable success will require ongoing training of a model using input from humans and technology.

20

☆☆☆

Example 2 – Putting the AI in FOIA

Example 2 – FOIA Customer Experience (CX) - FOIA Assistant
- June 2023 – February 2024

- Anticipate customer/requester needs; Human Centered Design
- FOIA Website Improvements - https://centerforplainlanguage.org/2022-federal-plain-language-report-card/

Customer Experience: Automate data-driven interactions with requesting customers
- Find and direct customer to existing released documents (i.e., as a request is being typed possibly responsive records are identified)
- Automate customer engagement early in request process

21

☆☆☆

Example 3 – Putting the AI in FOIA

Example 3 – Grouping Similar FOIA Cases and Parts of Cases – One Search for Many Cases
- June 2023 – February 2024

- Primary objective – reduce duplication of efforts
- Adequacy and Speed of Search: Ensure that responsive documents are relevant and accelerate the search/review process
- Use Natural Language Processing (NLP) models to transform incoming FOIA request letters into effective queries
- Check the relevance of the queries run compared to the FOIA request
- Pilot identification of classified/sensitive material and/or other FOIA exempt document types in search results
- Pilot identification of triangulated sensitive search results and automate any necessary responses

22

☆☆☆

Lessons Learned Themes To Date

- Managing data and records well is critical to success.
- Understand available data; develop standards now for future use.
- Learn about AI.
  - Recognize what you do and do not know.
  - Do research or reach out to colleagues – Chief FOIA Officer Council Technology Committee is an available resource for federal employees.
- Scope – Start small, explore one question in pilot(s).
- Identify manual processes that could be automated.
- Take (controlled) risks – don't be afraid to fail.
- Consider quality controls such as human reviews before you start.
- Consider bias, privacy, sensitivity of information, legal requirements at the start of your project.
- Discussions about integration of various IT tools using AI and machine learning capabilities.
- Resources are required for pilots and to implement changes.
- Be flexible and open to a variety of results.
- Be open to sharing successes and failures.
- Results in one pilot or project may not be applicable to other programs…but they may.

23

☆☆☆

Discussion

- Contact information – Eric F. Stein, SteinEF@state.gov

24

☆☆☆

Appendix

25

☆☆☆

Example 1 – 2023 Chief FOIA Officer Report

Source - https://foia.state.gov/Learn/Reports/Officer/2023.pdf (pages 19-21)

3. Does your agency currently use any technology to automate record processing? For example, does your agency use machine learning, predictive coding, technology assisted review or similar tools to conduct searches or make redactions? If so, please describe and, if possible, estimate how much time and financial resources are saved since implementing the technology.

26

☆☆☆

Example 1 – Chief FOIA Officer Report

Source - https://foia.state.gov/Learn/Reports/Officer/2023.pdf (pages 19-21)

Yes, the Department's eRecords Archive leverages machine learning to tag emails as "personal" or as "news clippings" when searches are being conducted for responsive records. Being able to eliminate these types of materials, as appropriate, during the initial search reduces the time and effort needed locate responsive agency records and reduces agency response time.

The Department also had a successful pilot from October 2022 through January 2023 using machine learning and technology assisted review in its Declassification Program, separate from FOIA. This pilot, conducted by A/GIS/IPS in partnership with the Department's Bureau of Information Resource Management (IRM) and Center for Analytics and led by the Department's Chief Data Officer, trained a model to conduct declassification reviews of electronic cable records (i.e., communications between Washington and overseas posts such as embassies and consulates) by using past declassification decisions from human review from 1995-1997. The model was trained on human review decisions to identify cable features that are typically indicative of information that is released and that which is exempt from release. The results were reviews that were 97%-99% in agreement with the human reviews. In 2023, the Department plans to leverage this model to complete the 25-year review of cables from 1998. The manual review process takes an entire year; the machine learning review takes 20-30 minutes to assign a declassification decision to every cable. In the 2023 review, over 72,000 cables (63% of the annual total) were assigned confident decisions by the model, requiring only minimal human quality control. The remaining cables will be decided by human review. This process also includes several quality-control steps and reviews of what the technical model says can be declassified and exempt, as well as additional controls to look for highly classified or sensitive information. Leveraging this model will include both technology and human review moving forward, not just one or the other.

The machine learning work for the Department's Declassification Program has been extended for at least the rest of 2023 to explore additional records that could undergo review in addition to Department cables. Additionally, A/GIS/IPS, IRM, and the Center for Analytics are developing a new pilot for FOIA to explore machine learning searches of centralized records and potential initial responses for newly received requests for information that has already been processed in the past by the Department to improve FOIA response times and customer experience.

27