Roadmap for the Adoption of Artificial Intelligence in Diplomatic Security

Indiana University Cybersecurity and Global Policy Program in partnership with the Diplomatic Security Office of Protective Intelligence Investigations

A Diplomacy Lab Project

05/03/21

Table of Contents

Executive Summary

Artificial intelligence is a field of technology focused on creating machine systems capable of completing tasks that typically require human intelligence. When used effectively, AI can analyze data quickly, identify trends in data, and make predictions about future trends. This type of intelligence is sought after by the Department of State and the United States Federal Government for strategic use and allocation of resources.

The Department of State's Bureau of Diplomatic Security prompted the Indiana University Spring Diplomacy Lab with the question of how to utilize artificial intelligence in its security operations. We set out to identify AI's potential for benefitting officers in investigative procedures, to streamline the flow of information crossing intelligence officers' desks and to help officers identify actionable information critical to security operations quickly and efficiently.

Intelligence officers are required to analyze data in various forms and locations, ranging from emails, to word documents, to cables and to SPOT reports. We recognize an opportunity for AI implementation in the DS' SPOT reporting system. To help the DS utilize AI in its SPOT reporting system and then in further security operations, we offer a guiding roadmap.

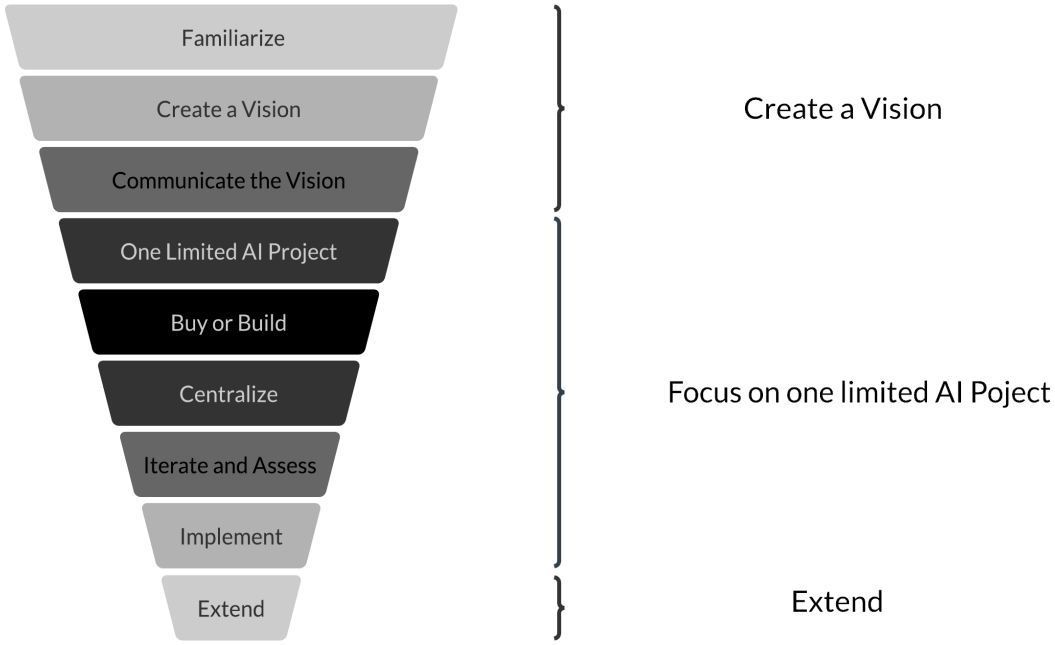
This roadmap is to be used to outline a strategy for the effective implementation of artificial intelligence and machine learning within DS security operations. The roadmap addresses phases from preparation to implementation. Implementing AI for the first, second, third, or hundredth time could require navigating unfamiliar territory. We make three high-level recommendations:

Create a vision: It is vital to begin with a clear vision of what AI is and is not, its intended use within the bureau, and a plan for sharing that vision with a team.

Focus on one limited AI project: Focusing on one limited AI project provides ample opportunity to test, learn from, and build in a controlled setting without setting overwhelming expectations and working with resources at hand.

Extend into further AI initiatives: After creating a clear vision and testing one limited AI project, the bureau will be ready to assess the strengths and weaknesses of the project and launch other initiatives to accomplish different goals.

These three recommendations are laid out in a series of 9 steps:



1. Familiarize

Before launching an AI project, the Bureau of Diplomatic Security must become familiar with artificial intelligence, understanding the technology's capabilities and limitations along with its ethical, structural and security implications.

2. Create a Vision

To fully leverage the potential of AI systems, strategy and planning are essential. A strategy is effective when it has defined and measurable goals, a clear timeline, and coordination among individuals.

3. Communicate the Vision

Successful AI implementation requires employee familiarity with and education in its capabilities. Along with understanding AI, employees must understand how AI helps the department achieve its goals and how the technology is going to be implemented in the workplace.

4. One Limited AI Project

We recommend focusing on one limited AI project to build, test, and learn from in a controlled setting. An example we offer is an AI algorithm that produces an automated SPOT report warning level scoring system. This algorithm will analyze SPOT report text to calculate a predicted warning level. By simply allowing a customized algorithm to read each SPOT report as they are sent in, the system will be able to assign each a 'predicted warning level' that will indicate exactly how threatening the computer thinks each report is.

5. Buy or Build

One of the core decisions in launching an AI project is how to effectively build and maintain an AI tool. The Bureau will decide to either purchase the tool or build it with resources internal to the Bureau. This decision depends on a number of organizational considerations and constraints. Before one can make a decision on whether to buy or build, they must determine the focus of the tool and the goals of the organization.

6. Centralize

In preparing for a new AI project, the bureau needs to centralize its data, or prepare data to be analyzed by the AI system. Datasets matter because they are used by machine learning algorithms to continuously improve themselves over time. It is data that AI assesses and it is data that trains an AI system to do so. In centralizing, the Bureau can prepare its data inputs by measuring them in terms of completeness, accessibility, quality, connectivity, quantity, and validity.

7. Iterate and Assess

The AI project will also require an iterative design process and a plan for auditing. Iterative design, meant for flexibility and adaptability, is ideal for evaluating the algorithms used in AI tools. The algorithm used in the AI tool is the most important component to evaluate as it is what tells the AI computing system what to do. A plan for assessment and auditing will ensure that the algorithm being utilized is resulting in an AI tool that functions in accordance with the intentions of the project.

8. Implement

The bureau must approach implementation with a plan for determining the AI tool's readiness and for reassessing the tool as time passess. The tool must be capable of working at the equivalent of a "human or better" before being implemented into the real world. After implementing the tool must be reassessed for its effectiveness.

9. Extend

After creating a vision and testing one limited AI project, the bureau will be ready to launch other initiatives to accomplish different goals. The small-scale project developed in previous steps should be viewed as a building block for further AI implementation throughout the team's operations.

1 Familiarize

The Bureau of Diplomatic Security must become familiar with artificial intelligence and ensure employees' familiarity through education. This means understanding what the technology is capable of doing and what it's limitations are. While the DS works to understand AI and its capabilities, there are some major considerations to keep in mind. First, artificial intelligence is a field with many resources for educating and equipping professionals for its use. Such resources will serve to benefit the bureau as its teams prepare for an initial AI project. Second, with AI familiarity comes recognition of the skills needed for its implementation. It will be natural for the Bureau to assess the skills needed to take on an AI project and fill any skill-gaps that exist. Third, artificial intelligence is a field of technology with ethical, structural and security implications that are essential to understand before a project launch rather than after.

Defining Artificial Intelligence

Artificial Intelligence (AI) is a computer system that is able to complete tasks that previously required human intelligence. AI is used to spot the patterns in trends in massive amounts of data, then predict these same trends and values based on all of the previously collected data points. These predictions can be used by employees to make practical decisions with filtered data. This technology is a large and interdisciplinary field that encompasses other fields such as psychology, linguistics, and statistics. It also has numerous applications across various industries such as healthcare, education, marketing, etc.

Employee Familiarity

This broad overview of AI generally leads people to two mentalities: glorifying AI and vilifying it. The glorification of AI leads people to believe it will solve all of the world's problems while the vilification of AI draws attention to potential bias and ethical issues. Both of these conclusions are inaccurate and cause more harm to AI projects than good, which is why it is

important to educate employees on what exactly AI is and what to expect working alongside it.

For the Department to get the most they can from AI, it is important to help agents understand the new technologies better. AI allows computers to complete the more monotonous tasks, giving human employees more time to concentrate on tasks at which they can outperform AI. Agents not only need to be trained on the basics of what AI is, but need to be trained on the benefits of AI and its limitations. People who don't know the "why" or the "how" artificial intelligence works may also not understand the limitations and capabilities of this technology. For example, AI can excel when it is given a large quantity of reliable data and can give answers within parameters that have been defined. However, problems that are novel must be solved by humans. A computer cannot comprehend certain human factors. Acknowledging biases and ethical limitations of AI is important to avoid misuse.

In order to thrive in the growing age of artificial intelligence, AI must be used correctly in the workplace. Successful organizations will recognize opportunities for AI implementation that allow their employees to save time and do their jobs more efficiently. Computers must work alongside human employees, encouraging human specialties of communication and team management that are essential in a strong workplace.

Identifying training resources

According to an MIT article, one of the biggest problems companies face when trying to educate employees about AI and ML tools is the lack of time and sponsorship given during the training process. To fix this, increasing the amount of educational resources could lead to better employee development and create new skills with these tools. For example, Google and Amazon have invested millions of dollars on AI research facilities internationally that have led employees to access more tools and to an increase of necessary training. A way this training would potentially be implemented in diplomatic security would be having agents participate in daily training. This training would consist of learning ways to integrate AI into their own jobs either in technical or non- technical ways using online resources such as Udemy, Udacity, or Coursera. Each of these websites offer various AI courses that range from beginner to advance, allowing agents with any skill set to learn. In 2018, IBM launched its own AI skills academy to all employees. This program was to develop skills in AI, learning how to use AI in their jobs, as well as collaborate with clients to use AI in their businesses. The program had two tracks, which were either technical and non-technical, with four levels in each track ranging from basic to expert. In 2019, over 4,000 employees completed all four levels which has led to their company becoming one of the top tech companies in AI development.

Assessing Skills Needed

Another problem organizations are having when implementing AI is finding the right employees with the necessary skills. This has led to there being a skill gap that's caused

companies to lack qualified AI professionals and organized structure. One consideration to resolve this could be recruiting individuals with business and technical backgrounds and put more emphasis on cross-organizational communication. Another thing would be creating a test or survey on AI applications for all agents, and based on the results, divide them into different skill groups. From there, each group can start training in the necessary areas and advance from there.

PWC developed an AI toolkit of frameworks that focus on creating a new way of thinking about responsible AI, business development, and strategic execution. This framework is based on five major concepts: Governance, interpretability & explainability, Bias & fairness, Robustness & Security, and Ethics & Regulation. Each creates a path that will help companies design and deploy responsible AI applications when training employees. The significant thing about this framework is that the entire business, rather than the technology and development teams must follow and collaborate in order for there to be success. On top of that, in the next 4 years, PWC plans on upskilling all employees on topics such as managing AI, data analytics, autonomous vehicles and more.

Ethical, Structural, and Security Implications

There are three aspects to consider in the familiarize and educate phase; the ethical, structural, and security considerations of AI implementation. Starting off with understanding the ethical aspect of AI, since this technology is becoming more integrated with our society, policies are being created to ensure that the development of AI follows civil and human rights standards. The structural aspect of AI is understanding all of the components that go into it before getting started. For example, this could be the cost of development, types of education/certifications needed, bandwidth, and tools to create maximum efficiency. Regarding the security aspect, this is very important when educating yourself with AI because as with most technology now, cyber threats are becoming more common and targeting AI systems is almost guaranteed. Hackers can exploit technologies in unforeseeable ways, so understanding how to keep these AI tools up to date and secure is important in order to protect them. Each of these considerations are crucial when in the education phase because each has a large role in making sure the tools are used and developed properly and follow the laws of our society.

2 Create a Vision

In order to leverage the full potential of AI systems and tools, it is essential to strategize and create a clear plan. A strategy is effective when it has a clear purpose and identifies a distinct target outcome.

Elements of a Good Strategy

1. Identify goals: In order for a strategy to be effective, it must have a specific solution it is aimed at achieving. Because of the wide variety of AI uses, there are a wide variety of strategies and goals. When the government of Mexico implemented AI, one goal was to "Develop tools for continued education in AI (current and next administration). People who are in the workplace will need to demonstrate that their skills are up to date. There are many providers of training, but quality can be difficult to assess." Based on this goal, they were able to implement a system that maintains a list of recommended courses to uptake employee knowledge. In regards to DS, the purpose of AI primarily relies in effective data management. The purpose of AI implementation is to assist employees in making decisions regarding security quickly and efficiently. It also serves as a way to organize and manage information the Bureau receives.
2. Create a clear timeline : A timeline encourages everyone allowed in the implementation to be held accountable for accomplishing their assigned tasks. It also encourages those involved to think realistically about what can be accomplished in a time frame. AI implementation can take months to implement, in order to make the transition as smooth as possible, goals must be set periodically. For example, it may be advantageous to have employees complete training by a certain week. A bigger goal would be a "go live date" in order to test the new technology and assess what needs adjustments.
3. Define measurable goals : This allows you to break your goals down into measurable elements so you have a clear focus on what aspect of the project is on track and what may need more attention. As mentioned previously, a goal of DS may be to have employees educated on the new technology. It is important to be specific and recognize what aspect of AI they must be fluent in or if they should take a class on the basics of AI. A measurable goal would say 90% of employees have completed a training course on AI by a certain date. This goal ensures that all employees are receiving the same knowledge and are ready to interact with new technology and fully utilize it.
4. Coordinate: There are many entities involved in the implementation of AI. Coordination allows for a united action to achieve a common goal. All sectors must be coordinated in order to gain all the benefits of a new system and have it operate smoothly. Having weekly or bi-weekly meetings is one way to make sure everyone knows upcoming goals that need to be accomplished. It is also a way for employees to address questions and concerns they may have with the new technology.

3 Communicate the Vision

Successful AI implementation requires employee familiarity with and education on AI's capabilities. Along with understanding AI, employees must understand how AI helps the department achieve its goals and how the technology is going to be implemented in the workplace. Familiarity and education combats the over-glorification and villainization of AI,

reducing hostility to the new technologies. Goal orientation generates buy-in. Space for feedback and questions from employees will also provide for transparency, a workplace characteristic ideal for successful AI projects.

AI is capable of collecting and using data to train machines how to act like humans. AI is being used in everyday life (ex: Netflix movie recommendations and voice recognition softwares) as well as in companies and governments (ex: chat bot assistants and transcribing audio). This broad overview of AI generally leads people to two mentalities: glorifying AI and vilifying it. The glorification of AI leads people to believe it will solve all of the world's problems while the vilification of AI makes people think it will steal everyone's jobs. Both of these conclusions are inaccurate and cause more harm to AI projects than good.

The glorification of AI is harmful because it sets an expectation for instant success which leaves little space for the "iterate and assess" cycle that is essential to the growth of AI and its capabilities. If the stakeholders of the project participate in glorifying AI, they will likely lose faith in the project as soon as a mistake is made, and it is inevitable that mistakes will occur. It is important to clearly communicate to the team that mistakes will be made. Communicating this earlier rather than later will help the team build the "iterate and assess" mentality referenced in step seven of the roadmap.

On the other hand, people who believe AI is untrustworthy may be hesitant to support AI pilot projects. According to an article by Karen Bannon, more than 25% of Americans believe that technology like AI is threatening their jobs; however, it is unlikely that that is the case, and it will most likely improve their day-to-day work. AI can take away the "mindless tasks that should be automated" to leave more time to accomplish company goals. For example, Poland started using AI in 2014 to filter unemployment applications into classes that dictated which citizens needed more/less resources to re-enter the job market. Using AI to sort and classify applications left government employees more time to work with the people who needed assistance rather than looking at their applications. This natural language processing technology used in Poland can be implemented in the DS SPOT reporting system, but instead of analyzing levels of unemployment, it will analyze reports to calculate predicted threat levels. Similarly, it will give employees more time to handle threats rather than identifying threats.

Project leaders can avoid the glorification and vilification of AI by educating their team members and employees on the true capabilities of AI. It is important that employees understand AI as a tool to use rather than a system to rely on. The DS can look to other U.S. government agencies (such as the Customs and Border Protection and the Social Security Administration highlighted in Play Four) to exchange information regarding how these agencies communicated goals and familiarized employees with AI.

Generating Buy-In

Artificial Intelligence can assist in achieving DS' vision of being an agile, proactive, and flexible security and law enforcement agency by leveraging digital tools to enhance security

operations and transform the way DS detects threats, analyzes security incidents, identifies trends from standardized reporting, and forecasts staffing needs.

These benefits can be seen through the proposed AI reliant SPOT reporting system. Currently, the SPOT reporting system is reliant on email, so information sharing is not streamlined or efficient. AI systems require structured data systems to operate. This system will allow the SPOT reports provided to be more easily accessible which will in turn allow information to be disseminated faster. AI will also assist in rapidly evaluating this information to better identify, assess, and mitigate threats. Overall,

Using AI to revamp existing systems, such as the SPOT report system, will help reduce human error, assist in time-consuming jobs, and increase overall productivity. While AI comes with many benefits, it also comes with high upfront costs and learning curves that can deter people from implementing the technology into the workplace.

Benefits of AI	Downsides of AI
• Reduction in Human Error • Available 24/7 • Assists with repetitive, time consuming jobs • Increased speed & efficiency	• High upfront costs • Learning curve • Security and ethical concerns

Maintaining Transparency

Besides understanding the benefits of AI, it is important for employees to understand the strategy and plan behind it. AI thrives in environments where constant feedback is present, so it is important to ensure there are mechanisms in place that allow all employee thoughts and concerns to be addressed.

Case studies

Employee Perceptions of Effective AI Principle Adoption: This article showcases employee perceptions of AI and demonstrates how important it is to communicate effectively throughout the process of implementing AI.

Understanding Artificial Intelligence - ethics and safety: This article identifies ethical concerns of AI technologies being adopted by the British Government.

4 One Limited AI Project

Automated SPOT report warning level scoring

We recommend an AI algorithm that produces an automated SPOT report warning level scoring system. This algorithm will analyze SPOT report text to calculate a predicted warning level. In its current state, the organizational process of receiving and sorting through SPOT reports leaves something to be desired. The reports are unsorted, unassessed, and by manually searching through them, the system is left entirely to human error. AI will help with this problem. By simply allowing a customized algorithm to read each SPOT report as they are sent in, the system will be able to assign each a 'predicted warning level' that will indicate exactly how threatening the computer thinks each report is. This rating would then be used to sort the order of reports' importance for a human to check.

How does this help DS with its handling of SPOT reports?

By integrating this algorithm into the bureau's SPOT report submission pipeline, the trained algorithm would report its predicted warning level score for a given SPOT report. That warning level can be used as a point of reference when analyzing which reports to prioritize. This additional point of reference could function as a handle by which to speed up and streamline the SPOT report sifting process.

Where will the algorithm go?

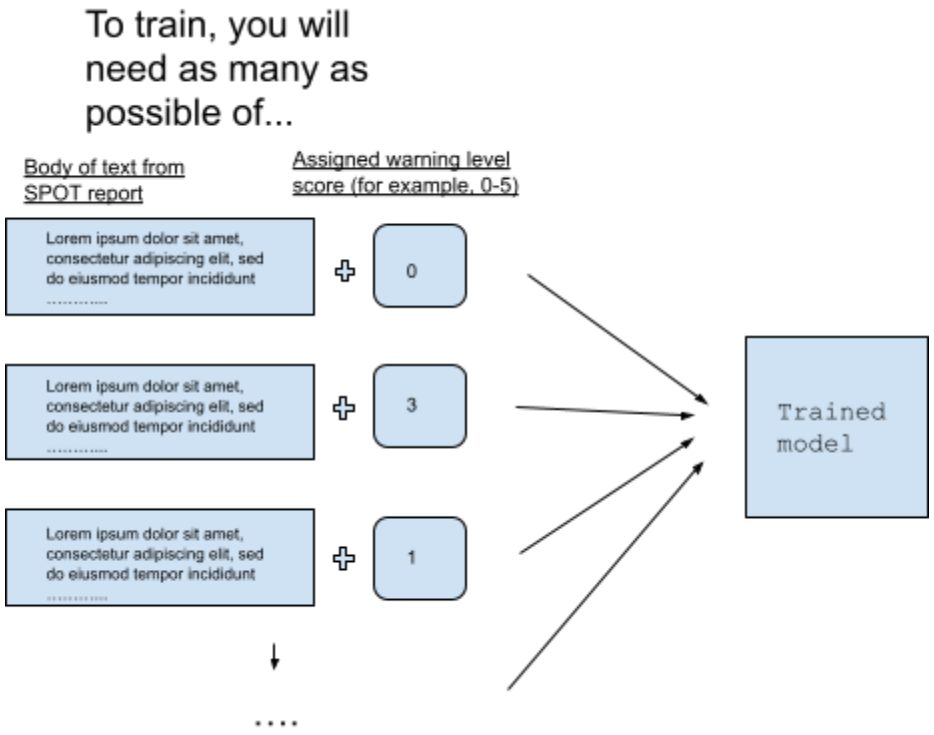
Whichever system is being used to collect incoming SPOT reports, whether that be at an embassy/field level or part of a central/unified system, this algorithm could be inserted and applied. In its trained state, the resulting model is simply a digital tool that may be distributed and used by any level of the department that deals with SPOT reports.

Required to get this working

In getting started the algorithm will need to be trained. Training is a process of teaching the AI tool to recognize cues within the data that match the different warning level scores.

Needed:

- body of training examples (As many as possible of: a real past SPOT report in text form, + a corresponding true warning level score) See image below.



How do we implement this?

The following code, made available by Patrizia Castagno, is a working example of training for text message spam classification, and the same concepts/methods would apply to our SPOT report warning level classification if we were to have the data. This example works well as a proof of concept because, despite the changes in what the input/output represents, the underlying training and classification method functions the same.

This method will work, but will also have a lot of technical improvements once SPOT report data is collected. The following steps will elaborate more on if the Bureau should outsource the resources for implementing this algorithm, how the Bureau can prepare data to train the tool's algorithm, how the Bureau should test and audit the tool's utility, and how the Bureau can practically begin the implementation phase of using the tool.

5 Buy or Build

One of the core challenges to any organization building an AI strategy is determining what method would be most effective for building and maintaining an AI tool. This dilemma often boils down to the decision to "build," meaning to use internal technical personnel and build up internal technical infrastructure to construct the tool in-house, or "buy," meaning to purchase an AI tool through standard government procurement processes. According to a study conducted by Stanford and NYU, 53% of the AI initiatives launched by the largest 142

government agencies were developed in-house, as opposed to around 30% of these being developed by a contractor. This decision is dependent on a number of different organizational considerations and constraints. Before one can make a decision on whether to buy or build, they must determine the focus of the tool and the goals of the organization.

Guiding Questions

The first consideration that must be addressed in the buy or build dilemma is identifying the need of the organization. It must be determined what problem is being solved with the proposed AI strategy in order to justify the use of AI in this context. In DS's case, the need is to more effectively analyze internal documents like SPOT reports in order to better inform agents on the ground of possible threats to State Department assets.

Next, the specific inputs and outputs of the AI tool must be considered. What material is being pumped into this proposed tool and what information are we getting out? This is crucial as the inputs being fed into the AI will determine its activity and the desired output determines utility. In this case, the input would be internally produced documents such as SPOT reports and the output would be the prioritization of certain reports in terms of severity to educate the focuses and decisions of DSS agents.

The organization must then ask itself what tool can solve this problem, and after determining that, whether that tool already exists. If a tool already exists which can solve DSS's problem, then it may be worthwhile to consider buying off the shelf, but if not, contracting a private tech firm or building in-house should be considered. Narrowing down these initial objectives and opportunities is crucial in making an educated decision to buy or build.

After assessing the needs of the organization, decision criteria must be determined. One of these criteria is time constraints. The time by which this tool is needed will impact your decision to develop the tool or purchase it off the shelf. If the tool is needed quickly, it might be necessary to purchase the tool outright to avoid the lengthy process of internal technical development, but if the timeline is longer or non-existent, in-house development may be more plausible. Another consideration is cost constraints. The amount of money the organization is willing to spend on a tool will play a significant role in determining whether to build or buy. Depending on the foci and priorities determined above, both buying and building can have monetary advantages, but in general, off-the-shelf AI tools are cheaper but risk being less effective or specific to organizational goals than the more expensive in-house products.

Another decision criteria is the maintenance required to make this AI tool effective. If this tool is being employed to complete a dynamic task, that is one where the algorithm must take into account new and changing modes of wrongdoing and subject matter, consistent maintenance is essential to maintaining an effective tool. This is because the tool must be constantly updated to account for new unfamiliar inputs or ineffective outputs in order to keep the algorithm up to date and functioning properly. This would be best addressed by an

in-house technical team that can consistently update the algorithm to suit the changing needs and realities of the organization and its work.

Once decision criteria are determined, the organization should begin addressing the pros and cons of buying or building in the context of the general considerations which accompany this decision. These considerations include budgetary constraints, manpower, pre-existing technical capacity, the focus of the tool, flexibility, and security.

When it comes to budgetary constraints, a government organization building an AI tool in-house can bring with it significant hurdles to overcome. Much of the cost of building an AI tool in-house can be attributed to the highly technically trained staff that must be hired to construct and maintain such a tool. Hiring suitable talent can be difficult within government organizations due to compensation and hiring limitations imposed by civil service laws. Private tech firms tend to be able to provide higher salaries, especially compared to organizations where technological innovation is not prioritized, resulting in possible difficulty in hiring technical staff. Externally sourced governance tools can be cheaper than internally produced ones, and academic literature has long concluded that the private sector will often produce goods and services at lower costs due to better-incentivized workforces and tighter managerial control. The inability to offer incentive-based compensation in the public sector paired with limits on hiring and firing can make hiring talent for in-house development more difficult. However, external sourcing can also impose heavy monitoring and transaction costs that would not accompany an internally sourced AI tool. This is because an AI tool produced externally requires consistent monitoring by the organization to ensure it aligns with its priorities and organizational constraints. Additionally, purchasing an AI tool off the shelf can bring with it possible hurdles in the way of procurement. Purchasing these tools requires the formation of a relationship with a tech firm capable of creating this tool, creating a coherent and effective plan to create this tool and implement it within the organization, and then spending the actual sum of money to purchase the tool.

Creating an AI tool for your organization requires significant technical expertise. The manpower available to your organization must be assessed before making any decisions on whether to buy or build. If an organization has a preexisting technical team capable of creating an AI tool to fulfill the organization's goals (and this tool does not already exist in an effective capacity) then it makes sense to pursue an internally developed AI tool. This is because the tool could be developed with relatively little additional cost to the organization. However, if the organization's technical team does not have this capacity, constructing the tool in-house would necessitate the hiring of new employees. As outlined above, this process has several obstacles both logistically and financially. The alternative to this hiring process would be to purchase a tool off the shelf or contracting a tech firm to create the tool. Both of which have the potential to be very costly, but these costs would be more short-term than an internal hiring process. Included in this manpower consideration is also the necessity to maintain the tool once implemented. An AI tool, especially when dealing with dynamic tasks,

may require extensive maintenance to remain effective. Building an internal technical staff would allow maintenance to be conducted more consistently and wouldn't require consistent private contracting to complete.

The tasks being completed by this AI tool and how nuanced that task is can be very important in determining whether to buy or build. Included in this is the consideration of flexibility once a tool is implemented. If the task being completed by the AI tool is especially complicated, that is a task necessitating quality monitoring, discretion, and organizational understanding, building internal technical capacity may be essential. An example of organization-specific needs resulting in a more effective AI product can be seen in the Social Security Administration's application of AI. Their tool was developed in-house utilizing employee expertise, and because employees were the ones determining how the tool would be productive, it ended up being extremely effective. It produced flags on documents where they are most useful for adjudicators, making the work of reviewing social security claims more efficient. Off-the-shelf solutions to complex, organization-specific tasks often invite corner-cutting by contractors who lack a nuanced understanding of the organization and the problems which the AI will be addressing. Procurement may yield more technically sophisticated tools than government organizations, but they can fall short in terms of personalization to the organization and the work it is doing. The threats facing DSS are rarely static, and having internal technical expertise which can create priorities in AI application and continuously update algorithmic tools to incorporate new risks and strategies employed by adversaries into calculations is crucial to maintaining an effective and up to date AI tool.

Another key consideration, especially in the context of a security agency like DSS, is the security of your AI tool. An adversary being able to understand how the tool works and create strategies to avoid triggering said tool would wholly undermine its purpose. Private contractors have the potential to be less secure with specifics about a government-employed AI program than the government itself would be. Private contractors often make it known that their tools are in use by government agencies, which opens the door to adversaries purchasing that technology and analyzing its technical components to determine how to avoid triggering said tools. Additionally, because these private contractors do not serve the American people, there can be an increased risk of direct leakage pertaining to the details of an AI tool to an adversary.

Upon determining your needs, decision criteria, and the benefits and detriments of each strategy in key consideration areas, you can make the decision to buy or build your AI tool. Both strategies can be effective, but it is key to make this decision based on the goals and constraints of your organization.

Examples

Customs and Border Protection (CBP), another federal law enforcement agency, has developed several different but supplementing AI initiatives ranging from facial recognition

to risk prediction on individuals entering the United States. CBP, and its parent organization the Department of Homeland Security, have consistently chosen to use contractors for the development of its AI/ML tools, and appear to have combined products from companies that likely employ different models. These tools appear to be quite successful and CBP has reinvested to improve these tools, but as it expands its use it will "likely have to build greater internal agency AI/ML expertise to address the knowledge gap created by reliance on contractors". In one internal report CBP admitted that it was unable to explain failure rates in one of its facial recognition tools due to proprietary technology being used. If CBP cannot understand its own technology it is vulnerable to adversarial attacks and other unknown threats.

(Cuellar, Engstrom, Ho, Sharkey; 30)

The Social Security Administration is likely the largest adjudication agency in the western world, taking on more than 2.5 million disability claims annually. This caseload created a significant backlog of claims and hearings, making the organization seriously inefficient. In order to address these challenges the SSA began efforts to improve its data infrastructure and develop staff with technical competence to improve efficiency and accuracy when processing claims. It has now prototyped three different AI tools relying on in-house expertise and the building of internal technical capacity. These tools include clustering algorithms to improve case processing, AI which can predict which cases have a high/low probability of receiving benefits, and Natural Language Processing to identify weaknesses in draft opinions. The results of this case are yet to be seen, but its success will largely depend on the improvement of data quality, methods and evaluation, and capacity. One obstacle to in-house AI development illustrated in this study are how conventional organizational boundaries (such as hiring) can impede innovation. In this case one of the leaders of the initiative was only authorized to hire attorneys to work on this project, limiting his ability to prototype AI tools by using technology staff.

(Cuellar, Engstrom, Ho, Sharkey; 37)

Recommendations

The course of action which best suits DS's AI development plan would be to first complete an extensive review of available AI tools with the department's procurement division to determine whether a tool which fulfills our objectives exists. If a tool of this sort is available off-the-shelf then it should be purchased and a technical team should be built around it to build up data capacity/organization and conduct maintenance on the new tool in order to maintain its effectiveness and suitability to the organization and its tasks. If an adequate off-the-shelf tool is not available, then the department should look to build its technical team and develop the tool in-house to ensure it fits with organizational goals, is utilizing DS expertise to improve its utility, and minimize oversight costs and security risks.

6 Centralize

In preparing for a new AI project, the bureau needs to centralize its data, or prepare data to be analyzed by the AI system. Datasets are used by machine learning algorithms to continuously improve themselves over time. Data can be prepared through measurements of completeness, accessibility, quality, connectivity, quantity, and validity.

Earlier, we recommended one limited AI project that will assess SPOT reports. That tool uses a natural language processing algorithm. Such an algorithm turns large amounts of text into matrices. Words are turned into vectors and then the computer performs analysis in a manner that allows the machine to learn from these vectors and matrices. The machine then detects patterns in these vectors and matrices to be able to confidently predict future patterns in the data. So, in this case, the data is large amounts of text. But other algorithms for other AI projects may require other forms of data. That data will be analyzed by AI and to train AI to analyze such data, the same types of inputs need to be used to create the AI tool.

Datasets have six categories that will determine if they are high-quality or not. Just because the data is high-quality - doesn't mean it is perfect. This is absolutely okay, as data doesn't actually need to be the very best (i.e perfect). The six categories that define a high-quality dataset are:

- Completeness - Are there missing records?
- Accessibility - Can the data be accessed?
- Quality - Are the records correct or do they contain mistakes?
- Connectivity - Can the different data sets be joined?
- Quantity - Is the amount of data limited?
- Validity - Does the data contain outliers?

According to the former head of data science at 2021.AI, Benjamin Biering, these six categories all ask very important questions. Of course, in these categories, some weigh more heavily than others. Some categories would be easier to fix than others. For example, it would be easier to fill in missing records than to make sure the data won't be limited in quantity.

Completeness

In dealing with limited or incomplete data - the question arises of how much is enough? There are many variables that can determine the smallest size of a dataset (goals, complexity, time frame). It is best to start off with smaller, more simplistic models that don't require a bunch of datasets before moving on to more experienced data which may require a huge dataset. If someone is running short on data - there are also different strategies to handle that. Data augmentation dictates gathering new data based upon the datasets you already own. Data synthesis dictates creating new data points by indulging in more complex sampling techniques. This would mean using methods that fully utilize the power of deep learning to generate this new data..

Accessibility

According to the Massachusetts Institute of Technology (MIT) and assistant professor of IT at HEC Montréal, Gregory Vial, data accessibility refers to the availability of the data or how easily retrievable it is. Data being accessible to the system or organization at large is a crucial aspect but one that becomes easier with a large network that is connected to the same database. Accessibility hinges on the ability of Connectivity and Completeness, but if the infrastructure to maintain an AI is already in place, then the infrastructure for accessing that data is also accounted for. Accessibility is too often looked at as an issue with IT, but issues with Data Accessibility can also be accounted for in how the AI is constructed and how different attitudes towards Accessibility can harm it.

The MIT Sloan Review looked at 6 North American Companies and found that a startling trend between Business Executives (Non-AI Decision Makers), who want to collect large amounts of data to access, Data Scientists, who are more concerned with the quality of data, and the Data Engineers who build the AI infrastructure, all three of these groups have a low-to-medium interest in data accessibility. This overall lack of care given to accessibility caused the construction of AI that was not properly prepared for the data that was given to it, thus harming the accessibility of the data. This means that for Accessibility to not become a problem in hindering AI development and implementation:

- Know what kind and quantity of data is being collected.
- Know the expectations of the AI and properly factor those into its creation/implementation
- Properly construct an AI that can read and organize the data without being overburdened
- Know the expectations of the AI so that the implementation/creation will not hinder Accessibility once the AI is implemented.

Quality

AI is a data-driven technology, but not all datasets are created equally. The effectiveness of the AI is determined by the Quality of the data provided. Proper precautions are needed to ensure that bias doesn't take over the dataset and corrupt it. While there is no outright cure for it, not taking the proper precautions can risk potentially devastating consequences. This is why using high-Quality training data is important. In order to ensure data is actually high-quality data - one must make sure to have a

- Diversity of Data
- Clean Data
- Clearly Annotated Inputs.

Diversity of data dictates that when data isn't diverse, AI models will show bias. This bias was shown in 2015 when Amazon's algorithm that was used for hiring was found to be biased against women. The reason for this was since the dataset for the algorithm was based

upon the number of resumes submitted over the past 10 years (majority men), the AI was trained to favor men over women. The lesson here is that one should use data from a variety of sources that will limit bias.

Clean data dictates that irrelevant data (missing values, typos) will not be helping the AI's learning abilities - especially in the training data. The training data is incredibly foundational to the AI model being built. The training data will be the first information that the AI model will learn on. The lesson here is that it needs to be clean and clear, and you should remove any corrupted information that will induce huge amounts of bias into the model.
Clearly annotated inputs dictate that without clear, relevant labels, the AI model will unlikely learn how to make the correlations that are being asked for. The lesson here is good labels give your model information it needs to make correlations in the real world, where inputs aren't labeled.

Connectivity

Connectivity hinges on how data sets can be connected to each other and this can be accomplished with a standardized, universal method of implementing data sets into the AI so that patterns can be seen by the program and therefore be connected. Connectivity matters during the construction of the AI, but standardizing how information and data are inputted into the system by whoever can easily allow the AI to connect different data sets.

The key ideas for Connectivity are:

- Standardized input for data sets
- Optimizing AI for connections between data sets as required
- Implementing a system for anyone to input data that the program can read and make connections/combinations with other data sets in the program already

Quantity

As mentioned in the quality section above, not all data sets are created equally. This Idea returns here with the fact that not every kind of data, and not every data source is useful or high quality for the AI models to learn. In general, more data should ultimately lead to much better results. At some point though, you will reach a crossroads of diminishing returns. This means that no additional data will be needed as the dataset is already broad enough to get the most out of the AI model. Over time, the cost can creep on the organization and thus, makes it less sustainable. The lesson here is a dataset only needs to be so big before it's truly representative of the whole.

Validity

Outliers in AI data collection can seem like an unfortunate side effect of data collection that threatens to distort the results of a data set. And outliers can indeed have this effect on a data set if they are too common and not filtered out as outliers. Since data sets for this application can be harder to filter out, the methods of detecting these outliers will be harder than simply making a point graph and seeing which points fall outside the range that the rest

of the points stay in. In order to reduce outliers, one commonly used method in machine learning is one-class classification that involves fitting a set with normal data to be expected and seeing if the new data conforms to that set's definition of 'normal' or 'anomaly'. This method is not perfect as it only uses the data it was initially fed in order to determine if the incoming data is 'normal' or not. These problems multiply with the number of features in the data being analyzed. This phenomenon is High-Dimensional Outlier Detection. In order to get around both of these problems, the AI should have a set that can easily detect outliers and normal data, but one that updates itself over time with new information deemed 'normal' in order to help the program better understand data with more accuracy and efficiency.

7 Iterate and Assess

We recommend implementing an iterative design process to continually audit and assess the AI tool's usefulness and accuracy. Iterative design allows for flexibility and adaptability as a project is evaluated for strengths and weaknesses. One of the main components to be evaluated is the programming algorithm of the AI tool. The algorithm is what tells the AI computing system what to do. Auditing is a process through which the Bureau can ensure the algorithm being utilized is resulting in an AI tool that functions in accordance with the intentions of the project.

Iterative Design

Iterative design processes are used in business and many other fields to continually improve upon the solution and design of one's creation or product. You may have heard of some iterative design process frameworks such as scrum, agile, kanban, and lean. Iterative design starts with a prototype and over iteration becomes the solution. The value is in between the solution and the prototype the design is being iterated against over and over to help get closer to the final outcome. This is the heart and core of iterative design, which is to get one step closer to the end state. Iterative design can be used to tune, tweak, improve, and remove bias from algorithms used for AI tools.

Natural Language Processing Algorithms turn large amounts of text into matrices. Words are turned into vectors and then the computer performs analysis in a manner that allows the machine to learn from these vectors and matrices. The machine then detects patterns in these vectors and matrices to be able to confidently predict future patterns in the data. Unless the data is both thorough and consistently updated, results of the program will not be optimal. The performance and agility of the development environment should be assessed often.

Iterative design was created with this main premise in mind; "the desire to remain flexible and adaptable in the face of uncertainty and complexity." Compared to the traditional "linear" approach which may be thought of as a "one and done" approach, where everything is

completed once and never iterated against again. The feedback loops in iterative design is what makes it ideal for adoption and why it adds so much value to organizations who use it. This is also how we narrow down a wide scope and wide goal of what we want to accomplish into our desired goal and desired scope.

Without an iterative environment for the algorithm to work in, there is almost no point in having one, to begin with. The only reason would be to make those assessments once and never again, which would be the linear approach. The point of the iterative environment is so that the algorithm can constantly assess, as well as be able to incorporate new data points the department may decide to explore depending on the department's operations and workflows. Iterative methodologies make the use of AI more dynamic and develop a "reusable" algorithm as needed.

One of the main reasons Healthcare.gov failed initially was due to the lack of efficient project management. This also coincided with the "large scale of this project, the complexity of it, the large number of stakeholders, and also the shift in requirements." In hindsight we can speculate from this that an iterative design methodology would have significantly reduced the likelihood of failure. This is due to the fact that these major reasons for failure would have been constantly assessed and evaluated, which is something that wasn't done. This would have enabled that "flexible" environment that was demanded to ensure a successful implementation initially.

Auditing Mechanisms

Auditing mechanisms help prevent unwanted consequences and inform decisions about the algorithm and its output. Unwanted consequences could look like unfair bias and other severe risks. Auditing also helps the algorithm operate without anything unfair, inequitable, and discriminatory from happening. For successful auditing, the Bureau must identify what it deems important, and what can be shared with the auditor. Then one can decide what the auditing process will look like and how to go about it.

We suggest developing a learning goal with the auditor and developing guidelines for handling low and high level audits. It is important that the auditor knows the extent to which the algorithm is being audited. With a clear understanding of what they need to know for a successful re-evaluation, the auditor will be set up for success. Examples of information they may need to know include parameters, objective functions, and input data.

We also recommend employing "mitigation strategies." These strategies are implemented to constantly assess your algorithm. There are two ideologies behind this auditing approach: the "white-box" where the auditor may know all of the information in the model and the technical details of the model, or the "black-box" where the auditor is given a partial idea of

what the model consists of. Auditing and re-evaluation depend on the auditors' access to information.

We will discuss in our next step what exactly an audit is composed of,but for now it is important to keep in mind examples of what could be assessed. In the process of implementing the AI tool, one will need to assess performance, accuracy, and precision of the tool. There will be more characteristics to assess but those characteristics are up to the Bureau to decide in line with the intended purpose of the tool . This is to make sure that the AI tool you are using is making decisions the way bureau employee would or even better.

Some non-technical aspects of the algorithm one can audit are the robustness, privacy, and fairness of the algorithm. For example, in fairness, we can examine the bias of the algorithm and adjust accordingly. This is dependent upon the auditor's findings and the level of audit being conducted. For example, the audit could go in ample technical detail of how this algorithm is producing its bias, or it could be a simple sheet or checklist.

8 Implement

The process of implementation focuses on two areas: readiness and reassessment. Readiness is the ability for the department to have reached or surpassed their required standards and metrics in the previous trial stages. The tool must be capable of working at the equivalent of a "human or better" before being implemented into the real world. Without hard testing metrics, the tool may be released to agents in a stage that is not properly equipped for real-life application. Following the transition from testing to real-life, the department must constantly assess then reassess the effectiveness and usability of the tool. As time surpasses and priorities change, the tool must be adapted to the current needs. To achieve this, the department should maintain a tech team with the skills to work on maintenance and updates. In summary, the department must not approach the implementation without determining its readiness or without a plan for reassessment.

Readiness

The movement from a testing environment to a live tool is the core change during the implementation step. Metrics are required to deem the AI project acceptable for real-world application. Artificial intelligence is typically not ready for live usage until it has the capabilities of a human or better. As an example, if a human were given a bucket of apples and oranges, and was tasked to sort them out, they would be perfect or near perfect in the task. In this situation, if an AI tool was created for sorting apples and oranges, it would not be able to move from testing to application until it reached the abilities of a human. Without the ability to perform as well or better than a human, AI will not be beneficial to the organization as it slows down or harms the current process.

AI is typically evaluated by these principles:

- Precision: refers to closeness of two or more measurements to each other
    - When the model predicts it is an apple, it must be sure it is an apple

- Accuracy: defined as the closeness of a measured value to a true value
    - The rate the model performs perfectly

- Recall: how often something was wrongly classified as not true
    - This displays when the tool says apple when it was an orange

Moving from the apples and oranges example, it is imperative to establish what is "human or better" within the department. Although the concept is much more complex, the same principles are applied in both AI tools. To understand what qualifies the AI tool as "human or better", the department must first understand what the metric of "human" is. This will require in depth metrics covering how agents react regarding SPOT reports. When looking at this, areas like speed, accuracy, frequency, legitimacy of the agents' responses must all be considered. After the baseline of "human" is correct and the AI tool has proven its ability, it will become time to implement into daily workflows.

Reassessment

The second aspect of implementation is the constant reevaluation of the AI tool. This step can be similar in ways to the testing phase, but it differs because the tool has already proved its ability. Reevaluation is essential for any successful AI tool because of the speed that technology and needs change. While there are many options for a system of reevaluation, Agile has proven to be an effective method. The strategy of Agile consists of four parts:

1. Small incremental variables
2. Allowing customers to be involved (agents in this case)
3. Continuous Q&A during the process
4. Building around the firm requirements

Applying these ideas into the reevaluation of the AI tool will allow the team to constantly give feedback and suggestions for ways of improvements. Since the tool will be reevaluated, the tool will be able to be changed and altered for the current needs of the department. Throughout the reevaluation security must be continually adapting. For example, without a strong security system in place, the department can leave itself vulnerable for potential hackers to gain information. Overall, reevaluation is important because it allows the system to adapt and stay current with changing needs.

9 Extend

After creating a clear vision and testing one limited AI project, the bureau will be ready to assess the strengths and weaknesses of the project and launch other initiatives to accomplish different goals. The process should be sustainable, cost-effective, and incrementable.

Once you have completed the steps of familiarizing, strategizing, and implementing, you can evaluate the effectiveness of the first initiative. Using the AI strategy developed in the first steps, you can begin development of a secondary initiative to build upon the first. If the implemented initiative is currently accomplishing x, how can y and z be accomplished through extension. The small-scale project developed in previous steps should be viewed as a building block for artificial intelligence throughout the team's operations. There are numerous benefits to scaling quickly- efficiency is optimized when implementations are extended past the pilot stage. Generating returns or savings within DS via the initiative will help to fund later stages of development.

Considerations

How can feedback from team members contribute to a larger technology strategy inside of the Bureau of Diplomatic Security?

Case Studies:

As exemplified by a massive digital transformation undertaken by a SmartIndustry affiliate, the desire to start small can bring impressive financial gains. One large industrial company prioritized small inventory management and capacity optimization initiatives in an effort to conserve capital. Within the first nine months after implementation, the firm generated $20 million in added value and drafted a roadmap outlining their next ten data-transformation initiatives, looking to generate up to $200 million over the next five years and raise their EBITDA margin up to 4%. The same firm found that data-driven companies that have employed AI within their R&D, supply chain, marketing, sales, and manufacturing have improved their EBITDA up to 30% over their competitors. The Boston Consulting Group recommends executing multiple small projects in phases at a time to share a common infrastructure. While reinventing core-IT systems can cost upwards of millions of dollars, taking a pragmatic, disciplined approach to data management will result in similar time and cost savings in the long term.

Authors in Alphabetical Order

Austin Warren
Colene Short
Gavin Henderson
Henry Bobeck
Jackson Sherman
Jason Esquivel
Kai Takesue
Kate Riordan
Mac Cannon
Maren McClelland
Max Krajacic
Natalie Rioux
Rebecca Prushinski
Sam Chapman
Sunny Gandhi