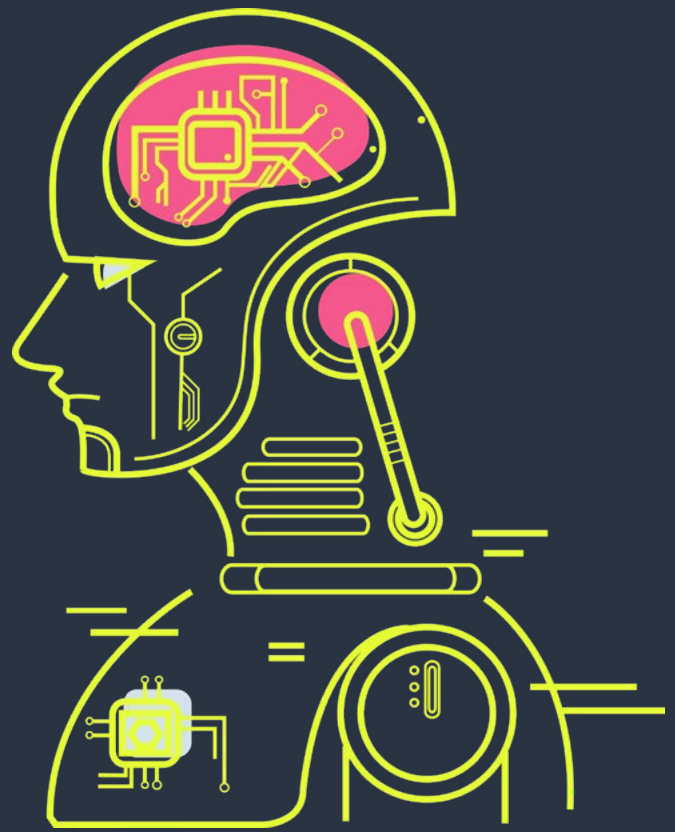
Beyond the Hype of Machine Learning



GOVLOOP
E-BOOK
2019

aws

govloop

Executive Summary

About halfway into their first year, babies begin to learn object permanence. If you've ever wondered why the game peek-a-boo's shock factor dwindles for infants, it's because babies begin to understand – slowly but surely – that even when something is obscured or falls out of sight for an instant, it still exists.

Babies develop their understanding of the world in learning environments fostered by instructive adults and growing neural networks. And similarly, artificial intelligence (AI) has blossomed over time – albeit not rivaling the miracle of mankind's brilliance – because of the ability to learn.

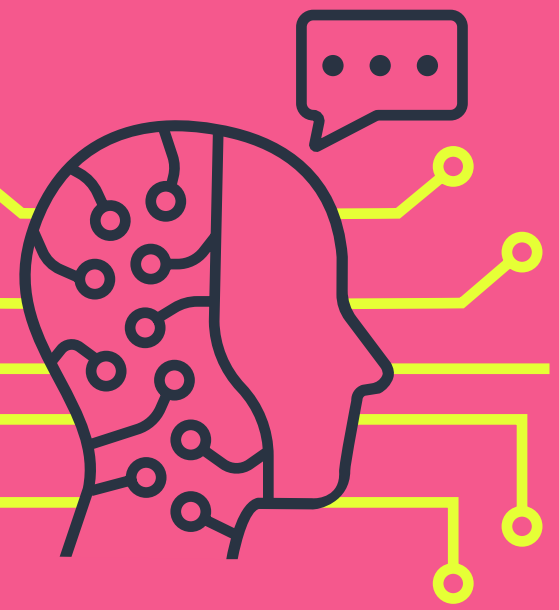
As defined by the International Organization for Standardization (ISO), artificial intelligence is "the ability of a system to acquire, process, and apply ... facts, information, and skills acquired through experience or education." Artificial intelligence can carry out tasks that till now were only associated with human intelligence by automating, understanding and generating real-world processes.

Machine learning (ML) is a subset of artificial intelligence that improves upon its performance automatically by using data sets and algorithms to train itself. The idea of machine learning is that the technology can teach itself based on inputs to understand deeper patterns and generate complex models that consider a variety of factors, including impressive volumes of data that humans might not be able to process.

Machine learning is greatly changing government. Automated systems that can incorporate human intelligence fit into just about every job description imaginable, leading the way for workers to carry out more important tasks and deal directly with customer service. AI and ML are also helping blaze the path to unimagined business improvements, incorporating predictive analytics and streamlining cumbersome processes.

The following pages will focus on how machine learning works, what its impact is for government and how top government agencies see the technology. There are also two Q&A interviews with leaders in the federal AI and ML space, as well as a timeline and statistics about the continuing progression of this technology in government and beyond.

The Overview of Machine Learning and Artifical Intelligence



$973 million
is how much the 2020 federal budget requests for non-defense AI programs.

50%
of the White House's proposed AI budget would go to National Science Foundation research and related activities.

99%
of the time, wildfires were accurately identified by NASA AI that taught algorithms how to recognize fires from remote-sensing images.

7,437
publications about machine learning appeared on PubMed in 2019 from January through October vs. 543 publications about machine learning on PubMed in 2009.

1%
of state chief information officers (CIOs) said that artificial intelligence is being used widely throughout their state.

90%
More than 90% of the most severe acute kidney injury cases can be identified 48 hours earlier with the use of a machine learning model, a crucial timeframe for injuries of this kind.

60
employees work at the Defense Department's recently formed Joint Artificial Intelligence Center.

The top three areas where AI is expected to make a measurable improvement for state CIOs are cybersecurity; fraud, waste and abuse detection and management; and citizen-facing digital services.

"Continued American leadership in artificial intelligence is of paramount importance to maintaining the economic and national security of the United States."
– President Donald Trump

"We will harness the potential of AI to transform all functions of the Department positively, thereby supporting and protecting U.S. servicemembers, safeguarding U.S. citizens, defending allies and partners, and improving the affordability, effectiveness, and speed of our operations."
– Summary of the 2018 Department of Defense Artificial Intelligence Strategy

The ‘How’ and ‘Why’ of Artificial Intelligence

By now, many people have artificial intelligence burnout. They’ve seen the movies, heard the promises and watched the commercials, but despite it all, their day-to-day has not changed all that much.

But for those working in government, artificial intelligence should be exciting in several ways.

Public sector employees are often overworked and under-resourced, both spaces where AI can help. Artificial intelligence automates activities, removing busy tasks from employees’ duties, and analyzes large sets of data, enabling employees to make more data-based decisions.

Additionally, while government agencies often find themselves reacting to new market trends, responding to innovations as they come instead of actively pushing forward the change, with AI and ML, government is on the forefront – believe it or not. For decades, the U.S. government has funded, researched and advocated for the advancement of cutting-edge AI and ML.

Timeline of Artificial Intelligence in Government

- **1950:** Alan Turing publishes his seminal paper, “Computer Machinery and Intelligence”
- **1963:** The Defense Advanced Research Projects Agency funds AI at MIT
- **1986:** Carnegie Mellon University builds the first autonomous car
- **1997:** The world’s chess champion, Garry Kasparov, is defeated by IBM’s Deep Blue chess-playing program
- **2011:** Apple unveils Siri, its intelligent voice response system
- **2018:** The Defense Department launches the Joint Artificial Intelligence Center
- **2019:** President Donald Trump signs an executive order on “maintaining” America’s advantage in AI

Source: Harvard

The ‘How’ and ‘Why’ of Artificial Intelligence

But as machine learning continues to percolate throughout government discussions, it’s important to know what’s what when it comes to ML.

First, as noted above, machine learning is a subset of artificial intelligence. Therefore, all machine learning projects are subsumed under the umbrella of artificial intelligence, but not all artificial intelligence projects have to do with machine learning.

Second, machine learning does not run rogue. Developers give ML programs algorithms and data sets that dictate how they operate, what their scope is and where they access information. Then, machine learning generates models based on its inputs, picking up patterns, writing models and executing actions.

Noteworthy at this stage is that although it is based on the continuous refinement of inputs, machine learning is far from immune to bad data. If you’ve seen internet bots go off the rails or predictive models malfunction, it’s likely because of biased or incorrect data inputs.

Third, machine learning can run supervised or unsupervised. In supervised formats, machine learning receives categorized items and attempts to sort or distinguish them by patterns and logic. In unsupervised learning, the inputs are not itemized at all and just exist in large sets. Then, unsupervised machine learning programs attempt to ascertain defining characteristics.

Finally, machine learning is not confined to one realm. Robots will not be more rational than humans on every occasion, but machine learning can, in specific activities, go beyond the limits of people by processing incredible amounts of data. After all, that’s why the world’s best chess player is a machine. Machine learning also exists for simpler tasks, such as data processing, trend analysis and itemization, when the sheer size of demands is too much for human employees to tackle.

As a result, the use cases for machine learning are seemingly limitless. Machine learning can detect when machinery will start to experience wear and tear – known as predictive maintenance for government – or run predictive analytics to forecast budgets. It can also face forward to citizens, allowing governments to reach their constituents using voice recognition systems and chatbots.

But keep in mind that successful machine learning projects require a robust framework behind them. A correct selection of tasks, clear definition of algorithms and strong pipeline of data are vital for ML’s success in any agency.

Machine Learning's Impact in the Public Sector

How AI and Machine Learning Revolutionize Cybersecurity

The ability of AI and ML to process large volumes of data is particularly important when it comes to cybersecurity. Agencies at all levels are pelted with thousands of cyberattacks on a daily basis, and with limited staff and resources at their disposal, agencies lack the time to respond to each and every incident.

When the Office of Personnel Management (OPM) announced in 2015 it had been infiltrated by cyberattacks, the agency was caught flat on its feet. Meanwhile, the data it held – and the data of 21.5 million federal employees, applicants and their families – spilled into the open, even as the attack was unbeknownst to OPM.

Now, cybersecurity is a priority for agencies across government, with prominent mandates from the federal government demanding a higher level of security. While attacks increase in frequency, agencies are expected to do more.

Machine learning can come to the rescue here. Although machine learning can't keep attacks away from networks, it can prevent networks from letting them in – all without disrupting user productivity.

For example, if a regular, day-shift employee suddenly starts clicking on suspicious links, accessing the network at odd times or using multiple IP addresses, a machine learning defense system could suspend the account and prevent it from doing any further damage. Then, the user could be confronted.

Unlike other forms of automation, machine learning could also consider the user's profile and history. Rather than suspending an account and blocking an action just because of one late night assignment, ML-enabled cyber defenses don't get in the way of productivity. Instead, they look for trends of dangerous behavior and sniff them out before an issue arises.

Sometimes, however, networks will be breached, and then everything turns to damage control. In these cases, machine learning can patch systems and track down information that might be lost in real time by rapidly evaluating the health of systems.

The benefits of having machine learning as a security blanket are far more than insulation from negative headlines. The benefits are also internal, allowing IT and security to think more proactively and enable more systems.

Chatting With a Bot That Listens With Machine Learning

At work or at home, good listening is the key to positive relationships. Unfortunately, government agencies have a reputation for providing a depersonalized and unresponsive customer experience. Chatbots can change that.

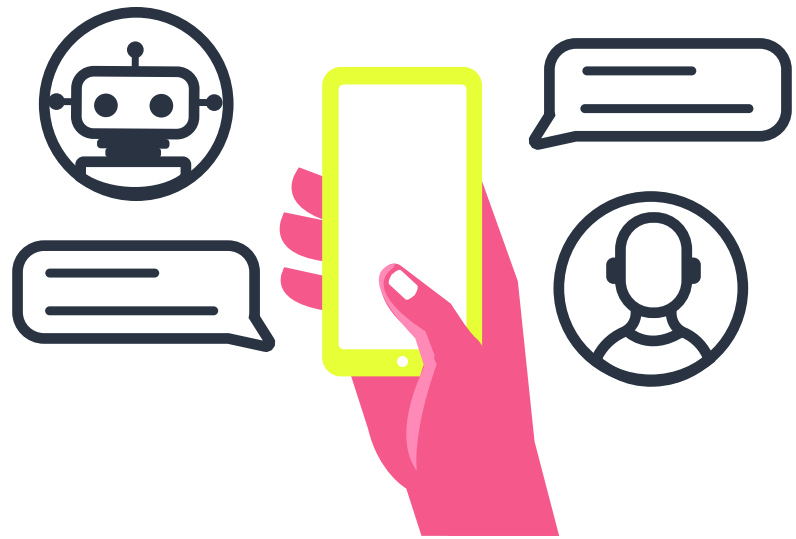
Chatbots listen all the time, picking up on the most important cues that users input and responding with mapped-out answers. Machine learning can revolutionize this experience by learning from similar questions how to produce the *best* answer.

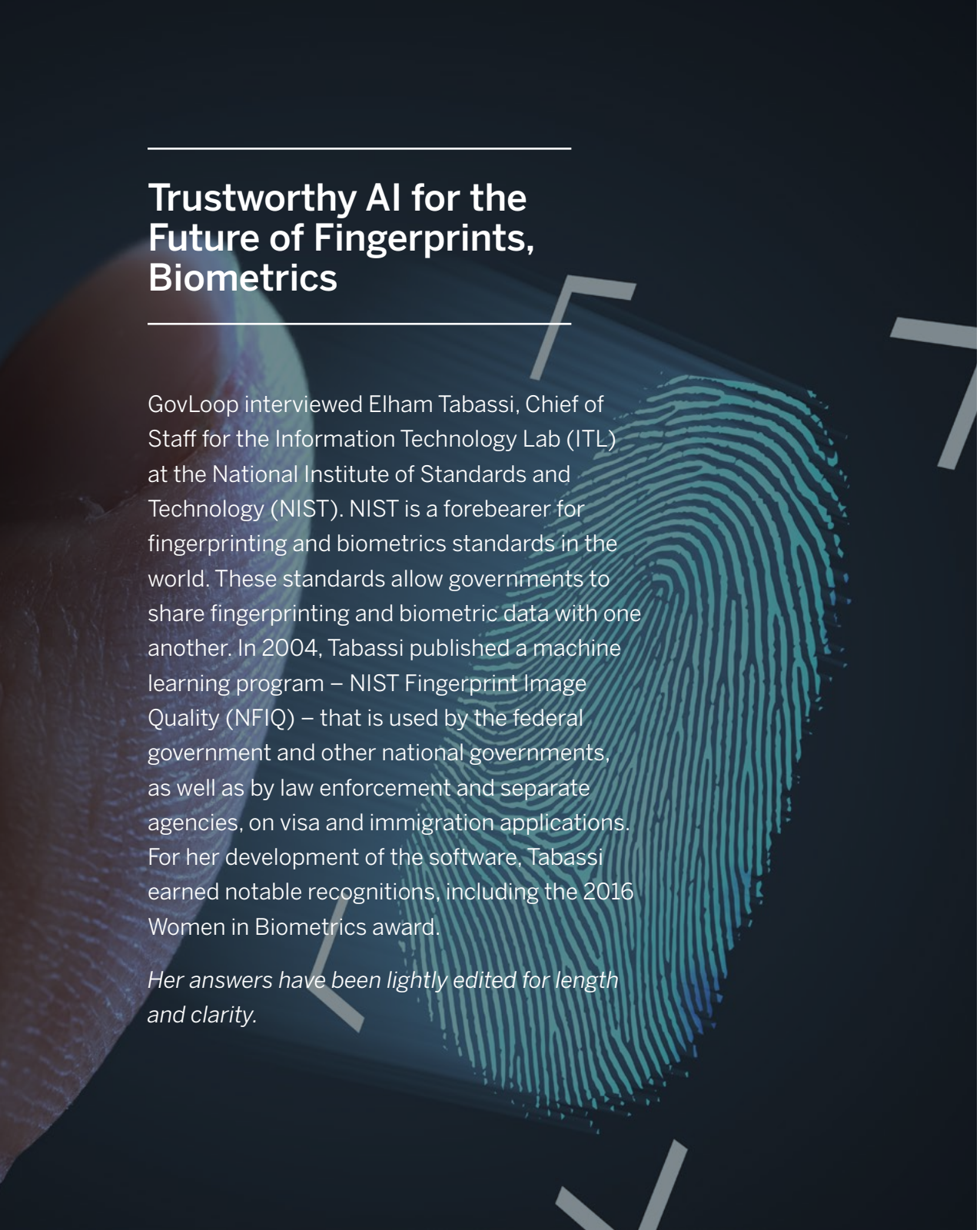
For example, if a citizen types into a chatbot asking for parking availability at a nearby park, the chatbot automatically delivers the "right" answer, that is, a list of nearby parking lots. With machine learning, though, the chatbot can consider the location of the user, the amount of parking spaces regularly available and the busyness of the day – and provide the most helpful answer.

By considering an amalgam of factors from prior interactions, machine learning would be able to provide a more nuanced, fact-based response than otherwise would have been available through an automated response.

Chatbots can also work based on reinforcement learning, a type of ML whereby reviews of the service provided influence how the system will operate in the future. For example, if a chatbot asks, "Was this answer helpful?" and the answer is no, then it will be less likely to answer the question with the same response, weighing other options' helpfulness instead.

In this way, machine learning listens to and responds to feedback – and again, does so immediately.





Trustworthy AI for the Future of Fingerprints, Biometrics

GovLoop interviewed Elham Tabassi, Chief of Staff for the Information Technology Lab (ITL) at the National Institute of Standards and Technology (NIST). NIST is a forebearer for fingerprinting and biometrics standards in the world. These standards allow governments to share fingerprinting and biometric data with one another. In 2004, Tabassi published a machine learning program – NIST Fingerprint Image Quality (NFIQ) – that is used by the federal government and other national governments, as well as by law enforcement and separate agencies, on visa and immigration applications. For her development of the software, Tabassi earned notable recognitions, including the 2016 Women in Biometrics award.

Her answers have been lightly edited for length and clarity.

GovLoop: What are some of the use cases for AI and ML at NIST?

Tabassi: NIST has six different laboratories; Information Technology Laboratory is the one that I'm from. I want to start by saying use cases in other laboratories at NIST. These include materials, robotics, wireless networking. They all are using AI techniques to advance their measurement sciences. In Information Technology Laboratory, we do a lot of standards and evaluation for AI. There is a group working on biometrics, and they do evaluations of spatial cognition algorithms. While we don't do spatial cognition algorithms, we evaluate other ones.

I saw you have quite a story, winning awards and serving as the principal architect of a watershed technology in fingerprinting and biometrics. Would you mind elaborating on that history?

Prior to being a chief of staff at Information Technology Laboratory, I was doing fingerprint work at NIST. Our fingerprint work at NIST goes back to the '60s. What happened in the mid-'60s is computers are being recognized as new resources for doing computations much more efficiently than the humans. And FBI at that time decided to use computers to automate some of its fingerprinting.

What they realized is the format of the files were all different, and there was a need for a sort of a standard formatting and, if you wish, language for fingerprint exchange. And then, they came back to NIST again, and the group that was part of that, before becoming the staff here, in collaboration with the whole community, worked on the standard.

And that was a standard for storage and exchange of fingerprint data. That standard since '86 has been revised several times [and] has expanded to include many other modalities. Right now, I think it has 16 or 17 different modalities – fingerprint, iris, face, DNA, bite mark, many others.

Fast forward to after 9/11, the formation of the Department of Homeland Security and passage of the Patriot Act. The US-VISIT Program requires all the visitors to U.S., when they apply for the visa, they will be fingerprinted, the fingerprint will be sent to U.S. for background checks, and if the visa is issued to them, when they arrive to U.S., they get fingerprinted again. They check the fingerprint at the time of arrival to match it with the fingerprint that is submitted at the time of the request for the visa to make sure that the person that was issued the visa and the person who is arriving to U.S. are the same person.

So, the question was fingerprint quality. What's the likelihood of this fingerprint being matched correctly? And that project landed on my desk, and out of that I developed a software, NIST Fingerprint Image Quality, that when a fingerprint is captured, you look at it and say the likelihood. And how it's useful, it's the old saying of "garbage in, garbage out." If you can prevent the poor-quality fingerprints [from getting] into the system, we have improved the performance of the system.

How does machine learning come into play here?

You present, you develop, you build a model that's from the patterns and information characteristic that you present to them. So, if I want to make it an easier example, if you want to predict the price of a house from number of the bedrooms, number of the bathrooms and the ZIP code that the house is located in, you can get a lot of examples of houses. If you give enough of this and write a computer program to learn the function from the input fit, in this case it's the number of the bathrooms, bedrooms and the ZIP code, you can try to predict or compute the price of the house.

So, in the case of the fingerprint image quality, I gathered those types of information about the fingerprint. Fingerprints usually are this minutia, you know. Fingerprint is just this print, the pattern of the black and white of the lines, and that's sometimes a property and sometimes, diversion to many different, more than two of the ridges. So, I computed those types of information. And from the learned model, it comes up with the prediction with a likelihood of the match.

Such a big part of AI is trust in the system. How do we improve trust in AI?

The full potential of AI is going to be achieved only if users can have the confidence and are able to trust the technology for their use. We try to break it down into elements of, or aspects of, what trustworthy properties that the trustworthy AI should exhibit. And then for each of them, go and figure out how to measure them.


FREEDOM TO INNOVATE
Data Integrator
Reporting Data Mart Amazon Redshift
AMAZON QUICKSIGHT
Business Users
CAPACITY
Data Integrator + Big Data
Dev DW Amazon Redshift
Data Lake Amazon S3
SQL
ETL
Amazon Athena
C-SAT
Data Analysts + Data Scientists
AMAZON REDSHIFT
The federal government trusts the cloud with the most functionality, innovation, and experience.
aws is how
www.BuildOn.aws

Enabling a Smarter, Better Business Enterprise Using Machine Learning in the Cloud

An interview with Balaji Iyer, Business Manager for Amazon Web Services Machine Learning

The use cases for machine learning are endless, and your favorite digital platforms use machine learning to make experiences as seamless as possible. Many people use Netflix to stream movies, Venmo to reimburse friends or Airbnb to travel, and you can bet that all have well-defined machine learning portfolios.

Now, citizens and employees demand the same service that they receive in their day-to-day lives – where they interact with fluid and adaptive portals – to apply to their interactions with government. If they can pay for a flight with one click, see a historical user profile and receive recommended next trips, why shouldn't you be able to pay taxes the same way?

For those working in government, it's no secret that aging IT and inadequate budgets are the reasons why. But smooshed between organizational constraints and mounting demands, governments need a hand from technology.

"AWS transformed something very specialized such as data center operations into a utility model, where infinite resources are available at your fingertips, making innovation tangible for everyone," said Balaji Iyer, Business Manager for Amazon Web Services (AWS) Machine Learning.

GovLoop spoke with Iyer about how machine learning and artificial intelligence are transforming modern public sector enterprises and how cloud fits into the transformation. AWS offers ready-to-go machine learning and artificial intelligence that can fit into most IT environments.

Several lynchpins turn machine learning from an aspiration to an effective, central business element. First, data needs to be made available, as machine learning programs base their models and decisions off data. Therefore, the data not only needs to be present, but clean and accurate to avoid biased or flat-out wrong decisions.

Second, organizations need the resources and frameworks to incorporate machine learning. While self-made and piecemeal models can disrupt systems or require constant updates and tweaking, machine learning can also come ready to go when incorporated as part of the cloud.

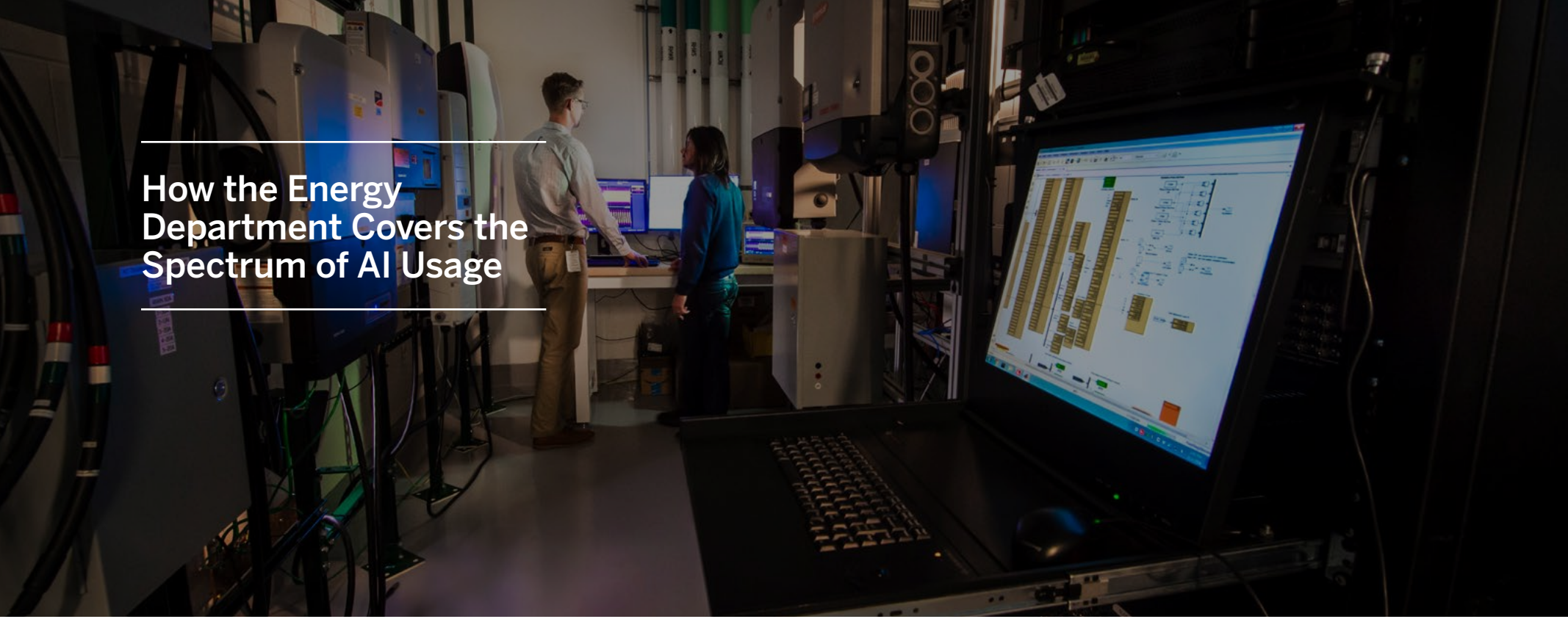
Finally, machine learning requires a level of expertise. While some organizations might be experienced with artificial intelligence, other agencies will have never used it before. In this case, agencies need to reskill and prepare their workforces to accompany ML.

AWS offers a training program with more than 30 self-paced digital courses that can be used to train anyone, from developers to business leaders, on ML. With Machine Learning University and the AWS cloud, developers, data scientists and business leaders can collaborate to power AI projects with real mission impacts.

"There's no one-size-fits-all when it comes to solutions, so our goal is to provide the broadest and deepest set of machine learning capabilities for builders across all levels of expertise, saving them money and time," Iyer said.

Cloud is ideal for achieving all of these tenets of ML. The cloud can open up data throughout an enterprise, come with ready-made capabilities and fit flexibly into IT environments – all culminating in a platform for successful machine learning.

Takeaway: Machine learning is not something that can just be airmailed into an enterprise. Take time to consider data quality, system fits and expertise, and invest in programs and technologies that will boost readiness in those areas.



How the Energy Department Covers the Spectrum of AI Usage

GovLoop interviewed Pam Isom, Deputy Chief Information Officer for Architecture, Engineering, Technology and Innovation at the Energy Department (DOE). Recently, GovLoop had reported that DOE is standing up an Innovation Community Center to advance the use of emerging technologies, such as machine learning, throughout the enterprise. That story can be read here.

Isom's answers about AI and ML have been lightly edited for length and clarity.

GovLoop: What are the main uses of AI at the Energy Department?

Isom: I know you're aware that we have a lot going on with artificial intelligence, and I'm sure you heard the announcement about the work that we're doing with AI and how we have stood up the Artificial Intelligence Technology Office. Some of the areas that we are focusing on as a department as a whole are energy, of course; medical; health care; transportation; [and] agriculture.

In the health space, Secretary [Rick Perry] had been speaking about work that we're doing to help discover what contributes to a traumatic brain injury, as an example. So that's not new news in the industry, but that work is continuing. And then Oak Ridge [National Laboratory], for instance, is involved with scanning clinical records to understand contributors to cancer, and hopefully we can find a cure for that deadly disease. So those are examples of things that we're doing in the medical and health care space.

But we're covering energy as a whole, and what can we do to boost energy efficiency in homes, in buildings, around the world. And then of course there is cyber and what things we can do to apply AI to thwart attacks against the energy grid, and then just cyber in general to help prevent cyberattacks, across not only the grid and energy infrastructure but cyberattacks as a whole.

So this isn't just cyberattacks internally, right? This is going into all different sectors?

All the different sectors. Within the Department of Energy, we cover various sectors, because we are the Department of Energy, so we have a broad space.

You also mentioned Oak Ridge for cancer research. Could you expand on that?

Our labs are involved with lots of different areas, and one of the areas that they're involved with is helping to figure out contributions to cancer. So we're scanning clinical reports to look for hidden causes of cancer.

Generally, how do you approach AI and ML? What sorts of frameworks need to be in place?

We use the standard tools that are out there, but because we have the national labs and work with those, they also are creating AI and ML tools. So it's very possible that working with them, because we have the research and development arm, that they are coming up with tools as well.

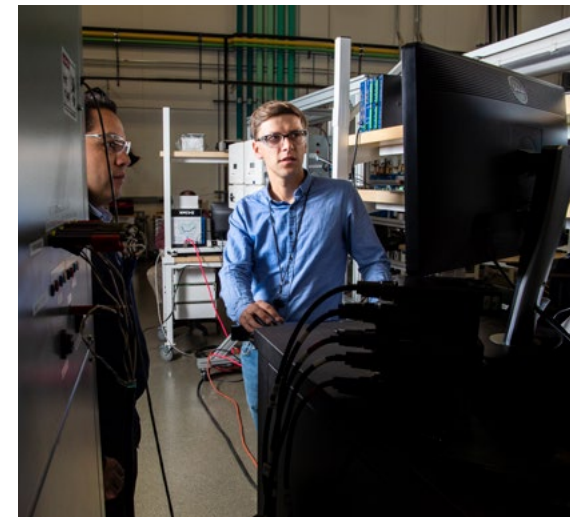
Typically, we follow the motto of ensuring that we have a diverse set of contributors to the algorithms and parameters, so that we are ensuring that we are actually learning and can apply machine learning in an unbiased way.

We know, for instance, that there's data and then there's knowledge and then there's AI. AI is dependent upon the data, and then using AI, we turn that information into knowledge that we can make decisions that are meaningful with. We use frameworks like that.

Make sure that the data is trusted. In order to do so, we're making sure that we are bringing in the right resources to help with the analysis and the algorithms that go into the AI models. Then what we're doing is, we are augmenting our AI with human intelligence as well. So those are kind of some guardrails that we're putting into place. And mostly we are making sure that we are doing things with no biases built in.

There's such a central importance of data to these efforts, as you were saying. What documents guide data governance and management at DOE?

We are contributors to the Federal Data Strategy, but we also lean on the insights from the Federal Data Strategy. Within the department itself, the Chief Data Officer is building an integrated data strategy for DOE.



Conclusion & Next Steps

From driverless cars to spreadsheets, the possibilities are endless when considering how machine learning will impact government. But the public sector must be ready. With a framework set up for data quality, continuous education and devoted research, the United States can propel its ML and AI to achieve incredible impacts for every citizen.

You're not alone on your AI journey. What's next if you want to bring AI to your agency?

1. Study up on AI standards and best practices. The bedrock of AI is good data, so promoting standardization within your agency is a must.

- *What is your agency's data governance and data management policy?*

2. Research use cases. When there's a will, there's a way. AI might be your way forward, but there's no way to know unless you define the problem first.

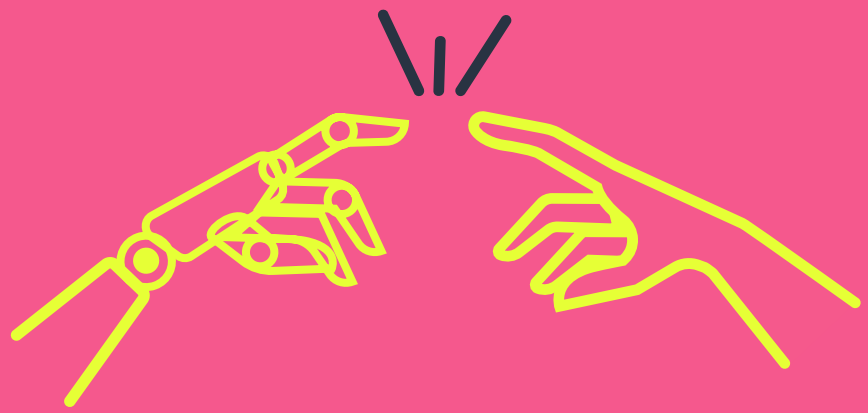
- *What challenges do you encounter day to day?*

3. Pilot the project. You don't need to save millions or save the world all at once. Start small, but think big.

- *How can I help solve one of those many challenges?*

4. Advocate for AI. Build a team to support you in your mission, engaging leadership, business users and IT professionals. Present a vision and see how far you can take it.

- *Who's with you?*



Thank you to Amazon Web Services for their support of this valuable resource for public sector professionals.

About GovLoop

GovLoop's mission is to inspire public sector professionals by serving as the knowledge network for government. GovLoop connects more than 300,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington, D.C., with a team of dedicated professionals who share a commitment to the public sector.

For more information about this report, please reach out to info@govloop.com.

govloop.com | @govloop