## Module 1: Introduction to New Technologies (DRAFT)

### Welcome

Welcome! Before we dive into new ways of working in government enabled by technology, we need to have a basic understanding of some of the technologies that make these methods possible.

Thus, in this module we introduce you to five key new technologies, namely:

1. Big data
2. Artificial intelligence
3. Machine learning
4. Collective intelligence, and
5. Blockchain

We are going both to define what these technologies are and how you can use them to solve public problems. Given the short time we have, we won't be able to go into all the ethical risks of these new technologies. However, we will point you to further resources and readings to learn more.

### Learning Objectives

We hope that by the end of this module you will be able to:

1. Describe five key technologies for solving public problems, and
2. Describe specific examples of how public officials are using these technologies to deliver better solutions to public problems.

Let's get started!

First, we'll consider big data.

Big data refers to extremely large data sets that are too big to be stored or processed using traditional means.

Today new collection, storage, transmission, visualization, and analytic technologies have triggered a massive proliferation of datasets collected by public and private entities about everything from health and wellness to phone and purchase records. This data is powerful raw material for problem-solving and the creation of tools that can further the public interest.[1]

[1] https://www.thegovlab.org/static/files/publications/Rights-Based-and-Tech-Driven-Noveck-YHRDLJ-vol-19.pdf

What makes data big data?

1) First, big data reflects increasing data volume: There is an ever-increasing quantity of data being generated. In 2015, 12 zetabytes -that's 12 x $10^{21}$ of data -- was created worldwide. That number is forecast to reach 163 zetabytes in 2025.[2] Just for comparison, the entirety of the library of congress is only 15 terabytes and it takes a billion terrabytes to get to one zettabyte.

2) Second, we are accelerating data velocity - the speed at which data is generated, analyzed and utilized. Today, data is generated in near-real time. Whether it is human generated data like point-of-sale credit card sales data and social media interaction data, or machine generated data like RFID tags and sensor data, people are interacting thousands of times a day with data collecting sensors and devices.

3) Third, big data is accumulating data variety: Data comes in a wide variety of formats including numeric, text, images, voice and video among others. Data can be:

    a) Structured data - this is data that can be pre-organized in traditional databases such as fields for phone numbers, zip codes and credit card numbers.

    b) Alternatively, more and more data is unstructured data - data that does not come pre-organized in traditional spreadsheet-style format. According to some estimates, unstructured data accounts for more than 95% of all data that is generated today. Yet with unstructured data, contemporary analytical methods make it possible to search, sort and spot patterns even without a predefined idea of what to look for.

The volume, velocity and variety of big data are making it more feasible and valuable for improving how we govern.

By analyzing data that government generates, collects, and shares, policymakers can understand past performance of public policies and services, evaluating both their efficiency and how they impact different populations.

For example, Economists Raj Chetty, Nathaniel Hendren, and Lawrence Katz studied twenty years of income records from families that moved neighborhoods using the housing choice voucher program. By analyzing that data, they discovered that the families that used housing choice vouchers to relocate earned significantly higher incomes, attained more education, and

[2] https://www.forbes.com/sites/andrewcave/2017/04/13/what-will-we-do-when-the-worlds-data-hits-163-zettabytes-in-2025/#52059e0e349a

were less likely to become single parents when compared to their peers who stayed in their neighborhoods. Citing this research, the Department of Housing & Urban Development overhauled the formula that it had been using for four decades to calculate rental assistance, and increased opportunities for families to move to low-poverty areas from high-poverty areas.

Larger quantities of data also enable the delivery of more tailored interventions in the present by helping governments at every level match people to the benefits to which they are entitled or the assistance of which they are in need.

For example, Louisiana's Department of Health uses Supplemental Nutrition Assistance Program (SNAP) enrollment data to sign people up for health benefits. Out of nearly 900,000 SNAP recipients, Louisiana has proactively enrolled 105,000 in Medicaid without a separate application process - instead, they are using a four question yes-or-no survey. This approach has helped some of the state's poorest residents get access to the benefits which they were entitled to, while also saving the state approximately $1.5 million in administrative costs.

Better access to data even helps with forecasting future outcomes, such as predicting who is likely to be a frequent visitor to the emergency room, thereby enabling more targeted interventions and treatment.

Having access to data and predictive models can help governments to prevent and deal with outbreaks and stem the spread of pandemics by answering questions like: how many people will a disease potentially infect; how far and how quickly will the disease spread; what areas and people are at highest risk; and when are they most at risk.

Indeed, big data is becoming an invaluable tool in determining which of our policies and services are working and for whom. Investing in policies, programs and services that work can have a dramatic impact on real peoples' lives.

This is why, in a rare moment of bipartisanship in 2016, Congress passed and the President signed a bill creating the Commission on Evidence-Based Policymaking calling for greater use of data to evaluate the efficacy of government programs.

Now let's take a look at Artificial intelligence (or AI), one of the key tools for quickly processing vast quantities of big data that can provide data-driven insights to address public needs.

AI can generally be defined as the programming of a computer to make decisions, and learn and perform tasks typically reserved for human intelligence. Examples of these tasks include the ability to understand and monitor visual and auditory information, reason and make predictions,

interact with humans and machines, and continuously learn and improve.[3] This technology enables machines to enhance and automate the work done by humans.

For example, Netflix and Spotify algorithms that recommend what movie to watch or song to listen to are powered by AI. The more you use them, the "smarter" they become.

There are at least four AI capabilities relevant for governing.

- First, AI can help policymakers make predictions in a way that is more comprehensive and less subject to human bias.

For example, in New York City, the FireCast project leverages AI to analyze data from across the city government to help the Fire Department identify buildings with the highest fire risks.

- Second is the detection of patterns and the ability to spot outliers. For example, AI can support policymakers to identify tax evaders or spot corruption, by identifying abnormal financial activities or trends.

- The third capability is computer vision. AI allows the collection, processing and analysis of information from digital images and videos. For example, the US Postal Service uses machine vision methods to recognize handwriting on envelopes to automatically route letters.

- Fourth is natural language processing, which is essentially the digital interpretation of spoken or written language. For example, the Australian Government uses a virtual assistant called "Alex" to field general inquiries from the public about taxation. Alex can understand a normal, conversationally-phrased question and direct people to online resources that can assist them, reducing the demand on call centres to answer basic inquiries. As a result of implementing this same technology, Australia's intellectual property agency has reduced calls from 12,000 per month to 5,000 per month using Alex. [4]

Now let's take a look at Machine Learning. Machine learning is a subset of AI. It is the science of teaching computers to learn. It is focused on creating systems that are capable of learning by themselves and is used to make processes more autonomous, efficient, and effective. It has become so popular that it is now practically synonymous with AI.[5] Familiar devices like Siri, Alexa and Google Home are all powered by machine learning.

[3] Mehr, Hila, 2017: "Artificial intelligence for citizen services and government,"Ash Center for Democratic Governance and Innovation - Harvard Kennedy School., August: pp. 1-12.
[4] https://www.zdnet.com/article/ip-australias-alex-is-more-than-just-a-chatbot/
[5] https://www.brookings.edu/research/what-is-machine-learning/

In 2017, a research team at Carnegie Mellon University and the University of Pittsburgh were able to train a machine to analyze brain images and identify people with suicidal thoughts.[6]

Kansas City, Mo., has developed a machine learning algorithm to help predict when potholes will form on city streets.

Joshua Blumenstock, a researcher at the University of California, Berkeley, called a random sample of 1,000 residents culled from a database of 1.5 million mobile phone users in Rwanda. His team used what they learned from the phone survey to develop and train a machine learning model to predict wealth and applied it to the complete calling data from those 1.5 million users to create a detailed map of the wealth levels of the whole country. When mapped, his model very closely approximated the government's national Demographic and Health study previously created through manual surveys. However, Blumenstock's approach achieved these results 10 times faster and 50 times cheaper, paving the way to apply machine learning models to accelerate demographic research.

Fourth on our list of technologies is the category of tools known as Collective intelligence. Although it receives less attention than AI, collective Intelligence describes how groups of people and machines assemble in ways that lead to advances in intelligence.

Let's start with an example: Since being diagnosed with type 1 diabetes at the age of 14, Dana Lewis had to tolerate the lack of innovation by conventional medical-device firms.

When her dissatisfaction finally boiled over in 2013, she created an artificial, do-it-yourself insulin system that administers the right amount of insulin automatically. Later, she decided to make the technology available to all those with the illness who were willing to build their own system. The resulting Internet community now has 400 'DIY diabetics' who share readings online and collaboratively improve the device over time.

Just as members of the public contribute to writing Wikipedia entries or restaurant reviews on Yelp, they can and will share their collective intelligence to solve public problems. New technologies are making it even easier to tap into the expertise and intelligence of the crowd to add legitimacy and effectiveness to policies and service delivery. " In a world where technology allows for global collaboration...there are few barriers to getting the right people together to work on some of our most pressing problems.".[7]

Risk Map is an example of collective intelligence. As Hurricane season batters the East Coast of the United States, Risk Map helps residents and government officials track flooding in Broward County, Florida. Residents affected by flooding can add information to a publicly available map

[6] https://www.nih.gov/news-events/nih-research-matters/machine-learning-identifies-suicidal-youth
[7] https://sciencedisrupt.com/the-wisdom-of-the-crowd-is-what-science-really-needs/

via popular social media channels.  Officials can then assess damage and re-route traffic. Risk Map aggregates flooding reports from thousands of individuals and uses them to create a real-time map of flooding to aid in relief efforts in Indonesia, India and Florida for millions.

Of course, there are risks associated with disseminating information in such a manner, but let's save those risks for a later discussion.

Finally, the last technology we will explore is Blockchain.

Blockchain is what is known as a distributed ledger technology that can make it possible to store user identity and transaction information both more securely and more openly. With this technology, a single database is replaced by a large number of identical databases, each hosted and maintained by a different party. When changes are entered in one copy, all the other copies are simultaneously changed, making discrepancies transparent and enabling changes to be tracked to their source.

In government, blockchain is being used for projects that address regulatory compliance, contract management, identity management and records management. It offers a secure, inexpensive, and most importantly transparent method of conducting several government and personal affairs.

For example, because blockchain makes it transparent when records are changed,  is very useful for projects requiring individuals to authenticate their identity online. The 2016 Massachusetts Democratic State Convention used the blockchain based Voatz app to create a more secure voting platform.

Blockchain also makes it possible to see changes in records, which is why it has become so popular for use with land registries.

WISeKey and Microsoft are working together to provide the Rwandan Government with blockchain platforms that support identity verification practices - especially as they relate to the management of national land registry systems.

The Illinois Blockchain Initiative - a collaborative effort among a number of state and county agencies in Illinois - is piloting the use of blockchain to make it possible for a citizen to share his or her birth records with multiple entities seamlessly and safely. In its blockchain-based birth registration system, once a birth is verified, the record is stored in the blockchain and can be shared with businesses and government institutions via encrypted access, simplifying the process of applying for a driver's license or bank account.

The Ukrainian Land Registry is using a blockchain system application that protects auctions from black market controls and instead, offers ways to even out the land price slide and increase income for farmers.

Blockchain also enables self-executing contracts known as smart contracts, where the software automatically executes the transaction -- such as selling or buying -- when the conditions are met without the need for human intervention, thereby decreasing opportunities for fraud or bribery.

**Conclusion**

That concludes our whirlwind introduction to some of the new technologies being used by governments around the world to improve public services and solve public problems. As we've seen, the rapid growth in data, combined with the use of technology to aggregate and analyse data in new ways, opens up numerous opportunities for governments to be more efficient, be more responsive to residents' needs and plan for the future more effectively.

To further explore some of the technologies and concepts we've discussed today, make sure to check out the readings associated with this training module.

Also, we encourage you to do your own research as well - there's a plethora of content out there on the technologies we have discussed today, so keep exploring!