FEDInsider

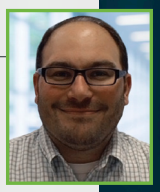
THE POWER OF DATA BEHIND ADVANCED ANALYTICS IN GOVERNMENT

Government, academia and private industry are partnering to get the most out of data by leveraging high-performance computing and artificial intelligence.

FEATURING:

> **Ron Thompson**
Associate CIO, Transformation & Data, NASA

> **Frank Indiviglio**
Acting Deputy Director, HPC, NOAA



> **Dr. Amanda Randles**
Assistant Professor, Biomedical & Computer Sciences, Duke University

> **Curt Smith**
VP & GPU Datacenter Architect, NVIDIA

> **Randall Berger**
Distinguished Chief Technologist, Hewlett Packard Enterprise

> **Chris Sexsmith**
Program Director, Data Science & Edge Computing, Red Hat

Billions of connected devices around the world are generating trillions of megabytes of data each day. As these massive datasets are transferred and stored in cloud data centers around the world, insights are drawn and used to change the way we do business.

How is this data being managed, and what advancements are being leveraged to solve the nation's most pressing challenges? Several data experts talked with FedInsider during a recent webinar to explore how the power of high-performance computing is being used for everything from climate-based modeling to biometrics.

The following are some of the most important aspects of those efforts.

Advanced Analytics at Work

In the biomedical space, high-performance computing is being used to advance the role of personalized blood simulations to study and treat cardiovascular disease. Amanda Randles, assistant professor of biomedical and computer sciences at Duke University, said she's working on the development of a code that takes clinical data from medical imaging like MRI scans to develop simulations of someone's vasculature, coronary arteries and blood flow.

These models can be used to diagnose a patient, explore a treatment option or predict possible procedure outcomes using non-invasive methods — and running them requires some of the biggest supercomputers in the world.

Similar model-based methods are also being used to understand long-term changes in climate and its impact on the Earth. Frank Indiviglio, acting deputy director of High Performance Computing and Communications for the National Oceanic and Atmospheric Administration (NOAA), said the agency is taking artificial intelligence and advanced analytics even further than climate-based modeling.

NOAA is focused on improving physical parameterization for weather and all Earth systems to enhance their computational performance and efficiency. Observational improvement is a big part of this area. The weather forecast is determined by observations collected through various means, like satellites and sensors on the field and in space. Getting the right data into the model at the right time provides a more accurate forecast, and AI can do quality control of those observations.

Tapping into the Power of Shared Data and Collaboration

Prior to the 2019 Executive Order on Maintaining American Leadership in Artificial Intelligence, Randall Berger, distinguished chief technologist at Hewlett Packard Enterprise, said the federal government was only at 1% utilization of AI and machine analytics.

The executive order mandated putting more federal funding into these areas, which helped implement machine learning, AI and analytics into government business and missions. The COVID-19 pandemic also accelerated advanced technology adoption, and agencies and industry are working together to share resources and solve critical problems.

The Energy Department's National Laboratories diverted a lot of resources with supercomputers and high performance computers last year to vaccine development, for instance. Berger suggests

federal utilization of AI could be "getting close to 50% where most government agencies are really relying on or looking at machine learning and AI today."

NOAA also works with agencies like the U.S. Geological Survey and NASA to advance scientific initiatives, collaborate on technological exchanges and optimize resources. "That keeps us current, that keeps our partners current, especially in a rapidly developing field like AI," Indiviglio said. "Our partners' innovations become the consortium, and we can share those across the agency divide."

Consortiums benefit cross-mission, too. During the ventilator shortage crisis in 2020, a team at Duke University created a brand-new device that could split a ventilator between two patients. In order for doctors to tune the ventilator to specific patients without the compute power needed for every simulation, Randles and her team worked with a consortium and partnered with Microsoft to run all their models on Azure. They completed 800,000 compute hours' worth of time to run every simulation they needed in a few days. Now, the simulation is ready for doctors to use.

The key to this was a science-based approach and interdisciplinary support. "We came in and said this is the scientific problem that we have. We don't care where we run it or how we run it, we want to get this done as quickly as possible," Randles said.

According to Curt Smith, Vice President and GPU Datacenter Architect at NVIDIA, this consortium effort can be replicated and is already being used by various branches of government as they begin to work together on implementing AI methods.

"There is more collaboration across different disciplines and part of that is going to be necessary in order to get ethical bias out," Smith said. "I'm seeing the world come together in different discipline areas, in science and in the military."

Overcoming Advanced Analytics Hurdles

A lack of trained data scientists along with limited budgets and resources can be a challenge for government agencies when advancing and implementing AI efforts. "A big take away and realization over the past couple years for many agencies is how crucial it is to understand the scope of the problem and the meaning of the data collected," said Chris Sexsmith, Program Director of Data Science and Edge Computing at Red Hat.

Sexsmith said false starts can consume a lot of an agency's resources – often, the perception that just because an agency has hoarded so much data, they can throw it all at a system at once to gain insights. "More often than not, you end up with faulty conclusions because you don't really understand that data," he said.

It's important to consider operationalizing AI models to avoid technical debt. First, understand the scope of the problem and how to attack it properly. Then, apply the technology and data needed. Smith said another challenge is processing the amount of incoming data. "It's difficult for both the bandwidth and the throughput of that data to keep up with the need to shorten the time to make decisions," Smith said.

Cleaning, sorting and ensuring the quality of incoming data is an operational challenge. Smith works with customers to implement large amounts of quickly accessible storage, advanced networking for computational and modeling purposes, and sound data governance to identify data bias.

How to Advance Computational Power

Smith said a culture of experimentation is needed to get reliable results from AI-driven data analytics, including the infrastructure to support the life cycle, the platform for accelerated computing and GPU networking and storage.

Organizations must also have the technical knowledge, an AI-trained workforce and a commitment to have data everywhere. "You need to understand the limitations. More data, more accurate data. Explainability. Ethical biases. You get all of that in the beginning," Smith said.

And a culture of experimentation must be integrated early into the cycle with incorporated user feedback and usability testing. "That culture of experimentation, you are going to make mistakes and you need to have the mindset of 'we are going to learn from our mistakes'," Smith said.

FEDInsider

Hosky Communications Inc.
3811 Massachusetts Avenue, NW
Washington, DC 20016

(202) 237-0300
Info@FedInsider.com
www.FedInsider.com
@FedInsiderNews
Linkedin.com/company/FedInsider/
@FedInsider

carahsoft

Carahsoft
14993 Sunset Hills Road, Suite 100
Reston, VA 20190
Contact: Michael Adams

(703) 871-8681
michael.adams@Carahsoft.com
Carahsoft.com/solve/machine-learning-artificial-intelligence/
Facebook.com/Carahsoft
Linkedin.com/company/Carahsoft
@Carahsoft