AGENCIES MAP OUT AN ADVANCED ANALYTICS FUTURE

Each new implementation reveals new possibilities for leveraging the power of data.

THE RAPID DEVELOPMENT of the advanced analytics field is being driven in part by technology and in part by imagination. Without a doubt, the last several years have brought significant leaps in the capabilities of tools and methodologies for capturing and analyzing massive amounts of data of all types.

Equally important, however, is the manner in which government leaders are beginning to understand the vast potential for advanced analytics in nearly every aspect of government operations and services. Chief data officers (CDO) and other data experts increasingly find their task is to help agency leaders imagine what is possible. Instead of saying, "Here's what advanced analytics can do for you," they are saying, "Tell me the problem you want to solve, and let me see how I can help."

"This is a people job, first and foremost," says Daniel Morgan, CDO at the Transportation Department, speaking at a June 30 event organized by the Advanced Technology Academic Research Center. "As a chief data officer, you just can't live behind your desk."

A CDO also can help agency leaders understand how much data they have on hand—and its value. For example, last year, Linda Powell, CDO at the Consumer Financial Protection Bureau (CFPB), had her team create a catalog of CFPB data for the bureau's internal stakeholders.

Rapid Response

At other times, however, the value of data becomes most apparent when the need is most pressing. That is the case right now with the fight against Zika, a mosquito-borne virus that has been spreading across the Americas. Public health officials hope advanced analytics can provide more insight into where and how it is spreading.

One such effort was the Austin Zika Hackathon, held in May. The event, hosted by Cloudera, assembled more than 50 data scientists, engineers, and students from the University of Texas Austin. The compute power was provided by Wrangler, a supercomputer at the Texas Advanced Computing Center (TACC).

During the hackathon, participants looked at how they might combine different data sets—outbreak reports, information on potential breeding grounds, social media feeds—to see what patterns might emerge, according to a TACC report.

"If you can see where all the water sources are and then overlay how the reports of outbreaks are happening, then you can create a model for how it's spreading and how it will spread in the future based on where the water sources are," says Ari Kahn, human translational genomics coordinator at TACC.

The summer flooding in Louisiana also brought a rapid response from data experts. Both NASA and the National Oceanic and Atmospheric Administration already collect and analyze data related to weather. As the situation unfolded, both agencies provided response teams with valuable real-time data analysis.

Last year, NASA created a rapid-response team to support emergency responses. This team works with officials at FEMA and other organizations to figure out what types of data might be most useful in a response (see GCN, "NOAA, NASA support flood response with data," Aug. 22).

However, government agencies also are increasingly interested in how advanced analytics could help address larger-scale problems, particularly in the social arena. For example, the Obama administration launched the Data-Driven Justice Initiative in June. This aims to help state and local agencies develop strategies for reducing the number of low-level, low-risk offenders sitting in jail cells.

Among other goals, the initiative aims to "[combine] data from across criminal justice and health systems to identify the individuals with the highest number of contacts with police, ambulance, emergency departments, and other services, and, leverage existing resources to link them to health, behavioral health, and social services in the community," the plan states.

More to Come

The recent surge in activity in advanced analytics clearly is due in part to advances in the technology itself, with each new generation of tools adding more capabilities while also growing easier to deploy and use. Expect that pattern to continue.

Another pattern has emerged as well: the "aha" moment. It often seems that once agency leaders see a compelling, real-life example of advanced analytics in action, their imaginations begin running wild.

In a recent report, International Data Corp., a market research and consulting firm, described big data and analytics as "game changers" for government agencies. Agencies, the report states, "need to effectively evolve their big data abilities as these capabilities are becoming increasingly critical to achieving mission outcomes."

ADVANCED ANALYTICS: THE ROAD AHEAD

The field of advanced analytics is likely to expand in ways agencies can't even yet imagine. Still, new and emerging tools suggest some important new directions.

ENTERPRISE DATA HUBS

When it comes to launching new initiatives, there's no reason to start from scratch. An enterprise data hub provides a reliable platform for managing operations and security of advanced analytics initiatives.

LOCATION-BASED INTELLIGENCE

Social media feeds have become important data sources in many advanced analytics initiatives. In many cases, the value of social data increases significantly when enriched with user-defined location data.

BIG TEXT

From the early days of big data, experts have always discussed the value of mining text, but early solutions often failed to deliver meaningful results. A new generation of technology and methodology has begun to deliver on that promise.

VISUAL SEARCH

Images and video often contain a wealth of information that goes untapped because of the sheer labor involved in sifting through all that material. That is changing, thanks to tools that automate the process of analyzing, indexing, and tagging visual content.

AUTOMATED DATA PREPARATION

The intensive work needed to prepare data for analysis usually requires a data expert, which often creates a nearly insurmountable logjam. The more that process can be automated, the more quickly an agency can get initiatives off the ground.

DATA INTEGRATION

Advanced analytic initiatives are often stymied because potentially valuable data is spread across so many different silos. Enterprise-level NoSQL provides a means of storing, querying, and searching a wide array of structured and unstructured data.

OPEN-SOURCE SOLUTIONS

Solutions based on open source provide agencies with a way of building on cost-effective, proven technology without getting locked into proprietary technology.

IN-MEMORY DATABASE

The sheer volume of data involved in today's advanced analytics initiatives can overwhelm traditional compute platforms. An in-memory database overcomes that obstacle, improving performance while reducing the data footprint and hardware and operational costs.

MACHINE-GENERATED DATA

Many agencies don't realize the value or the volume of data generated by their IT infrastructures. This data can provide unparalleled insight into the performance and security of the IT enterprise.