COVID Times Call for COVID Measures

With AI models rendered useless by the pandemic, data jocks are scurrying to make sense of a changed world. CIOs, meanwhile, need to walk a tightrope between facilitating all the new fact-finding and maintaining security and governance to keep data – as well as the company and its customers – safe. That's easier said than done in normal times.

by Mike Feibus

How do you model a world turned upside down? That's the existential challenge thrust upon those in business intelligence and data science, two intertwined spheres tasked with taming streams of data to help their organizations make sense of the present and predict the future. For them, big data has never looked so small.

So what do you do when your once-bankable dashboards and predictive models throw up their virtual hands?

Broadly, it means more data. And more models. More rapidly. For the moment, at least, that means days filled with:

- Hustling to create new dashboards and algorithms to solve problems, the likes of which weren't ever considered as recently as March.
- Crafting and revising forecasts for what tomorrow will bring based on an algorithm-busting history of just a few months.
- Scurrying to identify and lock down security holes that new data and machine learning relationships can cause.
- Amidst all the frenetic activity, it is the CIO who must weigh the benefits of expanded data access with its impact on security and governance. Because any fines, lawsuits and data hacks that might one day come out of this will be yours – and yours alone – to clean up.

More to Come

And in the long run, data jocks can expect more of the short run. Because the world isn't going to one day right itself. Which means that every passing day is adding a new round of data that is helping to tune the dashboards and models that will drive business and government decisions going forward.

"We've all seen the many charts, with five years of very clean historical data, and then a massive drop when the world changed," Christian Kleinerman, Senior Vice President of Product at cloud data platform provider Snowflake, told me. "So now all forecasts, all predictive models, all machine learning models, all of those are all out of whack. Because the technology is based on predicting what's going to happen based on what's happened.

"I don't know that anyone could have predicted that all predictions would be so far off."

By upending data usage patterns, Kleinerman said, the crisis has highlighted the benefits of the cloud's flexible pay-for-what-you-use model over on-premises data centers for many customers, and accelerated migrations. And as customers mix and match data sets stored in different public clouds, they are more interested in performing workloads on platforms that can access data across data platforms, he said.

Ali Ghodsi, CEO of analytics platform provider Databricks, agreed, adding that AI is every bit as important to customers since the pandemic, though the priorities for it are changing.

"When we talk to our customers, the CFOs are looking for ways to cut their budgets," he said. "So now, the focus is more on using AI, machine learning, and automation to save money."

Wendy M. Pfeiffer, CIO at Nutanix, seconded that assessment. "Companies are trying to preserve cash and do more with less," she said. "My peers running global IT organizations are actively adopting machine learning tools in order to increase the effective capacity of their teams and improve productivity."

On the revenue side, Ghodsi said, AI used to be more about growing the business, whether by expanding existing customers' spend or by finding new business. Now, companies are more intent on building models to hold onto what they've got.

"They're worried they're going to lose business," he said. "So with machine learning, they can actually see, based on all the data they have, which customers are likely to churn. So they can reach out and be more proactive with them."

Fleet of Foot
If there's an underlying imperative to all this, it's speed. "Speed and agility, that's definitely one of the trends we're seeing in the age of COVID," said Arun Ulag, Vice President of the Business Intelligence Platform at Microsoft. "Clarity at this time is critically important because businesses are having to make decisions very, very quickly."

Ulag pointed to Swedish Health Services, a nonprofit health system in the Seattle area, which quickly developed an app with assistance from Microsoft to help assess the impact of COVID-19 on resources, and quickly respond.

The app, called the COVID-19 Emergency Response App, or CERA, gave front-line workers a way to quickly report information like the number of patients, the percent of COVID-19 patients, hospital bed utilization and emergency room crowds. The information all rolled up into dashboards that decision makers could use to manage staff, supplies and care between the system's facilities, including five hospitals and two standalone emergency departments.

"These tools they created on the Power platform in a matter of days, not weeks or months."

COVID Times Call for COVID Data

It's not just healthcare providers that are pining for virus-related information. The pandemic is impacting just about any pursuit you can imagine, from how much fertilizer to stock to how many fans are safe to admit to a ball game. Which is why virtually every public and private entity is trying to get its virtual hands on as much data as it can.

Many cloud-related service providers are offering customers access to collections of publicly available virus-related data from sources like Johns Hopkins University and the World Health Organization. Databricks' collection, for example, is available to the public. Snowflake has integrated Starschema's collection of databases and is making that available for free.

"At this point, it's the most requested data set we've ever published," Kleinerman said. Out of about 3,500 customers, Snowflake has had more than 2,000 requests for the information.

Cloud platform provider Cloudera is also making publicly available health data accessible to its customers. As well, it's working with visual analytics company SynerScope to provide what they call a "Virus Operations Center." Architected in the spirit of a security operations center, the platform culls an impressive array of disparate data sets to help human resources and operations decisionmakers keep employees safe while maintaining workflow as much as possible. KPMG has said it is using the application with its own customers.

Unfamiliar Minefield

The pandemic is making for strange bedfellows, at least from an informational perspective. Data jocks are comingling heretofore unrelated data sets – like worker skill sets and infection rates by ZIP code, for example – into fresh models for special-purpose, pandemic-related dashboards. And it's happened at a rate that has frightened some CIOs, CISOs and other IT decision makers tasked with safeguarding digital assets.

"I was talking to a customer today, and he said he stopped counting how many COVID-related dashboards his company has spun up after it passed 4,000," Anupam Singh, Chief Customer Officer at Cloudera, told me. "And there are upwards of 150,000 people logging into some of the dashboards.

"It's a CIO's nightmare."

"This was sprung on them very quickly," Singh continued. "And as they've opened the aperture for more people to see the data, they've had to completely rethink how to secure it."

Singh said customers find it is easier to safeguard data on a full suite than for teams that built mix-and-match platforms with different providers for real-time data streaming, warehousing and machine learning.

"In a DIY-type setup, one vendor changes their code, and it impacts the security and governance you put in place for the whole system," he said. "It's much more difficult and time consuming to maintain, at a time when teams just don't have the bandwidth."

Business As Unusual
Just when it started to feel as though dashboard builders had tamed big data to help make organizations safer and more efficient, the pandemic rendered many seemingly omniscient models useless. And though it may have the urgent, transient feel of a firefight, piecing together new models based on novel interdependencies while executives demand instant insights is the pace of things to come.

For CIOs, that adds an extra burden as the arbiter of access, security and governance. So be sure your platforms are robust enough to sustain all the jostling. Because when the dust settles, and business intelligence and data science have made sense of this post-pandemic world and moved on, you'll be the one left behind to sweep up any customer lawsuits, regulatory fines and data hacks.

Because while data might seem unusually small these days, your responsibility to safeguard it is as big as ever.

FT

CIO Columnist Mike Feibus is principal analyst at FeibusTech, a Scottsdale, Arizona, market strategy and analysis firm focusing on mobile ecosystems, client technologies, AI and data management. Reach him at mikef@feibustech.com. Follow him on Twitter @MikeFeibus.

Reprinted on behalf of Cloudera, Inc.