Data Management is Vital for AI Implementation, Health Officials Say

Healthcare industry is adopting AI tools and techniques thanks to data accessibility, but to truly reap the benefits, data needs proper governance.

Amanda Ziadeh

Thu, 12/13/2018 - 17:37

Roselie Bright, George Chambers, Sezin Palmer and Dr. Don Rucker speaking on the health panel at the CXO Tech Forum: AI and Big Data in Government Dec. 12.

Artificial intelligence is impacting the healthcare industry in a big way, from provider to physician and patients alike. And though talks of AI and health IT aren't new, the vast amounts of computing power and data generation today means more accuracy and intelligence in algorithmic models.

The healthcare industry is taking advantage of that amount of data. "Not only are we able to measure many more things with many different modalities, but we're

able to measure it to levels of resolution that we could never achieve before," said Sezin Palmer, mission area executive for national health at Johns Hopkins University Applied Physics Lab. She spoke on a panel at the Dec. 12 [CXO Tech Forum: AI and Big Data in Government](#) in Arlington, Va.

Palmer sees a huge shift in the way data availability has impacted areas like neural networks, deep learning and AI in the ability to answer questions beyond "yes" and "no."

"You're able to look across varying complexities and look at a multidimensional state and try to pull out patterns around disparate data centers," she said.

Incoming "Big Data" and AI Applications

Data availability is also helping the Food and Drug Administration expedite approvals for drugs and medical devices.

In November, the FDA issued a "[Prescription Drug-Use-Related Software](#)" notice, which announced a docket asking for public input on the agency's proposed framework for prescription drug-use-related software. The FDA is hoping this approach brings in a software that won't require FDA review before dissemination.

And prior to this notice, the FDA initiated the Digital Health Software Precertification Pilot [Program](#), intended to "inform the development of a regulatory model to assess the safety and effectiveness of software technologies without inhibiting patient access to these technologies."

With this program, the FDA envisions a future with a more streamlined regulatory oversight model of software-based medical devices developed by manufacturers that demonstrate quality, excellence and a commitment to monitoring real-world performance of their products on market. This regulatory model will focus on software developers first, rather than the product.

According to Roselie Bright, epidemiologist in the FDA's Office of Health Informatics, this prescription program is "a little bit of a pilot," she said in the panel with Palmer. FDA is still exploring how this would work, as questions remain about statutory authority.

"The idea is . . . if a company is mainly just producing software that's its medical

product, can we trust that if they have good processes in place that they're going to be good actors, that their software is going to be, basically, reliable and not hurt people?" Bright said, adding the company should take initiative if something is detected as going wrong.

Bright said the FDA questioned whether a company can be left alone to put software on the market based on the monitoring and requirements the FDA and its partners develop during this pilot.

This is still in its beginning stages, Bright said, but in terms of using "big data" for pre-market applications for any kind of product, an initiative called [real-world evidence and real-world data](#) would come into play.

Also in its early stages, the initiative is used to "monitor postmarket safety and adverse events and to make regulatory decisions." This data is used to help guideline decision support tools used in clinical practice, and medical product developers use it to support clinical trial designs.

But Bright said the FDA is still thinking about extending the current use of the initiative for supporting randomized clinical trials.

George Chambers, deputy chief information officer of the Health and Human Services Department, sees the evolution of AI and data availability from an infrastructure and operational standpoint. But as agencies begin to deploy tools and robotic process automation in major applications, Chambers said to step back first and ensure the proper configurations and testing are in place.

"I change applications, I have a management processes that make sure, similar to interfaces — so as they proliferate, how am I making sure that my configurations are all still in place, that I've tested it appropriately, and that I can deploy it in a major production environment that have multiple systems all potentially using these AI in all these different ways?" Chambers questioned in the panel.

And the challenge right now is that there are no control functions or software tools that allow an operational group to manage all the incoming proliferated applications.

But First, Data

AI provides visualization and predictive analytics tools that can be applied to an

application or data set. But without data integrity, the outcomes won't be useful or accurate.

"Is the same data element coming from this database comparable in terminology to this one over here?" Chambers questioned. "AI can help us do that, but once again, these are the challenges from an infrastructure and support standpoint, especially with an organization 88,000 strong, like HHS."

In addition to challenges in data management and governance, there's the challenge of structured and unstructured data.

Palmer and the APL's partners at Johns Hopkins Medicine focus on how best to leverage all the unstructured data that is now in electronic health records and on how to pull useful data from the electronic health record and combine it with other datasets, like imaging, genomics data and wearable tech data.

An area Palmer and APL are working on is natural language processing and generalization. Meaning, if an algorithm is developed to pull one piece of data, it can be trained to work well. But can that algorithm be taken, with minimal customization and put to work against a completely different dataset?

The APL has done this with an algorithm originally focused on prostate cancer. After a few tweaks to the base algorithm, it was made to work against some parameters of critical importance for multiple sclerosis.

Dr. Don Rucker, HHS' national coordinator for health IT, said structured and unstructured data have different uses. "If you want to structure data, you have to pay somebody to structure it, if it's not coming from a machine," he said on the panel.

The challenge also comes from being able to look at medical data documents as a whole, which shows core data elements, notes and machine-generated data.

HHS has been working on getting the notes coordinated into a different system through a process called the [Clinical Document Architecture Standard](#).

The department is also working on getting those notes and documents to reflect what happened with a patient. "If you actually look at the notes, what is not maybe immediately obvious, if you're not in healthcare, is most of the information and notes is actually templated nonsense," Rucker said, like billing documents. So,

ensuring those notes are meaningful patient information would be key, especially as that information becomes accessible and available to open APIs.

This, of course, comes back to data quality. And in order for the healthcare industry to truly benefit from the potential of AI and machine learning, the availability of data, data management and governance, and data integrity will be vital.

[View printer friendly version](#)
[HHS](#)
[data](#)
[AI](#)
[robotic process automation](#)
[health](#)
[CXO Tech Forum Eye AI Big Data Government](#)