VA Looking to Expand Usage of Artificial Intelligence Data

The agency is looking at how to best apply curated data sets to new use cases.

Adam Patterson

Fri, 03/27/2020 - 12:18

Photo Credit: iStock/aislan13

The Department of Veterans Affairs is closer to expanding its use of artificial intelligence and developing novel use cases.

In looking back on the early stages of the VA's newly launched artificial intelligence program, the department's Director of AI Gil Alterovitz noted ongoing questions about how to best leverage AI data sets for secondary uses.

"One of the interesting challenges is often that data is collected for maybe one reason, and it may be used for analyzing and finding results for that one particular reason. But there may be other uses for that data as well. So when you get to secondary uses you have to examine a number of challenges," he said at AFCEA's Automation Transformation conference.

Some of the most pressing concerns the VA's AI program has encountered include questions of how to best apply curated data sets to newfound use cases, as well as how to properly navigate consent of use for proprietary medical data.

Considering the specificity of use cases, particularly for advanced medical diagnostics and predictive analytics, Alterovitz has proposed releasing broader ecosystems of data sets that can be chosen and applied depending on the demands of specific AI projects.

"There's a lot to think about data sets and how they work together. Rather than release one data set, consider releasing an ecosystem of data sets that are related," he said. "Imagine, for example, someone is searching for a trial you have information about. Consider the patient looking for the trial, the physician, the demographics, pieces of information about the trial itself, where it's located. Having all that put together makes for an efficient use case and allows us to better work together."

Alterovitz also discussed the value of combining structured and unstructured data sets in AI projects, a methodology that Veterans Affairs has found to provide stronger results than using structured data alone.

"When you look at unstructured data, there have been a number of studies in health care looking at medical records where if you look at only structured data or only unstructured data individually, you don't get as much of a predictive capability whether it be for diagnostics or prognostics as by combining them," he said.

Beyond refining and expanding these data applications methodologies, the VA also appears attentive to how to best leverage proprietary medical data while protecting personally identifying information.

The solution appears to lie in creating synthetic data sets that mimic the statistical parameters and overall metrics of a given data set while obscuring the particularities of the original data set it was sourced from.

"How do you make data available considering privacy and other concerns?" Alterovitz said. "One area is synthetic data, essentially looking at the statistics of the underlying data and creating a new data set that has the same statistics, but can't be identified because it generates at the individual level a completely different data set that has similar statistics."

Similarly, creating select variation within a given data set can serve to remove the possibility of identifying the patient source, "You can take the data, and then vary that information so that it's not the exact same information you received, but is maybe 20% different. This makes it so you can show it's statistically not possible to identify that given patient with confidence."

Going forward, the VA appears intent on solving these quandaries so as to best inform expanded AI research.

"A lot of the data we have wasn't originally designed for AI. How you make it designed and ready for use in AI is a challenge and one that has a number of different potential avenues," Alterovitz concluded

[View printer friendly version](#)
[Veterans Affairs](#)
[Artificial Intelligence](#)
[Data Analytics](#)
[Standard](#)