

## Machine Learning, Text Analytics Aid in Food Safety at FDA

The emerging technologies can predict, anticipate and protect against dangerous chemicals in the U.S. food supply.

[Faith Ryan](#)

Tue, 09/03/2019 - 04:34



Photo credit: monsitj/iStock

A new automated data analytics program is crucial for the early detection of signals and predications for regulated chemicals that may pose highly hazardous health risks at the Food and Drug Administration.

The agency's Center for Food Safety and Applied Nutrition first initiated the project, called the Emerging Chemical Hazard Intelligence Platform. It allows the center to anticipate potential chemicals associated with adverse health events before they get out of control, explained its Office of Food Additive Safety's Informatics and

Information Systems Lead and Senior Policy Advisor Ernest Kwegyir-Afful.

"Every time we have one these big [food safety] incidents, we have to drop everything so we can actually deal with it," said Kwegyir-Afful at SAS' Unleash Analytics: Making AI & Analytics Real event Aug. 20. This includes U.S. food supply chemical incidents, such as detecting products that increase the production of melanin in babies to measuring arsenic toxicity, he said.

The center is responsible for tracking the significance of exposure to chemicals or the level of consumption of each chemical as well as tracking what foods are currently in the market. It also analyzes whether there are contaminants in the food supply. (Food is classified based on its chemical structure and the interactions with other chemicals and the human body.)

Instead of reacting to regulated chemical outbreaks after they occur, the center adopted its intelligence platform to anticipate through text analytics to discover chemical risks earlier in the signal lifecycle and to better manage potentially harmful public health events.

In preparation of the solution, the FDA established a Food Advisory Committee comprised of toxicologists, chemists, public health officials and other experts to look at the different data streams of information that could be used to create predictive models and train them using human-curated data. These models would determine which signals indicated health risks in the supply chain. The majority of the data streams pulled were text-based sources, such as current event news articles about chemical spills and scientific journal articles, to individual reports that appear unrelated except for the individuals all having consumed the same food product.

In addition to requiring a platform that could analyze text, the team needed a system to compile predictive analytic results, identifying patterns of events that occur over a period of time and create those into visualizations that would be simpler to understand for those without a data analytics background — text analytics paired with machine learning was the end-to-end solution.

Once these data analytics tools were tested, they helped reduce time-consuming processes and were able to find critical patterns in the unstructured data, explained Kwegyir-Afful. The automated, visualized information also help public health officials make well-informed, data-driven decisions.

However, there are big challenges to implementing machine-learning solutions, Kwegyir-Afful said, such as garnering C-level support, a lack of confidence in decision-making from derived machine-learning solutions and attracting data scientists under the current federal general service pay scale.

Still, Kwegyir-Afful suggested officials start from the top down.

“Find champions higher in the hierarchy,” such as CIOs or directors, he said, and clearly explain the problem and solution.

He also recommended to look for an already established process and figure out a way to make it more efficient. For instance, he said, could a process that takes 100 hours be reduced to anything lesser? Lastly, start small with a problem you can solve in six months or less to prove it is possible and gain trust and appropriate executive-level backing.

[View printer friendly version](#)

[data](#)

[FDA](#)

[machine learning](#)

[predictive analytics](#)

[emerging tech](#)

[data analytics](#)

[Health IT](#)

[Standard](#)