WHITE PAPER

Delivering Mission Impact

# AI-Enhanced Risk Profiler “Sentinel”

## Revolutionizing FDA Inspections in the Prior Notice Process

*Andrew Zeswitz, Jeff Myers, Ali Mirzaeian*

**DECEMBER 2023**

## TABLE OF CONTENTS

***Executive Summary, 3***

***New Potential Uses of Artificial Intelligence to Improve the Safety of Food and Drugs Used by Americans, 3***

***Summary of Our Solution Concept: "Sentinel", 4***

***Setting the Context, 4***

***Solution Concept Details, 5***

***Technologies Used, 8***

***Data Sources, 9***

***Results of our Model and Benefits of AI, 9***

***Conclusion, 10***

***About the Authors, 10***

## Executive Summary

The US Food and Drug Administration (FDA) has a rapidly growing need for improving the inspections process given the increasing volume of regulated products being imported to United States. Thus, FDA's inspection program could benefit from improved targeting, to strengthen deterrent and preventive effects. Consequently, the FDA hopes to improve inspection results by using more, different, and better information and data to enhance the efficiency and effectiveness of its inspection program amidst expanding workload.

FDA's Division of Food and Defense Targeting (DFDT) currently conducts inspections based upon:

- Risk associated with a product
- Product history (past violations)
- Manufacturer, shipper, importer history (past violations)
- Routine surveillance

## New Potential Uses of Artificial Intelligence to Improve the Safety of Food and Drugs Used by Americans

FDA can leverage Artificial Intelligence (AI) to better identify risky products, manufacturers, shippers, and importers, empowering the FDA to strategically allocate time and resources to areas of concern, and areas where inspection efforts can best deter and detect products that may harm consumers. That will allow the FDA to maximize the impact of regulatory efforts in the pursuit of public health and safety.

REI Systems' hypothesis is that AI can dramatically improve the targeting and impact of FDA's regulatory inspection efforts, both from using data already held by the FDA, and by leveraging data available from outside sources. We believe that AI can significantly improve the effectiveness of FDA's constrained inspection resources.

The purpose of this white paper is to describe the approach we are taking, and the results we anticipate. Specifically, REI Systems is exploring four areas that we expect will bring the benefits of AI to support the FDA's mission.

**01** **Decision Support:**
Experimenting to supplement human judgement with machine learning as applied to data the FDA already holds and collects regarding manufacturers, shipments, product lines, etc.

**02** **Media Scraping:**
Scraping social and published media to identify concerns/risk associated with products, manufacturers, countries/regions, etc.

**03** **Harvesting and Analyzing Partner Data:**
Gathering data (i.e., black-lists, inspection failures, etc.) from state departments of health and agriculture, foreign government drug inspection agencies, US Customs and Border Protection (CPB), USDA's Animal, Plant and Health Inspection Service, etc.

**04** **Analyzing Data Shared by Cooperating Producers:**
Cooperating, preferred producers, and manufacturers may provide access to data from their own databases and quality control mechanisms which the FDA could use to help inform the agency's inspection and action priorities.

We developed a conceptual model for each of these four, but the first two are furthest along in the development lifecycle. Thus, today we are prepared to share architecture for that opportunity, along with a demonstration (**using synthesized data**), and wireframes for how they would work. Upon request, we would be happy to discuss our concepts and timelines for when a "proof of concept" for the remaining three options could be demonstrated.

## Summary of Our Solution Concept: "Sentinel"

REI Systems' AI Predictive Solution, called "Sentinel," will harness FDA and other public data sources to assess every product line consumed in the US, helping to detect unsafe products. In initial phases of our AI proof of concept, REI Systems' AI model uses 4-6 critical data points from FDA's Prior Notice system, to classify products as high, medium, or low risk and can add publicly available information (initially from media outlets) to help determine a risk profile. A subsequent third phase will seek to add data from partners and from cooperating producers – all of which will add detail and precision over time, with feedback from inspection results being leveraged to strengthen the effectiveness of FDA inspections and enforcement by further training and strengthening capabilities of the AI predictive model.


Sentinel will aid the FDA in pinpointing high-risk products, aspiring to at least a 90% accuracy rate (measured by F-1 score). This will empower the FDA to more efficiently prevent hazardous products from harming Americans and will help the FDA to allocate its inspection and enforcement resources on the products that pose the highest risk to Americans' health.

## Setting the Context

Under provisions of the Food, Drug and Cosmetic Act (the FDC Act), importers of food products intended for introduction into U.S. interstate commerce are responsible for ensuring that the products are safe, sanitary, and labeled according to U.S. requirements. In addition to meeting the requirements of U.S. food regulations including food facility registration, importers must follow U.S. import procedures as well as the requirements for **Prior Notice**.

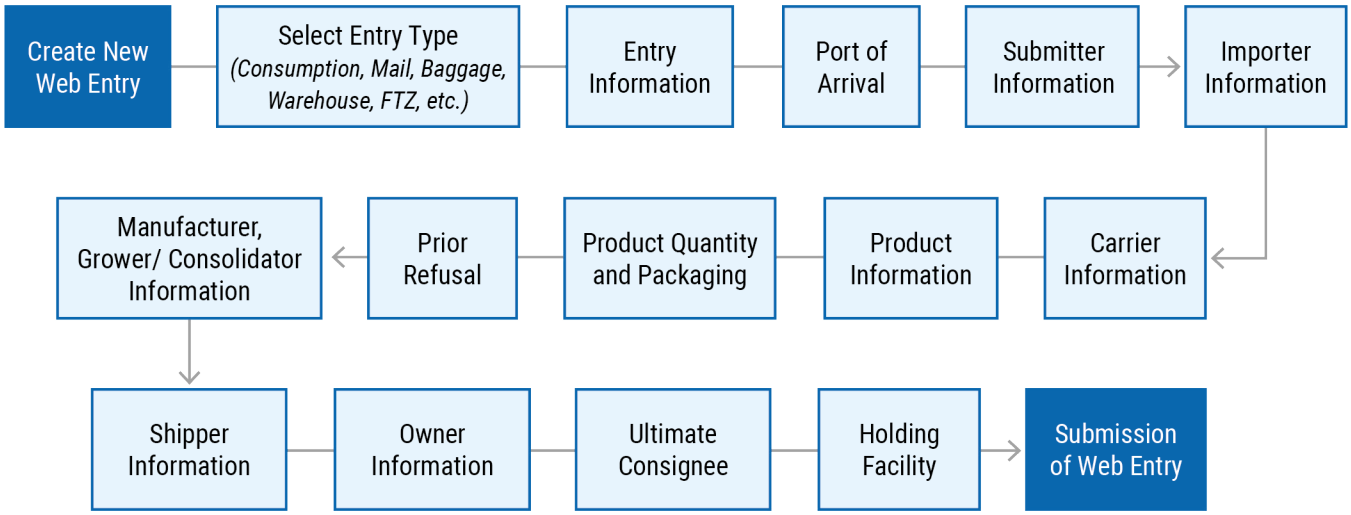
Along with other provisions, the FDC Act requires that producers and manufacturers provide the FDA with prior notification of products offered for import into the United States. Advance notice of import shipments allows FDA, with the support of the CBP, to target import inspections more effectively and help protect that nation's food supply against terrorist acts and other public health emergencies.

## CURRENT FDA INSPECTION PROCESS

Individuals or companies can choose to submit Prior Notices directly through the Prior Notice System Interface (PNSI), or through the Automated Broker Interface (ABI) to the CBP.

Prior Notice includes information about the product, quantity, and packaging, and related facilities, such as the manufacturer, shipper, owner, and ultimate consignee. Information required varies by entry type.

**Figure 1: Prior Notice Submission Steps**




# Solution Concept Details

In Phase 1 of our AI Proof of Concept, we focus on the individuals and companies looking to import goods into the United States. Our solution draws on relationships between various entities like owner, manufacturer, shipper, filler and assign a score based on various data sets which can be accessed internally within the FDA. (Eventually, in subsequent phases, we will rely on additional datasets from outside the FDA.)

Our solution for Phases 1 and 2 breaks down scores into further granularity, including:

1

**PHASE 1:** *Initial Proof of Concept (Dec 7)*

1. **Track Score** (leveraging FDA's historical records about entities)
2. **Media Score** (based on negative mentions in social media and published media)

**2** PHASE 2: *Refined Proof of Concept (Q1 2024)*

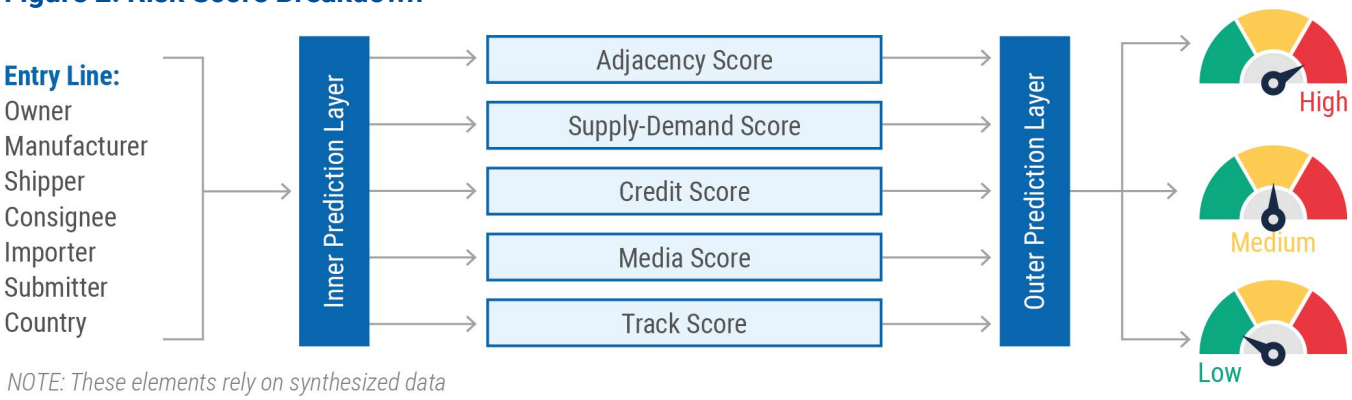
3. **Adjacency Score** (proximity/affiliation with entities that have previously experienced adverse inspection results)
4. **Supply-Demand Score** (an otherwise low-risk entity's reliance on production inputs from a high-risk supplier)
5. **Credit Score** (obtained from third party credit reporting agencies)

**3** PHASE 3: *Anticipated Future / Aspiration (Timeline TBD)*

6. **Partner Data** (i.e., harvest, analyze and incorporate blacklists, inspection failures, etc. from CBP, USDA APHIS, state departments of health and agriculture, and foreign government agencies)
7. **Data shared by trusted, cooperating producers** (from databases held by food and drug producers that FDA has determined to be trustworthy)

The final risk score is obtained using a two-layer predictive model. In the first layer, five types of risk scores are generated, and these scores are then used in the second predictive model to generate the final risk score that will be high, medium, or low:

**Figure 2: Risk Score Breakdown**

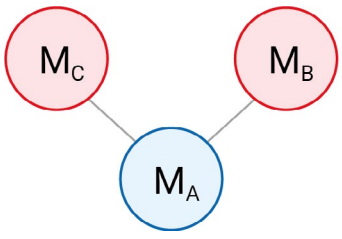



*NOTE: These elements rely on synthesized data*

## ADJACENCY SCORE

The FDA can benefit by establishing relationships between different entities of the same type (i.e., manufacturers in these cases). For example, in Figure 3, three manufacturers are shown. Manufacturer A ($M_A$) has a low-risk score, while both Manufacturer C ($M_C$) and Manufacturer B ($M_B$) have high-risk scores. In the current risk score implementation, only $M_A$ is considered individually. However, in the adjacency score, both $M_B$ and $M_C$ also influence the risk score of $M_A$ because the manufacturers are adjacent based on corporate ownership or geography.

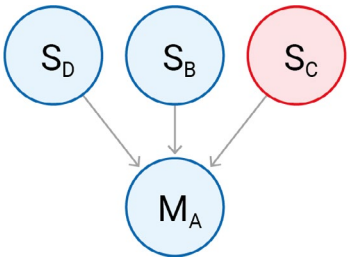
**Figure 3: Illustration of Adjacency Risk**



## SUPPLY-DEMAND SCORE

Another risk factor that is not considered in the current risk profile is supplier relationship structures. This is a specific case of adjacent entities (explained above) where they are in a supplier-procurer relationship within the supply chain. Figure 4 illustrates, showing three suppliers, C, B, and D, that support production by manufacturer A ($M_A$). $M_A$ has a low risk if we don't consider the risk of its suppliers. However, if we consider the supply demand network, the fact that it relies upon known high-risk Supplier C changes the risk score of $M_A$.

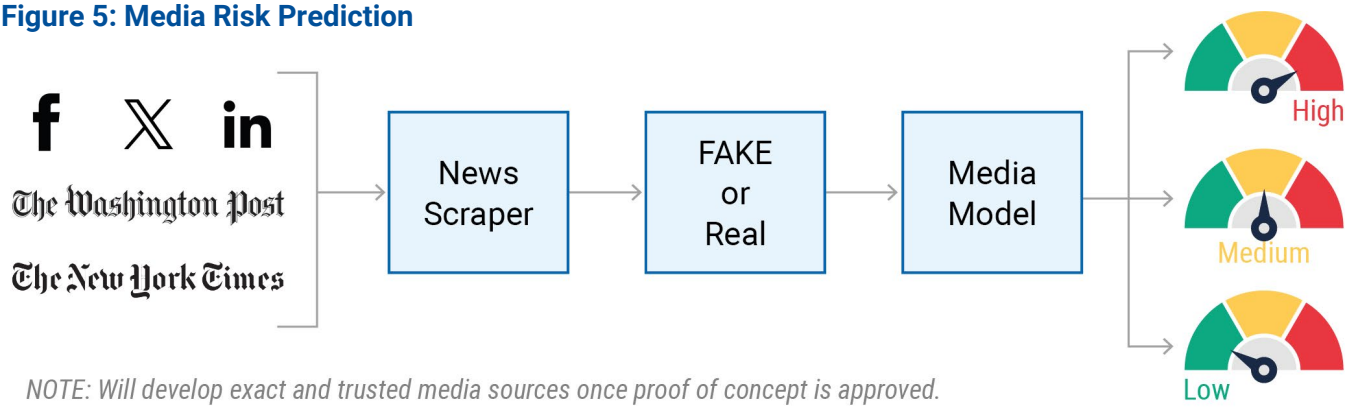
**Figure 4: Illustration of Supply-Demand Risk**




## MEDIA SCORE

Social media (LinkedIn, Facebook, and X formerly Twitter), as well as traditionally published media (Washington Post and New York Times) can provide the first reports of a problem with a product. Thus, our AI model will scrape social media and publicly available published media reports that mention manufacturers/producers and products, using sentiment analysis to determine if the mention is positive or negative. Our AI model will use machine learning to refine the frequency and character of mentions that are associated with inspections that find non-compliant or harmful regulated products.

**Figure 5: Media Risk Prediction**




*NOTE: Will develop exact and trusted media sources once proof of concept is approved.*

## CREDIT SCORE

A third-party credit score for the manufacturer/producer will help identify entities which are under financial duress, and which thus may be more inclined to take shortcuts in procuring ingredients, in good manufacturing practices, in quality control, or other facets of manufacturing/production. Our AI model uses machine learning to draw on this data, and to explore relationships between credit scores and other FDA data.

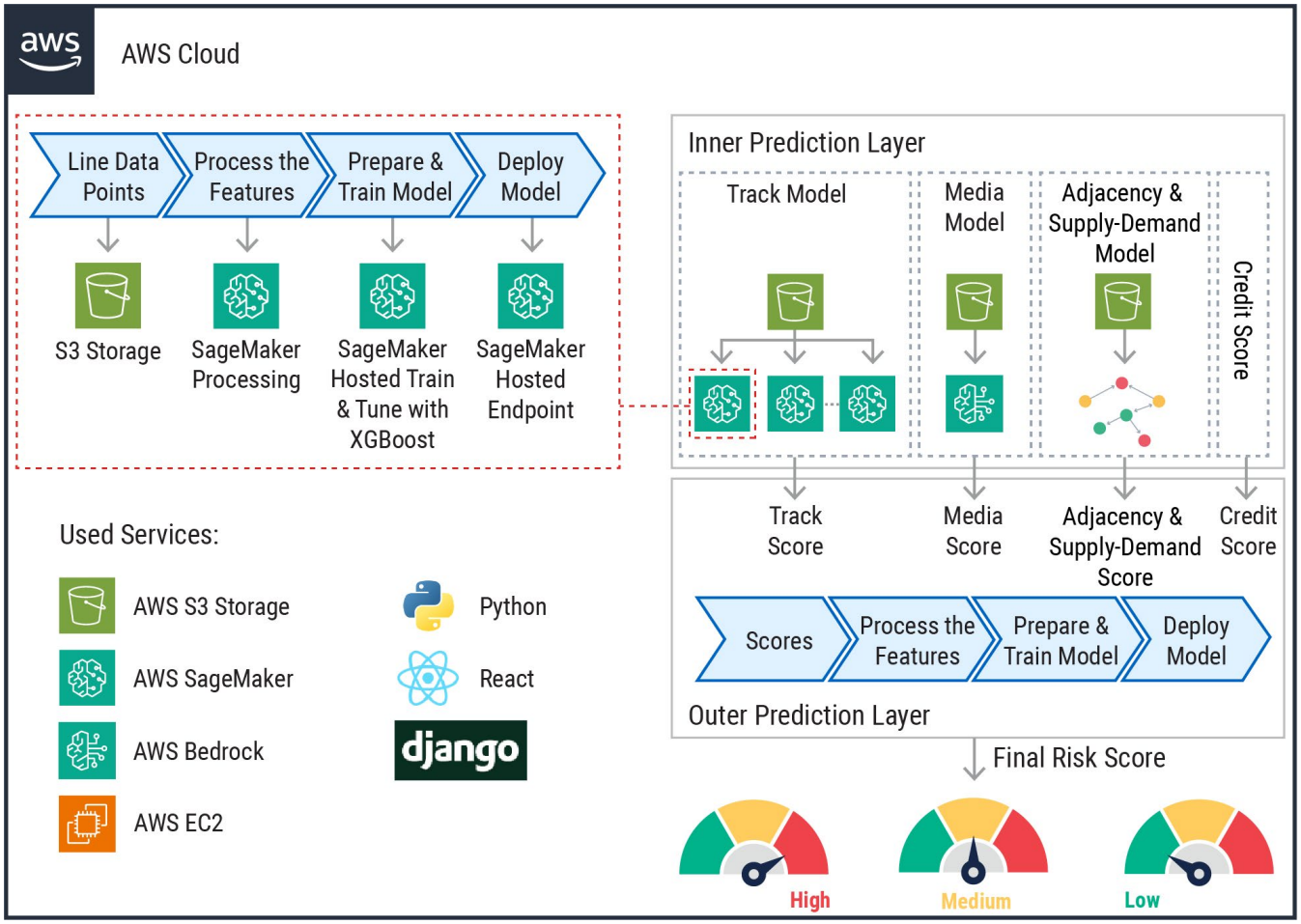
## TRACK SCORE

Leveraging the historical records of each entity, our model assesses and assigns a tracking score. For instance, if Manufacturer A has a documented history of importing products without encountering any issues, it is deemed to have a lower risk compared to Manufacturer B, which lacks any such recorded history. This predictive approach allows us to evaluate and categorize entities based on their past performance, offering valuable insights into potential risks associated with each.

## FINAL RISK SCORE

Figure 6 below summarizes the pipeline of our approach to predicting the Final Risk Score for each input line:

**Figure 6: Media Risk Prediction**




**The Final Risk Score indicates whether FDA inspection is warranted for a regulated product coming in as an import for American consumption.**

## Technologies Used

REI Systems has used the following technologies to date to develop our proof of concept (POC).

- AWS EC2 Instance
- AWS S3
- AWS SageMaker
- AWS Bedrock
- Python Programming
- Django Framework
- React Library

## Data sources

For our initial AI model, we used the following data sources:

- **Synthesized Data to Approximately Replicate Data from the FDA's Prior Notice System**
  Please note that we have synthesized the FDA Prior Notice data to use in creating and testing our model so that it can be publicly demonstrated, examined, and improved. We would be pleased to collaborate with the FDA to use real data in place of the synthesized data.
- **Social Media Data**
  Our media scraping tool currently ingests data from LinkedIn, Facebook, X, the Washington Post, and the New York Times. In collaboration with the FDA, we would be happy to expand media data sources.
- **Data.gov**
  Sentinel can use multiple data sets published by other federal, state, and universities to build the knowledge and trends that can help improve risk assessments. Some of the key dataset that can be leveraged from the Department of Agriculture (US Agriculture Trade Data, US Food Imports dataset, and International Food Security), the Department of Commerce for trade related information, the Department of Justice for foreign entities and restrictions, the Department of State for advisories, and world events.

## Results of our Model and Benefits of AI

### HIGHEST POSSIBLE ACCURACY RATE FOR IDENTIFYING HIGH RISK IMPORTS

Should the FDA decide to collaborate with REI Systems to evaluate and improve our AI model, using real data, the model should be trained and then applied to a randomly selected group of import lines. Following that application, we recommend that the FDA inspect all of the selected import lines (not just those with high final risk scores), to determine the accuracy of the model, and determine how best to improve the model to achieve the desired result of a high (and continuously improving) accuracy rate (or another target, as determined by the FDA).

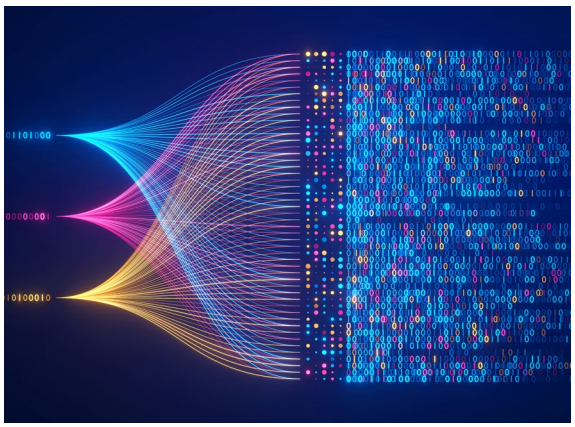
REI Systems' deep understanding of Prior Notice and its associated data enables us to develop the model more efficiently and swiftly, leveraging our extensive experience.

### REDUCE FDA OPERATIONAL BURDENS

The AI model will help the FDA improve health and safety as increasing volumes of food and drugs are consumed, despite limited numbers of inspectors and inspections – because FDA will be much better at targeting high risk products.

### ACCELERATE SUPPLY CHAIN FLOW

AI may allow the FDA to reduce the burden of inspections on producers and manufacturers, as their products can move more quickly through the supply chain.



## DISCOVER UNANTICIPATED PATTERNS TO BETTER PROTECT AMERICAN CONSUMERS

The benefit of AI will be its efficiency as it ingests data from multiple sources, discovers unanticipated relationships between data from those multiple sources, and creates the ability to better identify products that are harmful to American consumers.

## ALIGNMENT WITH THE AI EXECUTIVE ORDER

The design of the model and its implementation will ensure that the potential production-level system meets the requirements of the Section 10.1(b) of the AI Executive Order and the prospective OMB guidance to agencies in terms of transparency, equitability, and risk management.

# Conclusion

In conclusion, REI Systems' "Sentinel" represents a significant advancement in FDA inspection processes, leveraging AI to enhance risk profiling. By analyzing valuable data from Prior Notice combined with other diverse data sources, Sentinel could accurately identify high-risk imports, streamline inspections, and safeguard public health.

**To view a demo or learn more about REI's Sentinel proof of concept, contact AI@reisystems.com**

# About the Authors

**Andrew Zeswitz** is REI Systems' Chief Technology Officer. He has helped Federal agencies to develop and use IT for over 25 years.

**Jeff Myers** is a Senior Director at REI Systems. He has advised local, state, and federal governments on strategy and performance as a consultant and as a government employee.

**Ali Mirzaeian** is a Data Scientist and AI/ML Researcher at REI Systems. He holds a PhD in Computer Engineering, and has taught and conducted research at both George Mason University and Virginia State University.