A Digital Immune System for Early Biological Threat Detection

Interactive Multi-Signal Biosecurity Dashboard

Isaias Martinez

July 14, 2026

1 Abstract

This document describes the design and functioning of an interactive Digital Immune System dashboard that integrates three real-time public-health indicators: wastewater viral load, hospital admissions, and environmental air-quality signals. The system standardizes and fuses these independent data streams to produce a unified risk index and alert tiers. Users can upload their own data or explore predefined outbreak scenarios. Future extensions, including machine learning and automated data ingestion, are also outlined.

2 Introduction

Biological threats—whether natural pandemics, engineered pathogens, or environmental hazards—often emerge long before they are formally detected. Most surveillance systems act only *after* a signal becomes obvious. The goal of the **Digital Immune System** is to invert this paradigm: detect weak anomalies early, synthesize multiple evidence streams, and provide actionable early warnings.

The interactive dashboard described here demonstrates a prototype early-warning architecture using three synthetic signals:

- **Wastewater viral load** (e.g., SARS-CoV-2 RNA copies)
- **Hospital admissions** (respiratory-driven surge indicators)
- **Environmental air quality** (PM2.5 and AQI patterns)

All data sources are processed, standardized, fused into a risk metric, and visualized.

3 System Overview

The application is built in Python using the `streamlit` framework. The system performs the following pipeline:

1. User uploads wastewater, hospital, and environmental CSVs
2. The system identifies date and value columns
3. Each signal is smoothed with a rolling mean
4. Z-scores are computed relative to recent historical baselines
5. Signals are fused using adjustable weights
6. A risk score and alert tier are produced
7. The dashboard renders trend lines, standardized plots, and alerts

The entire process mimics how an immune system aggregates multiple danger signals.

4 How the Dashboard Works

The dashboard consists of several functional layers:

4.1 Data Input Layer

The user has two options:

- **Upload three CSV files**: wastewater, hospital, environment
- **Select a predefined scenario**: baseline, outbreak, bioterror

The system automatically identifies the date column and numeric data column, unless multiple numeric columns exist, in which case the user selects one.

4.2 Signal Processing Layer

For each metric $x(t)$, a rolling z-score is computed:

$$z(t) = \frac{x(t) - \mu_{\text{roll}}(t)}{\sigma_{\text{roll}}(t)}$$

where μ_{roll} and σ_{roll} are rolling mean and standard deviation using a user-selected window (e.g., 28 days).

4.3 Fused Risk Layer

Each signal contributes to a fused risk score:

$$R(t) = w_{ww} \cdot z_{ww}(t) + w_{hosp} \cdot z_{hosp}(t) + w_{env} \cdot z_{env}(t)$$

Weights are controlled in the sidebar and automatically normalized.

4.4 Alert Tiers

Alert tiers are determined by thresholds:

$$\text{High if } R(t) \geq \tau_{\text{high}}, \qquad \text{Medium if } \tau_{\text{med}} \leq R(t) < \tau_{\text{high}},$$

Values below zero represent baseline activity.

Each scenario loads a different set of synthetic CSVs packaged with the system.

5 Dashboard Outputs

The dashboard presents:

5.1 Raw Trend Lines

Three separate charts:

- Wastewater viral load (copies per liter)
- Hospital admissions (daily or weekly)
- Environmental pollutant level (AQI or PM2.5)

Each chart displays the uploaded or scenario-generated data as a time series.

5.2 Standardized Signal Plot

A combined z-score line chart shows:

- Wastewater z-score
- Hospital z-score
- Environment z-score
- Fused risk score

This chart highlights synchronized anomalies across systems.

5.3 Alert Table

A chronological table displays:

- Date
- Fused risk
- Alert tier

A downloadable CSV is automatically generated.

6 What the Negative and Positive Scores Mean

A negative z-score means the signal is *below* recent baseline (normal or declining). A positive z-score indicates above-baseline activity.

A fused risk score greater than zero suggests emerging anomalies; strong values indicate synchronized spikes across multiple systems.

7 Future Work

The current version of the digital immune system is a functional prototype, but there are several important directions for expansion and refinement. One major area for future development is the incorporation of additional signals beyond wastewater, hospital admissions, and environmental air quality. Integrating clinical laboratory positivity rates, syndromic surveillance data such as influenza-like illness and wearable-derived cough frequency, genomic sequencing reads from wastewater or clinical isolates, and even social-media-based symptom signals would provide a more comprehensive and resilient multi-layered view of emerging biological threats. Each additional data stream adds another independent dimension of evidence, substantially improving sensitivity to weak or early outbreak signals.

Machine learning also represents a substantial opportunity for enhancement. Although the current system relies on rolling z-scores and weighted fusion, more advanced anomaly detection methods—such as Isolation Forests, Local Outlier Factor, or deep-learning-based autoencoders—could detect subtle deviations from normal patterns. Multi-modal fusion architectures, including transformer models or graph neural networks designed to integrate diverse time series, could learn relationships across signals automatically. Furthermore, short-term forecasting models could provide not only real-time alerts but also predictions of near-future outbreak trajectories, enabling proactive rather than reactive decision-making. These ML additions would allow the system to adapt its baselines over time and produce probabilistic risk scores with quantified uncertainty.

Another critical area for improvement is automatic data ingestion. In the current prototype, the dashboard requires manual CSV uploads or static synthetic datasets. A production-strength system would automatically collect data from wastewater APIs, hospital capacity and admissions feeds, EPA AirNow or OpenAQ air-quality endpoints, and regional surveillance systems. Automated nightly ingestion, cleaning, and optional retraining would transform the dashboard from a retrospective analytics tool into a true real-time surveillance platform capable of continuous monitoring without human intervention.

Finally, several engineering refinements would be required to harden the prototype into an operational system. Deploying the dashboard on secure cloud infrastructure would allow it to scale and integrate with multiple data providers. Real-time alerts could be delivered through SMS, email, or secure enterprise communication systems such as Slack. Access control and user-level permissions would be necessary for handling sensitive clinical or governmental datasets. Containerizing the application with Docker would simplify deployment, reproduction, and long-term maintenance. These engineering improvements would strengthen reliability, security, and usability as the system evolves toward operational readiness.

8 Conclusion

The Digital Immune System dashboard demonstrates a unified multi-signal biosecurity framework that approximates innate immune sensing. By integrating wastewater, hospital, and environmental indicators, the system produces early warning signals that can support public-health action, disaster response, or national security surveillance.

The modular pipeline allows for expansion into more complex models and real-time data feeds, offering a roadmap toward next-generation biological threat detection systems.