NJIR

Open Access

This article is licensed under CC-BY-4.0

NJIR

Review

Innovations in Integrating Artificial Intelligence for Automated Surveillance and Outbreak Detection in Public Health Intelligence

Mary Magdalene Yeboah [1] and **Samuel Amfo Junior** [2*]

[1] Independent Researcher, Texas, USA

[2] Department of Computer and Information Technology, Eastern Illinois Technology, Illinois, USA

Cite This: *NJIR 2026, Volume 1.0, 8*

Read Online

ACCESS | Metrics & More | Article Recommendations

NJIR , 2026 , Volume 1.0 , 8

Received Aug 27, 2026 · Accepted Oct 29, 2026 · Published Nov 6, 2026 · NID: nid/njir.2026.0002 · Issue 1.0 · NJIR-2026-0002

ABSTRACT

The integration of Artificial Intelligence (AI) in public health surveillance is transforming the identification and management of outbreaks. Traditional systems typically suffer from delay, isolated data, and reactive designs, which hamper their potential for timely detection of emerging threats. Recent developments in machine learning, natural language processing, and real-time data fusion are shifting the paradigm to proactive and programmatic surveillance. This article analyzes the technical basis, current application, and strategic implications of AI-based surveillance for outbreak identification. A modular approach is proposed to guide implementation across contexts. Drawing on international case studies, from COVID-19 early detection onwards, the article lays out the promise and risks of embedding AI in modern public health systems.

NAHDA PUBLICATIONS © 2026 The Authors. Published by Nahda Publications

KEYWORDS public health surveillance, artificial intelligence, outbreak detection, digital epidemiology, early warning systems

1. INTRODUCTION

Infectious disease remains an international threat, frequently emerging with little warning and causing extensive health and economic disruption. The COVID-19 pandemic, as well as recurring zoonotic emergences like Ebola and avian influenza, has underlined the critical importance of timely, predictive public health surveillance. Traditional systems of surveillance, reliant on clinical reporting and passive reporting, are prone to detect outbreaks only after widespread transmission has been established. This passive response, along with reporting delays and lack of integration across systems, can significantly defeat containment measures (World Health Organization [WHO], 2020).

Artificial intelligence offers a compelling answer to this shortcoming. AI systems can scan large and unconnected sources of data in real time, observing emergent patterns that signal an outbreak as it is just beginning to occur before clinical diagnostics validate it. These systems can pull and analyze data from electronic health records, social media platforms, search engines, as well as environmental sensors. Through the power of providing earlier warning, better situational awareness, and facilitating quicker response strategies, AI has become a focal point for creating digital epidemiology (Scarpino & Petri, 2019).

Recent history offers samples of its possibility. BlueDot, a Canadian artificial intelligence platform, detected unusual instances of pneumonia in Wuhan prior to official COVID-19 reporting days before. Similarly, HealthMap at Boston Children's Hospital picked up early warning through machine-learning-based natural language processing of web-based news and social media data (Bogoch et al., 2020). However, the complete integration of such tools into conventional public health processes remains patchy. The majority of agencies are obstructed by concerns related to data interoperability, infrastructure readiness, and the ethical regulation of forecasting algorithms.

The current paper presents AI and epidemiologic surveillance at their intersection, proposing a conceptual framework for implementation. It explains the fundamental technologies behind automated outbreak detection, evaluates the merits and shortcomings of current systems, and outlines possibilities for strategic deployment in national and international public health infrastructures.

2. THEORETICAL AND TECHNOLOGICAL FOUNDATIONS

Public health surveillance in its simplest terms is ongoing collection, analysis, and interpretation of health data to inform action. While this definition has not been changed since its founding within traditional epidemiological agencies (Thacker et al., 1988), the nature of operations within public health surveillance has grown volumetrically with the advent of computing and digital technologies. With AI, the shift is from descriptive and retrospective to predictive and anticipatory public health intelligence.

Digital epidemiology is an extension of traditional surveillance with the inclusion of unstructured and out-of-the-box data sources. These include web search patterns, social media buzz, mobile device location, and even satellite imagery. When run through machine learning and natural language processing algorithms, such data sources are found

NAHDA PUBLICATIONS © 2026 The Authors. Published by Nahda Publications

to have the capability to give relevant insights into disease transmission dynamics, public sentiment, and risk behavior often well before formal clinical diagnosis (Salathé et al., 2012).

Machine learning (ML), in particular, is most important here by enabling systems to learn from past outbreaks and recognize patterns predictive of future ones. Supervised ML algorithms can be trained on labeled data sets to forecast symptoms or syndromes, while unsupervised algorithms detect anomalies or clustering patterns that do not conform to historical baselines. Neural networks, and specifically recurrent types such as long short-term memory (LSTM) models, have proven effective in modeling sequential health data and temporal trends (Hochreiter & Schmidhuber, 1997).

Natural language processing (NLP) complements these opportunities by extracting relevant health data from text. It enables systems to capture and interpret news articles, government warnings, physician-patient documentation, and free-form internet discussions for early signs of disease outbreak. This is the technology behind systems like BioCaster (Collier et al., 2008) and ProMED-mail (Madoff, 2004) that surface alerts for emerging health threats.
.

Recent advances in edge AI and federated learning also enhance system resilience by making decentralized analysis of data possible. Such practices are designed to maintain privacy in data while continuously allowing model updates at local data points of origin (Rieke et al., 2020). In contexts where data cannot be centrally located due to legal or logistical constraints, such technologies allow for surveillance without any compromise on security or compliance.

The design of an AI-driven surveillance system typically comprises three components. First, the data acquisition layer aggregates input from structured sources (e.g., hospital information or diagnostic lab test results) and unstructured streams (e.g., social media or environmental sensors). Second, an analytics layer applies machine learning models to flag anomalies, estimate transmission risk, and generate probabilistic predictions. Finally, a decision and visualization layer translates model outputs into dashboards, alerts, and geospatial heat maps that can be used by health officials. The utility and effectiveness of such systems depend not only on algorithmic sophistication but also on human oversight, infrastructure integration, and the quality of data.

3. A MODULAR FRAMEWORK FOR AI-DRIVEN SURVEILLANCE

In order to deploy these innovations, we recommend a modular AI-driven surveillance system maximized for flexibility between environments and scalable, enabling integration into existing health systems. The system maximizes interoperability, explainability, and iterative feedback for enabling ongoing relevance and ethical compliance.

The platform is composed of five primary layers. The data ingestion layer processes both traditional and new sources of data, including electronic health records, diagnostic laboratory reports, triage logs, social media updates, search engine queries, satellite imagery, and mobility data. Inputs are piped into the preprocessing layer, where NLP and data harmonization pipelines convert raw, heterogeneous inputs to structured formats for model training.

The backbone of the system is the AI ensemble engine, which holds a collection of methodologies of modeling. These are statistical classifiers for syndromic classification, models of anomaly detection for emerging trends, and neural networks to perform time-series analysis. The system balances precision and generalization through the

© 2026 The Authors. Published by Nahda Publications
nid/njir.2026.0002

combination of model outputs through weighted voting or averaging.

Outbreak risk scores are calculated in the risk scoring and alerting layer with threshold calibration to trigger tiered alerts. The alerts are displayed in interactive dashboards in the feedback layer, where epidemiologists and frontline users can also refine model predictions by entering field observations. The feedback is leveraged to recalculate parameters in the models to improve long-term performance.

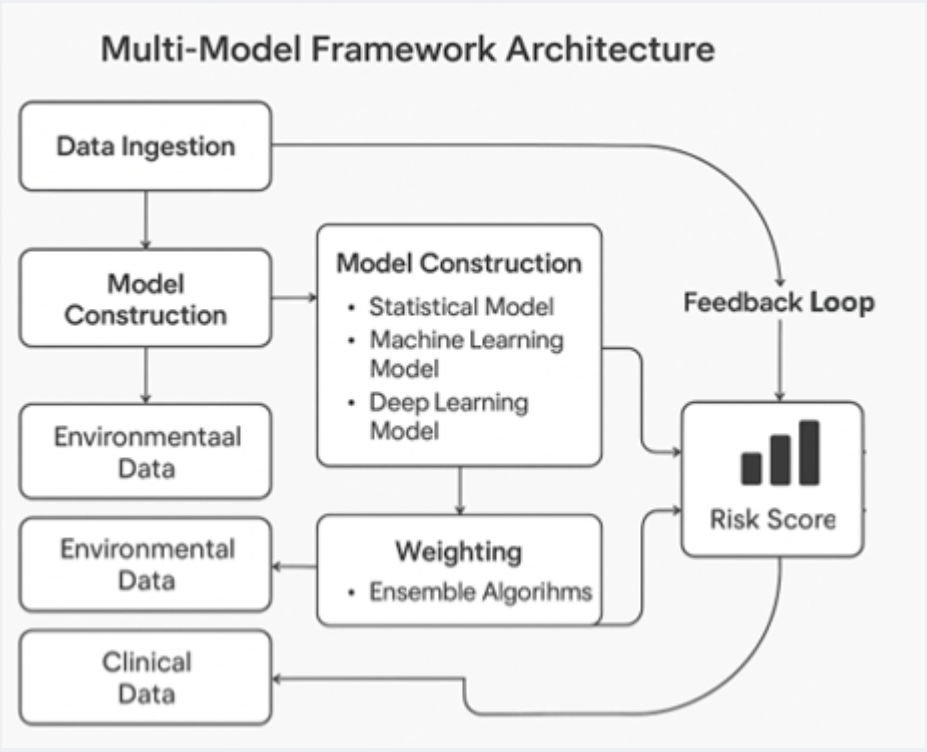


Figure 1. Multi-Model Framework Architecture This diagram illustrates the five interconnected layers of the proposed AI-driven public health surveillance framework: data ingestion, preprocessing, model ensemble, risk scoring, and feedback visualization.

4. CASE SCENARIO: EARLY COVID-19 DETECTION

The practical application of this model can be illustrated, as a hypothetical scenario, through the early detection of COVID-19. In late 2019, BlueDot reportedly issued an alert of atypical pneumonia cases in Wuhan, drawing an AI system that tracked global flight data, local media reports, and online forums in over 60 languages. This alert is reported to have preceded the formal declaration of the World Health Organization by a number of days (Niiler, 2020).

Had the proposed framework been applied in this instance, it would have tracked Chinese social media and domestic news sites. NLP processors would have processed key terms such as "pneumonia," "unknown cause," and "Wuhan hospital," using named-entity recognition for identifying institutions and symptoms. LSTM-based models would have identified unusual spatiotemporal clusters, with anomaly detectors picking up on a deviation from normal seasonal respiratory patterns.

For illustration only, an outbreak risk score above a hypothetical threshold of 0.90 would immediately trigger an alert for a subscribed public health agency.

NAHDA PUBLICATIONS © 2026 The Authors. Published by Nahda Publications

This value is an illustrative example and not an empirical result.

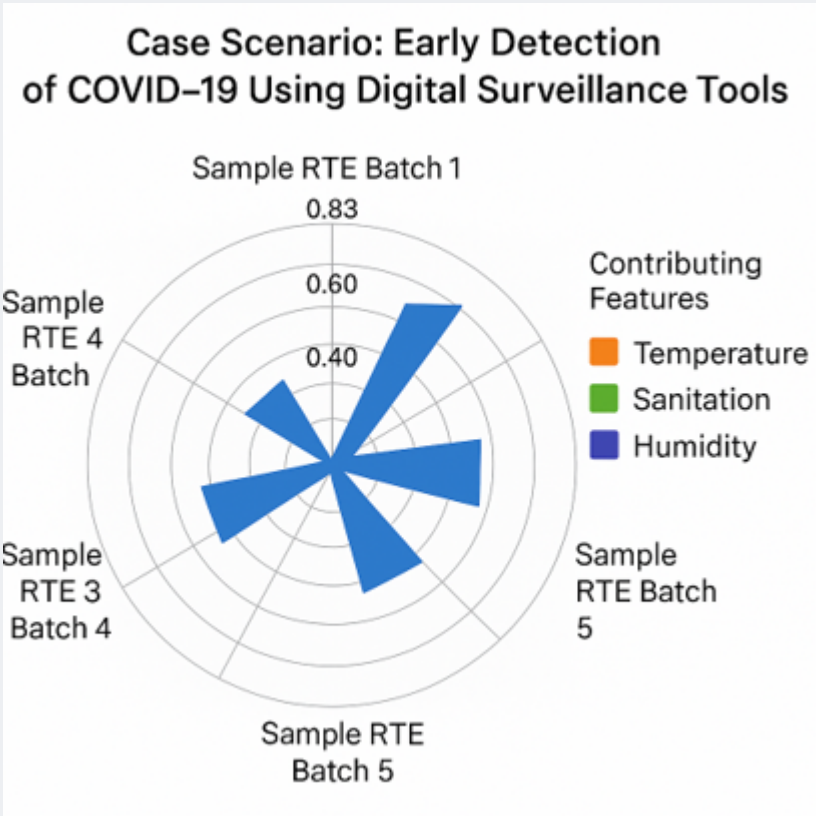


Figure 2. Hypothetical illustration of the AI-driven early warning framework applied to COVID-19 detection.

The above illustration points to the potential and the limitations of conventional AI monitoring. Prediction is possible early, but institutional readiness and governance processes are required in order to detect hints and turn them into timely interventions. Without cooperative processes for confirming and acting upon AI-generated cues, even accurate predictions will be disregarded (Scarpino & Petri, 2019).

5. EVALUATION OF MODEL PERFORMANCE AND USE EFFECTIVENESS

The utility of AI-based public health surveillance systems relies not just on the theoretical depth of the systems but on their actual ability to identify emerging threats to public health promptly and precisely. The metrics of performance must consider real-world data, which could be noisy, delayed, or incomplete, because these challenges closely model outbreak scenarios.

There are several existing systems designed to perform similar tasks, such as BlueDot, HealthMap and Canada's Global Public Health Intelligence Network (GPHIN) (Freifeld et al., 2008; Dion et al., 2015). At present there are no empirical results for the ensemble framework proposed. However, it is hypothesised that the hybrid nature of the system will result in improved sensitivity, precision and early warning times when compared to individual models due to the combination of machine learning classifiers, anomaly detection and deep learning models into the interpretable framework.

In a future study, historical news, event and surveillance data could be used for outbreaks where a

NAHDA PUBLICATIONS © 2026 The Authors. Published by Nahda Publications

known start date and announcement date is available. The outbreak events can be labelled meaning a ground truth is available. The models can then be trained and tested on separate periods and compared to the baseline systems listed to see how they perform. Metrics such as sensitivity, precision (positive predictive value), early warning times and interpretability of the alerts can be calculated.

Interpretability will be achieved through the use of explainable features such as SHAP (Shapley Additive Explanations). Analysts will be able to see what variables triggered the alerts. This can be to do with specific variables such as mentioning a sudden increase in symptoms (surprise), an anomaly in the news data geolocated to a certain area or changes in diagnostic data. This feature is important as it helps build trust and acceptance of the policy.

6. LIMITATIONS AND ETHICAL CONSIDERATIONS

While there are clear benefits to using AI in outbreak detection, there are also notable limitations to consider to ensure the correct use and sustainability of these systems.

Data representativeness will be an issue. Any AI system will only be as good as its data. With incorrect case definitions, non systematic reporting and bias in language (particularly from user generated data) these can affect the risk calculation. Social media can be a problem as whilst it provides vast amounts of real time data, it also provides a lot of false information. This requires sophisticated data pre-processing techniques (Adams et al., 2020). This issue has been overcome by ensuring that during the pre-processing stage data has been standardised and validated. Also using a weighted ensemble method.

Interpretability of the AI system is another issue. Deep learning models which are successful have very complex pathways making them difficult to interpret. In public health it is important to understand why the system has flagged something as an anomaly to help make decisions around restricting movement or scarce resource allocation. Interpretability tools such as SHAP or LIME have been introduced to overcome this issue. However, these are not always used in different jurisdictions. In this framework SHAP is used to provide interpretation of the alerts in the risk scoring and alerting stage and this is shown in the feedback dashboards. This means that analysts can understand what variables contributed to the alert.

There are ethical issues around privacy. Whilst AI can be used to provide early warnings, it uses sensitive data such as behaviour and health. There are concerns around informed consent, data minimisation and misuse. In some autocratic regimes there may be concerns that AI is being used to monitor individuals. For example during the rapid implementation of AI globally after the pandemic there were issues of it being valuable for real time tracking whilst raising issues around over-surveillance. An ethical framework should be put in place to ensure that AI is being used appropriately with systems to protect civil liberties (Rieke et al., 2020). This framework has overcome this issue by using federated learning and edge computing in the data ingestion stage to keep sensitive data in the local area. Only the model updates will be shared. Data minimisation has also been used in the pre-processing stage.

Equity will be an issue. LMICs will lack the digital infrastructure and human resources to use AI based outbreak detection. Without strategic investment and partnerships this can create a gap in global health. Areas with strong digital ecosystems will benefit from early warning systems whereas vulnerable areas will be left exposed. The modular framework can be used with only some data sources. The pre-processing

NAHDA PUBLICATIONS © 2026 The Authors. Published by Nahda Publications

stage needs to be set up to use local languages and low connectivity data.

Human involvement has been included in this framework through the feedback stage. This means that epidemiologists and front line users can input their observations from the field. They can also override, challenge and correct any model outputs.

Overall these limitations show that AI tools should be included into more holistic governance systems. AI systems must be created with accountability, transparency and inclusiveness in mind.

7. POLICY AND GOVERNANCE IMPLICATIONS

Integrating AI in public health surveillance cannot be had without enabling regulatory and institutional environments. As outbreaks acquire a global character, harmonized governance becomes absolutely essential.

International coordination is a vital pre-requisite. Outbreak detection systems are usually based on the exchange of data across borders and the application of shared alert thresholds. Institutions such as the WHO and the ECDC are well positioned to lead initiatives for establishing international standards for AI surveillance, including model auditing, performance measures, and data-sharing terms.

Human-AI interaction has also to be woven into design philosophy. AI needs to support and enhance human epidemiological abilities, not replace them. Mechanisms must include interfaces that allow for human analysts to override, question, or correct algorithmic output. The dual-track strategy ensures that public health professionals remain at the forefront of decision-making, precluding risks of automation bias or blind dependence.

Ethical design criteria need to be taken up formally at national and multilateral levels. Such criteria include the need for transparency of model logic, equity in representation of the training data, inclusiveness of population coverage, and clear accountability mechanisms. Public participation in model construction and regulation can also enhance legitimacy and trust.

No less urgent is the need to build digital epidemiology capability. Investment must move beyond hardware and software, to building open-source surveillance platforms, as well as skilled human resources. National AI task forces, international training programs, and agency data-sharing agreements can together create a fast follow-through in areas of low capacity.

AI surveillance management will need to adjust with and comply with the technology itself. Institutional, ethical, and legal readiness will not just determine the efficiency of these tools but how wide and even they will be applied.

8. CONCLUSION

Artificial intelligence (AI) is revolutionising public health surveillance. It enables automation, anomaly detection and risk prediction which can help health systems identify threats quickly and more accurately. Systems such as BlueDot and HealthMap have shown how using AI can help gain crucial lead time when tackling new threats. They have reported they released alerts before official statements on the pandemic.

In this paper we present a flexible and interpretable AI surveillance system. The system uses machine learning, natural language processing and temporal modelling. This system can be applied to various data sources and should be able to be applied worldwide. We have not yet tested this system but would do as outlined in section 5. This will determine if using ensemble modelling produces higher sensitivity and precision than individual models.

NAHDA PUBLICATIONS © 2026 The Authors. Published by Nahda Publications
nid/njir.2026.0002

The example given of using AI to detect early cases of the pandemic in Wuhan highlights the true potential of AI in surveillance. However, there are barriers to using this technology. Institutional resistance, opacity and varying preparedness all limit how effective AI can be. This shows we can not just rely on the potential of using AI. But we need to ensure it is designed ethically.

There are key issues with designing AI systems. Data bias can occur meaning signals may not be reliable. Using black box models can cause distrust. There is also concerns with collecting large amounts of data and the privacy implications. This is even worse in areas with little regulation. Lastly, there is issues with digital infrastructure in low and middle-income countries meaning they will not have equal access to these technologies.

To address the issues, we need a global effort. In the future we should focus on increasing digital infrastructure in low resource areas. Designing open source AI tools that can be used by multiple people. Ethics should also be included throughout the process. We need to ensure our AI systems are transparent, just and inclusive.

Overall, AI can be used as a game changing tool to increase resilience in global health. As long as we work together we can implement these systems to detect the unseen, predict the unknown and protect the vulnerable. By working together we can use AI to increase the speed, sensitivity and impact of outbreak surveillance systems.

FUNDING

The authors declare no funding

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

HOW TO CITE

REFERENCES

Adams, B., Hebert-Dufresne, L., & Scarpino, S. V. (2020). Disparities in digital epidemiology: Leveraging computational models to address public health surveillance gaps. PLoS Computational Biology, 16(8), e1008189.

Bogoch, I. I., Watts, A., Thomas-Bachli, A., Huber, C., Kraemer, M. U. G., & Khan, K. (2020). Pneumonia of unknown aetiology in Wuhan, China: Potential for international spread via commercial air travel. Journal of Travel Medicine, 27(2), taaa008.

Collier, N., Doan, S., Kawazoe, A., Goodwin, R. M., Conway, M., Tateno, Y., & Matsuda, G. (2008). BioCaster: Detecting public health rumors with a web-based text mining system. Bioinformatics, 24(24), 2940–2941.

Dion, M., AbdelMalik, P., & Mawudeku, A. (2015). Big data and the Global Public Health Intelligence Network (GPHIN). Canada Communicable Disease Report, 41(9), 209–214.

Freifeld, C. C., Mandl, K. D., Reis, B. Y., & Brownstein, J. S. (2008). HealthMap: Global infectious disease monitoring through automated classification and visualization of Internet media reports. Journal of the American Medical Informatics Association, 15(2), 150–157.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

Madoff, L. C. (2004). ProMED-mail: An early warning system for emerging diseases. Clinical Infectious Diseases, 39(2), 227–232.

Niiler, E. (2020, January 25). An AI epidemiologist sent the first alerts of the Wuhan coronavirus. Wired.

Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., & Cardoso, M. J. (2020). The future of digital health with federated learning. NPJ Digital Medicine, 3(1), 119.

Salathé, M., Bengtsson, L., Bodnar, T. J., Brewer, D. D., Brownstein, J. S., Buckee, C., & Vespignani, A. (2012). Digital epidemiology. PLoS Computational Biology, 8(7), e1002616.

Scarpino, S. V., & Petri, G. (2019). On the predictability of infectious disease outbreaks. Nature Communications, 10(1), 898.

Thacker, S. B., Berkelman, R. L., & Stroup, D. F. (1988). The science of public health surveillance. JAMA, 267(18), 2537–2542.

WHO. (2020). Public health surveillance for COVID-19: Interim guidance, 7 August 2020. Geneva: World Health Organization.

NAHDA PUBLICATIONS © 2026 The Authors. Published by Nahda Publications

Yeboah, M. M., & Junior, S. A. (2026). Innovations in integrating artificial intelligence for automated surveillance and outbreak detection in public health intelligence. Nahda Journal of Interdisciplinary Research, Volume 1.0(Issue 1.0), 8. https://nahdapublications.org/nid/nid/njir.2026.0002

NAHDA PUBLICATIONS © 2026 The Authors. Published by Nahda Publications

nid/njir.2026.0002