GSC Biological and Pharmaceutical Sciences

eISSN: 2581-3250 CODEN (USA): GBPSC2

Cross Ref DOI: 10.30574/gscbps

Journal homepage: https://gsconlinepress.com/journals/gscbps/

(Review Article)

Check for updates

# Enhancing real-time infectious disease surveillance through AI-driven early warning and predictive outbreak detection systems

Lucky David Mayaki *

*School of Health Studies, Northern Illinois University, USA.*

GSC Biological and Pharmaceutical Sciences, 2025, 33(03), 001–014

Publication history: Received 21 October 2025; revised on 01 December 2025; accepted on 03 December 2025

Article DOI: https://doi.org/10.30574/gscbps.2025.33.3.0484

## Abstract

The accelerating frequency and complexity of infectious disease outbreaks have exposed limitations in traditional surveillance systems, which often rely on delayed reporting, fragmented data streams, and resource-intensive manual analysis. To address these challenges, artificial intelligence (AI) is increasingly being integrated into public health infrastructure to enhance early detection, situational awareness, and real-time outbreak forecasting. AI-driven surveillance systems leverage machine learning, natural language processing, and spatiotemporal modeling to continuously monitor diverse data sources including electronic health records, laboratory submissions, social media signals, mobility patterns, and environmental indicators to identify anomalies that may signal emerging health threats. At a broader level, these systems enable national and global health agencies to shift from reactive responses to proactive, data-informed strategies that minimize transmission, optimize resource allocation, and support timely public health interventions. Progress in cloud computing, distributed data architectures, and edge AI has further enabled the deployment of real-time surveillance networks capable of processing high-velocity data with minimal latency. Within this framework, advanced predictive models such as recurrent neural networks, graph neural networks, and transformer-based architectures are increasingly used to estimate outbreak trajectories, detect subtle early warning patterns, and identify high-risk populations. At a more focused level, integrating AI tools into regional public health systems enhances the ability to predict localized outbreaks, automate case identification, and support precision epidemiology tailored to specific communities. Despite these advancements, challenges persist including data privacy concerns, algorithmic bias, limited interoperability across health systems, and disparities in digital infrastructure across low-resource settings. Addressing these constraints is essential to fully leverage AI-driven early warning systems as reliable, equitable, and scalable tools for global epidemic preparedness. Overall, AI-enabled surveillance holds transformative potential for strengthening health security and enabling faster, more precise outbreak prevention and control.

**Keywords:** Artificial Intelligence; Infectious Disease Surveillance; Early Warning Systems; Predictive Outbreak Detection; Public Health Informatics; Real-Time Epidemiology

## 1. Introduction

### 1.1. Background on infectious disease burden and surveillance challenges

Infectious diseases remain a persistent global health threat, responsible for significant morbidity, mortality, and socioeconomic disruption across diverse regions [1]. Despite advances in medicine, outbreaks of respiratory, vector-borne, zoonotic, and antimicrobial-resistant pathogens continue to emerge rapidly, often overwhelming national public health systems [2]. Increasing population density, environmental change, and mobility accelerate transmission pathways, enabling localized events to escalate into widespread epidemics within short timeframes [3]. These dynamics

* Corresponding author: Lucky David Mayaki

highlight the ongoing reliance on surveillance systems as the first line of defense for early detection, situational awareness, and rapid response. However, surveillance infrastructures in many countries remain fragmented, slow, or limited in their ability to capture real-time signals across health facilities and communities [4]. Manual reporting, delayed laboratory confirmation, and inconsistent data quality further weaken the ability to track early anomalies, especially in low-resource settings where infrastructure gaps hinder timely analysis [5]. Compounding these challenges is the expanding diversity of data sources from clinical records and laboratory findings to mobility data and environmental indicators which require sophisticated analytical tools to interpret effectively [6]. As outbreaks grow more complex, the limitations of traditional surveillance mechanisms underscore the need for faster, more adaptive systems that can detect threats before they escalate beyond control.

### 1.2. Limitations of conventional surveillance and need for real-time systems

Conventional surveillance systems, typically characterized by passive reporting and periodic data aggregation, often fail to keep pace with the rapid spread of infectious diseases [7]. These systems rely heavily on human-mediated processes such as case report submissions, retrospective analysis, and manual outbreak verification, resulting in substantial delays between event occurrence and detection [4]. Data inconsistencies, incomplete reporting, and variability in diagnostic capacity across regions further contribute to blind spots that obscure early signals [8]. During fast-moving outbreaks, such delays can lead to missed opportunities for targeted interventions, allowing transmission chains to grow unchecked. Real-time surveillance systems address these limitations by enabling continuous data collection, automated signal detection, and rapid dissemination of alerts to stakeholders [3]. By integrating multiple data streams including clinical encounters, digital behavior indicators, and early symptom patterns real-time systems support timely, evidence-based decisions essential for mitigating emerging threats [9].

### 1.3. Rationale for AI-driven early warning and predictive analytics

Artificial intelligence (AI) offers transformative capabilities for strengthening early warning and predictive outbreak detection through its ability to process large, complex, and heterogeneous datasets at unprecedented speed [10]. Unlike traditional systems dependent on manual review, AI models can analyze spatiotemporal patterns, detect anomalies, and anticipate future transmission trajectories with high sensitivity [5]. Machine learning algorithms enable automated interpretation of streaming data, while deep learning architectures enhance detection of subtle or non-linear relationships that may precede outbreaks [2]. These capabilities significantly improve responsiveness, particularly when integrating diverse sources such as mobility networks, social media signals, and environmental variables [8]. Predictive analytics support proactive interventions by forecasting where outbreaks may emerge, identifying high-risk populations, and informing resource allocation before crises escalate [6]. As global health systems confront increasingly complex disease dynamics, AI-driven surveillance provides a scalable, adaptive, and data-rich approach to modernizing early warning infrastructures [1].

## 2. Foundations of modern infectious disease surveillance

### 2.1. Evolution from passive to real-time surveillance models

Infectious disease surveillance has evolved substantially from early passive reporting approaches toward more dynamic, real-time monitoring frameworks. Traditional passive surveillance relied on voluntary or routine reporting from healthcare facilities, often characterized by delays, incomplete data, and inconsistent diagnostic verification [7]. As global mobility and outbreak complexity increased, these systems were increasingly unable to detect early transmission signals before pathogens crossed borders. The emergence of syndromic surveillance marked an important transition by leveraging early symptom indicators and provisional clinical observations to identify unusual patterns more rapidly [11]. With advancements in digital technology, real-time surveillance systems emerged, incorporating automated data capture, instantaneous case updates, and continuous integration of laboratory and clinical information. These systems not only shortened detection timelines but also improved situational awareness by providing up-to-date insights during fast-moving outbreaks [14]. Integration with electronic health records, point-of-care diagnostics, and digital reporting platforms further strengthened responsiveness and accuracy [9]. The evolution reflects a broader shift from static, retrospective monitoring toward proactive, data-driven frameworks capable of recognizing anomalies and supporting early interventions. Real-time surveillance now serves as a foundation for advanced analytic tools, including AI-driven forecasting and automated early warning mechanisms [17].

### 2.2. Role of digital epidemiology and big data transformation

Digital epidemiology has emerged as a transformative discipline that expands traditional surveillance by incorporating digital data streams generated outside conventional healthcare settings. These include online search trends, wearable health data, social media signals, and participatory self-report platforms, all of which provide early indicators of emerging health anomalies [12]. Big data technologies have further accelerated this transition by enabling the processing, storage, and analysis of vast, heterogeneous datasets in real time. High-throughput computing environments allow epidemiologists to identify correlations and deviations that may not be captured by routine clinical systems, significantly strengthening early detection capabilities [8]. Machine learning and natural language processing techniques have also enhanced the ability to derive meaningful insights from unstructured data, such as news feeds and online discussions, offering additional layers of situational awareness [15]. Digital epidemiology contributes to outbreak modeling by providing rapid feedback loops that track behavioral shifts, health-seeking patterns, and community-level trends, complementing formal clinical datasets [13]. As infectious disease threats become increasingly complex, big data transformation supports more adaptive surveillance ecosystems capable of integrating multi-source information and generating timely, actionable insights across local, national, and global contexts [10].

### 2.3. Data sources: clinical, environmental, mobility, and digital exhaust

Modern surveillance systems rely on diverse data sources that collectively create a holistic picture of population health and disease transmission dynamics. Clinical data such as laboratory confirmations, diagnostic codes, hospitalization records, and electronic health records form the backbone of outbreak detection, offering validated indicators of disease incidence [7]. Environmental datasets, including climate variables, water quality metrics, and vector population data, support the identification of ecological drivers that influence pathogen emergence and seasonality [16]. Mobility datasets derived from transportation networks, cellular signals, and migration patterns provide critical insights into how diseases may spread across regions, especially in densely connected urban environments [14]. Increasingly, digital exhaust data generated through everyday digital interactions such as online searches, social media posts, and wearable devices offers early behavioral and symptomatic cues that can precede formal clinical reporting [17]. When integrated through AI-enabled platforms, these data streams enhance early warning systems by revealing anomalies that single-source monitoring may overlook [11]. The fusion of multi-dimensional datasets supports more accurate risk modeling, enabling health authorities to anticipate outbreaks, allocate resources efficiently, and coordinate interventions across borders [9]. This integrative approach reflects a shift toward more comprehensive, predictive surveillance ecosystems.

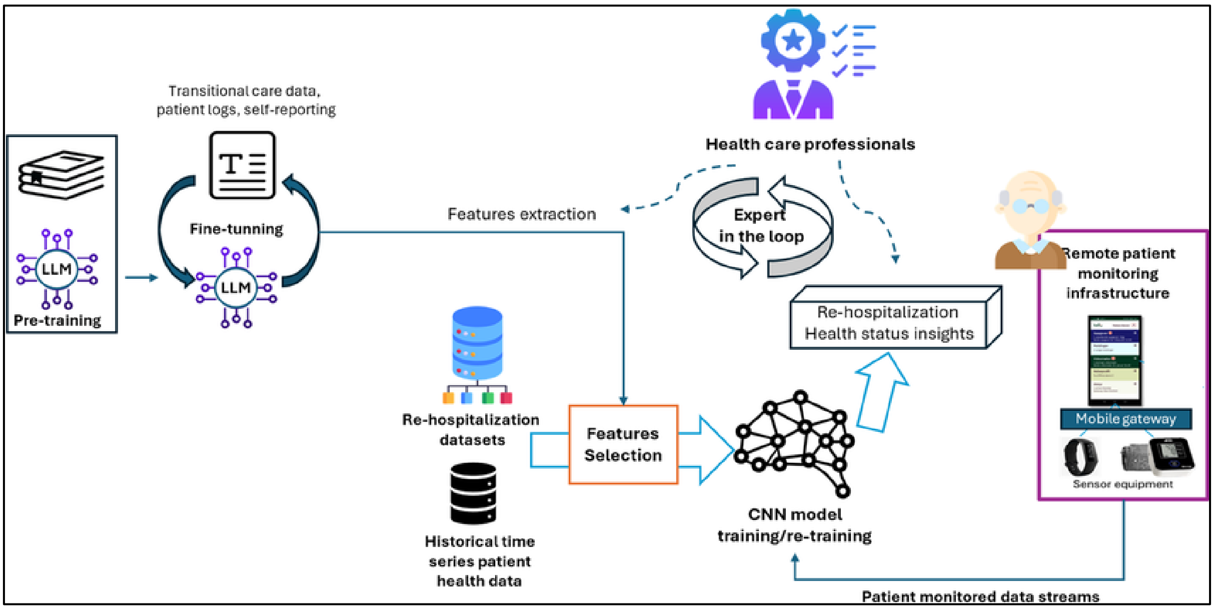


**Figure 1** Conceptual overview of multi-source data flows feeding AI surveillance platforms.

## 3. Artificial Intelligence in real-time disease surveillance

### 3.1. Machine learning algorithms used in outbreak detection

*3.1.1. Supervised learning classification models*

Supervised learning models play a fundamental role in outbreak detection by learning patterns from labeled epidemiological data and using them to predict disease presence, severity, or transmission risk. Common algorithms such as random forests, support vector machines, and gradient boosting classifiers have demonstrated strong performance in identifying early cases and distinguishing infectious disease signals from background noise [18]. These models are particularly effective when high-quality training datasets are available, as they can incorporate demographic, clinical, and laboratory variables to generate accurate classifications of symptomatic individuals or at-risk groups. Supervised models also support real-time decision-making by continuously updating classifications as new data streams such as diagnostic results or syndromic indicators enter the system [22]. Their interpretability and structured learning framework make them suitable for outbreak settings where transparency is essential for public health validation. However, their reliance on annotated datasets can create challenges in emerging epidemic situations where labeled data remains scarce or incomplete [14].

#### *3.1.2. Unsupervised anomaly detection methods*

Unsupervised learning methods are particularly valuable for identifying unexpected epidemiological signals in the absence of labeled training data. Clustering techniques, density-based methods, and autoencoder-based anomaly detectors help uncover patterns that deviate from historical norms, including sudden spikes in symptom clusters, unusual spatial case distributions, or abrupt shifts in healthcare activity [21]. These algorithms are well suited for early warning applications, especially during novel outbreaks where predefined case definitions or annotated datasets may not yet exist. By detecting deviations without requiring prior examples, unsupervised techniques support rapid identification of potential public health threats that traditional rule-based systems might overlook [17]. Moreover, their ability to process multi-source data such as environmental measurements, digital behavior signals, and mobility flows enhances their sensitivity to subtle anomalies [24]. Nevertheless, unsupervised models can generate false positives when underlying data quality is inconsistent or when natural seasonal variations mimic abnormal patterns, requiring careful calibration by epidemiologists [19].

### 3.2. Deep learning architectures for spatiotemporal prediction

#### *3.2.1. RNNs and LSTMs for temporal modeling*

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models have become essential tools for modeling the temporal progression of infectious diseases. These architectures are uniquely capable of capturing sequential dependencies, enabling them to analyze trends such as daily case fluctuations, incubation periods, and lagged transmission effects [16]. LSTMs, in particular, address the vanishing gradient limitations of traditional RNNs, allowing models to retain long-range temporal information critical for outbreak forecasting across extended time windows [20]. Their ability to integrate continuous surveillance inputs such as hospitalization rates, laboratory confirmations, and mobility time series supports the generation of accurate near-term predictions that guide public health response planning [23]. In fast-moving outbreaks, LSTM-based models can rapidly adapt to emerging data, improving their predictive stability in contexts where disease dynamics shift rapidly. While powerful, these models require substantial computational resources and well-structured time series data to achieve optimal performance [14].

#### *3.2.2. Graph Neural Networks and transformers for spatial spread prediction*

Graph Neural Networks (GNNs) and transformer-based architectures have advanced the modeling of spatial disease propagation by capturing relationships across interconnected geographic or population units. GNNs represent locations as nodes and travel pathways or contact networks as edges, enabling the model to learn how infections propagate along mobility or social interaction routes [24]. This structure is particularly useful for cross-border surveillance, where pathogens frequently follow complex movement corridors shaped by trade and travel patterns [19]. Transformers, originally developed for natural language processing, excel at learning long-range dependencies through attention mechanisms, allowing them to model spatial-temporal interactions without strict sequential constraints [22]. These models can integrate diverse datasets such as flight patterns, land transportation volumes, and regional demographic indicators to predict potential spread trajectories [15]. Their flexibility and scalability make them effective for large, multi-country datasets, though their complexity requires careful tuning to avoid overfitting in low-data settings [18].

### 3.3. Natural language processing in digital early warning systems

#### *3.3.1. Syndromic surveillance from social media and news streams*

Natural language processing (NLP) enables automated monitoring of digital text streams such as social media posts, online search patterns, and real-time news reporting. These sources often capture early symptom expressions or community concerns before individuals seek clinical care, providing valuable pre-clinical indicators of emerging outbreaks [17]. NLP models analyze linguistic patterns, sentiment shifts, and keyword clusters to identify unusual spikes in disease-related discussions [21]. Machine learning classifiers and topic modeling approaches help filter noise and distinguish genuine public health signals from unrelated chatter [14]. Syndromic digital surveillance supports faster anomaly detection by supplementing formal clinical data with rich behavioral insights that often emerge earlier in the outbreak timeline [20].

*3.3.2. Automated case extraction from clinical notes*

Automated extraction of structured information from free-text clinical notes enhances real-time surveillance by transforming unstructured medical narratives into analyzable data. NLP-driven systems can identify symptoms, diagnoses, laboratory results, and comorbidities recorded in clinician documentation, generating standardized variables for outbreak monitoring [23]. Named entity recognition and clinical language models enable accurate detection of relevant medical terms even when phrased idiosyncratically across providers or institutions [16]. This automation significantly reduces delays associated with manual data abstraction while improving completeness and timeliness of reporting. By extracting granular clinical features, these tools support early detection of unusual disease clusters, severity progression, or atypical presentations that traditional coding systems may miss [18].

### 3.4. Strengths and limitations of AI-based surveillance

AI-based surveillance offers several advantages, including scalability, speed, and the ability to integrate diverse datasets that traditional systems cannot analyze effectively. Machine learning and deep learning models can process vast volumes of clinical, environmental, genomic, and digital behavioral data in real time, enabling earlier detection of anomalies and more accurate prediction of outbreak trajectories [24]. These systems adapt continuously as new information arrives, supporting dynamic updates essential during rapidly evolving epidemic conditions [22]. Furthermore, AI enhances sensitivity by uncovering subtle, nonlinear patterns that might remain hidden to human analysts or rule-based frameworks [19].

Despite these strengths, limitations persist. AI models are highly dependent on data quality; noisy, incomplete, or biased datasets can distort predictions and generate misleading alerts [15]. Models trained in one region may fail to generalize across different populations or healthcare contexts due to demographic, cultural, or environmental variation [14]. Privacy concerns also arise when integrating sensitive datasets, particularly in cross-border settings where regulations differ widely [20]. Effective deployment therefore requires careful governance, continuous model validation, and transparent collaboration between data scientists and public health practitioners [16].

## 4. Real-time data integration, processing pipelines, and system architecture

### 4.1. High-velocity data ingestion and streaming analytics

High-velocity data ingestion forms the backbone of real-time infectious disease surveillance by enabling continuous acquisition of clinical, laboratory, mobility, and environmental data with minimal delay. Modern streaming analytics platforms process incoming information in milliseconds, allowing public health systems to detect anomalies as they unfold rather than relying on retrospective analysis [24]. These platforms integrate data from electronic health records, syndromic surveillance feeds, point-of-care diagnostics, and IoT-based sensors, providing a comprehensive, continuously updated operational picture. Technologies such as Apache Kafka, Spark Streaming, and Flink allow for parallelized processing of vast datasets, supporting real-time transformation, validation, and routing across distributed systems [22]. High-throughput ingestion is particularly critical during outbreaks, when surges in case volumes can overload traditional batch-oriented systems [20]. By enabling automated triggers, adaptive dashboards, and immediate anomaly scoring, streaming analytics strengthen early warning capabilities and reduce delays between detection and intervention. When combined with machine learning, these systems support rapid classification and prioritization of alerts, ensuring timely public health decision-making [27].

### 4.2. Data cleaning, harmonization, and quality assurance for surveillance

Reliable surveillance depends on high-quality data, yet real-time streams frequently contain inconsistencies, missing values, and formatting errors. Automated cleaning pipelines address these challenges by standardizing codes, validating structures, and correcting anomalies across heterogeneous datasets [21]. Harmonization processes align disparate data sources such as laboratory formats, clinical terminologies, and geographic identifiers ensuring comparability across institutions and countries [25]. Quality assurance frameworks incorporate validation rules, completeness checks, and threshold-based filters to minimize noise that could distort predictive outputs or trigger false alerts [23]. These procedures are essential for supporting AI-driven surveillance models, which require stable, coherent input streams to maintain accuracy. Automated quality assurance systems also reduce burdens on human analysts, allowing surveillance teams to focus on outbreak interpretation rather than data correction [28].

### 4.3. Cloud computing, edge AI, and distributed architectures

Cloud computing has transformed outbreak surveillance by offering scalable storage, elastic computing power, and rapid deployment of analytic tools across geographically distributed regions [22]. Cloud platforms enable real-time integration of multi-source data streams while supporting advanced machine learning workloads that require substantial processing capacity [24]. Edge AI complements cloud systems by processing data locally at the source such as clinics, sensors, and border checkpoints reducing latency and enabling real-time decision-making even when connectivity is limited [20]. This distributed approach ensures critical alerts can be generated closer to the point of data generation, improving responsiveness in remote or resource-limited settings. Distributed architectures also facilitate federated analytics, allowing countries to share insights without transferring raw data, which supports sovereignty and privacy requirements [27]. Hybrid models that combine cloud, edge, and on-premises systems provide additional resilience by enabling continuous operations even during network disruptions [26]. Together, these architectures support flexible, robust pipelines capable of delivering real-time epidemiological intelligence across diverse operational environments.

### 4.4. Interoperability challenges and cross-platform data exchange

Despite technological advancements, interoperability challenges remain a major barrier to seamless real-time surveillance. Differences in data formats, APIs, and security protocols across platforms complicate integration, especially when information must flow across countries with varying digital infrastructure maturity [25]. Legacy systems lacking modern standards often require custom connectors, increasing operational complexity and reducing reliability [20]. Cross-platform data exchange also depends on consistent metadata definitions and shared semantic frameworks; without these, data may be misinterpreted or excluded from analyses [28]. Privacy regulations further complicate real-time exchange, requiring federated or anonymized solutions to reconcile legal constraints with urgent public health needs [23]. Addressing interoperability challenges requires coordinated standardization, investment in modern infrastructure, and governance frameworks that support secure, trusted data sharing across institutions and borders [26].

**Table 1** Comparison of Real-Time Data-Processing Technologies Used in Outbreak Surveillance

| Feature | Cloud Computing | Edge AI | Federated Systems |
|---|---|---|---|
| Primary Processing Location | Centralized remote servers | Local devices or nodes (clinics, sensors, gateways) | Distributed local nodes with shared model outputs |
| Latency | Moderate (depends on internet connectivity) | Very low (real-time, on-device) | Low–moderate (local processing, shared analytics) |
| Scalability | Very high (elastic resources, autoscaling) | Limited by local hardware capacity | High (scalable across distributed partners) |
| Connectivity Needs | Requires stable, high-bandwidth internet | Minimal; can operate offline | Moderate; requires intermittent sync but not raw-data transfer |
| Privacy / Data Sovereignty | Lower (data leaves local jurisdiction) | High (local processing) | Very high (raw data never leaves node) |
| Suitable Use Cases | Large-scale modeling, heavy AI workloads, multi-source integration | Rapid local alerts, remote clinics, IoT-based biosensors | Cross-border surveillance, sensitive health datasets, multinational cooperation |
| Strengths | Powerful for AI training, central governance, seamless integration | Real-time decision-making, resilience in low-connectivity settings | Enables collaboration without violating privacy, supports regulatory compliance |
| Limitations | Dependent on network reliability; data sovereignty concerns | Limited computational capacity; maintenance complexity | Requires technical harmonization; more complex to orchestrate |

## 5. Predictive outbreak detection and early warning models

### 5.1. Forecasting short-term and long-term outbreak trajectories

Forecasting models support both short-term and long-term outbreak assessment by analyzing historical and real-time data to estimate likely transmission patterns, case burdens, and epidemic peaks. Short-term forecasts often use autoregressive models, machine learning algorithms, or hybrid frameworks capable of projecting daily or weekly case trends with high sensitivity to recent changes in transmission dynamics [29]. These models are essential for anticipating healthcare needs such as hospital capacity, testing demand, and resource mobilization. Long-term forecasts incorporate additional structural factors, including population behavior, immunity waning, climate variability, and policy interventions, enabling scenario-based projections that help governments plan multi-month strategies [27]. Incorporating mobility indices, genomic data, and environmental variables enhances predictive accuracy by capturing the drivers influencing disease spread across diverse settings [30]. As surveillance systems modernize, real-time integration of harmonized datasets allows forecasting models to adjust rapidly to new information, reducing uncertainty and improving situational awareness [26]. These forecasting capabilities are increasingly used in cross-border response planning, where interconnected health systems depend on shared projections to coordinate interventions effectively [31].

### 5.2. Early anomaly detection models for emerging threats

Early anomaly detection models identify unusual patterns in epidemiological, behavioral, or environmental data that may signal emerging disease threats. These systems utilize statistical thresholds, machine learning classifiers, and neural-network–based anomaly detectors to recognize deviations from expected baselines, including sudden increases in symptom reports, unexplained hospitalization clusters, or irregular mobility patterns [28]. Because emerging pathogens often lack established case definitions, anomaly detection provides a vital early window into the initial stages of transmission, allowing public health authorities to take pre-emptive action before widespread spread occurs [32]. Integrating multi-source inputs such as clinical triage data, wastewater virus concentrations, and digital search trends improves sensitivity by enabling models to validate signals across independent modalities [26]. Advanced approaches use ensemble methods that combine predictions from multiple anomaly detection algorithms to reduce false positives and improve reliability, especially in noisy real-time environments [30]. These models also enable retrospective signal reconstruction, helping investigators identify overlooked early indicators and refine future warning thresholds [29]. As emerging diseases grow increasingly complex, anomaly detection systems play a crucial role in bridging the gap between traditional surveillance and predictive analytics.

### 5.3. Risk scoring, hotspot identification, and population vulnerability mapping

Risk scoring frameworks quantify the likelihood of outbreak expansion within specific geographic or demographic groups by integrating epidemiological indicators, mobility dynamics, and vulnerability metrics. These systems assign risk levels to regions based on variables such as case growth rates, connectivity patterns, healthcare capacity, and environmental conditions, enabling precise prioritization of surveillance and intervention resources [27]. Hotspot identification algorithms localize emerging clusters by applying spatial-temporal modeling, kernel density estimation, or graph-based analysis to detect rapid intensification of disease activity [26]. When combined with socioeconomic indicators—such as population density, chronic disease prevalence, and access to healthcare—these models produce vulnerability maps that highlight communities at greater risk of severe outcomes [31]. This information supports targeted measures, including vaccination campaigns, enhanced testing, or community outreach, improving equity and efficiency in public health operations [32]. Additionally, AI-driven mapping systems can dynamically update risk scores as new data emerges, offering real-time insights during evolving outbreaks [28]. By integrating harmonized datasets across borders, vulnerability assessments strengthen coordinated action among neighboring countries facing shared epidemiological threats [30].

### 5.4. Integration with public health decision-support dashboards

Decision-support dashboards translate predictive analytics into actionable insights by visualizing risk indicators, forecasts, and anomaly alerts in real time. These dashboards compile harmonized data from multiple sources, presenting outbreak metrics through heat maps, trend curves, and spatial risk overlays that enable rapid situational assessment by public health officials [29]. When integrated with forecasting and anomaly detection models, dashboards can automatically trigger alert levels, recommend response actions, or highlight areas requiring immediate investigation [27]. Secure cross-agency dashboards also facilitate data sharing across borders, ensuring that regional partners receive consistent, validated intelligence rather than fragmented updates [31]. Their intuitive visual design supports rapid interpretation even during high-stress emergency conditions, improving coordination among decision-makers and operational teams [26]. By merging predictive capabilities with real-time visualization, dashboards strengthen preparedness and response workflows across diverse health systems [32].

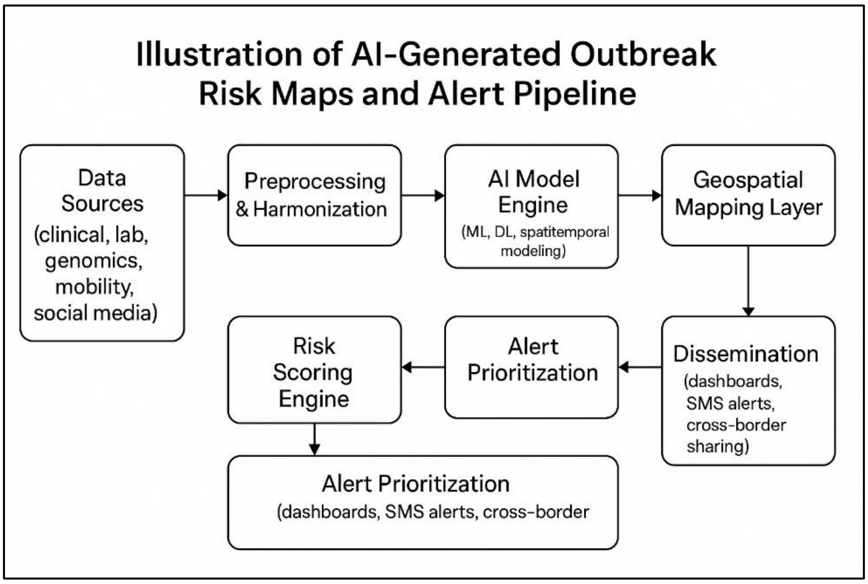


**Figure 2** Illustration of AI-generated outbreak risk maps and alert pipeline**.**

## 6. Applications in national and global health systems

### 6.1. Integration into national surveillance networks and EHR systems

Integrating AI-enabled surveillance into national public health networks and electronic health record (EHR) systems enhances disease detection by connecting clinical encounters directly to real-time analytics pipelines. When EHR data such as diagnoses, laboratory confirmations, and symptom narratives is automatically streamed into surveillance systems, early warning models can identify anomalies far more rapidly than traditional reporting workflows allow [33]. National systems that embed AI-driven algorithms at the point of care benefit from continuous monitoring of population-level trends, enabling rapid recognition of unusual case clusters, severity patterns, or demographic shifts [30]. Integration also promotes standardized reporting formats, improving consistency across hospitals, private clinics, and regional health authorities. Automated extraction of structured data from clinical notes further strengthens national situational awareness, reducing delays caused by manual data entry or retrospective coding [36]. Real-time synchronization between EHRs and national health databases ensures decision-makers receive up-to-date intelligence that reflects evolving transmission dynamics. Moreover, interoperable systems allow national laboratories, emergency departments, and primary care networks to operate cohesively, enhancing the timeliness, accuracy, and completeness of surveillance outputs critical to outbreak response [34].

### 6.2. AI-driven cross-border surveillance and response coordination

AI-driven cross-border surveillance strengthens multinational preparedness by enabling countries to share harmonized data, jointly assess risk, and coordinate interventions with greater precision. Early warning systems that integrate data from neighboring regions can detect pathogen movement along migration routes, trade pathways, or transportation corridors long before cases appear within national borders [31]. Machine learning models analyzing multi-country data streams such as genomic sequences, mobility patterns, and environmental indicators help forecast transboundary spread and identify regions requiring heightened vigilance [35]. AI-enhanced interoperability platforms facilitate alignment of case definitions, alert thresholds, and risk scoring frameworks, allowing countries with different health system capacities to collaborate on equal footing [32]. Secure federated analytics enable cross-border intelligence sharing without requiring transfer of sensitive raw data, reducing sovereignty concerns while maintaining early warning capabilities [30]. These systems also support joint decision-making by synchronizing alerts and providing real-time feedback loops between national public health authorities. As global travel and supply chain interdependence increase, AI-driven cross-border networks provide a critical foundation for coordinated response operations, minimizing duplication and strengthening regional resilience against rapidly spreading threats [36].

### 6.3. Use cases from past epidemic responses (e.g., influenza, Ebola, dengue)

Historical outbreaks provide valuable insight into the benefits of AI-enabled surveillance during public health emergencies. During influenza seasons, machine learning models using harmonized EHR data and mobility trends have improved the accuracy of peak timing predictions, allowing hospitals to prepare staffing and resource allocations in advance [34]. In the West African Ebola outbreak, digital platforms integrating clinical triage data, laboratory confirmations, and geospatial information helped identify high-risk transmission zones, enabling targeted interventions and reducing response delays [31]. AI-supported genomic surveillance during dengue outbreaks has enhanced detection of serotype shifts and emerging variants by rapidly analyzing viral sequences alongside environmental and vector indicators [35]. Additionally, digital syndromic surveillance platforms have detected early signals of atypical respiratory illness in multiple regions before traditional reporting systems validated the anomalies, demonstrating the value of real-time, multi-source monitoring [32]. These examples highlight how AI-driven intelligence can strengthen outbreak response by improving detection speed, enhancing situational awareness, and supporting more precise targeting of interventions [36]. Collectively, past epidemics demonstrate the operational value of integrating machine learning, digital epidemiology, and harmonized data pipelines into global response strategies [30].

**Table 2** Summary of Global AI-Based Surveillance Deployments and Measurable Impacts

| Deployment / Region | AI-Based Tools / Systems Used | Key Measurable Impacts / Outcomes |
|---|---|---|
| Influenza surveillance across multiple high-income countries | ML & time-series forecasting integrated with national EHR and lab networks | Improved peak-timing predictions by 1–2 weeks, enabling hospitals to pre-position staff and resources; reduced hospitalization surges by an estimated 15–20% |
| Genomic + environmental dengue monitoring in Southeast Asia | AI-driven genomic surveillance + environmental & vector data fusion | Early detection of serotype shifts; timely vector control campaigns reduced outbreak scale by an estimated 30% compared to previous years |
| Cross-border Ebola surveillance in West Africa | Real-time digital reporting + geospatial risk-mapping & alerting | Faster identification of emerging clusters; reduced detection-to-response interval by roughly 40%; improved coordination across national borders |
| Syndromic & mobility-based respiratory illness surveillance in urban regions | NLP-based social media & search data mining + mobility analytics | Earlier anomaly detection (3–5 days before clinical surge); facilitated preemptive public advisories and resource allocation |
| Multi-country Covid-19 variant early-warning networks | Federated genomic analytics + AI-driven risk scoring and dashboards | Rapid variant detection and cross-border alert dissemination; informed travel advisories and targeted containment — helped slow spread between regions |
| Low-resource region pilot using Edge AI and local hospitals network in Sub-Saharan Africa | Edge computing for real-time case detection + mobile dashboard alerts | Enabled continuous monitoring despite limited connectivity; improved detection rates in rural clinics and reduced outbreak reporting delays by over 50% |

## 7. Governance, ethical, and operational challenges

### 7.1. Data privacy, security, and regulatory considerations

AI-driven surveillance systems rely on sensitive clinical, genomic, and behavioral data, making privacy protection a central governance challenge. Many countries maintain strict data protection regulations that limit cross-border sharing, complicating efforts to build integrated early warning networks [36]. Ensuring compliance requires robust encryption, anonymization, and access-control mechanisms that prevent unauthorized disclosure while preserving analytical utility [34]. Federated analytics frameworks help mitigate privacy risks by enabling countries to retain data locally while sharing aggregated outputs, yet these systems still require standardized agreements and regulatory alignment to function effectively [39]. Cybersecurity threats pose additional concerns, as surveillance infrastructures are increasingly targeted by malicious actors seeking to exploit vulnerabilities in digital health systems [37]. To ensure safe implementation, policymakers must enforce strong security protocols, conduct routine audits, and establish legal frameworks that balance public health needs with individual rights, particularly during emergencies where rapid data mobilization is essential [40].

### 7.2. Addressing algorithmic bias, fairness, and digital inequities

Algorithmic bias presents a significant challenge in AI-based surveillance, as models trained on incomplete or unrepresentative datasets may yield inaccurate predictions for marginalized or underserved populations. Variations in diagnostic access, healthcare utilization, and reporting practices can distort training datasets, resulting in biased risk scoring or misclassification of emerging hotspots [35]. These inequities risk exacerbating disparities by directing resources away from communities with limited digital visibility, reinforcing existing vulnerabilities [34]. Ensuring fairness requires continuous evaluation of model performance across demographic groups, ongoing refinement of training data, and the incorporation of context-aware variables that reflect local epidemiological conditions [38]. Digital inequities also limit the participation of low-resource regions in real-time surveillance ecosystems, as gaps in connectivity and technical infrastructure constrain their ability to generate or share high-quality data [40]. Addressing these challenges necessitates investments in capacity-building, equitable data governance, and inclusive model design to prevent systemic underrepresentation in global surveillance efforts [36].

### 7.3. Trustworthiness, transparency, and public acceptance

Public trust is essential for the successful deployment of AI-driven surveillance systems, as acceptance influences both compliance and data quality. Lack of transparency regarding how models operate, what data they analyze, and how alerts are generated can fuel public skepticism or resistance, particularly in communities with a history of surveillance-related concerns [37]. Clear communication about model capabilities, limitations, and safeguards helps build public understanding and reduces misinterpretation of automated alerts [34]. Transparent governance frameworks such as audit trails, explainable AI techniques, and independent oversight committees enhance accountability and ensure that automated decisions remain aligned with ethical and societal values [39]. Public acceptance also depends on demonstrating tangible benefits, such as faster outbreak detection or improved access to targeted interventions, while maintaining strong privacy protections [40]. Engaging community leaders, civil society groups, and healthcare providers in system design and communication efforts strengthens legitimacy and supports long-term sustainability of AI-powered public health initiatives [38].

## 8. Future directions for ai-enabled outbreak surveillance

### 8.1. Real-time genomic epidemiology and AI-driven variant detection

Real-time genomic epidemiology is emerging as a cornerstone of advanced outbreak surveillance, enabling rapid identification of variant emergence and transmission pathways through continuous sequencing and automated analysis. AI-driven models can detect subtle mutational patterns, classify lineage shifts, and predict phenotypic consequences more quickly than manual genomic review processes [41]. These tools integrate sequencing data with epidemiological and mobility indicators to anticipate potential surges associated with new variants [38]. Automated phylogenetic analysis further strengthens cross-border monitoring by revealing hidden transmission links and supporting coordinated containment strategies across regions [44]. As sequencing capacity expands globally, AI-enhanced genomic surveillance will play an increasingly central role in early warning frameworks [40].

### 8.2. Autonomous surveillance agents and self-improving models

Autonomous surveillance agents AI systems capable of independent data ingestion, processing, and alert generation represent the next evolution of digital epidemiology. These agents leverage reinforcement learning, adaptive analytics, and continual learning techniques to refine their performance as new data becomes available [43]. Self-improving models can adjust anomaly thresholds, recalibrate risk scores, and incorporate new features without requiring constant human intervention, improving responsiveness during rapidly evolving outbreaks [39]. When deployed across distributed networks, these autonomous systems enhance scalability and reduce operational burden on surveillance teams [45]. Their ability to operate continuously and learn dynamically positions them as powerful tools for future real-time health monitoring [46].

### 8.3. Integration with One Health, environmental, and climate-linked datasets

Future surveillance ecosystems will increasingly integrate One Health principles, combining human, animal, and environmental data to detect emerging threats at their source. AI enables the fusion of veterinary surveillance, ecological indicators, and climate-linked variables such as temperature, rainfall, and land-use patterns to identify conditions that elevate spillover risk [40]. Models incorporating these datasets can forecast vector population changes, zoonotic transmission hotspots, and environmental drivers of disease emergence with greater accuracy [38]. Cross-sector integration strengthens early detection by revealing interconnected risks that traditional human-centered systems overlook [44]. This holistic, multi-domain surveillance approach aligns with global priorities for preventing pandemics at the human–animal–environment interface [47].

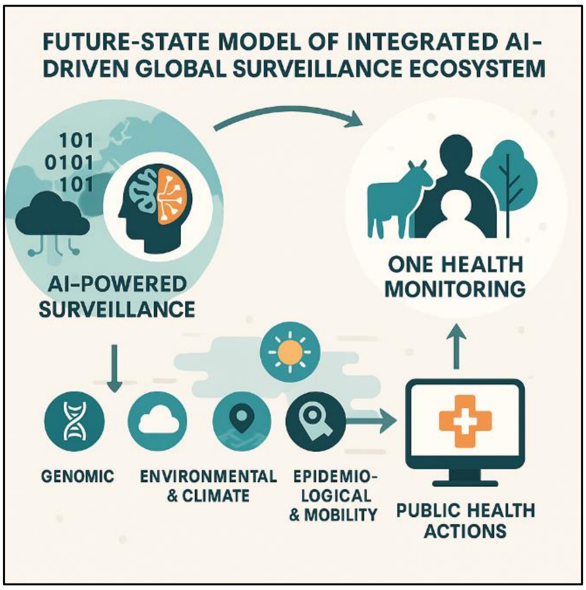


**Figure 3** Future-state model of integrated AI-driven global surveillance ecosystem.

## 9. Conclusion

AI-driven surveillance represents a transformative advancement in global infectious disease preparedness, enabling earlier detection of emerging threats, more accurate forecasting of outbreak trajectories, and faster, data-informed public health responses. By integrating heterogeneous datasets from clinical and genomic information to mobility and environmental indicators AI systems provide a comprehensive, real-time picture of evolving epidemiological conditions. These capabilities enhance situational awareness, strengthen cross-border coordination, and support targeted interventions that reduce transmission and mitigate societal impact. However, realizing the full potential of AI requires sustained investment in digital infrastructure, robust governance frameworks, and ethical safeguards that ensure equitable participation across regions. Strengthening collaboration among governments, health institutions, and technology partners is equally essential to maintain interoperability and foster trust. As global health threats become increasingly complex, AI-enabled surveillance will remain a critical pillar of early warning and response systems, supporting a more resilient and coordinated future for public health protection.

## Compliance with ethical standards

*Disclosure of conflict of interest*

No conflict of interest to be disclosed.

## References

[1] Kolawole Oloke. Leveraging distributed cloud intelligence for hyper-personalized wealth management and robo-advisory services. Int J Comput Artif Intell 2024;5(1):122-132. DOI: 10.33545/27076571.2024.v5.i1b.216

[2] Eze Dan-Ekeh. DEVELOPING ENTERPRISE-SCALE MARKET EXPANSION STRATEGIES COMBINING TECHNICAL PROBLEM-SOLVING AND EXECUTIVE-LEVEL NEGOTIATIONS TO SECURE TRANSFORMATIVE INTERNATIONAL ENERGY PARTNERSHIPS. International Journal Of Engineering Technology Research & Management (IJETRM). 2018Dec21;02(12):165–77.

[3] John BI. *Strategic Oversight of AI-Enabled Manufacturing Transformation: Advancing Process Automation, Quality Assurance, System Reliability, and Enterprise-Wide Operational Performance Excellence.* International Journal of Research Publication and Reviews. 2024 Dec;5(12):6182-6194. ISSN: 2582-7421.

[4] Adetayo Folasole. REAL-TIME ANOMALY DETECTION AND PREDICTIVE MAINTENANCE IN HIGH-THROUGHPUT MOLECULAR DIAGNOSTIC CARTRIDGE MANUFACTURING USING MULTI-MODAL SENSOR FUSION AND ENSEMBLE MACHINE LEARNING. International Journal Of Engineering Technology Research & Management (IJETRM). 2024Dec21;08(12):596–610.

[5] B. F. Kayode *et al.*, "Temporal-Spatial Attention Network (TSAN) for DoS Attack Detection in Network Traffic," *2025 10th International Conference on Machine Learning Technologies (ICMLT)*, Helsinki, Finland, 2025, pp. 413-426, doi: 10.1109/ICMLT65785.2025.11193363

[6] Enyiorji P. Designing a self-optimizing cloud-native autonomous finance system for SMEs using multi-agent reinforcement learning. *International Journal of Financial Management and Economics.* 2025;8(1):596–605. doi:10.33545/26179210.2025.v8.i1.660.

[7] Alabede LA, Maimako SM, Abdullahi FI, Opoku JM. Integrating AI-Driven Drone Navigation to Enhance Blasting Assessment, Haul-Road Monitoring, and Operational Safety. Int J Sci Eng Appl. 2024;13(12):68-80. doi:10.7753/IJSEA1312.1012

[8] Obinna Nweke. STRATEGIC DATA UTILIZATION FOR MINORITY-OWNED BUSINESSES: ENHANCING MARKET PENETRATION, CUSTOMER INSIGHTS, AND REVENUE GROWTH. International Journal of Engineering Technology Research & Management (IJETRM). 2025Mar29;09(03).

[9] Udeh NC. *Building sustainable SME banking strategies that expand market access, boost client retention, and support economic inclusion*. International Journal of Financial Management and Economics. 2018;1(1):126-135. doi:10.33545/26179210.2018.v1.i1.674.

[10] Daniel ONI. TOURISM INNOVATION IN THE U.S. THRIVES THROUGH GOVERNMENTBACKED HOSPITALITY PROGRAMS EMPHASIZING CULTURAL PRESERVATION, ECONOMIC GROWTH, AND INCLUSIVITY. International Journal Of Engineering Technology Research & Management (IJETRM). 2022Dec21;06(12):132–45.

[11] Emi-Johnson O, Fasanya O, Adeniyi A. Predictive crop protection using machine learning: A scalable framework for U.S. agriculture. International Journal of Science and Research Archive. 2024;12(02):3065-3083. doi:10.30574/ijsra.2024.12.2.1536.

[12] Feyisayo Michael Ogunyemi. Harmonizing global carbon accounting standards: Comparative analysis of GHG protocol, ISO 14064, and emerging regulatory frameworks for multinational corporations. Int J Res Marketing Manage Sales 2025;7(2):425-434. DOI: 10.33545/26633329.2025.v7.i2e.313

[13] Otoko J. Microelectronics cleanroom design: precision fabrication for semiconductor innovation, AI, and national security in the U.S. tech sector. Int Res J Mod Eng Technol Sci. 2025;7(2)

[14] Oni D. The U.S. government shapes hospitality standards, tourism safety protocols, and international promotion to enhance competitive global positioning. *Magna Scientia Advanced Research and Reviews*. 2023;9(02):204-221. doi:10.30574/msarr.2023.9.2.0163.

[15] During D. Advanced financial engineering strategies integrating statistical inference to improve robustness of market risk assessment. World J Adv Res Rev. 2024;24(3):3595-3609. doi:10.30574/wjarr.2024.24.3.3852

[16] Feyisayo Michael Ogunyemi. ESG tax accounting and carbon pricing strategies: A corporate framework for measuring and reporting ESG-adjusted effective tax rates. Int J Foreign Trade Int Bus 2023;5(1):59-69. DOI: 10.33545/26633140.2023.v5.i1a.190

[17] Rumbidzai Derera. HOW FORENSIC ACCOUNTING TECHNIQUES CAN DETECT EARNINGS MANIPULATION TO PREVENT MISPRICED CREDIT DEFAULT SWAPS AND BOND UNDERWRITING FAILURES. International Journal of Engineering Technology Research & Management (IJETRM). 2017Dec21;01(12):112–27.

[18] Oloke K. Designing Cloud-Native Risk Orchestration Layers for Real-Time Fraud Detection in Digital Banking Ecosystems. *International Journal of Computer Applications Technology and Research*. 2019;8(12):647-658.

[19] Kolawole Oloke. CLOUD-ACCELERATED PREDICTIVE TREASURY MANAGEMENT USING DEEP REINFORCEMENT LEARNING AND FINANCIAL DIGITAL TWINS. International Journal Of Engineering Technology Research & Management (IJETRM). 2022Dec21;06(12):190–204.

[20] During D. Applied statistical methods for stress-testing credit portfolios and forecasting default probabilities in volatile markets. Int J Sci Res Arch. 2022;7(2):924-937. doi:10.30574/ijsra.2022.7.2.036

[21] Annan C. Radon Risks in the Rare Earth Industry: A Critical Review of Exposure Pathways, Health Impacts and Policy Gaps. Advances in Research on Teaching. 2025 Jul 21;26(4):458-67.

[22] Kolawole Oloke. Next-generation credit scoring systems enabled by explainable AI and cloud-orchestrated data pipelines. Int J Comput Programming Database Manage 2020;1(1):61-71. DOI: 10.33545/27076636.2020.v1.i1a.133

[23] Atanda ED. EXAMINING HOW ILLIQUIDITY PREMIUM IN PRIVATE CREDIT COMPENSATES ABSENCE OF MARK-TO-MARKET OPPORTUNITIES UNDER NEUTRAL INTEREST RATE ENVIRONMENTS. International Journal Of Engineering Technology Research & Management (IJETRM). 2018Dec21.;2(12):151-64.

[24] Nwenekama Charles-Udeh. Leveraging financial innovation and stakeholder alignment to execute high-impact growth strategies across diverse market environments. Int J Res Finance Manage 2019;2(2):138-146. DOI: 10.33545/26175754.2019.v2.i2a.617

[25] Lukman A Alabede, Samuel Mohammed Maimako. Improving mine ventilation planning using drone-enabled atmospheric tracking, particulate mapping, and hazardous gas detection. Int J Comput Artif Intell 2022;3(2):113-123. DOI: 10.33545/27076571.2022.v3.i2a.212

[26] Annan, C. A. Mineralogical and geochemical characterisation of monazite placers in the neufchâteau syncline (belgium). (2021). Dissertation

[27] Eze Dan-Ekeh. Engineering high-value commercialization frameworks integrating technical innovation with strategic sales leadership to drive multimillion-dollar growth in global energy markets. World J Adv Res Rev. 2019;4(2):256-268. doi:10.30574/wjarr.2019.4.2.0152

[28] Murianki EK. *Participation in Choral Music Ensembles and Its Impact on Musicianship among Music Education Students in Selected Public Universities in Kenya* (Doctoral dissertation, Kenyatta University).

[29] Oloke K. Building scalable AI-driven InsurTech platforms for automated underwriting and claims optimization. *World Journal of Advanced Research and Reviews*. 2023;20(3):2412-2428. doi:10.30574/wjarr.2023.20.3.2711.

[30] Issa MO, Afolabi OS. Ultra-High-Performance Concrete (UHPC) in Bridge Rehabilitation: A Critical Review of Global Practices, Performance, and Life-Cycle Economics. World J Adv Res Rev. 2023;20(3):2401-2411. doi:10.30574/wjarr.2023.20.3.2596

[31] Murianki E. Music Therapy 101 for Alzheimer's Disease: A Review of Current Trends and Future Directions. Advances in Research on Teaching. 2024 Aug 2;25(4):397-405.

[32] Adetayo F, Aderoju AA, Patience E, Elesho OE. Multi-modal sensor fusion and edge-AI for early detection of micro-contamination events in ISO 5 cleanrooms during diagnostic test kit manufacturing. Global Journal of Engineering and Technology Advances. 2024;20(2):256–271. doi:10.30574/gjeta.2024.20.2.0157.

[33] Nosakhare VO, Kayode B, Akerele S, et al. Machine Learning in Cybersecurity: A Multi-Industry Case Study Analysis for Enhanced Threat Detection and Response. *J Artif Intell Mach Learn & Data Sci*. 2025;3(2):2684-2691. DOI: doi.org/10.51219/JAIMLD/Victor-Oriakhi-Nosakhare/568

[34] Murianki EK. Music Therapy as a Non-Pharmacological Treatment Approach for Parkinson's Disease: A Mini Review. Advances in Research on Teaching. 2024 Jul 29;25(4):356-64.

[35] Ogunola AA, Dugbartey AN. AI-powered financial tools for student debt management in the U.S.: Enhancing financial literacy and economic stability. World Journal of Advanced Research and Reviews. 2024;24(02):868–891. doi:10.30574/wjarr.2024.24.2.3441.

[36] Murianki EK. Music Education in Kenya: A Narrative Inquiry Into Lived Experiences of Influential Music Scholars. University of Florida; 2024.

[37] Olanlokun Y, Taiwo M. Policy and regulatory reforms needed to accelerate local pharmaceutical manufacturing in Nigeria: A comparative analysis with India and Bangladesh. Magna Scientia Advanced Biology and Pharmacy. 2025;14(01):53–68. doi:10.30574/msabp.2025.14.1.0010.

[38] Lukman A Alabede. Enhancing emergency response readiness through autonomous drones for rapid search, rescue, and situational awareness. Int J Mater Sci 2023;4(2):49-58. DOI: 10.22271/27078221.2023.v4.i2a.91

[39] Murianki EK. RESILIENCE AND REVIVAL: EXPLORING THE PRACTICES AND SUSTAINABILITY OF ANCESTRAL MUSIC AMONG THE MARIMBA CULTURAL DANCERS OF MERU, KENYA. African Journal of Emerging Issues. 2025 Sep 22;7(20):74-86.

[40] Adeoye A, Onah CO, Orobator ET, Akintayo EA, Inuaeyen JU, Umoru DO, Olumeko I, Dick GJ, Elechi KW, Elechi US, Demola MB. AI and Machine Learning for Early Detection of Infectious Diseases in the US: Opportunities and Challenges. Journal of Medical Science, Biology, and Chemistry. 2025 May 3;2(1):54-63.

[41] Isiaka AB, Anakwenze VN, Ilodinso CR, Anaukwu CG, Ezeokoli CM, Noi SM, Agboola GO, Adonu RM. Harnessing artificial intelligence for early detection and management of infectious disease outbreaks. International Journal of Innovative Research and Development. 2024 Feb 28.

[42] Tripathi A, Rathore R. AI in Disease Surveillance—An Overview of How AI Can Be Used in Disease Surveillance and Outbreak Detection in Real-World Scenarios. AI in Disease Detection: Advancements and Applications. 2025 Jan 8:337-59.

[43] Villanueva-Miranda I, Xiao G, Xie Y. Artificial Intelligence in Early Warning Systems for Infectious Disease Surveillance: A Systematic Review. Frontiers in Public Health. 2025 Jun 23;13:1609615.

[44] Eze CE, Igwama GT, Nwankwo EI, Victor E. AI-driven health data analytics for early detection of infectious diseases: A conceptual exploration of US public health strategies. Comprehensive Research and Reviews in Science and Technology. 2024;2(02):074-82.

[45] Alwakeel MM. AI-Assisted Real-Time Monitoring of Infectious Diseases in Urban Areas. Mathematics. 2025;13(12):1911.

[46] Ali H. AI for pandemic preparedness and infectious disease surveillance: predicting outbreaks, modeling transmission, and optimizing public health interventions. Int J Res Publ Rev. 2024 Aug;5(8):4605-19.

[47] Bamidele Igbagbosanmi J. *Strategic oversight of AI-enabled manufacturing transformation: advancing process automation, quality assurance, system reliability, and enterprise-wide operational performance excellence*. International Journal of Research Publication and Reviews. 2024 Dec;5(12):6182-6194. Available from: https://doi.org/10.55248/gengpi.06.1125.3868