Disease Prevention and Epidemiology

https://dpe.cultechpub.com/dpe

Cultech Publishing

Review

Artificial Intelligence for Public Health Surveillance: Promises, Pitfalls, and an Ethics-Equity Roadmap for Deployment

Gavin Chibundu Ikechukwu, Simeon Ikechukwu Egba*, Promise Chibuike Paul, Virginus Agozie Umeh

Department of Biochemistry, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

*Corresponding author: Simeon Ikechukwu Egba, egba.simeonikechukwu@mouau.edu.ng

Abstract

Artificial intelligence (AI)–such as machine learning (ML), natural language processing (NLP) and large multimodal models-is transforming the field of public-health surveillance, facilitating earlier outbreak detection, the combination of heterogeneous streams of data, and high-scale signal triage. All these capabilities have potential to enhance timeliness, sensitivity and geographic coverage of the surveillance systems, especially when used together with nontraditional sources of information like social media, news media, electronic health records, mobility data and environmental (e.g. wastewater) indicators. Though, actual implementations reveal significant traps: biased training data may enhance health inequities, bad data management and inadequate models may harm trust and responsibility, and incomplete validation may lead to false alarms or missed events. Moreover, the use of surveillance also presents more ethical issues, including privacy and consent to surveillance and the likelihood of surveillance increasing social evils. The review summarizes the existing evidence and practice in the technical, ethical and governance aspects of AI in surveillance of public-health. We list recent achievements and opportunities, methodological and operational issues, and trace the most significant risks. Based on global best practices and new reporting recommendations, we come up with a realistic roadmap towards ethical, equitable, and successful AI surveillance implementation. Among the recommendations, there are: strict pre-deployment validation on representative datasets; model documentation and performance reporting (where applicable) based on TRIPOD+AI/CONSORT-AI, continuous concept drift monitoring and measures of fairness, robust data governance and privacy-preserving architecture, stakeholder involvement, including affected communities, and building capacity in low- and middle-income countries (LMICs). We conclude by outlining priority research and policy measures that are required to ensure the translation of the technical promise of AI into benefits to the public-health with minimal harms.

Keywords

Artificial intelligence, Public health surveillance, Epidemiology, Ethics, Algorithmic fairness, Governance, TRIPOD+AI

Article History

Received: 1 December 2025
Revised: 24 December 2025
Accepted: 7 January 2026
Available Online: 12 January 2026

1. Introduction

Public health surveillance can be defined as the systematic and constant gathering, examination, and assessment of health-related information that is required to plan, implement, and examine the public-health interventions [1]. It is the foundation of disease prevention and response to epidemics, providing the authorities with the means to discover emerging threats, track trends, and distribute resources effectively. Traditional surveillance systems-including laboratory-confirmed case reporting, clinician notifications, sentinel networks, and mortality registries- are impossible to do without [2]. These systems are however prone to delays in reporting, incomplete capture of data as well as lack of spatial resolution especially in low resource environments or when dealing with rapidly changing outbreaks. With the rising physical population, mobility, and the complexity of epidemiology of infectious and non-communicable disease, conventional methods are becoming less and less effective at delivering timely and granular information [3]. It is on this background that the last twenty years have been characterized by the blistering development of artificial intelligence (AI) and machine learning (ML), which provide new tools to enhance and supplement the classical surveillance mechanisms [4]. AI-based systems are capable of processing a wide variety of diverse and large streams of data, including electronic health records (EHRs), syndromic surveillance feeds, environmental, genomic and mobility and digital trace data, at scales that would be impractical with a manual process [5]. First mover event-detection systems like HealthMap and BlueDot had shown the ability of natural language processing (NLP) and real-time web crawling to detect unusual disease signs ahead of official warnings, even in the early phases of recent outbreaks of diseases worldwide [6]. Similarly, wastewater metagenomic sequencing may be combined with ML-based nowcasting models to demonstrate the potential of tracking the level of prevalence of each pathogen at the community level, in order to detect trends related to transmission patterns early, and optimize the use of resources [7,8]. In spite of these developments, there are methodological, operational, ethical and sociopolitical concerns that AI introduction into surveillance systems evokes. Issues related to data quality and representativeness, bias in algorithms, privacy concerns, the lack of transparency of complicated models, and possibilities of the unintended harms or misuse are raised [9]. Unless managed responsibly and designed with equity in mind, AI-based surveillance could either support current inequities or destroy trust in the government. Thus, this review generalizes the existing information on the use of AI in public health surveillance, outlining its prospects, constraints, and gaps in implementation. Our additional suggestion is a comprehensive roadmap of ethics-equity which focuses on responsible design, high-quality validation, transparent reporting, community involvement, and strong supervision. It specifically targets population-level surveillance systems, and not clinical decision-support systems and gives more emphasis on real-world utility, fairness, and sustainable governance.

2. Scope and Methods

In this case, a narrative review design was chosen as the solution to the fast changing, interdisciplinary, and policy-oriented focus on AI in the field of public health. This method allows synthesis of conceptually heterogeneous evidence types, such as policy documents, methodological frameworks, empirical studies and illustrative case examples, which do not lend to systematic review or meta-analysis.

This review was based on four primary domains of literature: (1) international guidance documents on AI and health; (2) recent methodological and empirical reviews of AI in public health; (3) reporting, evaluation, and governance standards of AI (e.g., TRIPOD+AI); and (4) case studies and commentaries on AI in high-impact academic and public health journals.

PubMed, WHO repositories, major public health journals and digital health journals, along with selected grey literature sources, such as Healthmap and BlueDot, were searched specifically to identify the literature. Search terms entailed AI plus ML combined with such keywords as public health, disease surveillance, ethics, bias, governance, and reporting standards. They focused on the recent publications to capture the dynamic policy environment.

The inclusion criteria included English-language publications that covered AI applications that were relevant to population health, surveillance, ethics, evaluation, or governance. The criteria of exclusion were: the studies were restricted to clinical decision-making with no public health consequences, the highly technical articles that did not provide the contextual analysis, and irrelevant or obsolete sources.

As the narrative review approach implies, source quality was determined qualitatively by the authority of the authority issuing it or the journal, reporting transparency, applicability to real-life health application, and input to the knowledge of ethical, methodological, or governance issues.

Figure 1 illustrates how an AI system, built on ML, NLP, and large multimodal models, connects to global data sources. Insights from these AI tools are then delivered to a clinician working on a computer, supporting clinical decision-making.

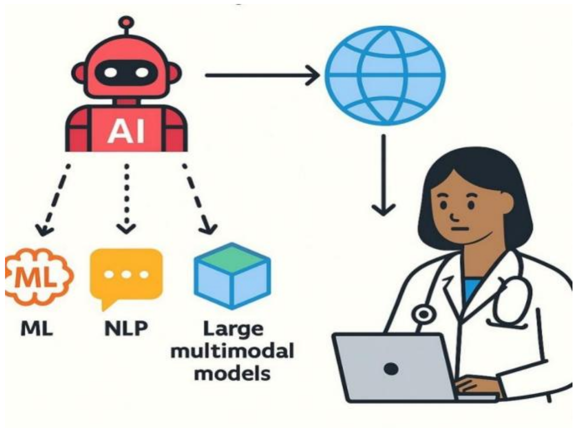

Figure 1. Conceptual overview of AI-enabled public-health surveillance and clinical decision support.

3. What is "AI" in Surveillance? Data, Models, and Pipelines

AI in public-health surveillance denotes effective computational systems that can consume, process, and make inferences about large, heterogeneous, and usually unstructured data to signal disease, predict trends, or detect anomalies [10]. In contrast to classical statistical surveillance tools that use fixed indicators and structured data, AI tools can use heterogeneous data and identify nonlinear, complex trends that might be antecedents to the occurrence of observable epidemiological changes [9,10]. Figure 1 illustrates the transformation of public health surveillance by AI.

3.1 Data Sources

Surveillance systems assisted by AI are based on a wide range of streams of data [11]. The high-specificity information available on confirmed or likely cases is provided by clinical sources, such as EHRs, the outcomes of laboratory tests, the patterns of pharmacy dispensing, and radiological results [12]. Earlier, although less specific indicators of symptoms at population level, are syndromic data, e.g., emergency department (ED) chief complaints, telemedicine consult logs, or over-the-counter medication sales [13]. In addition to conventional health-sector data, the new surveillance is using more and more digital trace data, such as internet search queries, social-media posts, news stories, and online engagement statistics, which can represent an emerging popular alarm, symptom reporting, or behavioural shifts [14]. Indirect environmental data, especially the wastewater viral loads, metagenomic-sequencing data, meteorological, and satellite-based remote-sensing indicators, give valuable but indirect information on pathogen distributions and ecology of vectors [15]. Moreover, the mobility and transportation data, including anonymized mobile-phone location history or airline passenger traffic, can be used to train models that describe the dynamics of spreading the disease and define the risky transmission corridors [16]. Formatted public-health reports such as notifiable disease registries and outbreak summaries are also considered a backbone and even can be used as a ground-truth to train or validate models [17]. Although these nontraditional data sources provide better sensitivity in detection, nontraditional datasets may be highly noisy, biased in demographics, or even time-varying, which requires preprocessing and calibration [18].

3.2 Models and Approaches

Modern surveillance is based on a number of AI and ML practices [19]. Case prediction, nowcasting or classification Case prediction Case prediction Case prediction using supervised learning models, including gradient boosting machine, random forests, or deep neural networks, are commonly applied in situations where training data labelled with case outcomes are available [20]. The approaches that are not controlled such as clustering algorithms and anomaly-detection framework are particularly helpful in the detection of surprising patterns without a predetermined case label-important to clarify emerging pathogens [21]. To identify early outbreak indications, the NLP techniques are used to extract event-relevant data, which has been obtained on unstructured text sources in order to enable automated scanning of global news feeds, scientific publications, and social-media posts [22]. The latest developments in transformer-based architectures and large language models have made significant gains in the capabilities of detecting nuanced semantic hints, inference of epidemiologically significant entities, and a combination of multilingual data streams [23]. New hybrid models that combine mechanistic epidemiological models with the uses of ML enhance the accuracy of predictions without compromising on the interpretability and domain relevancy. As an example, mechanistic models are used to describe the dynamics of transmission and ML modules are used to learn residual patterns (e.g. behavioral fluctuations or reporting delays) [24]. Multimodal deep-learning networks also increase the power of analysis by integrating text, numeric, genomic, and image data, making it possible to have multidimensional situational awareness [25].

3.3 Operation Procedures and Pipelines

AI surveillance systems must be based on powerful end-to-end pipelines to be able to operate in the real-worlds of the public-health industry [26]. These start with data ingestion, harmonization and preprocessing such as de-duplication, normalization, outlier management and metadata integration. Rigorous cross-validation, external testing and calibration are then required to carry out model training and validation to deal with sampling bias and temporal drift [27]. Human-in-the-loop triage is generally included in operational deployment, as epidemiologists study AI-generated alerts, perform threshold adjustments and provide feedback, which is used to improve the model iteratively [28]. Constant checking is required to identify the declining performance caused by altering reporting practices, new variants, or the sociotechnical change [29]. Last but definitely not least, the computational outputs are translated into actionable information by means of accessible interfaces-dashboards, APIs, automated alert systems-that are consumed by the professionals working in the fields of public-health [30,31]. In general, AI in surveillance is a developing system of data ecosystems, analytics, and operational procedures, which should be well combined to produce timely, reliable, and equitable intelligence on public-health. Figure 2 illustrates this.

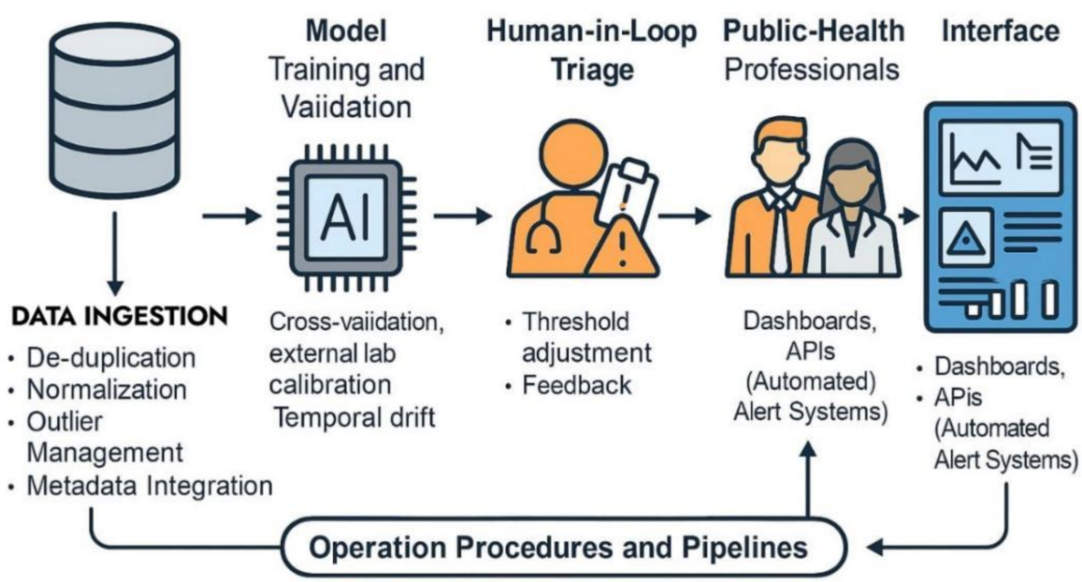

Figure 2. End-to-end human-in-the-loop AI framework for public health surveillance and decision support. The pipeline begins with data ingestion (de-duplication, normalization, outlier management, metadata integration), followed by model training and validation with attention to cross-validation, external calibration, and temporal drift. Model outputs then undergo human-in-the-loop triage and are delivered to public-health professionals through dashboards and automated alert systems, all governed by overarching operational procedures and pipelines.

4. The Promises: Why AI Matters to the Field of Public Health

Traditional public-health surveillance systems have been based on manual reporting, sentinel networks, and laboratory confirmation that, although dependable, are all necessarily limited by delays and lack of coverage and resource-lightness [32]. AI provides a possibility to complement these traditional methods with increasing the speed, scale, and depth of the population-health monitoring [33]. The increased interest toward AI-enabled surveillance by public-health agencies is based on a number of core promises as highlighted in Figure 3.

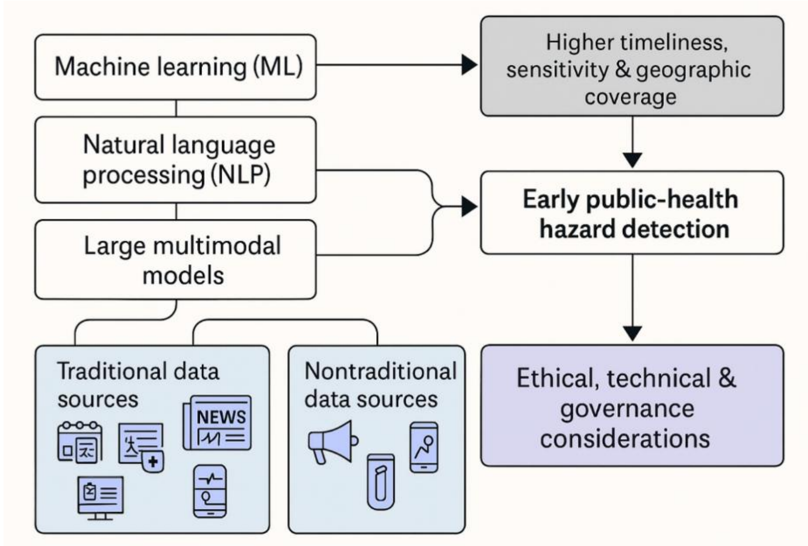

Figure 3. AI-enabled data integration for early public health hazard detection and governance. The integration leverages ML, NLP, and large multimodal models draw on both traditional and nontraditional data sources to improve the timeliness, sensitivity, and geographic coverage of surveillance. These capabilities enable early detection of public-health hazards, which must be accompanied by careful ethical, technical, and governance considerations.

4.1 Swift Detection and Increased Coverage of Observation

With AI systems, it is possible to detect events much earlier, and the coverage of observations is significantly increased, which is impossible with the use of traditional methods only. In contrast to traditional reporting systems, which require providers to submit reports or periodic surveys, AI is able to constantly take in and process vast nonhomogeneous data streams including news media, social-media noise, clinical data, emergency-department reports, digital search histories, and sensor data of the environment [34]. These systems have the capability of extracting weak and distributed signals that would be concealed by humans in the process of manually text mining, anomaly detection, as well as trend extraction. In reality, hybrid systems which incorporate automated analytics with expert epidemiological assessment have demonstrated that, in many cases, this type of model may predict upcoming events days or even weeks before the manual channels [35]. This time advantage can develop a time-critical moment of initial risk evaluation, quick field research, and pre-emptive communication, which may reduce the magnitude and effects of outbreaks.

4.2 Situational Granularity and Targeting

AI is associated with enhanced spatial, temporal, and population situation granularity. The contemporary machine-learners are capable of combining various streams of information, including human mobility patterns, meteorological information, ecological measurements, regional demographic characteristics, and case report information to produce finescale risk and transmission maps [32]. Such a degree of detail justifies more specific and commensurate action by the public health. As an illustration, a local vaccination plan, a work on vectors control, environmental clean-up operations, or a particular community outreach in particular neighbourhoods or sub-populations can be informed by risk-stratified models [36,37]. The changes in hotspots, seasonality, and context-specific vulnerabilities can be reflected in dynamic models that change over time as new data is received. This in principle enables the shift of public-health responses towards less general and more focused and context-specific interventions that are more preventive than reactive in nature.

4.3 Scale, Automation and Workload Offloading

Another potential benefit of AI is its ability to expand the surveillance capabilities and automate the routine analytic procedures. Since epidemic events tend to produce such huge amounts of digital data that people struggle to handle them, automated triage tools can be used to prioritize signals based on epidemiological relevance, raise concerns about anomalies to be examined by experts, and eliminate noise [9,38]. Such abilities are especially vital when large-scale events that have a lot of information occur like a pandemic, disaster related to climate, or a mass gathering where the incoming data volume and speed are higher than the processing capacity of human beings. Limited personnel are then able to be engaged in higher-order tasks like verification, contextual interpretation, communicating with the stakeholders and decision-making, than in cleaning up and coding and aggregating data manually [39]. Scalable AI tools can be used in resource-constrained environments to offer more scale to the analytic capacity to subnational levels, where trained epidemiologists are in short supply [40].

4.4 Data Combining and Identifying New Indicators

AI can as well be used to bring together data and detect novel indicators that would otherwise not be easily identified through manual techniques. The combination of wastewater monitoring, genomic sequencing, digital mobility information, pharmacy purchase information, wearable devices statistics, and population-level health indicators will allow building composite indicators that reflect more than one aspect of pathogen transmission or antimicrobial resistance [9,41]. This multimodal fusion has the potential to demonstrate the emerging biological or behavioural patterns, prior to their development into recognizable clinical disease. As a case in point, mixed genomic-epidemiological models can monitor the change of variants and the possible fitness benefits in near-time, whereas the use of wastewater-clinical fusion can detect silent transmission within the community when testing capacity is limited [41,42]. Equally, pattern recognition through AI can demonstrate abnormal prescribing patterns, syndrome clusters, or mobility patterns, which can indicate developing threats and prioritize investigation and devising adaptive surveillance countermeasures [10].

4.5 Planning and Allocation Resource Decision Support

Lastly, the AI can be used to provide strong decision-support solutions to the area of public health planning and resource allocation [43]. Under various conditions, forecasting models can predict the number of cases, the number of beds needed, or the geographic distribution to enable the health system to optimize the number of personnel required, bed capacity, and stock levels of critical commodities like oxygen, vaccinations, and therapeutics. At the population-health level, AI-based analytics would help coordinate cross-jurisdictional responses since this would indicate potential spillover between regions, where there is a weak point in a supply-chain, and where scarce public-health resources are rationally distributed during emergencies [32]. Tools based on scenarios may assist decision-makers to appreciate the trade-offs between alternative intervention strategies and predict downstream effects. Such tools when properly validated, clearly communicated and put into context can enhance the overall resilience and preparedness of public-health systems and enable timely and evidence based-decisions [44].

5. Practical Applications and Novel Applications

Public health surveillance based on AI has already moved past merely being a promise on paper, with numerous functioning systems demonstrating the ability of ML, NLP, and data-fusion systems to supplement current epidemiological systems [45]. These applications place emphasis on the possible and real-world challenges of implementing AI in various public health contexts. Some of these applications are shown in Table 1.

5.1 AI Surveillance Services Based on Events

One of the most prominent and oldest forms of AI usage in detecting outbreaks is an event-based surveillance platform. HealthMap and BlueDot among other systems automatically scan news sources around the world, social media and expert forums and unofficial coverage with NLP, geocoding, and rule-based or ML based relevance filters [34]. These platforms offer almost real-time situational awareness with the aggregation and categorization of signals across thousands of online sources, which supplements indicator-based surveillance. There is this operational usefulness of automating the detection of weak and distributed signals across the information ecosystem the way they performed during the emergence of COVID-19, including surfacing early signs before formal international announcements [46]. The same has been used to recognize zoonotic spillover reports, atypical clinical presentations, indicators of health-system stress, or indicators of public anxiety, which could be followed by official detection of clinical cases. More and more of such platforms are being scaled down to subnational capabilities, such as national surveillance dashboard integration and risk-assessment processes.

5.2 Wastewater Monitoring with AI Enhancers

Another quickly developing area is the wastewater monitoring with the support of ML. Wastewater surveillance has a population-level sample, which is an aggregation of symptomatic and asymptomatic infection, and is therefore a strong complement to clinical testing and syndromic surveillance. ML models have the ability to combine longitudinal data on wastewater concentration and environmental factors with sewer network attributes, demographical variables and past infection trends to produce now casts and short term predictions of prevalence rates of pathogens [47]. A number of the research works conducted in the context of the COVID-19 pandemic demonstrated that ML-enriched wastewater alerts tended to work out increasing transmission several days prior to case-based surveillance mechanisms, especially in environments with restricted diagnostic testing capacity or reporting delays [48]. With the expanding use of wastewater monitoring of pathogens, including influenza, RSV, antimicrobial-resistance (AMR) genes, and enteric viruses, AI-assisted analytics are likely to be highly valuable in the optimization of sampling strategies, inference of dynamics across a catchment, the field investigation of hotspots, and in early warning systems under both high-resource and low-resource settings [49].

5.3 Hybrid Epidemic Forecasting Systems

Another significant new application is hybrid epidemic forecasting systems. Ensemble models, integrating standard mechanistic epidemic models (e.g. SEIR or agent-based models), with ML parts to tune parameters, address systematic biases, or use exogenous data sources (e.g. mobility, climate variables and healthcare utilization) are used by many public-health agencies [50]. These combination methods have proven to provide best results with short-term predictions of influenza, dengue, and COVID-19 burden, which are translated into practical benefits to hospital capacity planning, oxygen and drug stocks, and emergency response organization. Concretely, hybrid systems may be integrated with the ordinary situation reports and decision dashboards to offer situation-based projections as well as uncertainty intervals [51]. ML has the potential to be integrated with mechanistic models so that the forecasts can be based on epidemiological theory but be more adaptable to changing conditions and data environments swiftly [51,52].

5.4 New AI-based Surveillance Applications

In addition to these set examples, there is a variety of new AI-aided surveillance applications being created:

Mobility-informed risk mapping: Combined mobile data, satellite image and ML-enhanced human-movement modelling can be used to detect probable transmission offshoots, cross-border connectivity, and regionally elevated vulnerability, notably where data touches on health facilities are scarce [53,54]. These outputs may inform goal-based vaccination, control of vectors or community outreach.

Genomic epidemiology pipelines: ML methods are being integrated in genomic surveillance to categorize lineages, predict the phenotypic features of variants (e.g. immune evasion or transmissibility proxies), and find abnormal mutational patterns, which may indicate an emerging threat [55]. These can be automated clustering and anomaly-detection tools used to prioritize sequences to be reviewed by experts and characterize them in the laboratory.

Digital phenotyping and sensor-based surveillance: Wearable devices, smartphone sensors, and passive behavioural data are under investigation to produce individual- and population-level health indicators that could be useful in early detection of respiratory infections, heat-related illness or mental-health stressors [56]. Although most of these systems are still in the experimental phase and present serious privacy issues, they demonstrate the widening boundary of the AI-powered data sources.

AI in AMR surveillance: AI development in AMR surveillance is being implemented via models to identify resistance in genomic sequences, forecasting resistance phenotypes, and used to predict AMR-risk in the regional scales [57]. Those systems can be used to complement the old system of laboratory networks by indicating the abnormal resistance profiles and directing empiric treatment or antibiotic-stewardship interventions.

5.5 Cutting Across Deployment Lessons

Combined, these illustrations depict the emergence of an ecosystem of AI-enabled tools that are an extension and, in other instances, replacement of traditional surveillance methods. There are operational maturity differences between uses cases: certain systems are integrated so deeply into regular business processes that they are no longer pilot projects or research prototypes. Some of the recurring cross-cutting themes are the necessity to have long-term data infrastructure, equity and representativeness, explicit data protection governance, and a human review and accountability mechanism. On the whole, AI in surveillance is a dynamic set of data ecosystems, analytical processes, and operationalized processes, which need to be thoroughly combined and contextually modified to produce the timely, reliable, and fair public-health intelligence [41, 58].

Table 1. Practical and novel applications of AI in public health surveillance.

Application area	Description/mechanism	Examples of AI methods used	Key public-health value
Event-Based Surveillance Systems	Automated scanning of global news, social media, and expert forums to detect early outbreak signals using NLP, geocoding and relevance filtering.	NLP pipelines, rule-based classifiers, ML relevance filters, geospatial tagging.	Near real-time situational awareness; early warnings; supplements indicator surveillance; supports national dashboards.
AI-Enhanced Wastewater Monitoring	ML analyzes pathogen concentration trends with environmental and demographic data to produce nowcasts and short-term prevalence predictions.	Regression models, time-series ML, data fusion, anomaly detection.	Early detection ahead of clinical reporting; hotspot identification; optimized sampling strategies.
Hybrid Epidemic Forecasting Systems	Mechanistic epidemic models combined with ML to correct biases, tune parameters, and integrate exogenous data such as mobility and climate.	Ensemble forecasts, ML-assisted SEIR models, neural networks.	Improved short-term forecasts; hospital and supply planning; uncertainty intervals for decision-making.
Mobility-Informed Risk Mapping	Mobile phone data, satellite imagery, and ML-based movement modelling to identify transmission corridors and high-vulnerability areas.	Geospatial ML, mobility modelling, remote-sensing analytics.	Supports targeted vaccination, vector control, and preparedness in low-data settings.
Genomic Epidemiology Pipelines	ML automates lineage classification and predicts phenotypic features of variants; detects unusual mutational patterns.	Sequence classifiers, clustering algorithms, anomaly detection.	Faster variant prioritization; strengthens genomic surveillance.
Digital Phenotyping & Sensor Surveillance	Wearables and smartphone data provide physiological and behavioral indicators for early detection of infections or stressors.	Signal processing, behavioral ML, sensor-fusion models.	Population-level and personalized early warning; expands non-clinical data sources.
AI for AMR Surveillance	ML predicts antimicrobial resistance from genomic data and identifies abnormal resistance trends regionally.	Genomic ML, phenotype prediction models, risk-mapping.	Complements lab networks; supports stewardship and empiric therapy decisions.
Cross-Cutting Deployment Lessons	Need for strong data infrastructure, equity, transparent governance, and human oversight across AI surveillance tools.	Governance frameworks, validation workflows, human-supervision models.	Ensures reliability, fairness, data protection, and sustainability.

Table 1 provides a structured overview of major practical and emerging applications of AI in public-health surveillance, summarizing eight domains ranging from event-based monitoring and wastewater analytics to hybrid forecasting and genomic epidemiology. For each domain, it outlines the core mechanism, the AI methods employed, and the specific public-health value delivered. The table highlights how AI complements traditional surveillance systems by enhancing early detection, prediction accuracy, and operational decision-making.

6. Key Pitfalls and Risks

Although there has been strong interest in the implementation of AI in public-health surveillance, the application in the real-world context is still plagued by challenging methodological, operational, ethical, and equity issues. These risks are not only hypothetical, as they have been proven several times throughout the health-AI literature and have the potential to destroy system performance as well as the trust of the population unless an active approach is taken toward them. Key types of risk gathered in the following subsections should guide the design, estimation, and regulation of AI-enabled surveillance systems. Table 2 illustrates these key pitfalls.

6.1 Data Quality, Representativeness and Sampling Bias

Surveillance based on AI frequently uses nontraditional, voluminous streams of data that is not based on randomized or population representative sampling frames, including social media posts, internet search history, mobility data, and wastewater signals [59]. Users of social-media are also younger, more urban and of higher socioeconomic status; geographic variations in internet penetration are also modelers of search trends meaning [60]. Environmental and mobility data sets can only represent certain groups of the population (e.g., smartphone users or customers who are on sewer connection) [61]. Models which are trained on such distorted inputs will run the risk of systematically characterizing disease burden or behavioral patterns. Unless addressed with some methodology (weighting, stratification, causal inference), the biases can also be carried over to subsequent decisions, which may underdetect underserved areas or overestimate well-monitored ones. The literature points out the fact that nonrepresentative big data such as their size might replicate or deepen inequities entrenched in social and infrastructural topography [62,63].

6.2 Algorithmic Bias and Harm is Detrimental

In addition to the data-level biases, algorithmic bias can be created by the model training decisions, label noise (i.e. incorrectly classified clinical cases), or optimisation functions that do not consider distributional fairness [64]. Using the example of anomaly-detection systems trained on dense urban data, clusters in cities will be regularly identified and sparse rural ones ignored [65]. On the same note, the performance of predictive models of influenza-like illness could be poor among people with poor access to healthcare since labels are based on clinical encounters [66]. The outcome of such systems leading to distributing resources may be an unequal allocation of vaccines, diagnostics, or interventions to control vectors. A review of the scholarly literature on health AI states that they find recurrent avenues of unfairness, including those related to data preprocessing and engineering features, threshold selection, and model updating, which highlights the importance of fairness auditing, subgroup performance assessment, and participatory design (in collaboration with affected communities) [67].

6.3 Interpretability and Trust

Deep neural networks and transformer models, which are considered high-capacity ML models, can also be highly predictive but do not provide end-users with any insight [68]. It is often needed by public-health analysts, epidemiologists, and policymakers to have clear reasoning, i.e., what data streams were used to raise an alert, how uncertainty was modelled, and what assumptions the model is driven by. The results produced by AI can be ignored, misunderstood, or even over relied on without information about the tool of interpretation like feature attribution, uncertainty estimates and model cards. Under-reliance and over-reliance are both dangerous: analysts can overlook early warning signs because they seem to be of black-box behavior, or may likewise assume spurious behavior without really questioning it [69,70].

6.4 Evaluation, Validation and Generalizability

Most models of AI surveillance perform well on retrospective or in-sample validations, but significantly worse when used prospectively. The concept drift-shifts are present in the real-world surveillance: changes in population behavior, pathogen characteristics, healthcare-seeking patterns, data collection processes, and environmental conditions [71]. The digital indicators can also be destabilized by changes in coding practices, laboratory workflows, or the interest of people in symptoms. Consequently, there is still a lack of externally validated evidence on most AI surveillance systems. Strategies based on the gold standard evaluation such as prospective trials, multi-site validation, calibration studies, and ablation studies, as well as compliance with reporting standards such as TRIPOD-AI or CONSORT-AI are required and uncommon [72,73]. Unless subjected to serious scrutiny, policymakers can choose the systems that have not been proved to be reliable in the infrastructures they will be implemented.

6.5 Privacy, Surveillance Creep and Consent

Surveillance in the interest of the health of the population has always been a matter of a delicate balance between the welfare of the people and their rights. The tension is compounded by AI which allows the incorporation of granular mobility traces, social-media content, genomic data, and behavioral indicators that can be sensitive or re-identifiable [74]. The results of models might indirectly indicate trends of minority or stigmatized groups even in cases where the data are aggregated. The absence of strong legal and governance structures poses a threat of the so-called surveillance creep whereby systems that were originally created to serve the public health interests are re-purposed to conduct policing, immigration enforcement, or even commercial exploitation [75]. Clear data reduction, differential privacy, and rigid purpose constraint are the main protective measures.

6.6 Operation and Equity Risks of LMICs

The countries of low and middle-income rank among those that have particular issues connected with the infrastructure discrepancies, a lack of digital literacy, unstable internet access, and absence of data-governance safeguards [76]. AI models that are trained off the shelf and trained in a high-income setting might not work when implemented in settings that have varying epidemiological distributions, health-seeking behavior, or reporting behaviors. Relying on external tools that are created and used may lead to intensifying the technological dependency and the global inequities. Workforce shortages can also contribute to the inability to monitor models, read outputs and respond to false alerts. Such systemic limitations cause the probability of model abuse, misuse, or rejection [77]. These pitfalls in combination point to the fact that effective implementation of AI in public-health surveillance needs more than technical expertise; it needs sociotechnical sensitivity, ongoing assessment, accountability systems, and equity-based governance strategies.

Table 2. Key pitfalls and risks of AI-enabled public-health surveillance and potential mitigation strategies.

Risk domain	Main issue	How it manifests in AI surveillance	Potential consequences	Possible mitigation/design considerations
Data quality, representativeness & sampling bias	Non-probability, skewed data sources	Reliance on social-media, search queries, mobility, and wastewater data that over-represent young, urban, higher-SES, smartphone or sewer-connected populations.	Systematic mischaracterization of disease burden and behaviour; under detection of underserved areas; overestimation in well-monitored settings; risk of amplifying existing inequities.	Use weighting/stratification; combine with representative surveillance; apply causal-inference methods; routinely audit data coverage and population subgroups; transparently report sampling limitations.
Algorithmic bias and harm	Model design and labeling choices that encode unfairness	Training on dense urban data leads anomaly detectors to flag cities but miss rural clusters; labels based on healthcare encounters lead to poor performance where access to care is limited; unfair thresholding and model updating.	Unequal allocation of vaccines, diagnostics, and interventions; systematic neglect of marginalized groups; reinforcement of structural inequities.	Fairness-aware objectives and constraints; subgroup performance assessment; fairness auditing along the pipeline (pre-processing, thresholds, updating); participatory design with affected communities.
Interpretability and trust	Black-box models limit understanding and appropriate use	Deep neural networks/transformers raise alerts without clear indication of data sources, assumptions, or uncertainty; lack of tools such as feature attribution, uncertainty estimates, or model cards.	Under-reliance (signals ignored as opaque) or over-reliance (spurious patterns accepted uncritically); misinterpretation of alerts by analysts and policymakers.	Embed explainability tools; provide model cards and documentation; communicate uncertainty explicitly; train users in interpretation; integrate human-in-the-loop review for critical decisions.
Evaluation, validation & generalizability	Limited prospective and external validation; concept drift	Models tuned to retrospective data perform poorly when underlying behaviour, pathogens, coding practices, or data workflows change; sparse use of prospective trials or multi-site validation.	Deployed systems may be unreliable in the settings where they are implemented; false reassurance or unnecessary alarms; difficult policy decisions based on weak evidence.	Conduct prospective and multi-site validation, calibration, and ablation studies; monitor for drift and recalibrate; follow reporting standards (e.g. TRIPOD-AI, CONSORT-AI); require evidence of external validity before scale-up.
Privacy, surveillance creep & consent	Use of granular, re-identifiable or sensitive data without robust safeguards	Integration of mobility traces, social-media content, genomic and behavioural data; models may infer trends in minority or stigmatized groups; systems later repurposed for policing, immigration, or commercial use.	Breaches of privacy and autonomy; stigmatization and discrimination; erosion of public trust; “surveillance creep” beyond public-health purposes.	Data minimization and reduction; privacy-preserving techniques (e.g. differential privacy, aggregation); clear purpose limitation and legal safeguards; independent oversight; transparent governance and consent processes.
Operational and equity risks in LMICs	Context misfit and capacity constraints in low- and middle-income countries	Off-the-shelf models from high-income settings applied in contexts with different epidemiology, health-seeking and reporting patterns; weak infrastructure, limited digital literacy, unstable internet, and fragile data governance; workforce shortages to monitor and act on outputs.	Poor performance, misuse or rejection of systems; increased technological dependency; widening global inequities in surveillance capacity and outbreak response.	Co-design and local adaptation of models; investment in infrastructure and digital literacy; building local data-governance frameworks; sustainable capacity for model monitoring and response; equity-focused global partnerships.

Table 2 outlines some of the key methodological, operational, ethical, and equity-related risks related to AI-enabled public-health surveillance in six areas including data quality, algorithmic bias, interpretability, evaluation, privacy, and implementation in LMICs. In each domain, the table describes how the risks in question are revealed in practice, the possible implications it has on populations and health systems as well as the signs of mitigation tactics to inform the design, deployment, and governance of AI surveillance tools.

7. Ethics and Equity Deployment Roadmap

Implementing AI in monitoring health issues in the society needs a structured governance system that is based on ethics and equity [78]. In contrast, compared to clinical decision-support tools, surveillance systems work at the population level, sometimes without individual consent and may affect policy decisions, resource allocation, mobility limitations and risk communication. All these features raise ethical responsibilities and require additional protection. Based on global principles, such as the WHO ethics and governance guidelines to AI in health and future AI reporting policies, this roadmap combines practical and viable actions to safeguard the responsible, equitable, and trustful use of AI in the field of communicating communicable diseases [79]. This is illustrated in Figure 4 and Table 3.

Table 3. Linking risk domains to mitigation actions and governance principles.

Risk domain	Key risks	Mitigation actions	Governance principles
Value and proportionality	Low added value; excessive data collection	Value-of-information analysis; data minimization	Benefit; proportionality
Bias and inequity	Unequal performance; unfair allocation	Subgroup evaluation; debiasing	Fairness
Opacity and misuse	Loss of trust; repurposing	Documentation; reporting standards	Transparency
Privacy and data misuse	Breaches; stigmatization	Privacy-preserving computation; security	Privacy
Automation bias	False alerts; missed events	Human-in-the-loop review	Accountability
Performance drift	Degraded accuracy	Monitoring; retraining	Evaluation
Social harm	Community mistrust	Engagement; communication	Participation
Dependency inequity	Lack of sustainability	Capacity building; partnerships	Equity

Table 3 provides an overview of the systematic correlation between key risk areas that are related to AI and data-driven systems to particular risks, the mitigation measures, and the guiding principles of governance. It also shows how technical, ethical, and social issues (bias, privacy, automation bias, and performance drift) can be resolved with specific operational interventions based on such principles as fairness, transparency, accountability, and equity. In general, the table offers a systematic system of coordinating the risk management habits and responsible governance in the medical field and AI-based medical practices.

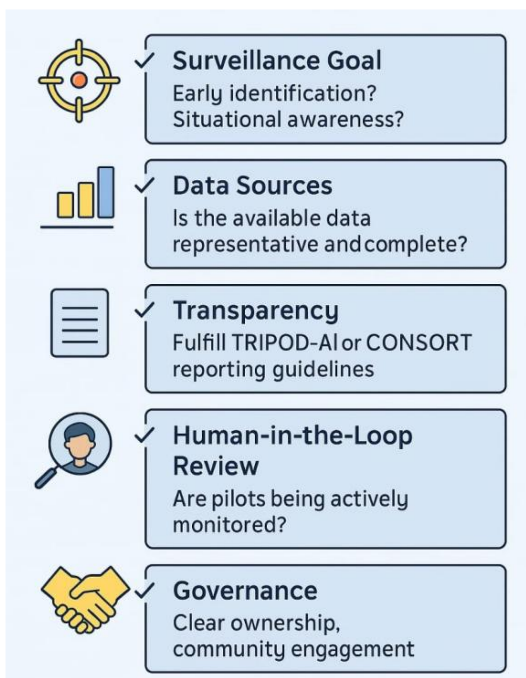

Figure 4. Governance, transparency, and oversight in AI-based surveillance systems. This checklist includes key considerations for implementing AI-enabled public-health surveillance systems, including clearly defining the surveillance goal, ensuring data sources are representative and complete, and maintaining transparency through adherence to reporting guidelines such as TRIPOD-AI or CONSORT. It also emphasizes the importance of human-in-the-loop review to actively monitor pilots and strong governance with clear ownership and community engagement.

7.1 Benefits, Proportionality and Necessity

The implementation of AI systems must be based on their ability to add value to current epidemiological practices [80]. The assessment of value of information and proportionality can be used to evaluate whether or not increased risks and costs are rewarded by timeliness, accuracy, sensitivity and coverage gains [81,82]. The collection of data must be kept to a minimum, simpler models used where sufficient, and an increase in the scope of surveillance must be clearly justified by the public health requirements.

7.2 Fairness and Non-Discrimination

Due to the fact that the results of surveillance are used to allocate resources, fairness is therefore necessary [83]. Early equity impact assessment must be done to highlight poorly represented or poorly defined populations. The performance of the model is to be assessed between sociodemographic groups and mitigation measures, including reweighting, debiasing, calibration, or threshold adjustment should be implemented in case the disparities are observed [84].

7.3 Transparency and Documentation

Transparency is required in trust, accountability and reproducibility. Implemented systems must have well documented information, explaining data sources, intended use, assumptions, limitations and failure modes [85,86]. Scientific rigor is enhanced by adherence to reporting standards like the TRIPOD-AI and the CONSORT-AI.

7.4 Privacy, Data Governance and Security

The surveillance made by AI should not violate epidemiological value but reduce the threat to privacy [87]. Federated learning and different privacy techniques like differential privacy can decrease the centralization of data [88]. There must be explicit data-use agreements, data controls and cybersecurity measures.

7.5 Human Oversight and Accountability

The AI must not be used to substitute epidemiological knowledge. Human-in-the-loop designs allow understanding the alerts in context and minimize the chances of responding to spurious events [89]. Accountability and learning are aided by the audit trails.

7.6 Evaluation and Continuous Monitoring

It is necessary to conduct post-deployment monitoring and pre-deployment validation. These should be performance, calibration, subgroup effects, false alerts and concept drift are to be checked regularly with formal retraining and version control policy [90].

7.7 Community Engagement and Public Transparency

Community participation, civil society, and involvement of frontline actors in the public health also aid in the process of identifying the risks surrounding communities and enhancing trust [91]. Good open communication regarding the role of the system and protection is important.

7.8 Capacity Building and Equitable Partnerships

The deployment of AI sustainably needs local capacity building, infrastructure investments, skills of the workforce, and controlling data especially in low- and middle-income nations. Fair partnerships need to have equal decision making and equal distribution of benefits [92-94].

8. Checklist Implementation

Public-health surveillance systems with AI potential should be introduced in a systematic, stepwise manner to maintain technical strength, ethical accountability, and functional viability, especially in low- and middle-income countries (LMICs), where resources and institutional capacity tend to be limited [95]. First, the implementers need to identify the purpose of the surveillance (e.g., early outbreak detection, situational awareness, or short-term prediction), and clearly explain the new value of AI compared to more basic rule-based or statistical methods. In LMIC context, this is essential to prevent opportunity costs in which scarce funding and human resources can instead be spent in other important public-health processes [96]. Capacity building is the key to successful checklist implementation. Instead of depending on vendor services or short-term contractors, LMICs can focus on modular training local epidemiologists, data managers, health informatics officers in domains like quality assessment of data, basic literacy in machine-learning, and model interpretation. Examples of these in practice are integrating AI training into Field Epidemiology Training Programs (FETPs), using regional centers of excellence, or using train-the-trainer models to facilitate transfer of skills between districts and ministries. Notably, the checklist is adaptable to fit the complexity of the models to the available expertise, and interpretable models should be used in situations where the capacity of advanced data science is scarce, it minimizes inequities in dependencies.

In regards to data governance, data sources like routine health information systems, laboratory reporting systems, mobile health applications, and community-based surveillance should be listed and evaluated based on completeness, representativeness, and bias. In LMICs where data fragmentation and missingness are the norm, the checklist ought to be emphasised on incremental improvement instead of the wide sacking of flawed datasets. There are concrete adaptations, such as minimum data principles, data-sharing accords between ministries and implementing partners, and privacy-by-design privacy designs, including data minimization, pseudonymization, and safe local data storage. These solutions will contribute to safeguarding vulnerable groups, as well as allow the utilization of AI in the event of a lack of elaborate national data-protection laws [97]. Model selection should focus on transparency and interpretability when the consequences of surveillance output are of high importance to the public-health or politics. It is especially applicable in LMICs, where the level of trust in digital systems might be low and the mechanisms of accountability evolve. Pre-registering analytic plans and reporting in a manner consistent with the TRIPOD-AI or CONSORT-AI guidelines increases the scientific rigor of the research and ease external review, even where the locally generated evidence is scarce [98]. The checklist can be adapted in the most practical way by providing the opportunity to comply with the reporting standards in phases with the primary focus on the main aspects of a model purpose, data sources, and performance metrics.

The partnership models will be relevant to maintaining AI-enabled surveillance in its limited environments. The co-development can be enforced by fair collaboration between governments, academic centers, regional public-health organizations and technologies developers instead of transfer of technology. Such cases are public-academic collaboration to validate models, South-South collaboration to share infrastructure, and agreements to be locally owned by data and algorithms. There should be a clear demand on the role clarity, data ownership, long term maintenance and exit strategy as the checklist to avoid vendor lock in and to achieve sustainability [99]. In terms of operation, human-in-the-loop review mechanism, audit trail, and prospective pilots are necessary but should be scaled accordingly. In LMICs, it can be the introduction of AI outputs into the current surveillance review meetings, the use of simple dashboards that can be viewed in low-bandwidth systems, and testing models in the selected districts and then rolled out nationwide. Marginalized groups (such as rural communities and informal settlements) should be explicitly named in equity and privacy evaluation to make sure that the deployment of algorithms does not worsen the current health inequalities.

Lastly, good governance structures (including a sense of responsibility in performing control roles, community involvement and feedback systems) become the cornerstone of responsible execution. Trust and legitimacy in the LMICs can be established by involving the local leaders, civil society organizations, and frontline health workers in the system design and evaluation. Through explicit translation of the requirements of checklists into local capacity, and emphasis on proportionality, participation, and equity, the AI-enabled surveillance systems can be realized with no significant increase in the existing inequities and a higher possibility of long-term benefit to the population [100].

9. Gaps in Research and Policy Priorities

Notable advancements have been made in the AI use of public-health surveillance, but a number of methodological and governance gaps remain that prevent the application of AI in ways that are safe, equitable, and effective [101]. It is critical to address these gaps in order to translate the technical innovation into the sustainable value of public-health.

There are still significant gaps in research. First, there is an urgent demand of methodological frameworks that allow sound fairness assessment in streamlining and real-time cases of surveillance. The current fairness metrics are generally based on the notion of static data, when the surveillance systems work in the dynamic environments with concept drift, changing quality of data, and changing behaviors of population [102]. Second, privacy preserving architectures need to be advanced, particularly the computationally light, affordable, and low resource feasible solutions [103]. Most of the current methods, including differential privacy and secure multi-party computation, are either complex or resource-sensitive in nature. Third, the area does not have standardized benchmark datasets that are privacy-sensitive, allowing a clear comparison of surveillance algorithms. The reproducible research may be made possible by synthetic or federated benchmark ecosystems, which would prevent sensitive health information to be disclosed.

Priority of policy is also critical. International norms and agreements that will limit the non-health use of surveillance produce is in urgent need to ensure that mission creep into policing, migration control or commercial profiling does not occur [104]. The mechanism of funding should also support mechanisms to build capacity in digital epidemiology, and open, interoperable infrastructure, which facilitates cross-border data sharing without undermining sovereignty should be supported by policy makers. Lastly, regulatory authorities should offer more explicit directions to the AI tools in the field of public health. Instead of strict, uniform rules, risk-based analysis, transparency, and ongoing monitoring should be based on the flexible good practice systems [105,106]. These are some priorities that will be used to make AI-powered surveillance reliable, fair, and consistent with the purposes of promoting health to the population.

10. Conclusion

The review has shown that AI has significant potential to enhance public-health surveillance through providing the ability to detect population-level health signals earlier, more granular, and responsive. Nonetheless, the results always demonstrate that technical performance per se is no longer a sufficient reason to deploy. In the absence of stringent validation, fair data usage, clear reporting, and responsible control, AI-based surveillance systems have the potential to increase the level of existing bias, weaken trust, and worsen health inequity. Operationalization thus must be done in a manner that is both technical design and ethically, legally, and socially appropriate. The priorities that were identified throughout the literature review are proper choice of problems, use of relevant and context-specific data, ongoing model validation, clear documentation, meaningful human oversight, strong privacy, and long term community involvement. Through an ethics-and-equity roadmap that is pragmatic and based upon accepted reporting standards (e.g. TRIPOD-AI and CONSORT-AI) and global principles of public-health (e.g. the use of systems), stakeholders will be able to transform emerging AI capabilities into surveillance systems that are not only innovative, but also trustworthy, fair, and fit to real-world practice of public-health. The AIs have revolutionary possibilities on the capacity of public-health surveillance to enable more extensive and timely and more detailed detection of population health indicators. However, technical workability is not enough; unless strict validation is implemented, fairness ensures, transparent reporting and responsible governance, the AI systems are likely to enhance negative effects and decrease social confidence. To put operationalized AI surveillance into practice, there should be coordinated technical, ethical and social solutions: proper problem choice, representative data and validation, transparent documentation, human control, privacy control and meaningful community participation. Adhering to an ethics-equity roadmap that is pragmatic and in line with reporting standards (e.g., TRIPOD +AI, CONSORT-AI) and international recommendations, practitioners can put the chances of AI becoming safer, fairer and more effective public-health practice higher.

Authors Contribution

Authors IGC, ESI, PPC and UVA contributed variously at different stages and aspects of the manuscript such as conception, literature search, writing of original draft and review/editing of the manuscript. All authors consented for the manuscript to be published in DPE.

Conflict of Interest

The authors declare no conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

References

[1] Shah HA, Househ M. Concepts, objectives and analysis of public health surveillance systems. Computer Methods and Programs in Biomedicine Update, 2024, 5, 100136. DOI: 10.1016/j.cmpbup.2024.100136

[2] Shen Y, Liu Y, Krafft T, Wang Q. Progress and challenges in infectious disease surveillance and early warning. Medicine Plus, 2025, 100071. DOI: 10.1016/j.medp.2025.100071

[3] Goswami N. A dual burden dilemma: Navigating the global impact of communicable and non-communicable diseases and the way forward. International Journal of Medical Research, 2024, 12(3), 65-77. DOI: 10.55489/ijmr.123202412

[4] Sadr H, Nazari M, Khodaverdian Z, Farzan R, Yousefzadeh-Chabok S, Ashoobi MT, et al. Unveiling the potential of artificial intelligence in revolutionizing disease diagnosis and prediction: a comprehensive review of machine learning and deep learning approaches. European Journal of Medical Research, 2025, 30(1), 418. DOI: 10.1186/s40001-025-02680-7

[5] Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, et al. Artificial intelligence for multimodal data integration in oncology. Cancer cell, 2022, 40(10), 1095-1100. DOI: 10.1016/j.ccell.2022.09.012

[6] MacIntyre CR, Chen X, Kunasekaran M, Quigley A, Lim S, Stone H, et al. Artificial intelligence in public health: the potential of epidemic early warning systems. Journal of International Medical Research, 2023, 51(3), 03000605231159335. DOI: 10.1177/03000605231159335

[7] Li G, Diggle P, Blangiardo M. Integrating wastewater and randomised prevalence survey data for national COVID surveillance. Scientific Reports, 2024, 14(1), 5124. DOI: 10.1038/s41598-024-55752-9

[8] Levy JI, Gangavarapu P, Pilz DA, Jesvaghane MA, Steedman A, Zeller M, et al. Real-time, multi-pathogen wastewater genomic surveillance with Freyja 2. medRxiv[Preprint], 2025, 2025.07.26.25332245. DOI: 10.1101/2025.07.26.25332245.
[9] 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, 13, 1609615. DOI: 10.3389/fpubh.2025.1609615
[10] Zeng D, Cao Z, Neill DB. Artificial intelligence–enabled public health surveillance-from local detection to global epidemic monitoring and control. InArtificial intelligence in medicine, 2021, 437-453. DOI: 10.1016/B978-0-12-821259-2.00022-3
[11] Nguyen MT, Truong LH, Tran TT, Chien CF. Artificial intelligence based data processing algorithm for video surveillance to empower industry 3.5. Computers & Industrial Engineering, 2020, 148, 106671. DOI: 10.1016/j.cie.2020.106671
[12] Alqahtani SS, Menachery SJ, Alshahrani A, Albalkhi B, Alshayban D, Iqbal MZ. Artificial intelligence in clinical pharmacy-A systematic review of current scenario and future perspectives. Digital health, 2025, 11, 20552076251388145. DOI: 10.1177/20552076251388145
[13] Dardour A, El Haji E, Begdouri MA. Video surveillance and artificial intelligence for urban security in smart cities: A review of a selection of empirical studies from 2018 to 2024. Computer Sciences & Mathematics Forum, 2025, 10(1), 15. DOI: 10.3390/cmsf2025010015
[14] Madamalla E. Digital epidemiology in action: A cross-platform review of social media and internet-based surveillance for infectious disease outbreaks. International Journal of Innovative Science and Research Technology, 2025. DOI:10.38124/ijisrt/25jul887
[15] McCall C, Wu H, Miyani B, Xagoraraki I. Identification of multiple potential viral diseases in a large urban center using wastewater surveillance. Water Research, 2020, 184, 116160. DOI: 10.1016/j.watres.2020.116160
[16] Serena L, Marzolla M, D'Angelo G, Ferretti Stefano. A review of multilevel modeling and simulation for human mobility and behavior. Simulation Modelling Practice and Theory, 2023, 127, 102780. DOI: 10.1016/j.simpat.2023.102780
[17] Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. Proceedings of the National Academy of Sciences, 2009, 106(51), 21484-21489. DOI: 10.1073/pnas.0906910106
[18] Ghozzi S, Zacher B, Ullrich A. Supervised learning for automated infectious-disease-outbreak detection. Online Journal of Public Health Informatics, 2019, 11(1), e62437. DOI: 10.5210/ojphi.v11i1.9770
[19] Siddique S, Haque MA, George R, Gupta KD, Gupta D, Faruk MJH. Survey on machine learning biases and mitigation techniques. Digital, 2023, 4(1), 1-68. DOI: 10.3390/digital4010001
[20] Chadwick F. AI for disease surveillance in the modern era: Early detection and rapid response. Global Journal of Medical and Biomedical Case Reports, 2025, 1, (1), 1-10.
[21] Babu Nuthalapati S, Nuthalapati A. Accurate weather forecasting with dominant gradient boosting using machine learning. International Journal of Science and Research Archive, 2024, 12(2), 408-422. DOI: 10.30574/ijsra.2024.12.2.1246
[22] Ghajari G, PK MK, Amsaad F. Hybrid efficient unsupervised anomaly detection for early pandemic case identification. NAECON 2024-IEEE National Aerospace and Electronics Conference, 2024, 279-284. DOI: 10.1109/NAECON61878.2024.10670679
[23] Clement M. Natural language processing (NLP) for document analysis. Artificial Intelligence, 2025.
[24] Akhtarshenas A, Dini A, Ayoobi N. ChatGPT or a silent everywhere helper: A survey of large language models. arXiv preprint arXiv: 2503.17403, 2025. DOI: 10.48550/arXiv.2503.17403
[25] Wu B, Luo S, Suh CS. A comprehensive review of propagation models in complex networks: From deterministic to deep learning approaches. arXiv preprint arXiv: 2410.02118, 2024. DOI: 10.48550/arXiv.2410.02118
[26] Al-Zoghby AM, Ismail Ebada A, Saleh AS, Abdelhay M, Awad WA. A comprehensive review of multimodal deep learning for enhanced medical diagnostics. Computers, Materials & Continua, 2025, 84(3), 4155-4193. DOI: 10.32604/cmc.2025.065571
[27] Adeniyi AE, Falola PB, Awotunde JB, Lawrence MO, Ubaru S, Randle O, et al. Edge and cloud-based deployment of AI-driven integration in smart healthcare. HealthTech Horizons: AI-Infused Metaverse Solutions for Smart Healthcare Systems, 2025, 457-489. DOI: 10.1007/978-3-031-99946-8_19
[28] Yates LA, Aandahl Z, Richards SA, Brook BW. Cross validation for model selection: a review with examples from ecology. Ecological Monographs, 2023, 93(1), e1557. DOI: 10.1002/ecm.1557
[29] Da'Costa A, Teke J, Origbo JE, Osonuga A, Egbon E, Olawade DB. AI-driven triage in emergency departments: A review of benefits, challenges, and future directions. International Journal of Medical Informatics. 2025, 105838. DOI: 10.1016/j.ijmedinf.2025.105838
[30] Özturan B, Quintana-Mathé A, Grinberg N, Ognyanova K, Lazer D. Declining information quality under new platform governance. Harvard Kennedy School Misinformation Review, 2025, 6(3), 1-29. DOI: 10.37016/mr-2020-176
[31] Schulze A, Brand F, Geppert J, Böl GF. Digital dashboards visualizing public health data: a systematic review. Frontiers in Public Health. 2023, 11, 999958. DOI: 10.3389/fpubh.2023.999958
[32] Stahlman G, Yanovitzky I, Kim M. Design, application, and actionability of US public health data dashboards: Scoping review. Journal of Medical Internet Research, 2025, 27, e65283. DOI: 10.2196/65283
[33] da Silva Mendes VI, Milhomens de Ferreira Mendes B, Moura RP, Mota Lourenço I, Ferreira Alves Oliveira M, Lee Ng K, et al. Harnessing artificial intelligence for enhanced public health surveillance: A narrative review. Frontiers in Public Health, 2025, 13, 1601151. DOI: 10.3389/fpubh.2025.1601151
[34] Henry A, Edward T. Utilizing artificial intelligence in population health management: Forecasting health trends and enhancing community well-being through data-driven insights. NatureBio, 2024, 1(1), 39-47.
[35] Methuku V. NLP and AI for public health intelligence: Automating disease surveillance from unstructured data. ICCK Transactions on Emerging Topics in Artificial Intelligence, 2025, 2(1), 43-56. DOI: 10.62762/TETAI.2025.222799

[36] Tuan DA, Uyen PV. Bridging the predictive divide: A hybrid early warning system for scalable and real-time dengue surveillance in LMICs. Acta Tropica, 2025, 107765. DOI: 10.1016/j.actatropica.2025.107765
[37] Zhang J, Wang X, Rong L, Pan Q, Bao C, Zheng Q. Planning for the optimal vaccination sequence in the context of a population-stratified model. Socio-Economic Planning Sciences, 2024, 92, 101847. DOI: 10.1016/j.seps.2024.101847
[38] Hongoh V, Hoen AG, Aenishaenslin C, Waaub JP, Bélanger D, Michel P, et al. Spatially explicit multi-criteria decision analysis for managing vector-borne diseases. International Journal of Health Geographics, 2011, 10(1), 70. DOI: 10.1186/1476-072X-10-70
[39] Aryffin HAK, Sahbudin MAB, Pitchay SA, Abhalim AH, Sahbudin I. Technological trends in epidemic intelligence for infectious disease surveillance: A systematic literature review. PeerJ Computer Science, 2025, 11, e2874. DOI: 10.7717/peerj-cs.2874
[40] Panteli D, Adib K, Buttigieg S, Goiana-da-Silva F, Ladewig K, Azzopardi-Muscat N, et al. Artificial intelligence in public health: promises, challenges, and an agenda for policy makers and public health institutions. Lancet Public Health, 2025, 10(5), e428-432. DOI: 10.1016/S2468-2667(25)00036-2
[41] Olawade DB, Wada OJ, David-Olawade AC, Kunonga E, Abaire O, Ling J. Using artificial intelligence to improve public health: a narrative review. Frontiers in Public Health, 2023, 11, 1196397. DOI: 10.3389/fpubh.2023.1196397
[42] Pennisi F, Pinto A, Ricciardi GE, Signorelli C, Gianfredi V. The role of artificial intelligence and machine learning models in antimicrobial stewardship in public health: a narrative review. Antibiotics, 2025, 14(2), 134. DOI: 10.3390/antibiotics14020134
[43] Lu D, Kalantar KL, Glascock AL, Chu VT, Guerrero ES, Bernick N, et al. Simultaneous detection of pathogens and antimicrobial resistance genes with the open source, cloud-based, CZ ID platform. Genome Medicine, 2025, 17(1), 46. DOI: 10.1186/s13073-025-01480-2
[44] Karaarslan E, Aydın D. An artificial intelligence-based decision support and resource management system for COVID-19 pandemic. Data Science for COVID-19, 2021, 25-49. DOI: 10.1016/B978-0-12-824536-1.00029-0
[45] Borchering RK, Healy JM, Cadwell BL, Johansson MA, Slayton RB, Wallace M, et al. Public health impact of the US scenario modeling hub. Epidemics, 2023, 44, 100705. DOI: 10.1016/j.epidem.2023.100705
[46] Ward P, Young A. Natural language processing and machine learning methods in public health surveillance: A narrative review. JMIR Preprints, 2020, 26351. DOI: 10.2196/preprints.26351
[47] Zhang Y, Chen K, Weng Y, Chen Z, Zhang J, Hubbard R. An intelligent early warning system of analyzing Twitter data using machine learning on COVID-19 surveillance in the US. Expert Systems with Applications. 2022, 198, 116882. DOI: 10.1016/j.eswa.2022.116882
[48] Thompson JR, Nancharaiah YV, Gu X, Lee WL, Rajal VB, Haines MB, et al. Making waves: wastewater surveillance of SARS-CoV-2 for population-based health management. Water Research, 2020, 184, 116181. DOI: 10.1016/j.watres.2020.116181
[49] Ai Y, He F, Lancaster E, Lee J. Application of machine learning for multi-community COVID-19 outbreak predictions with wastewater surveillance. PLOS One, 2022, 17(11), e0277154. DOI: 10.1371/journal.pone.0277154
[50] Armenta-Castro A, de la Rosa O, Aguayo-Acosta A, Oyervides-Muñoz MA, Flores-Tlacuahuac A, Parra-Saldívar R, et al. Interpretation of COVID-19 epidemiological trends in Mexico through wastewater surveillance using simple machine learning algorithms for rapid decision-making. Viruses, 2025, 17(1), 109. DOI: 10.3390/v17010109
[51] Chen S, Janies D, Paul R, Thill JC. Leveraging advances in data-driven deep learning methods for hybrid epidemic modeling. Epidemics, 2024, 48, 100782. DOI: 10.1016/j.epidem.2024.100782
[52] Ye Y, Pandey A, Bawden C, Sumsuzzman DM, Rajput R, Shoukat A, et al. Integrating artificial intelligence with mechanistic epidemiological modeling: a scoping review of opportunities and challenges. Nature Communications, 2025, 16(1), 581. DOI: 10.1038/s41467-024-55461-x
[53] Rahujo A, Atif D, Inam S A, Khan AA, Ullah S. A survey on the applications of transfer learning to enhance the performance of large language models in healthcare systems. Discover Artificial Intelligence, 2025, 5(1), 90. DOI: 10.1007/s44163-025-00339-0
[54] Inam SA. A review of artificial intelligence for predicting climate driven infectious disease outbreaks to enhance global health resilience. Discover Public Health, 2025, 22(1), 738. DOI: 10.1186/s12982-025-01167-4
[55] Lessani MN, Li Z, Jing F, Qiao S, Zhang J, Olatosi B, Li X. Human mobility and the infectious disease transmission: a systematic review. Geo-Spatial Information Science, 2024, 27(6), 1824-1851. DOI: 10.1080/10095020.2023.2275619
[56] Borkenhagen LK, Allen MW, Runstadler JA. Influenza virus genotype to phenotype predictions through machine learning: A systematic review: Computational prediction of influenza phenotype. Emerging Microbes & Infections, 2021, 10(1), 1896-1907. DOI: 10.1080/22221751.2021.1978824
[57] Sheikh M, Qassem M, Kyriacou PA. Wearable, environmental, and smartphone-based passive sensing for mental health monitoring. Frontiers in Digital Health, 2021, 3, 662811. DOI: 10.3389/fdgth.2021.662811
[58] Bilal H, Khan MN, Khan S, Shafiq M, Fang W, Khan RU, et al. The role of artificial intelligence and machine learning in predicting and combating antimicrobial resistance. Computational and Structural Biotechnology Journal, 2025, 27, 423-429. DOI: 10.1016/j.csbj.2025.01.006
[59] Ayesiga I, Yeboah MO, Okoro LN, Edet EN, Gmanyami JM, Ovye A, et al. Artificial intelligence-enhanced biosurveillance for antimicrobial resistance in sub-Saharan Africa Open Access. International Health (1876-3413), 2025, 17(5), 795. DOI: 10.1093/inthealth/ihae081
[60] Rilkoff H, Struck S, Ziegler C, Faye L, Paquette D, Buckeridge D. Innovations in public health surveillance: An overview of novel use of data and analytic methods. Canada Communicable Disease Report, 2024, 50(3-4), 93. DOI: 10.14745/ccdr.v50i34a02
[61] Hargittai E. Potential biases in big data: Omitted voices on social media. Social Science Computer Review, 2020, 38(1), 10-24. DOI: 10.1177/0894439318788322

[62] Nijs KD, Omodei E, Sekara V. Data bias in human mobility is a universal phenomenon but is highly location-specific. arXiv preprint arXiv:2508.00149, 2025. DOI:10.48550/arXiv.2508.00149

[63] Erfani A, Frias-Martinez V. A fairness assessment of mobility-based COVID-19 case prediction models. PLOS One, 2023, 18(10), e0292090. DOI: 10.1371/journal.pone.0292090

[64] Schlosser F, Sekara V, Brockmann D, Garcia-Herranz M. Biases in human mobility data impact epidemic modeling. arXiv preprint arXiv:2112.12521, 2021. DOI: 10.48550/arXiv.2112.12521

[65] Paulus JK, Kent DM. Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. NPJ Digital Medicine, 2020, 3(1), 99. DOI: 10.1038/s41746-020-0304-9

[66] Zhu Y, Ting KM, Carman MJ, Angelova M. CDF Transform-and-Shift: An effective way to deal with datasets of inhomogeneous cluster densities. Pattern Recognition, 2021, 117, 107977. DOI: 10.1016/j.patcog.2021.107977

[67] Wolk DM, Lanyado A, Tice AM, Shermohammed M, Kinar Y, Goren A, et al. Prediction of influenza complications: development and validation of a machine learning prediction model to improve and expand the identification of vaccine-hesitant patients at risk of severe influenza complications. Journal of Clinical Medicine, 2022, 11(15), 4342. DOI: 10.3390/jcm11154342

[68] Chandra J, Kaur R, Sahay R. Integrated Framework for Equitable Healthcare AI: Bias Mitigation, Community Participation, and Regulatory Governance. 2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT), 2025, 819-825. DOI: 10.1109/CSNT64827.2025.10968102

[69] Alqudah AM, Moussavi Z. A review of deep learning for biomedical signals: Current applications, Advancements, future prospects, interpretation, and challenges. Computers, Materials & Continua, 2025, 83(3), 3753-3841. DOI: 10.32604/cmc.2025.063643

[70] Wadhwani A, Jain P. Machine learning model cards transparency review: Using model card toolkit. 2020 IEEE Pune Section International Conference (PuneCon), 2020, 133-137. DOI: DOI: 10.1109/PuneCon50868.2020.9362382

[71] Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, et al. Model cards for model reporting. Proceedings of the Conference on Fairness, Accountability, and Transparency , 2019, 220-229. DOI: 10.1145/3287560.3287596

[72] Hinder F, Vaquet V, Hammer B. One or two things we know about concept drift-a survey on monitoring in evolving environments. Part A: detecting concept drift. Frontiers in Artificial Intelligence, 2024, 7, 1330257. DOI: 10.3389/frai.2024.1330257

[73] dos Santos Silva GF, Barcellos Filho FN, Wichmann RM, da Silva Junior FC, Chiavegatto Filho AD. Strategies for detecting and mitigating dataset shift in machine learning for health predictions: A systematic review. Journal of Biomedical Informatics, 2025, 104902. DOI: 10.1016/j.jbi.2025.104902

[74] Al-Ewaidat OA, Naffaa MM. Emerging AI-and biomarker-driven precision medicine in autoimmune rheumatic diseases: From diagnostics to therapeutic decision-making. Rheumato, 2025, 5(4), 17. DOI: 10.3390/rheumato5040017

[75] Aspell N, Goldsteen A, Renwick R. Dicing with data: the risks, benefits, tensions and tech of health data in the iToBoS project. Frontiers in Digital Health, 2024, 6, 1272709. DOI: 10.3389/fdgth.2024.1272709

[76] Mariner WK. Mission creep: public health surveillance and medical privacy. Boston University Law Review, 2007, 87, 347-395.

[77] Andrade JB, Fagundes TP, Katsuyama E, Silva GS. Digital health in low-resource settings: Comprehensive challenges and opportunities with a focus on stroke care. Stroke, 2026, 57(1), 245-253. DOI: 10.1161/STROKEAHA.125.050448

[78] Craig AT, Joshua CA, Sio AR, Donoghoe M, Betz-Stablein B, Bainivalu N, et al. Epidemic surveillance in a low resource setting: lessons from an evaluation of the Solomon Islands syndromic surveillance system, 2017. BMC Public Health, 2018, 18(1), 1395. DOI: 10.1186/s12889-018-6295-7

[79] Blasimme A, Vayena E. The ethics of AI in biomedical research, patient care and public health. Patient Care and Public Health, 2020, 703-718. DOI: 10.1093/oxfordhb/9780190067397.013.45 north_east

[80] Lee LM, Heilig CM, White A. Ethical justification for conducting public health surveillance without patient consent. American Journal of Public Health, 2012, 102(1), 38-44. DOI: 10.2105/AJPH.2011.300297

[81] Li Y, Zou Y, Xu H. From data to decisions: the integration of AI in epidemiological research. Frontiers in Computing and Intelligent Systems, 2024, 9(1), 23-29.

[82] De Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. Unique in the crowd: The privacy bounds of human mobility. Scientific Reports, 2013, 3(1), 1376. DOI: 10.1038/srep01376

[83] Morris G, Snider D, Katz M. Integrating public health information and surveillance systems. Journal of Public Health Management and Practice, 1996, 2(4), 24-27. DOI: 10.1097/00124784-199623000-00007

[84] Pratt B, Parker M, Bull S. Equitable design and use of digital surveillance technologies during COVID-19: Norms and concerns. Journal of Empirical Research on Human Research Ethics, 2022, 17(5), 573-586. DOI: 10.1177/155626462211181

[85] Kinney D. Aggregating concepts of fairness and accuracy in prediction algorithms. Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, 2025, 464-472. DOI: 10.1145/3715275.3732031

[86] Pushkarna M, Zaldivar A, Kjartansson O. Data cards: Purposeful and transparent dataset documentation for responsible AI. Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, 2022, 1776-1826. DOI: 10.1145/3531146.353323

[87] Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, Chan AW, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nature Medicine, 2020, 26(9), 1364-1374. DOI: 10.1038/s41591-020-1034-x

[88] Ganesh N, Reddy MT, Selvi GS. Privacy preserving machine learning: Secure training and deployment of models on sensitive data. 2025 International Conference on Intelligent Computing and Control Systems (ICICCS), 2025, 1199-1206. DOI: 10.1109/ICICCS65191.2025.10984800

[89] Oloyede J. Ethical reflections on AI for cybersecurity: building trust. Available at SSRN 4733563, 2024. DOI: 10.2139/ssrn.4733563
[90] Kim Y, Dán G, Zhu Q. Human-in-the-loop cyber intrusion detection using active learning. IEEE Transactions on Information Forensics and Security, 2024, 8658-8672. DOI: 10.1109/TIFS.2024.3434647
[91] Dolin P, Li W, Dasarathy G, Berisha V. Statistically valid post-deployment monitoring should be standard for AI-based digital health. arXiv preprint arXiv: 2506.05701, 2025. DOI: 10.48550/arXiv.2506.05701
[92] Buchbinder M, Juengst E, Rennie S, Blue C, Rosen DL. Advancing a data justice framework for public health surveillance. AJOB Empirical Bioethics, 2022, 13(3), 205-213. DOI: 10.1080/23294515.2022.2063997
[93] Pillai AS. Artificial intelligence in healthcare systems of low-and middle-income countries: requirements, gaps, challenges, and potential strategies. International Journal of Applied Health Care Analytics, 2023, 8(3), 19-33.
[94] Hutchinson J, Stilinovic M, Gray JE. Data sovereignty: The next frontier for internet policy? Policy & Internet, 2024, 16(1), 6-11. DOI: 10.1002/poi3.386
[95] Chassang G, Béranger J, Rial-Sebbag E. The emergence of AI in public health is calling for operational ethics to foster responsible uses. International Journal of Environmental Research and Public Health, 2025, 22(4), 568. DOI: 10.3390/ijerph22040568
[96] Tripathi A. Artificial intelligence in public health surveillance: A cross-disciplinary assessment of predictive analytics and ethical concerns. Edulogic International Journal for Multi Disciplinary Research, 2025, 1(1), 13-26. DOI: 10.63665/eijmr.v01i01.2
[97] Nandan Prasad A. Data quality and preprocessing. Introduction to Data Governance for Machine Learning Systems, 2024, 109-223. DOI: 10.1007/979-8-8688-1023-7_3
[98] Suter GW, Cormier SM. The problem of biased data and potential solutions for health and environmental assessments. Human and Ecological Risk Assessment, 2015, 21(7), 1736-1752. DOI: 10.1080/10807039.2014.974499
[99] Opara-Martins J, Sahandi R, Tian F. Critical analysis of vendor lock-in and its impact on cloud computing migration: a business perspective. Journal of Cloud Computing, 2016, 5(1), 4. DOI: 10.1186/s13677-016-0054-z
[100] Kumar S, Datta S, Singh V, Datta D, Singh SK, Sharma R. Applications, challenges, and future directions of human-in-the-loop learning. 2024, 12, 75735-75760. DOI: 10.1109/ACCESS.2024.3401547
[101] Hussain S, Xi X, Ullah I, Inam SA, Naz F, Shaheed K, et al. A discriminative level set method with deep supervision for breast tumor segmentation. Computers in Biology and Medicine, 2022, 149, 105995. DOI: 10.1016/j.compbiomed.2022.105995
[102] Khatri RB, Endalamaw A, Erku D, Wolka E, Nigatu F, Zewdie A, et al. Contribution of health system governance in delivering primary health care services for universal health coverage: A scoping review. PLOS One, 2025, 20(2), e0318244. DOI: 10.1371/journal.pone.0318244
[103] Alotaibi DA, Zhou J, Dong Y, Wei J, Ge XJ, Chen F. Quantile multi-attribute disparity (QMAD): An adaptable fairness metric framework for dynamic environments. Electronics, 2025, 14(8), 1627. DOI: 10.3390/electronics14081627
[104] De Cristofaro E, Durussel A, Aad I. Reclaiming privacy for smartphone applications. 2011 IEEE International Conference on Pervasive Computing and Communications, 2011, 84-92. DOI: 10.1109/PERCOM.2011.5767598
[105] McNabb SJ. Comprehensive effective and efficient global public health surveillance. BMC Public Health, 2010, 10(Suppl 1), S3. DOI: 10.1186/1471-2458-10-S1-S3
[106] Sedeeq FS, Arman P. Human-centric AI governance: An adaptive public international law framework for ethical and inclusive AI regulation in public health. Journal of Law, Medicine & Ethics, 2025, 1-2. DOI: 10.1017/jme.2025.10175