Journal of Computing Theories and Applications ISSN:3024-9104

Review Article

Bridging the AI Deployment Gap in Predictive Policing: An Analytical Review of Hybrid Multimodal Methods in Developing-Country Contexts

Mukhtar O. Omiyeniyi [1,*], Erastus O. Ogunti [1], and and Osekhonmen V. Abhulimen [2,*]

[1] Department of Computer Engineering, Federal University of Technology, School of Electrical Systems Engineering, Akure, Ondo State, Nigeria; e-mail : omiyeniyiolarewaju@gmail.com; eoogunti@futa.edu.ng
[2] Department of Information and Communication Engineering, Faculty of Engineering Technology, Elizade University, Ilara-Mokin, Ondo State, Nigeria; e-mail : bulioses@gmail.com
* Corresponding Author : Mukhtar O. Omiyeniyi ; Osekhonmen V. Abhulimen

Abstract: The rapid growth of crime in Nigeria and other developing countries has intensified the need for proactive, data-driven policing strategies. Predictive policing leverages historical crime data and artificial intelligence (AI) techniques to support crime forecasting, patrol planning, and resource allocation. However, most state-of-the-art predictive policing frameworks have been developed and evaluated in data-rich environments, whereas Nigeria and many developing countries continue to face fragmented crime records, partial digitization, narrative-heavy police data, limited digital infrastructure, and evolving AI governance. This paper presents an analytical review of AI-based predictive policing, focusing on Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), spatial-temporal modeling, graph-based prediction, Explainable Artificial Intelligence (XAI), and ethical AI deployment. The review introduces a taxonomy that organizes the literature according to problem class, model architecture, data modality, interpretability, and deployment requirements. The findings indicate that conventional ML remains effective for structured crime prediction, whereas DL and graph-based approaches provide superior capabilities for temporal forecasting and spatial crime-diffusion analysis. Despite its potential to exploit intelligence embedded in police narratives, witness statements, case files, social media, and other unstructured sources, NLP remains underutilized in operational predictive policing systems. Similarly, XAI is insufficiently integrated into current frameworks, limiting transparency, auditability, bias detection, and public trust. The review further identifies an AI Deployment Gap between developed and developing-country contexts, demonstrating that successful deployment depends not only on predictive accuracy but also on data readiness, local validation, explainability, governance, and institutional capacity. Overall, the review highlights the need for context-aware hybrid ML–NLP–spatial-temporal–XAI frameworks that are computationally efficient, ethically governed, and better suited to the operational realities of predictive policing in Nigeria and other developing countries.

Keywords: Artificial Intelligence; Crime Prediction; Explainable Artificial Intelligence; Machine Learning; Natural Language Processing; Predictive Policing; Spatial-Temporal Modeling; Smart Policing.

Received: April, 7th 2026
Revised: July, 13th 2026
Accepted: July, 18th 2026
Published: August, 6th 2026

1. Introduction

The advancement of predictive policing is driven not only by rapid developments in machine learning (ML) and computational intelligence but also by the socio-technical conditions under which these systems are deployed [1]. In many developing countries, particularly Nigeria and other resource-constrained environments, the primary challenge extends beyond the limited adoption of advanced AI models. Instead, it lies in what this review conceptualizes as the AI Deployment Gap—a set of interrelated constraints involving data readiness, governance maturity, and institutional preparedness that collectively influence the feasibility, effectiveness, and responsible deployment of AI-based predictive policing systems [2], [3], [4]. This perspective shifts the discussion from algorithmic performance alone toward the broader socio-technical ecosystem required for successful AI implementation.

A critical dimension of this deployment gap is the adaptability dilemma, in which the rapid advancement of AI technologies outpaces the development of localized legal, ethical, and regulatory frameworks [5]. Unlike many developed countries, where AI governance and algorithmic accountability are becoming increasingly institutionalized [6], [7], AI regulation in many developing countries remains fragmented and is often supported only by broader digital governance policies rather than dedicated AI regulatory frameworks [8]. Consequently, operational requirements such as auditability, accountability, transparency, and the "notice-and-explanation" principles expected in high-stakes AI applications remain difficult to achieve, particularly within law enforcement environments [6], [9].

A second dimension concerns data and infrastructural readiness. Crime data in many developing policing environments remain fragmented, inconsistently recorded, partially digitized, and frequently stored in narrative formats that are difficult to integrate into predictive AI pipelines [8], [9], [10]. These challenges are further compounded by limited digital infrastructure, unreliable power supply, unstable connectivity, and underdeveloped information management systems [2], [11]. Moreover, fragmented identification systems and disconnected crime repositories reduce data cohesion and interoperability [3], leading to an important technical consequence: distribution shift, whereby models developed using standardized, high-quality datasets often fail to generalize effectively when deployed in noisy, incomplete, or resource-constrained operational environments [12]. These observations suggest that deployment readiness depends not only on model capability but also on the quality and maturity of the underlying data ecosystem.

A third dimension relates to institutional readiness and socio-technical legitimacy [13], [14]. Predictive policing systems generally assume the existence of complementary institutional capabilities, including digital forensic infrastructure [2], analytical expertise [6], interoperable databases [3], and effective human oversight mechanisms [15]. However, these enabling conditions remain unevenly developed across many developing policing systems [8]. Beyond limiting deployment feasibility, such deficiencies also undermine public trust and institutional legitimacy [9]. As reported in [14], the effectiveness of smart policing is closely associated with service quality, transparency, and public confidence. In settings where trust in law enforcement is already fragile, opaque "black-box" AI systems may reinforce historical biases and weaken institutional legitimacy, particularly when predictive models are trained on socially biased crime datasets [3], [12].

Importantly, these socio-technical constraints also influence methodological choices in predictive policing research. In developed countries, recent studies have increasingly explored person-based predictive models [16]. However, growing concerns regarding algorithmic bias [17], stigmatization, privacy, and civil liberties have raised significant ethical challenges [7], [18]. By contrast, predictive policing research in many developing countries has predominantly emphasized place-based and hybrid spatial-temporal approaches [19], largely because detailed person-level geospatial information is often unavailable or subject to security restrictions. Consequently, hotspot-oriented prediction generally presents lower ethical, legal, and operational risks while remaining more compatible with existing data availability [2], [20]. These differences indicate that model selection is shaped not only by predictive performance but also by policy environments, data accessibility, governance requirements, and deployment constraints [21], [22].

Another important implication of the AI Deployment Gap is the loss of contextual intelligence in predictive systems that rely exclusively on structured data [23], [24]. Existing predictive policing research in Nigeria predominantly utilizes structured administrative records, including crime dates, geographic coordinates, and offense categories, while largely overlooking the rich semantic information contained in police narratives, witness statements, intelligence reports, and other unstructured textual sources [3], [4], [8]. This limitation has stimulated growing interest in hybrid ML–NLP architectures capable of integrating structured and unstructured crime intelligence into unified predictive frameworks [23], [25], [26], [27]. Within this context, Explainable Artificial Intelligence (XAI) should not be regarded merely as an auxiliary component but rather as a fundamental requirement for trustworthy, transparent, and accountable AI deployment, particularly in high-stakes and resource-constrained policing environments [9], [28], [29].

1.1. Positioning Relative to Existing Surveys and Contributions of This Review

Although predictive policing has been extensively reviewed through systematic reviews and survey studies, several important gaps remain regarding deployment challenges, methodological integration, and the unique characteristics of resource-constrained environments. Existing surveys have primarily focused on algorithmic performance, smart policing architectures, or ethical considerations within data-rich, developed-country settings, while providing limited discussion of predictive policing deployment in developing countries [11], [30], [31]. Consequently, the practical challenges associated with deploying AI under constrained socio-technical conditions remain insufficiently synthesized.

The first limitation relates to geographical and contextual bias. Most existing reviews are largely based on case studies from high-income countries, implicitly assuming the availability of mature digital infrastructure, interoperable institutional systems, and well-established AI governance frameworks. These assumptions are often difficult to transfer to policing environments in developing countries, where AI adoption frequently progresses faster than the development of supporting legal, ethical, and governance structures. As discussed in Section 1, this adaptability dilemma creates deployment challenges that differ substantially from those typically addressed in existing predictive policing surveys [5], [32].

A second limitation concerns methodological fragmentation. Previous technical surveys have documented significant advances in ML, DL, and Natural Language Processing (NLP) for predictive policing [22], [33]. However, these approaches are commonly reviewed as independent methodological streams rather than as complementary components of an integrated predictive ecosystem. This distinction is particularly important in developing-country contexts, where the exclusion of narrative-rich intelligence from predictive workflows may lead to substantial contextual information loss [8], [34]. Accordingly, this review considers NLP not merely as an auxiliary analytical tool, but as an essential component of hybrid predictive policing architectures that integrate structured and unstructured crime intelligence.

A third limitation concerns deployment feasibility under fragmented data environments. For example, [22] identified 64 ML techniques across 120 predictive policing studies, while [11] reported the continued dominance of conventional ML with only limited integration of DL and NLP approaches. Although these surveys provide valuable quantitative overviews of methodological trends, they primarily evaluate algorithmic prevalence rather than examining how data quality, infrastructural limitations, governance maturity, and institutional capacity influence model suitability, transferability, and operational deployment. This review extends previous work by analyzing predictive policing models through the broader perspective of AI Deployment Gap, incorporating issues such as distribution shifts, fragmented data ecosystems, and institutional readiness into the evaluation process.

Another important limitation involves the operational role of XAI. Existing surveys generally discuss explainability from a normative perspective, presenting transparency as a desirable characteristic of AI systems while providing limited analysis of its practical contribution to auditability, bias detection, operational accountability, and public trust [35]. This review expands that perspective by positioning explainability as a fundamental deployment requirement rather than an optional model attribute. In addition to interpretable predictive models, the review also examines post-hoc explanation techniques, including LIME and SHAP, within the broader context of socio-technical legitimacy and trustworthy AI deployment in high-stakes policing environments [9], [20].

Building upon these identified limitations, this review makes four primary contributions. First, it provides an integrated synthesis of ML, spatial-temporal modeling, NLP, hybrid predictive architectures, and XAI, rather than treating these research areas independently [19], [36]. Second, it introduces the AI Deployment Gap as a unifying analytical lens that connects predictive modeling with the realities of data readiness, governance, and institutional capacity in developing-country contexts [3], [5], [13]. Third, it evaluates state-of-the-art predictive models not only in terms of predictive performance but also with respect to distribution shifts, data limitations, deployment feasibility, and operational suitability [37], [38]. Finally, it synthesizes emerging research directions at the intersection of explainability, multimodal intelligence, and socio-technical AI deployment, thereby extending existing predictive policing surveys beyond algorithm-centric comparisons [6], [11].

Unlike previous surveys that primarily compare predictive policing algorithms or summarize applications across different jurisdictions, this review adopts the AI Deployment Gap as its central analytical framework. Rather than evaluating AI techniques solely based on predictive accuracy, the review examines how deployment-related factors—including data readiness, multimodal information availability, computational resources, explainability, governance, and institutional capacity—influence the selection, evaluation, and operational suitability of predictive policing models within developing-country contexts. Accordingly, the proposed taxonomy, research questions, evidence synthesis, and future research framework are all structured around understanding and addressing this deployment gap.

1.2. Motivation for Machine Learning, NLP, XAI, and Spatial-Temporal Models

The analytical focus of this review on ML, NLP, XAI, and spatial-temporal models is motivated not only by their individual predictive capabilities but also by their complementary roles in addressing the AI Deployment Gap. Collectively, these technologies respond to key deployment challenges, including non-linear crime dynamics, information loss in narrative-heavy police records, limited model interpretability, fragmented spatial intelligence, and the operational constraints commonly encountered in developing-country policing environments. Rather than representing isolated technological advances, these approaches collectively reflect the methodological evolution required to support context-aware and operationally feasible predictive policing systems.

Early predictive policing approaches relied primarily on statistical forecasting techniques such as Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Fractionally Integrated Moving Average (ARFIMA). Although these methods remain valuable for modeling temporal trends and seasonality, they are constrained by assumptions of linearity and stationarity and therefore struggle to capture the complex, non-linear, and context-dependent dynamics of crime [36], [39]. The transition toward ML was consequently driven by the need for models capable of learning high-dimensional relationships, accommodating heterogeneous predictors, and maintaining predictive performance under distribution shifts. This motivation is further supported by evidence from Nigerian crime forecasting studies, where Long Short-Term Memory (LSTM)-based models have demonstrated superior performance over conventional statistical approaches in modeling irregular urban crime patterns [39].

Beyond structured crime forecasting, the integration of NLP is motivated by the need to reduce information loss in predictive systems that rely predominantly on structured variables such as crime counts, timestamps, and geographic coordinates. In many policing environments, particularly across Africa, a substantial amount of operational intelligence remains embedded within unstructured sources, including incident narratives, police reports, witness statements, and case files. Ignoring these sources limits the contextual information available for predictive modeling and decision support. Recent advances in Large Language Models (LLMs) and NLP techniques demonstrate considerable potential for extracting actionable intelligence from these narrative-rich datasets. At the same time, existing studies highlight localization, language diversity, and transferability challenges, reinforcing the need for context-aware NLP integration rather than the direct adoption of generic language models [11], [12].

The inclusion of XAI is likewise motivated by considerations that extend beyond transparency alone. In predictive policing, explainability increasingly functions as a practical requirement for auditability, bias detection, accountability, and trustworthy deployment, particularly in high-stakes operational environments where opaque model predictions may undermine institutional legitimacy [40], [41]. This role becomes even more critical in developing-country contexts, where formal AI governance frameworks are still evolving and technical explainability can partially compensate for limited regulatory oversight [5]. Furthermore, recent studies have demonstrated that explanation techniques such as SHAP can improve the interpretability of crime-risk prediction by identifying measurable environmental factors associated with crime occurrence [9]. These findings reinforce the view that XAI should be considered an integral component of deployable predictive policing systems rather than merely an auxiliary feature.

The motivation for spatial-temporal and hybrid predictive models arises from the inherently dynamic, spatially correlated, and context-dependent nature of crime. Criminal activities evolve across both space and time and are influenced by complex interactions among environmental, demographic, and urban factors. Consequently, predictive policing research has gradually shifted beyond conventional hotspot analysis toward graph-based learning, deep spatial-temporal architectures, and hybrid predictive frameworks capable of modeling these complex dependencies [42], [43]. Importantly, this evolution has been driven not only by improvements in predictive accuracy but also by the need for greater robustness under noisy, fragmented, and resource-constrained data environments. Recent hybrid frameworks that combine clustering techniques, temporal learning, and graph-based reasoning further demonstrate the potential to improve model generalization across heterogeneous crime settings [19].

Taken together, these methodological pillars define the analytical scope of this review. ML addresses complex and non-linear crime dynamics, NLP enriches predictive systems by incorporating narrative-based intelligence, XAI enhances transparency and deployment accountability, while spatial-temporal and hybrid models improve contextual awareness and robustness. Their integration reflects not only the technological evolution of predictive policing but also the broader socio-technical requirements necessary for bridging the AI Deployment Gap in developing-country policing environments.

1.3. Research Questions

To examine the evolution of predictive policing from conventional statistical forecasting toward multimodal, explainable, and deployment-oriented AI systems, this review is guided by a series of interrelated research questions. These questions investigate the adoption of AI techniques, the integration of NLP and spatial-temporal modeling, the role of explainability and governance, and the practical challenges associated with deploying predictive policing systems in Nigeria and other developing-country contexts. Collectively, they are intended to explore not only which predictive approaches have been developed, but also how data readiness, institutional capacity, governance, explainability, and deployment feasibility influence the selection, evaluation, and operational implementation of predictive policing models. The review is structured around the following research questions:

- RQ1. What machine learning, deep learning, and spatial-temporal approaches have been applied in predictive policing, and what methodological trade-offs characterize these approaches?
- RQ2. How do hybrid and state-of-the-art predictive models address challenges associated with non-linear crime dynamics, spatial dependencies, and distribution shifts?
- RQ3. To what extent has Natural Language Processing been integrated into predictive policing systems, and what limitations constrain its adoption in narrative-heavy intelligence environments?
- RQ4. How do data readiness limitations, fragmented infrastructures, and institutional constraints influence the feasibility of predictive policing models in developing-country contexts?
- RQ5. How is Explainable Artificial Intelligence being utilized to support auditability, bias detection, and trustworthy deployment in predictive policing?
- RQ6. What ethical, governance, and deployment challenges influence the adoption of AI-driven predictive policing in Nigeria and other African contexts?
- RQ7. What research gaps and future directions emerge for developing context-aware, hybrid, and explainable predictive policing systems?

2. Foundations and Methodological Landscape of Predictive Policing

The methodological landscape of predictive policing has evolved from traditional statistical crime analysis to advanced ML, DL, NLP, spatial-temporal learning, and XAI. While previous reviews have primarily compared these approaches based on predictive accuracy and algorithmic performance, this review adopts the AI Deployment Gap as its primary analytical lens. Accordingly, the literature is synthesized not only according to methodological advances but also according to each method's suitability for deployment under varying conditions of data readiness, institutional capacity, governance, explainability, and computational feasibility. This perspective is particularly relevant to developing-country policing environments, where fragmented datasets, narrative-heavy intelligence, limited digital infrastructure, and evolving AI governance significantly influence the selection, implementation, and operational effectiveness of predictive policing systems. By shifting the focus from algorithmic superiority to deployment readiness, this review provides a broader understanding of how socio-technical conditions shape the practical adoption of AI in predictive policing.

2.1. Conceptual Foundations of Predictive Policing

Predictive policing refers to the application of data-driven analytical techniques to anticipate potential crime events, identify high-risk locations or individuals, and support proactive decision-making in law enforcement operations [11]. Although predictive policing is often evaluated in terms of forecasting accuracy, its successful implementation depends equally on deployment conditions, including data readiness, institutional capacity, governance, and operational feasibility. Consequently, this review examines the conceptual foundations of predictive policing through the AI Deployment Gap, emphasizing how socio-technical conditions influence the design, selection, and deployment of AI-based policing systems in developing-country contexts [13], [44].

The intellectual foundations of predictive policing are rooted in Intelligence-Led Policing (ILP) and Problem-Oriented Policing (POP). ILP emphasizes the systematic collection and analysis of crime data to support strategic and tactical decision-making, whereas POP focuses on identifying recurring crime problems and addressing their underlying causes [23], [45]. Predictive policing extends these policing philosophies by incorporating algorithmic forecasting models capable of uncovering latent patterns and anticipating future crime risks across multiple spatial and temporal scales [29], [46]. This evolution reflects the transition from reactive crime analysis toward proactive, intelligence-driven decision support.

From a theoretical perspective, predictive policing operates through two primary paradigms: place-based prediction and person-based prediction. Place-based predictive policing focuses on forecasting crime hotspots by analyzing spatial and temporal crime patterns, environmental characteristics, and situational contexts [33]. This approach aligns closely with Routine Activity Theory and Crime Pattern Theory, which propose that crime is spatially concentrated through the interaction of motivated offenders, suitable targets, and the absence of capable guardians [9]. In contrast, person-based predictive policing seeks to identify individuals or groups at elevated risk of offending or victimization using historical arrest records, social networks, and behavioral indicators [47]. Although person-based models have received increasing research attention, they remain highly controversial because of concerns related to algorithmic bias, stigmatization, discrimination, and civil liberties [11], [18]. Consequently, the choice between place-based and person-based prediction extends beyond predictive performance and increasingly involves ethical, legal, and governance considerations.

Early predictive policing systems relied primarily on statistical and heuristic approaches, including hotspot mapping, regression analysis, and time-series forecasting [3]. These methods provided valuable insights into crime trends, seasonality, and spatial concentration but were limited in their ability to capture complex nonlinear relationships and heterogeneous data sources [41]. As crime datasets became increasingly diverse and computational capabilities advanced, ML and AI enabled substantially more sophisticated predictive models. Consequently, predictive policing has evolved from a purely statistical forecasting approach into an AI-driven decision-support system capable of integrating multiple data sources and learning complex crime patterns [19].

Importantly, predictive policing is not intended to replace human judgment but rather to augment police decision-making by providing actionable intelligence for patrol allocation, resource optimization, and crime prevention planning [8], [15]. Contemporary research increasingly emphasizes that predictive outputs should be interpreted as probabilistic risk indicators rather than deterministic predictions, reinforcing the importance of human oversight and contextual interpretation [8], [45]. This distinction is particularly critical in environments where data quality, socio-economic complexity, and institutional constraints may substantially influence model reliability and operational performance [17]. Accordingly, successful deployment depends not only on predictive capability but also on responsible human-centered decision-making.

In developing regions such as Nigeria, the practical implementation of predictive policing is shaped not only by crime characteristics but also by deployment constraints, including fragmented data infrastructures, partial digitization, limited computational resources, and evolving governance frameworks [10]. These factors influence both the selection of predictive models and their operational feasibility, reinforcing the central premise of this review that successful AI deployment depends as much on socio-technical readiness as on algorithmic capability. Consequently, place-based approaches remain more practical than person-based systems because they rely on relatively accessible geospatial evidence while avoiding many of the ethical and legal concerns associated with individual profiling [19], [20].

Furthermore, recent theoretical developments have expanded the conceptual scope of predictive policing by positioning explainability, fairness, and accountability as fundamental design principles rather than optional system characteristics. Researchers increasingly argue that without transparency and interpretability, predictive policing systems risk undermining public trust, reinforcing historical biases, and reducing institutional legitimacy, particularly in marginalized communities [40], [41], [48]. This shift demonstrates that predictive policing should be understood not merely as a computational problem but as a socio-technical system in which technical performance, governance, ethics, and community trust are inherently interconnected.

In summary, the conceptual foundations of predictive policing extend well beyond algorithmic forecasting. While criminological theories, ILP, POP, and computational intelligence provide the methodological basis for prediction, their operational value ultimately depends on data readiness, governance structures, explainability, institutional capacity, and deployment feasibility. Collectively, these interrelated factors define the AI Deployment Gap adopted throughout this review and establish the analytical foundation for evaluating existing predictive policing methods, identifying research gaps, and motivating the proposed context-aware hybrid ML–NLP–spatial-temporal–XAI framework.

2.2. Evolution of Predictive Policing Models: From Statistical Methods to AI-Driven Systems

The evolution of predictive policing reflects not only advances in AI but also the changing requirements for operational deployment. Early research primarily focused on improving forecasting accuracy through statistical and ML techniques, whereas contemporary predictive policing increasingly emphasizes deployment readiness, including data quality, computational efficiency, explainability, governance, and institutional capacity. Accordingly, the methodological evolution presented in this section is interpreted through the AI Deployment Gap, highlighting how successive technological advances have progressively addressed—or, in some cases, exposed—the deployment challenges encountered in developing-country policing environments. Figure 1 summarizes this methodological evolution from traditional statistical approaches to contemporary hybrid AI-driven predictive policing frameworks.

Initial predictive policing models were predominantly based on classical statistical and econometric methods, in which crime was modeled as a temporally structured phenomenon characterized by trends, cycles, and seasonality. Time-series techniques, such as ARIMA and ARFIMA, were widely adopted because of their mathematical rigor, interpretability, and relatively low computational requirements [36], [39]. Research by [46] demonstrated that these models were effective for short-term crime forecasting, particularly at aggregated spatial scales. Their simplicity and transparency also made them attractive for operational environments with limited computational resources.

However, subsequent studies, particularly those conducted in developing countries, revealed that classical statistical models were inadequate for capturing the nonlinear, spatially heterogeneous, and context-dependent nature of crime. Crime dynamics are influenced by complex interactions among socio-economic conditions, environmental characteristics, population mobility, and offender behavior, factors that violate the assumptions of linearity and stationarity underlying conventional time-series models [9], [49]. Nevertheless, these approaches remain operationally relevant in resource-constrained environments because they require relatively small datasets, modest computational resources, and provide transparent decision-making processes. This trade-off illustrates an important aspect of the AI Deployment Gap, where operational feasibility may outweigh incremental improvements in predictive performance.

The emergence of ML marked a major turning point in crime prediction research by enabling models to capture complex nonlinear relationships beyond the capability of traditional statistical approaches [50]. Studies conducted in developed countries demonstrated that supervised learning algorithms, including decision trees, random forests, support vector machines, and ensemble learners, consistently outperformed statistical models in crime classification and hotspot prediction tasks [51], [52], [53]. Their flexibility, robustness, and moderate computational requirements make conventional ML models particularly suitable for policing environments where structured crime databases are available but computational infrastructure remains limited—a situation commonly encountered in many developing countries.

Consequently, ML represents an important balance between predictive capability and deployment practicality.

The subsequent adoption of DL architectures further transformed predictive policing by enabling hierarchical feature learning and long-term temporal dependency modeling. Recurrent Neural Networks (RNNs), particularly LSTM networks, became widely adopted for temporal crime forecasting because of their ability to capture complex sequential crime patterns [11], [37], [42]. Comparative studies conducted in developing-country settings consistently reported that LSTM-based models achieved higher forecasting accuracy than ARIMA and other statistical baselines, especially in large metropolitan environments [40], [44]. These findings substantially influenced subsequent crime prediction research worldwide. However, the improved predictive capability of DL comes at the cost of increased dependence on large, high-quality datasets and greater computational resources, limiting routine deployment in policing environments characterized by fragmented records, incomplete digitization, and constrained infrastructure.

As predictive policing research matured, researchers increasingly recognized that crime is inherently a spatial-temporal phenomenon, leading to the development of models capable of simultaneously capturing geographic interactions and temporal dynamics [54]. Early spatial analyses primarily relied on techniques such as Kernel Density Estimation (KDE) for hotspot identification [49], [55]. More recent studies have increasingly adopted clustering algorithms, particularly Density-Based Spatial Clustering of Applications with Noise (DBSCAN), because of their ability to identify irregularly shaped crime clusters while effectively handling noisy observations [15]. Furthermore, hybrid frameworks combining spatial clustering with DL models have demonstrated notable improvements in predictive performance [33], [44]. Nevertheless, their successful deployment depends heavily on the availability of accurate geospatial information and integrated spatial databases, conditions that remain inconsistent across many policing agencies in developing countries.

More recently, graph-based learning paradigms, particularly Graph Neural Networks (GNNs) and Spatio-Temporal Graph Neural Networks (ST-GNNs), have emerged as state-of-the-art approaches in predictive policing research. Unlike conventional spatial models that treat locations independently, graph-based architectures model crime locations as interconnected nodes, enabling the learning of spatial diffusion patterns and relational dependencies [44], [54]. Studies applying ST-GNNs have reported substantial improvements in crime forecasting accuracy and hotspot detection, representing a significant advancement over earlier predictive modeling paradigms [40]. Despite these promising results, the operational deployment of graph-based models requires continuously updated spatial networks, reliable relational data, and substantial computational infrastructure. These requirements remain difficult to satisfy in environments where crime information is fragmented, partially digitized, or inconsistently maintained, further illustrating the practical challenges associated with AI deployment.

The methodological trajectory observed in developing countries generally follows the same global evolution. However, the slower adoption of advanced predictive policing models should not be interpreted as evidence of technological lag. Rather, it primarily reflects differences in deployment readiness. Fragmented crime databases, partial digitization, inconsistent reporting standards, limited computational infrastructure, and evolving governance frameworks substantially constrain the operational feasibility of sophisticated AI models despite their demonstrated predictive capability. Consequently, researchers have increasingly explored hybrid architectures that seek to balance predictive performance with deployment practicality under local operational conditions [19], [32], [39]. This observation reinforces the central premise of this review that deployment success depends as much on socio-technical readiness as on algorithmic advancement.

Importantly, the distinction between developed and developing countries lies less in algorithmic sophistication than in deployment context. Although researchers across different regions increasingly investigate comparable AI techniques, differences in data readiness, institutional capacity, governance, explainability requirements, and computational infrastructure largely determine whether these models can be successfully deployed in operational policing environments. Thus, the AI Deployment Gap should be understood not as a technological disparity but as a socio-technical deployment challenge that shapes the practical adoption of predictive policing systems.

In summary, the evolution of predictive policing represents more than a sequence of algorithmic improvements; it reflects a progressive transition toward AI systems that must satisfy increasingly complex deployment requirements. From statistical forecasting to hybrid multimodal AI frameworks, each methodological transition addresses specific limitations in predictive capability while simultaneously introducing new challenges related to data readiness, computational demands, explainability, governance, and institutional capacity. Viewed through the AI Deployment Gap, these socio-technical considerations provide a more comprehensive basis for evaluating predictive policing models than predictive accuracy alone and establish the analytical foundation for the subsequent discussion of ML, NLP, spatial-temporal learning, and XAI. Figure 1 illustrates the evolution of predictive policing approaches, highlighting the transition from traditional reactive policing and statistical forecasting toward hybrid AI-driven frameworks that integrate ML, NLP, spatial-temporal learning, and XAI to support more effective and responsible predictive policing systems.

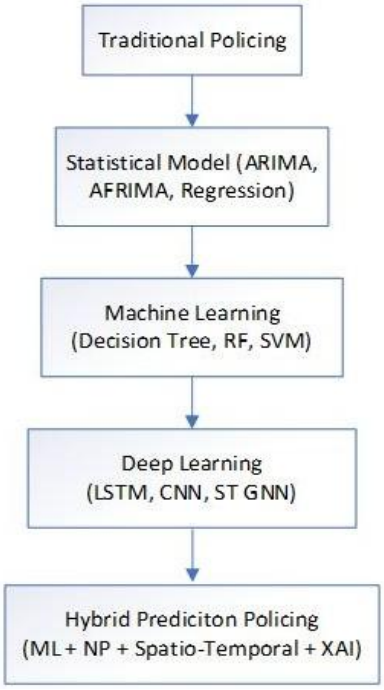

Figure 1. Evolution of Predictive Policing

2.3. Machine Learning Techniques for Crime Prediction

Machine Learning (ML) has become one of the dominant methodologies in predictive policing because it provides an effective balance between predictive capability and operational feasibility across diverse policing environments [22], [44]. Beyond improving crime prediction accuracy, ML enables the analysis of nonlinear, heterogeneous, and high-dimensional crime data [56] while remaining sufficiently flexible for deployment under varying levels of data availability and computational resources. Within the AI Deployment Gap framework adopted in this review, ML techniques are evaluated not only according to predictive performance but also based on their suitability for real-world deployment in developing-country policing contexts. Consequently, the discussion extends beyond algorithmic accuracy to consider deployment readiness, interpretability, computational efficiency, and operational sustainability.

Tree-based and ensemble learning algorithms remain among the most widely adopted ML approaches for crime prediction. Decision Trees (DTs) provide interpretable, rule-based decision structures that are well suited for operational policing because their predictions can be readily understood and justified. However, they may become unstable when trained on noisy or highly complex datasets [33]. Ensemble approaches, such as Random Forest (RF) and XGBoost, improve robustness by combining multiple learners and reducing overfitting [22], [51], [57]. Study [57] identified RF as the most frequently applied ML algorithm, appearing in 25 of 120 reviewed studies, while [44] reported that XGBoost achieved prediction accuracies of 94% for Chicago and 88% for Los Angeles. Evidence from Nigeria also demonstrates the strong predictive performance of DT and RF models [32]. Nevertheless, these improvements introduce an important trade-off: although ensemble methods generally achieve higher predictive accuracy, they reduce model transparency, making legal auditability, bias detection, and operational explanation more challenging [3]. Consequently, tree-based algorithms remain attractive for operational policing because they provide a practical balance among predictive performance, computational efficiency, interpretability, and deployment feasibility when structured crime databases are available.

Kernel-based and instance-based learning methods, particularly Support Vector Machines (SVMs) and k-Nearest Neighbours (k-NN), continue to be effective classification techniques, especially when handling sparse or high-dimensional datasets [51]. However, the predictive performance of SVM is highly dependent on kernel selection, parameter optimization, and class balance [52]. Evidence from [32] further indicates that SVM performance deteriorates under noisy or highly imbalanced local datasets, as demonstrated in the Maiduguri case study, where it underperformed compared with tree-based algorithms. These findings suggest that model selection should be driven by local data characteristics and operational requirements rather than by algorithmic popularity alone. Viewed through the AI Deployment Gap, deployment success depends on selecting models that align with the quality and characteristics of the available data rather than simply adopting the most advanced algorithm.

Deep learning approaches, particularly LSTM networks [19], [39], Convolutional Neural Networks (CNNs) [44], and hybrid CNN–LSTM architectures [37], have substantially expanded the capability of predictive policing by modeling complex temporal and spatial-temporal crime dynamics [58]. In Oyo State, [39] demonstrated that LSTM outperformed ARIMA-based forecasting, highlighting its effectiveness for modeling irregular urban crime patterns. However, these performance gains introduce a complexity–stability trade-off. While CNN–LSTM architectures capture intricate temporal features, [37] reported that they remain sensitive to outliers and exhibited elevated Mean Squared Error when trained on noisy datasets. Furthermore, DL models require substantially larger training datasets, greater computational capacity, and continuous model maintenance, reducing their operational suitability in many resource-constrained policing environments. Consequently, the practical adoption of DL depends not only on predictive capability but also on the availability of sustainable computational and data infrastructures.

Generative AI and LLMs represent a more recent frontier in predictive policing research [24], [59], [60], particularly for extracting actionable intelligence from unstructured police narratives. Models including GPT-3, GPT-4, BART, Baichuan, ChatGLM, and Qwen have demonstrated promising performance in crime classification, entity extraction, and police report analysis [23], [25]. However, [12] reported that a fine-tuned GPT-3 model achieved 97% accuracy on a San Francisco dataset but experienced a substantial performance decline when evaluated in another urban environment, illustrating the effects of distribution shift. This finding is particularly relevant for Nigeria and other developing countries, where police records are predominantly narrative-based, locally contextualized, and inconsistently digitized. Consequently, the successful deployment of LLMs depends less on the underlying model architecture than on localized training data, domain-specific adaptation, governance mechanisms, and continuous validation. Rather than directly adopting globally pretrained models, developing-country policing agencies are therefore more likely to benefit from domain-specific fine-tuning and locally adapted language resources.

Hybrid frameworks have emerged as an effective strategy for overcoming the limitations of individual predictive models by integrating clustering, classification, spatial analysis, and temporal forecasting. For example, [36] combined K-Means clustering with CNN-based analysis, while [19] integrated DBSCAN with an enhanced LSTM architecture for urban crime prediction in Nigeria. Similarly, forensic-oriented research by [61] demonstrated how clustering and probabilistic mixture models can be used to associate physical forensic evidence with criminal history databases. From the perspective of the AI Deployment Gap, hybrid architectures provide a practical deployment strategy because they exploit the complementary strengths of multiple algorithms while mitigating weaknesses associated with fragmented datasets, heterogeneous intelligence sources, and operational resource constraints. This flexibility makes hybrid systems particularly attractive for predictive policing in developing-country contexts, where no single model can adequately address all deployment challenges.

These ML has substantially advanced predictive policing by enabling the modeling of complex crime dynamics that extend beyond the capabilities of traditional statistical approaches. However, no single algorithm consistently satisfies the diverse operational requirements of predictive policing. Tree-based methods provide interpretability and computational efficiency; ensemble models improve predictive accuracy at the expense of transparency; SVM performs well under structured conditions but remains sensitive to data quality; DL effectively captures complex temporal dependencies but requires considerable computational resources; and LLMs unlock valuable intelligence from unstructured police records but demand localized adaptation, governance, and continuous validation. Viewed through the AI Deployment Gap, these trade-offs demonstrate that successful predictive policing depends not only on algorithmic performance but also on deployment readiness, data quality, explainability, institutional capacity, and operational feasibility. Accordingly, selecting an appropriate ML approach should be regarded as a context-dependent deployment decision rather than simply an optimization problem aimed at maximizing predictive accuracy.

2.4. Spatial-Temporal and Hybrid Crime Prediction Models

Crime is inherently a spatial-temporal phenomenon, emerging within specific geographic contexts while evolving through identifiable temporal patterns, spatial diffusion, and inter-area dependencies [62], [63]. Consequently, effective predictive policing requires models capable of jointly learning spatial and temporal relationships rather than treating them independently. Within the AI Deployment Gap framework adopted in this review, spatial-temporal approaches are evaluated not only according to forecasting accuracy but also in terms of computational efficiency, data requirements, scalability, and deployment feasibility in resource-constrained policing environments.

Traditional spatial analysis methods, including Geographic Information Systems (GIS), Kernel Density Estimation (KDE) [26], [55], [64], and the Standard Deviational Ellipse, remain valuable for hotspot visualization and supporting place-based policing strategies [33]. However, these techniques are primarily retrospective, rely heavily on mapping assumptions, and often prioritize visual representation over statistical inference [62]. DBSCAN addresses several of these limitations by identifying irregularly shaped crime hotspots while effectively handling noisy spatial data without requiring a predefined number of clusters [19]. In Nigeria, GIS-based hotspot mapping has been successfully applied to construct spatial crime inventories across multiple states, revealing strong interdependencies among neighboring administrative regions [63]. Although GIS- and KDE-based approaches remain attractive because of their simplicity, interpretability, and low computational requirements, their retrospective nature limits their ability to support proactive operational decision-making.

Deep learning has substantially expanded spatial-temporal crime prediction by enabling the joint modeling of temporal dynamics and spatial interactions [65]. RNNs and LSTMs effectively capture long-term temporal dependencies [56], whereas CNN–LSTM architectures integrate spatial feature extraction with temporal forecasting to improve predictive performance [37]. Nevertheless, these models introduce an important complexity–stability trade-off. CNN–LSTM architectures are capable of learning rich spatial-temporal representations but remain sensitive to noisy observations and outliers, whereas Dynamic Mode Decomposition (DMD) provides a more stable alternative for identifying low-rank spatial-temporal structures [44]. Consequently, selecting an appropriate spatial-temporal model requires balancing predictive capability with robustness under real-world operational conditions.

Graph-based learning represents a further methodological advancement by modeling crime locations as interconnected nodes rather than independent spatial entities. Architectures such as GNNs, ST-GNNs, Spatio-Temporal Graph Convolutional Networks (STGCN), Spatio-Temporal Sparse Hypergraph Networks (ST-SHN), and Spatio-Temporal Hypergraph Structure Learning (ST-HSL) capture crime diffusion, inter-location dependencies, multiplex spatial interactions, and label-scarcity challenges more effectively than conventional grid-based models [11], [40], [54], [58]. These capabilities have established graph-based learning as one of the current state-of-the-art approaches in predictive policing. However, their practical deployment depends on continuously updated spatial networks, high-quality relational data, and sufficient computational infrastructure, requirements that remain difficult to satisfy in many developing-country policing environments.

Recent state-of-the-art models have increasingly focused on reducing these computational constraints while preserving predictive performance. Transformer-based and sparse-attention architectures, including ACSAformer, LCRNet, and LGSTime, improve long-sequence modeling by incorporating sparse attention mechanisms, adaptive graph convolution, and computationally efficient learning strategies [66], [67], [68]. Similarly, MiST and DeepCrime integrate spatial, temporal, and categorical information into a unified latent representation to support citywide abnormal-event forecasting [65], [69]. These developments indicate that contemporary spatial-temporal prediction is evolving toward efficient multimodal fusion rather than merely increasing model complexity. From the perspective of the AI Deployment Gap, such efficiency-oriented architectures represent an important step toward practical deployment because they reduce computational demands while maintaining strong predictive capability. Nevertheless, successful operational implementation continues to depend on data quality, localized validation, institutional readiness, and governance rather than algorithmic efficiency alone.

Hybrid frameworks attempt to combine the complementary strengths of multiple modeling paradigms to improve both predictive robustness and deployment practicality. The Nigerian-focused CriPaaP framework integrates DBSCAN for hotspot clustering, an enhanced LSTM for temporal pattern learning, and ST-GNN for modeling global spatial relationships, resulting in improved hotspot localization and reduced forecasting error [19], [44]. Meanwhile, interpretable spatial-temporal approaches, including AIST and SHAP-based spatial explanation techniques, enhance auditability by identifying the points of interest, historical crime patterns, and environmental variables that contribute most strongly to high-risk predictions [6], [9]. These findings suggest that hybrid architectures not only improve predictive performance but also provide greater flexibility for addressing heterogeneous data sources and operational constraints commonly encountered in developing-country policing environments.

These spatial-temporal and hybrid models have significantly advanced predictive policing by representing crime as a dynamic, interconnected, and context-dependent phenomenon. Nevertheless, each modeling paradigm presents distinct deployment trade-offs. GIS and KDE remain computationally efficient and highly interpretable but are primarily retrospective; DBSCAN effectively identifies irregular crime hotspots yet lacks temporal forecasting capability; CNN–LSTM captures complex spatial-temporal dependencies but requires substantial computational resources and remains sensitive to noisy data; DMD provides stable forecasting but may overlook infrequent events; ST-GNNs effectively model crime diffusion but depend on high-quality relational data; whereas sparse Transformer architectures improve computational efficiency while requiring reliable multimodal inputs. Viewed through the AI Deployment Gap, these trade-offs demonstrate that the successful deployment of spatial-temporal predictive policing systems depends not only on predictive performance but also on data readiness, computational feasibility, explainability, governance, and institutional capacity. Accordingly, these observations reinforce the need for context-aware hybrid ML–NLP–spatial-temporal–XAI frameworks capable of balancing predictive accuracy with the practical realities of deployment in developing-country policing environments.

2.5. Natural Language Processing in Smart Policing

NLP has become an essential component of smart policing because a substantial proportion of law enforcement intelligence exists in unstructured narrative form, including police reports, crime diaries, case files, emergency call transcripts, intelligence briefs, and social media content [70]. Within the AI Deployment Gap framework adopted in this review, NLP plays a particularly important role by addressing one of the major deployment challenges in developing-country policing environments: the inability of conventional predictive systems to effectively utilize narrative-rich intelligence. Consequently, NLP extends predictive policing beyond structured crime records by transforming textual information into actionable intelligence that can support operational decision-making.

Early applications of NLP in policing primarily focused on information extraction, Named Entity Recognition (NER), and rule-based text processing. Techniques based on NER, Conditional Random Fields (CRF), SVM, lexicon-based systems, and rule-based General Architecture for Text Engineering (GATE) modules demonstrated that entities such as weapons, drug names, persons, locations, organizations, and incident descriptors could be extracted from police narratives [20], [33], [71]. Although these approaches achieved high precision and recall in relatively structured narrative settings, their performance deteriorated when police reports contained spelling errors, slang, incomplete syntax, inconsistent formatting, or locally specific expressions [72]. This sensitivity to noisy text remains a significant barrier to operational deployment. Nevertheless, their relatively low computational requirements make these methods attractive for policing agencies with limited digital infrastructure, despite their limited adaptability across jurisdictions and reporting styles due to their reliance on handcrafted rules.

Recent advances in Transformer-based architectures and LLMs have expanded the role of NLP from basic information extraction to semantic understanding and narrative reasoning. Models such as BERT, BART, GPT-3, GPT-4, Llama, Baichuan, ChatGLM, and Qwen have been successfully applied to crime classification, entity extraction, qualitative coding, and the interpretation of police narratives [27], [34], [60], [73]. Evidence indicates that BERT outperforms conventional classifiers in incident-type identification, while fine-tuned lightweight LLMs substantially improve F1-scores for entity extraction, particularly for challenging entities such as organizational names [27], [73]. These developments demonstrate that modern NLP is evolving beyond syntactic processing toward contextual intelligence generation. However, effective deployment depends less on the underlying model architecture than on localized language adaptation, domain-specific training, and continuous validation within operational policing environments.

Despite these advances, deploying NLP in predictive policing presents several important trade-offs. First, model performance is highly dependent on the characteristics of the training data. Study [12] showed that a fine-tuned model achieving high accuracy in one city experienced a substantial decline when applied to another urban context, illustrating the impact of distribution shifts caused by differences in reporting styles, terminology, crime categories, and local policing practices. Second, general-purpose pretrained models often fail to capture the legal, procedural, and tactical language commonly found in police narratives, making domain-specific adaptation essential [74]. Third, LLM-based systems continue to require human oversight. According to [73], LLMs perform effectively as screening and negative-filtering tools, but ambiguous or high-impact cases still require expert review to maintain predictive reliability, accountability, and socio-technical legitimacy [24]. Collectively, these findings indicate that successful NLP deployment depends as much on local institutional practices and data characteristics as on advances in language modeling.

Beyond predictive performance, NLP also introduces practical challenges related to computational cost, privacy protection, and operational integration [58]. Standard Transformer architectures typically require substantial computational resources, limiting their applicability in resource-constrained policing environments. This challenge has stimulated the development of lightweight language models, sparse-attention mechanisms, and computationally efficient approaches such as Simulated Annealing Sparsity [59]. Consequently, deployment strategies in developing-country policing environments increasingly favor lightweight, locally hosted, and domain-adapted NLP models that balance predictive capability with computational efficiency, privacy preservation, and long-term operational sustainability [75].

Overall, the principal challenge is not simply the limited adoption of NLP but its limited integration within operational predictive policing systems. In many developing-country policing environments, narrative intelligence remains isolated from structured crime databases, spatial-temporal models, and explainability mechanisms, thereby widening the AI Deployment Gap. From this perspective, NLP should be regarded as a foundational intelligence layer that transforms police reports, crime diaries, case files, and other unstructured records into structured predictive features that can be integrated with downstream analytical models. When combined with ML, spatial-temporal learning, and XAI, NLP enables richer situational awareness while simultaneously supporting deployment requirements related to contextual adaptation, transparency, privacy protection, and human oversight. Accordingly, the future of predictive policing lies not in adopting increasingly larger language models, but in integrating context-aware NLP within practical, explainable, and operationally sustainable AI ecosystems.

2.6. Explainable Artificial Intelligence and Ethical Considerations

As predictive policing evolves from statistical forecasting toward increasingly sophisticated ML, DL, LLMs, and graph-based architectures, explainability has become a fundamental requirement for operational deployment rather than merely a desirable technical feature. Within the AI Deployment Gap framework adopted in this review, predictive performance alone is insufficient to support operational decision-making. Police officers, supervisors, policymakers, and oversight bodies must understand why a model identifies a particular location, event, or crime pattern as high risk before its recommendations can be trusted, audited, and translated into policing actions [5], [6], [76]. Consequently, explainability plays a critical role in bridging predictive intelligence and responsible operational deployment.

A fundamental distinction in XAI lies between post-hoc explanation and interpretable-by-design models. Post-hoc approaches, including Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP), generate explanations after a model has produced its predictions [35], [77]. These techniques provide valuable insights into the input features that influence individual predictions. For example, [9] demonstrated that SHAP can identify urban characteristics, such as commercial facilities and visitor density, that contribute to elevated crime risk, enabling police agencies to move beyond generalized patrol allocation toward more targeted prevention strategies. While post-hoc explanations substantially improve model transparency, operational policing environments often require inherently interpretable models in which explanations are produced alongside predictions, thereby reducing uncertainty during real-time decision-making.

Interpretable-by-design models embed explainability directly within the predictive architecture. Attention-based approaches, such as the Attention-based Interpretable Spatio-Temporal Network (AIST), utilize attention mechanisms to reveal which regions, temporal intervals, traffic patterns, or points of interest contribute most significantly to high-risk predictions [29], [78]. Such capabilities allow decision-makers to trace predictions back to specific evidence, improving both interpretability and operational confidence. Similarly, techniques such as Layer-wise Relevance Propagation (LRP) can identify whether model decisions are driven by meaningful crime-related indicators or by spurious and potentially biased features [79]. These approaches enhance transparency while simultaneously supporting more reliable operational validation.

Bias detection represents another critical function of XAI and should be regarded as an operational necessity rather than solely an ethical concern. Predictive policing models are commonly trained using historical crime records that may reflect long-standing patterns of over-policing, underreporting, or selective law enforcement [18], [45]. Without appropriate safeguards, predictive models may reinforce these historical biases by repeatedly classifying previously over-policed neighborhoods as high-risk, thereby creating self-reinforcing feedback loops [3]. XAI provides mechanisms for determining whether predictions are driven by legitimate crime-related evidence or by indirect proxies associated with poverty, ethnicity, geographic disadvantage, or historically biased policing practices. Complementary fairness metrics, including demographic parity and equalized odds, further support the evaluation of whether predictive outputs disproportionately affect particular groups [6]. Together, these mechanisms strengthen deployment readiness by enabling investigators and operational commanders to verify the rationale behind AI-generated recommendations before operational decisions are made.

The importance of explainability becomes even more pronounced in Nigeria and many other developing-country contexts, where AI governance, regulatory oversight, and digital infrastructures continue to evolve [75]. Rather than reflecting technological limitations, these conditions increase the need for transparent, auditable, and human-supervised AI systems capable of supporting accountable operational deployment [5]. Consequently, explainability should be viewed as an essential mechanism for strengthening institutional trust and facilitating the responsible adoption of predictive policing technologies within resource-constrained environments.

Furthermore, explainability must be adapted to local legal, cultural, and institutional contexts. Predictive policing models developed in high-income countries may not adequately reflect African policing structures, customary justice systems, or community-based conflict resolution practices [3], [76]. As a result, technically accurate models may still encounter limited operational acceptance if their recommendations cannot be interpreted within local institutional and societal contexts. Effective deployment therefore requires XAI to be integrated with data quality assurance, human oversight, fairness assessment, and community-sensitive implementation strategies [7], [30], [80]. This broader perspective highlights that explainability is not only a technical capability but also an important component of trustworthy AI governance.

Overall, the reviewed literature demonstrates that XAI has evolved from an auxiliary explanation technique into a core architectural component of modern predictive policing systems. While advances in ML, NLP, and spatial-temporal learning continue to enhance predictive capability, their operational value ultimately depends on whether model outputs can be understood, validated, and trusted by investigators, supervisors, and policymakers. Accordingly, XAI serves as the critical bridge between algorithmic performance and responsible AI deployment, enabling predictive policing systems that are not only accurate but also transparent, accountable, and operationally sustainable across diverse policing environments.

2.7. Synthesis of Research Gaps

The literature synthesis indicates that the primary limitation of contemporary predictive policing research is no longer the absence of advanced AI models but the gap between algorithmic capability and practical deployment. Across the reviewed studies, increasingly sophisticated ML, DL, NLP, and spatial-temporal approaches consistently demonstrate substantial improvements in predictive performance. However, their operational adoption remains constrained by fragmented data ecosystems, limited institutional readiness, insufficient explainability, governance challenges, and weak integration of heterogeneous intelligence sources. Rather than representing isolated technical issues, these limitations collectively define the AI Deployment Gap that emerges throughout the reviewed literature. Accordingly, this review synthesizes the identified challenges into four interrelated dimensions: information integration, model generalization, explainability and governance, and deployment robustness.

The first dimension is the information integration gap. Most predictive policing systems continue to rely predominantly on structured variables such as crime type, date, time, and location, whereas narrative resources—including police pocket notebooks, crime diaries, case files, intelligence reports, and other textual records—remain only weakly integrated into predictive pipelines. This limitation is particularly evident in Nigeria and many other developing-country contexts, where policing data are largely narrative-driven. Consequently, valuable contextual information, including suspect intent, modus operandi, actor relationships, and emerging crime patterns, is often excluded from predictive models, reducing both situational awareness and operational effectiveness [24], [79]. These findings further highlight the importance of integrating NLP as a foundational intelligence layer rather than treating it as an auxiliary analytical component.

The second dimension concerns the model generalization gap. Many state-of-the-art predictive models are developed and evaluated using standardized, high-quality datasets collected from data-rich environments. In contrast, policing agencies across Africa frequently operate with fragmented databases, partial digitization, inconsistent reporting standards, and heterogeneous information systems [3], [5]. These conditions introduce distribution shifts, whereby models that perform well in one operational context may exhibit substantial performance degradation when deployed elsewhere. The reviewed literature therefore suggests that successful deployment depends less on algorithmic sophistication than on localized validation, domain adaptation, and the ability of predictive models to accommodate diverse operational environments.

The third dimension is the explainability and governance gap. Although many studies acknowledge the importance of transparency, explainability is still frequently treated as an additional feature rather than a core design principle. In operational policing, however, high-risk predictions must be auditable, contestable, and understandable to investigators, supervisors, and policymakers. Post-hoc explanation techniques, including LIME and SHAP, provide valuable interpretability, yet they often approximate the behavior of complex predictive models rather than directly incorporating explainability into the decision-making process. Consequently, the reviewed evidence indicates that explainability should be embedded throughout the AI lifecycle alongside governance, fairness assessment, and human oversight to support responsible operational deployment.

The fourth dimension is the deployment robustness gap. While sophisticated architectures such as CNN–LSTM, ST-GNN, and Transformer-based models substantially improve spatial-temporal prediction, their deployment remains sensitive to noisy records, sparse labels, outliers, and computational constraints. For example, [37] reported that CNN–LSTM architectures effectively capture joint spatial-temporal dependencies but experience increased prediction error when trained on noisy urban datasets, whereas Dynamic Mode Decomposition (DMD) provides a more stable alternative under such conditions. These findings demonstrate that practical deployment requires balancing predictive performance with computational efficiency, robustness, explainability, and institutional capacity rather than maximizing predictive accuracy alone.

Collectively, these findings indicate that the future of predictive policing depends not simply on developing increasingly sophisticated AI algorithms but on designing predictive systems that remain reliable, explainable, and deployable under diverse socio-technical conditions. The synthesis presented in this review demonstrates that information integration, model generalization, explainability, governance, and deployment robustness should be addressed as interconnected challenges rather than independent research problems. Consequently, future predictive policing systems are likely to benefit from context-aware hybrid ML–NLP–spatial-temporal–XAI architectures capable of integrating narrative intelligence, supporting localized model adaptation, embedding explainability throughout the prediction pipeline, and remaining computationally practical for deployment in developing-country policing environments.

3. Methodology

This study adopts a Systematic Literature Review (SLR) methodology guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework to ensure transparency, reproducibility, and methodological rigor throughout the identification, screening, selection, and synthesis of relevant literature. While PRISMA provides the procedural foundation for the review, the literature is analyzed through the AI Deployment Gap perspective developed in this study. Rather than comparing predictive policing algorithms solely on predictive performance, the selected studies are critically evaluated according to key deployment dimensions, including data readiness, multimodal information integration, model generalization, explainability, governance, institutional capacity, and operational feasibility in developing-country policing environments. Accordingly, this review focuses on studies applying AI, particularly ML, NLP, spatial-temporal modeling, and XAI, to crime prediction and smart policing, with an emphasis on Nigeria, other developing-country contexts, and comparative evidence from developed countries.

3.1. Review Design

The review was designed to examine how advances in AI for predictive policing translate into real-world operational deployment across different policing environments, with particular emphasis on developing-country contexts. Rather than contrasting the level of algorithmic sophistication between developed and developing countries, the review adopts the AI Deployment Gap as its primary analytical framework. Accordingly, studies from developed countries are considered methodological benchmarks that demonstrate state-of-the-art AI capabilities, whereas studies from Nigeria and other developing countries are synthesized to examine how differences in data readiness, institutional capacity, governance, explainability, and operational feasibility influence the design, evaluation, and deployment of predictive policing systems. This perspective enables the review to move beyond purely technical comparisons toward a broader understanding of deployment readiness across diverse socio-technical environments.

A qualitative evidence synthesis was considered more appropriate than a quantitative meta-analysis because the selected studies exhibit substantial heterogeneity in crime domains, datasets, geographic settings, AI techniques, evaluation metrics, and deployment objectives. Rather than statistically aggregating predictive performance, the review synthesizes evidence across studies to identify methodological trade-offs, deployment constraints, multimodal integration strategies, explainability practices, and emerging research trends spanning ML, DL, NLP, spatial-temporal modeling, graph-based learning, XAI, and hybrid predictive policing frameworks. This analytical approach facilitates the identification of the principal dimensions of the AI Deployment Gap while providing context-aware recommendations for AI deployment in developing-country policing environments.

3.2. Literature Search Strategy

The literature search was conducted across IEEE Xplore, ScienceDirect, Google Scholar, and supplementary sources, including ResearchGate and targeted searches of the International Journal of Computer Techniques (IJCT). These sources were selected to provide comprehensive coverage of both highly indexed AI and computing research and context-specific studies from developing countries. IEEE Xplore and ScienceDirect provided access to peer-reviewed publications on predictive policing, ML, spatial-temporal modeling, NLP, and XAI, whereas Google Scholar facilitated broader retrieval of multidisciplinary and regional studies. ResearchGate and targeted IJCT searches were additionally employed to identify relevant publications from Nigeria and other developing-country contexts that are not consistently indexed in major digital libraries. This multi-source strategy was intended to balance methodological breadth with contextual coverage, ensuring that both globally recognized advances and region-specific deployment challenges were adequately represented.

The search strategy was derived from the review objectives and the proposed AI Deployment Gap framework, enabling the retrieval of studies addressing both methodological developments and deployment-related challenges. Search terms included predictive policing, crime prediction, smart policing, Artificial Intelligence, Machine Learning, Deep Learning, Natural Language Processing, Explainable Artificial Intelligence, spatial-temporal crime prediction, crime hotspot prediction, graph neural networks, large language models, governance, explainability, bias, fairness, accountability, institutional readiness, operational deployment, Nigeria, Africa, and developing countries. Representative search strings included "predictive policing" AND "machine learning", "crime prediction" AND "deep learning", "smart policing" AND "artificial intelligence", "crime prediction" AND "natural language processing", "predictive policing" AND "explainable AI", "spatial-temporal crime prediction" OR "crime hotspot prediction", "crime prediction" AND "graph neural network", "crime prediction" AND "large language model", "predictive policing" AND ("Nigeria" OR "Africa" OR "developing countries"), and "predictive policing" AND ("governance" OR "bias" OR "fairness" OR "accountability" OR "institutional readiness").

3.3. Data Collection

The initial literature search retrieved more than 300 records from the selected digital libraries and supplementary sources. After duplicate removal and preliminary screening, studies were excluded if they did not address AI-enabled predictive policing, crime prediction, smart policing, XAI, NLP, spatial-temporal modeling, governance, or AI deployment challenges. Following title, abstract, and full-text screening, 142 studies were retained for detailed assessment, resulting in a final corpus of 81 studies that directly addressed the objectives of this review and provided sufficient methodological or deployment-related evidence for qualitative synthesis.

To better understand the geographical distribution of the literature, the selected studies were categorized according to their primary study context rather than the nationality of the authors. As illustrated in Figure 2, the final corpus comprised 14 Nigeria-focused studies, 7 China-focused studies, 13 studies from other developing or emerging-country contexts, 29 empirical or benchmark studies conducted in developed countries, and 18 global review, conceptual, or foundational AI/XAI studies. This regional categorization supports the comparative objective of this review by highlighting differences in deployment environments, data ecosystems, governance structures, institutional readiness, and operational constraints across diverse policing contexts.

The selected studies span the period from 2002 to 2026, capturing the methodological evolution of predictive policing from early GIS-based crime mapping, police report processing, and ILP to contemporary AI-driven frameworks incorporating deep learning, GNNs, LLMs, XAI, sparse-attention mechanisms, and multimodal spatial-temporal learning. This broad temporal coverage enables the review to examine not only advances in predictive methodologies but also the evolution of deployment-related challenges accompanying increasingly sophisticated AI systems.

Each selected paper was manually reviewed and coded according to publication year, country or region, study type, data modality, policing problem, AI methodology, reported performance, key limitations, and relevance to Nigeria, Africa, or other developing-country contexts. Additional coding captured whether a study addressed ML, deep learning, NLP, XAI, spatial-temporal prediction, ethical considerations, governance, or deployment feasibility. Based on this coding process, the literature was organized into the taxonomy presented in Table 1, which classifies studies according to problem class, model architecture, data modality, interpretability or governance function, and representative references.

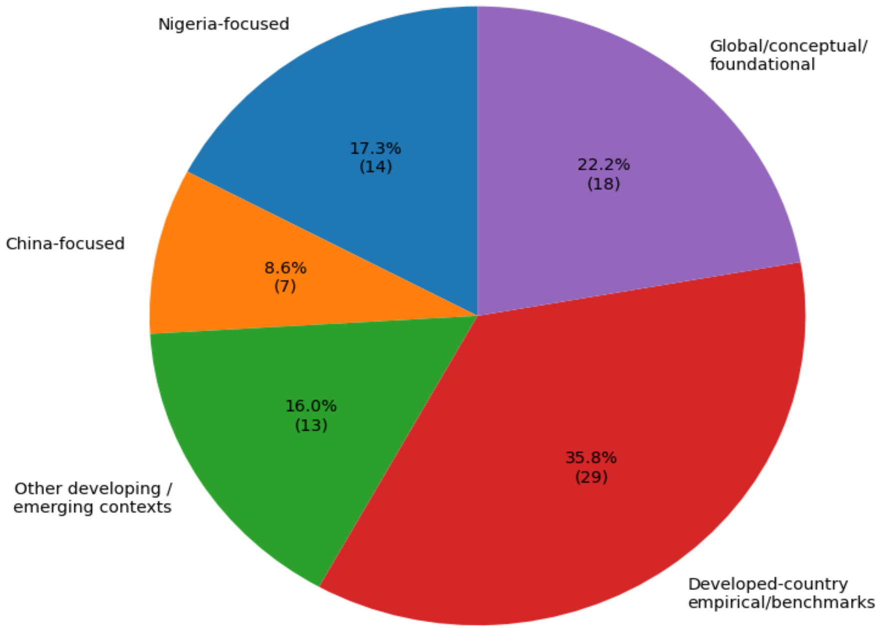

Figure 2. Regional contribution of the 81 selected studies.

Table 1. Taxonomy of AI-Based Predictive Policing Research

Problem Class	Representative Model Architectures	Primary Data Modality	Interpretability / Governance Role	Representative Studies
Crime forecasting and trend analysis	RF, DT, XGBoost, LightGBM	Structured administrative records, census data, crime logs	Moderate; feature importance and operational interpretability	Safat et al. [44]; Sridharan et al. [36]; Yusuf & Samuel [32]
Temporal crime forecasting	RNN, LSTM, GRU, N-BEATS, CNN–LSTM	Structured time-series crime data	Low–Moderate; nonlinear temporal modeling	Devi & Kavitha [56]; Adedamola & Ogundunmade [39]; Albino et al. [37]
Spatial hotspot mapping	GIS, KDE, DBSCAN	Geospatial coordinates and incident locations	High; intuitive hotspot visualization	Ferreira et al. [62]; Idhoko et al. [63]; Orekoya et al. [19]
Spatial-temporal crime diffusion	GNN, ST-GNN, STGCN, MTGNN	Spatial interaction networks	Low–Moderate; captures latent spatial dependency	Wu et al. [54]; Zhang et al. [81]; Orekoya et al. [19]
Long-range spatial-temporal forecasting	Sparse Transformers, Adaptive Graph Convolution	Long-sequence spatial-temporal data	Moderate; efficient long-range dependency modeling	Qin et al. [58]; Qin, Wei & Gao [43]
Information and entity extraction	NER, BiLSTM–CRF, Flair, BERT, RoBERTa	Police narratives, judicial documents	Moderate; extracts entities and event information	Chau et al. [20]; Ku et al. [72]; Juez-Hernandez et al. [74]; Park et al. [68]
Narrative reasoning and synthesis	LLMs, Instruction-tuned LLMs	Police reports, witness statements, emergency calls, social media	Moderate; semantic reasoning with human validation	Xing & Chen [25]; Relins et al. [73]; Liu et al. [23]; Zhang et al. [81]
Multimodal predictive fusion	ML–NLP, Graph–Text Fusion, Spatial–Temporal Fusion	Structured and unstructured crime data	Moderate; integrates heterogeneous crime intelligence	Huang et al. [69]; Huang et al. [65]; Lukmanjaya et al. [24]
Forensic identity mapping	GGMM, Binary Clustering	Fingerprints, blood samples, forensic evidence	Moderate; supports suspect linkage	Mande et al. [61]
Evaluative logic and auditability	LIME, SHAP, Attention-based explanation, PAS	Model outputs, operational rules	High; transparency and auditability	Ribeiro et al. [35]; de Santana Correia, & Colombini [78]; Kim et al. [9]; Lundberg et al.[77]
Ethical and deployment assessment	Mixed methods, Governance frameworks	Policy documents, interviews, surveys	High; trust, fairness, governance, deployment	Uduo & Obaji [13]; Asuquo & Sinha [3]; Halford [6]; Sam [5]

Regional coding further revealed clear differences in research emphasis. Studies conducted in developed countries predominantly establish methodological benchmarks and introduce state-of-the-art AI architectures, whereas research from developing countries increasingly focuses on fragmented data infrastructures, limited computational resources, governance challenges, and deployment constraints. In particular, Nigeria-focused predictive policing research became noticeably more active between 2024 and 2025, reflecting growing interest in AI-assisted policing in response to security challenges such as the Boko Haram insurgency, banditry, farmer-herder conflicts, kidnapping, communal violence, cybercrime, and rapid urban crime growth. Collectively, these trends indicate an increasing demand for predictive policing systems that extend beyond algorithmic accuracy to address practical deployment requirements, including multimodal information integration, explainability, governance, and computational efficiency.

4. Results and Thematic Synthesis

This section presents the findings of the systematic review by synthesizing evidence from the selected studies according to the taxonomy presented in Table 1. The synthesis is organized around the seven research questions (RQ1–RQ7), providing a structured interpretation of methodological developments, deployment challenges, and emerging research directions in AI-based predictive policing. Rather than summarizing individual studies, the analysis identifies recurring patterns, methodological trade-offs, and cross-study evidence that characterize the current state of predictive policing research. Across all research questions, a consistent theme emerges: the primary challenge is no longer the development of increasingly sophisticated AI models but their reliable, explainable, and context-aware deployment under diverse socio-technical conditions. This overarching observation provides the conceptual foundation for the thematic analyses presented in the following sections.

4.1. RQ1: What AI techniques have been applied in predictive policing, and what methodological trade-offs characterize these approaches?

The reviewed literature demonstrates that predictive policing has evolved into a diverse AI ecosystem rather than a single methodological paradigm. Four major categories of approaches consistently emerge across the selected studies: traditional ML for structured crime prediction, DL for modeling complex nonlinear temporal dynamics, graph-based and spatial-temporal learning for representing crime diffusion and spatial dependencies, and NLP supported by LLMs for extracting intelligence from unstructured policing narratives. As reflected in the taxonomy presented in Table 1, these approaches should be regarded as complementary components of the predictive policing pipeline rather than competing alternatives, illustrating the broader transition from conventional statistical forecasting toward hybrid, multimodal, and explainable AI systems.

The synthesis further indicates that each methodological category addresses different operational requirements while introducing distinct technical and deployment trade-offs. Traditional ML techniques remain attractive because they are computationally efficient, relatively interpretable, and well suited to structured crime records that are commonly available across policing agencies. However, their capability to capture complex spatial-temporal interactions and heterogeneous data sources remains limited. In contrast, DL architectures substantially improve the modeling of nonlinear crime dynamics and long-term temporal dependencies but require larger training datasets, greater computational resources, and stronger explainability mechanisms to support operational deployment. Similarly, graph-based and attention-driven models provide more realistic representations of crime diffusion and inter-location relationships, although their effectiveness depends heavily on reliable relational data, accurate graph construction, and mature digital infrastructures. Meanwhile, NLP and LLM-based approaches reduce information loss by transforming police reports, crime diaries, case files, and other narrative intelligence into structured predictive features. Nevertheless, their deployment introduces additional challenges related to language localization, privacy, domain adaptation, and the need for continuous human validation.

Beyond the technical characteristics of individual algorithms, the reviewed evidence suggests that model selection should be viewed primarily as a deployment decision rather than simply an optimization problem aimed at maximizing predictive accuracy. Across the reviewed studies, predictive performance is consistently influenced by factors extending beyond model architecture, including data quality, institutional readiness, computational capacity, governance requirements, explainability, and operational feasibility. These considerations become particularly important in Nigeria and other developing-country contexts, where fragmented datasets, partial digitization, limited computational infrastructure, and evolving governance frameworks restrict the direct adoption of sophisticated AI models originally developed in data-rich environments. Consequently, the suitability of a predictive model is determined not only by its algorithmic capability but also by its compatibility with local operational conditions.

A recurring finding across the reviewed literature is that no single AI technique adequately satisfies the diverse technical, organizational, and operational requirements of predictive policing. Instead, the evidence consistently favors hybrid predictive frameworks that combine the complementary strengths of multiple AI paradigms. Such frameworks integrate the computational efficiency and interpretability of traditional ML, the temporal modeling capability of DL, the spatial reasoning of graph-based learning, the contextual intelligence extracted through NLP, and the transparency provided by XAI. Rather than pursuing increasingly complex standalone algorithms, current research increasingly emphasizes balanced AI ecosystems that simultaneously address predictive performance, interpretability, robustness, governance, and operational feasibility. This shift represents one of the clearest trends identified in the reviewed literature and provides a practical direction for future predictive policing systems, particularly within resource-constrained developing-country environments.

4.2. RQ2: How do hybrid and state-of-the-art predictive models address nonlinear crime dynamics, spatial dependencies, and distribution shifts?

The reviewed literature indicates that hybrid and state-of-the-art (SOTA) predictive policing models have evolved primarily to overcome three persistent challenges that limit conventional crime prediction systems: nonlinear crime dynamics, spatial dependency, and distribution shifts. Rather than relying on a single predictive paradigm, contemporary approaches increasingly combine temporal learning, spatial reasoning, graph-based representations, attention mechanisms, and semantic feature extraction to capture the complex and interconnected nature of criminal activity. This evolution reflects a broader transition from standalone predictive models toward integrated AI frameworks capable of supporting decision-making under heterogeneous operational conditions.

The first major finding concerns the modeling of nonlinear crime dynamics. Crime patterns rarely follow stable or linear trajectories, as they are influenced by complex interactions among social, economic, environmental, and behavioral factors. Across the reviewed studies, deep recurrent architectures and attention-based models consistently demonstrated superior capability for capturing long-term temporal dependencies, irregular crime trends, and evolving behavioral patterns. Transformer-based models further enhance prediction by incorporating contextual information extracted from narrative intelligence, including offender intent, crime motives, and case-specific characteristics [56], [68]. Collectively, these findings suggest that integrating structured and unstructured information enables richer representations of evolving criminal behavior than traditional forecasting approaches.

A second recurring finding relates to spatial dependency. The reviewed literature consistently treats crime as a networked phenomenon rather than a collection of independent events. Graph-based and spatial-temporal models improve predictive capability by explicitly modeling interactions among neighboring locations, transportation corridors, and urban activity patterns, thereby capturing crime diffusion more effectively than conventional hotspot analysis. Evidence from Nigerian studies similarly shows that criminal activities frequently cluster across adjacent administrative areas, emphasizing the importance of learning geographic proximity, functional relationships, and spatial interactions when constructing predictive models [54], [61], [63]. These observations reinforce the growing importance of graph-based learning for representing the interconnected nature of urban crime.

The third and perhaps most significant finding concerns distribution shifts, which emerge as one of the primary barriers to real-world deployment. Across the reviewed studies, models trained using standardized datasets frequently experienced reduced performance when applied to jurisdictions with different reporting practices, crime taxonomies, socio-economic conditions, or data quality. These findings indicate that predictive performance alone provides an incomplete measure of model effectiveness. Instead, successful deployment requires localized validation, domain adaptation, robust feature engineering, and continuous model updating, particularly in developing-country policing environments characterized by fragmented records and inconsistent digitization [74]. Distribution shift therefore represents not merely a technical issue but a fundamental challenge to the operational transferability of AI models.

Taken together, the evidence suggests that the practical value of hybrid and SOTA predictive models arises not solely from their algorithmic sophistication but from their ability to accommodate diverse operational contexts. Their effectiveness is determined by the interaction between computational capability and deployment conditions, including data quality, institutional readiness, digital infrastructure, governance, and explainability. This finding also explains why models achieving state-of-the-art performance on benchmark datasets cannot be assumed to generalize directly to resource-constrained policing environments without appropriate adaptation and validation.

Overall, the reviewed literature indicates that the greatest benefits of hybrid predictive policing systems are achieved when methodological innovation is accompanied by deployment readiness. Rather than pursuing increasingly complex architectures in isolation, future research should prioritize context-aware multimodal frameworks that integrate temporal modeling, spatial reasoning, narrative intelligence, and explainability while remaining computationally efficient and operationally sustainable. Such an approach is particularly relevant for Nigeria and other developing-country contexts, where successful deployment depends on balancing predictive capability with local data characteristics, governance requirements, and institutional capacity.

4.3. RQ3: To what extent has NLP been integrated into predictive policing systems, and what limitations constrain its use in narrative-heavy intelligence environments?

The synthesis indicates that NLP has evolved from rule-based information extraction toward Transformer-based architectures and LLM-driven narrative intelligence, substantially expanding the capability of predictive policing systems to exploit unstructured textual data [60]. Despite these advances, NLP remains significantly underrepresented compared with conventional ML techniques. For example, [11] reported that traditional ML accounts for approximately 67% of crime prediction studies, whereas only about 2% integrate DL with NLP. This imbalance suggests that although law enforcement agencies generate large volumes of narrative information—including police pocket notebooks, crime diaries, case files, emergency call transcripts, intelligence reports, and social media content—much of this intelligence remains disconnected from operational predictive policing systems.

Across the reviewed literature, NLP has primarily been applied to four major tasks: information extraction, forensic text analysis, crime classification, and multimodal intelligence fusion. Early approaches relied on rule-based systems and lexical methods to identify entities such as weapons, vehicles, locations, drugs, and suspect identities. More recent Transformer-based and LLM architectures have substantially extended these capabilities by extracting richer semantic and contextual information from narrative records. For instance, [81] transformed more than one million judicial documents into a standardized spatial-temporal repository using LLMs; [68]reported strong F1-scores for homicide-related information extraction; [25] demonstrated that domain-adapted lightweight LLMs improved entity extraction from police reports by 10–50%; and [27] showed that BERT provides greater robustness when processing noisy policing narratives. Collectively, these findings demonstrate that modern NLP has progressed beyond simple entity recognition toward generating structured intelligence that can support downstream predictive analytics.

A recurring finding across the reviewed studies is that the greatest contribution of NLP lies not merely in extracting textual information but in enriching predictive models with contextual intelligence unavailable in conventional structured crime databases. Narrative policing records frequently contain valuable information regarding points of interest, mobility patterns, criminal hideouts, gang activities, escape routes, suspect behavior, community tensions, and modus operandi, all of which are difficult to represent using crime counts, timestamps, or geographic coordinates alone. The literature consistently suggests that integrating these semantic features with spatial-temporal and graph-based models improves contextual risk assessment, hotspot interpretation, patrol planning, and resource allocation by providing a more comprehensive representation of evolving crime dynamics.

Despite these promising developments, several barriers continue to limit the operational adoption of NLP in predictive policing. The reviewed studies consistently report that NLP models remain sensitive to noisy records, local dialects, incomplete syntax, inconsistent reporting styles, and domain-specific policing terminology. Furthermore, although LLMs perform well in extracting explicitly stated entities, they remain less reliable when required to infer offender intent, legal relevance, causal relationships, or other implicit contextual information that is not directly expressed in the text [38], [68]. Commercial LLMs may also generate procedurally plausible yet legally inaccurate outputs, while post-hoc explanations cannot always faithfully represent the underlying reasoning process. These limitations reinforce the continuing importance of human-in-the-loop validation, particularly for high-impact operational policing decisions [60], [73].

Beyond algorithmic performance, the synthesis identifies several practical deployment constraints. Standard Transformer architectures typically require considerable computational resources, whereas cloud-based LLM services introduce additional concerns related to privacy, governance, and data sovereignty when processing sensitive policing information. These challenges are especially pronounced in developing-country policing environments, where computational infrastructure, secure data management, and AI governance frameworks remain relatively limited. Consequently, recent research increasingly advocates lightweight, domain-adapted, and privacy-preserving NLP models incorporating sparse-attention and other efficiency-oriented mechanisms to improve operational feasibility within resource-constrained environments [25], [59].

Overall, the reviewed literature suggests that the principal limitation is no longer the capability of NLP itself but its limited integration within end-to-end predictive policing frameworks. In most existing studies, NLP remains an isolated text-processing component rather than an integral layer within multimodal AI systems, leaving valuable narrative intelligence disconnected from structured crime records, spatial-temporal reasoning, graph-based learning, and explainability mechanisms. The synthesis therefore indicates that the greatest opportunity for advancing predictive policing lies in seamlessly integrating NLP-derived intelligence with ML, spatial-temporal learning, graph-based models, and XAI. For Nigeria and other developing-country contexts, such integration should prioritize lightweight, locally adapted, and privacy-preserving NLP solutions supported by human oversight to ensure that narrative policing records can be transformed into reliable and operationally actionable intelligence.

4.4. RQ4: How do data readiness limitations, fragmented infrastructures, and institutional constraints affect predictive model feasibility in developing-country policing contexts?

The synthesis consistently indicates that the feasibility of predictive policing in developing-country contexts is determined less by the availability of advanced AI algorithms than by the readiness of the surrounding deployment environment. Across the reviewed studies, successful implementation depends on the combined influence of data readiness, digital infrastructure, institutional capacity, governance, and explainability rather than predictive accuracy alone. These findings suggest that evaluating predictive policing systems solely on benchmark performance provides an incomplete assessment of their practical value, particularly in resource-constrained policing environments.

A recurring challenge identified in the literature is limited data readiness. In Nigeria, policing information remains fragmented across crime diaries, police pocket notebooks, investigation case files, intelligence reports, and other narrative records, resulting in heterogeneous and only partially digitized datasets. Consequently, predictive models that rely exclusively on structured variables, such as crime type, date, time, and location, often fail to exploit valuable contextual information, including offender intent, modus operandi, witness observations, behavioral indicators, and localized risk factors [48]. Similar challenges have been reported in India, where national crime databases continue to exhibit class imbalance and underrepresentation of infrequent crime categories. In contrast, studies from China demonstrate how extensive digitization of judicial records has enabled LLMs to construct large-scale spatial-temporal repositories that support richer predictive analytics [25], [60], [81]. Collectively, these findings indicate that the maturity and quality of the underlying data ecosystem are more influential than algorithmic sophistication in determining the practical applicability of ML, NLP, and spatial-temporal prediction models.

The literature also identifies digital infrastructure as a major determinant of deployment feasibility. In Nigeria, fragmented information systems, limited interoperability among databases, inconsistent electronic record management, unreliable electricity supply, and constrained network infrastructure continue to impede the development of reliable predictive policing pipelines [13]. These limitations restrict the deployment of computationally intensive DL, Transformer-based, and real-time spatial-temporal architectures that depend on continuous data availability and substantial computing resources [2]. Similar evidence from Bangladesh demonstrates that predictive performance is closely associated with computational capacity, dataset quality, and resilience to noisy observations [44]. Accordingly, many recent studies advocate lightweight ML models, computationally efficient sparse-attention mechanisms, and optimized hybrid architectures that can operate effectively under resource-constrained conditions instead of directly transferring computationally demanding models developed in high-resource environments.

Beyond technical considerations, the reviewed evidence highlights institutional readiness and governance as equally important determinants of successful deployment. Implementing predictive policing requires organizations capable of interpreting, validating, auditing, and governing AI-generated recommendations throughout the decision-making process. Nigeria currently lacks comprehensive AI-specific policing policies, standardized digital reporting procedures, and formal mechanisms supporting algorithmic transparency and contestability, thereby limiting accountability and public confidence. Comparable studies from India emphasize the importance of human-in-the-loop decision support for preserving operational oversight [3], [36], while research conducted in Bangladesh demonstrates that interpretable models improve operational acceptance by allowing officers to understand why particular locations are identified as high-risk [29]. These findings collectively indicate that organizational preparedness and governance structures are essential components of operational AI deployment rather than complementary considerations.

Another recurring observation is that the deployment barriers identified across developing-country contexts are fundamentally socio-technical rather than purely computational. Advanced predictive models rarely achieve sustainable operational impact when deployed within fragmented data ecosystems, limited digital infrastructures, or weak governance environments. Consequently, successful implementation depends on adapting AI systems to local operational realities instead of assuming that models developed using standardized benchmark datasets will generalize across jurisdictions without modification. This finding reinforces the importance of contextual validation as a prerequisite for responsible AI deployment.

Overall, the reviewed literature suggests that predictive policing systems intended for Nigeria and other developing-country contexts should be designed around deployment readiness rather than algorithmic complexity. Effective implementation requires integrated frameworks that combine NLP-based narrative intelligence extraction, interoperable data pipelines, lightweight spatial-temporal learning, localized model validation, XAI-supported human oversight, and governance mechanisms aligned with institutional and cultural realities. These findings reinforce one of the central conclusions of this review: the long-term success of predictive policing depends on building AI systems that are not only technically effective but also operationally feasible, institutionally supported, and contextually appropriate for the environments in which they are deployed.

4.5. RQ5: How is Explainable Artificial Intelligence being used to support auditability, bias detection, and trustworthy deployment in predictive policing?

The reviewed literature indicates that XAI has evolved from a model interpretation technique into a fundamental enabler of trustworthy predictive policing. As predictive policing systems become increasingly dependent on complex ML, deep learning, and graph-based architectures, explainability has emerged as a prerequisite for operational deployment rather than an optional technical enhancement. Across the reviewed studies, the value of predictive models is consistently shown to depend not only on predictive accuracy but also on whether their outputs can be understood, audited, challenged, and governed throughout the decision-making process. Within the taxonomy proposed in this review, XAI therefore serves as the interpretability and governance layer that bridges algorithmic prediction with operational accountability.

A recurring finding across the literature is that XAI supports predictive policing through three complementary functions: auditability, bias detection, and trustworthy deployment. For auditability, post-hoc explanation techniques such as Local Interpretable Model-Agnostic Explanations (LIME), SHapley Additive exPlanations (SHAP), attention-based mechanisms, and inherently interpretable models identify the variables, spatial locations, and temporal factors that contribute to high-risk predictions [9], [29], [35]. These explanations improve transparency by allowing officers, supervisors, and analysts to examine the rationale underlying model outputs before operational decisions are made. However, the reviewed studies also consistently acknowledge an important limitation: post-hoc explanations provide approximations of model behavior and may not fully reflect the internal reasoning processes of complex deep learning and graph-based architectures [28], [78]. Consequently, explainability should be considered an integral component of model design rather than solely a post-processing step.

The literature further demonstrates that XAI plays an important role in bias detection and algorithmic accountability. Because predictive policing systems are typically trained using historical crime records, they may unintentionally reproduce patterns of selective enforcement, over-policing, or broader socio-economic disparities if these biases remain undetected [17], [18], [80]. Explainability techniques enable analysts to distinguish whether predictions are driven by legitimate criminogenic factors or by indirect proxies associated with neighborhood characteristics, historical enforcement practices, or other socially sensitive variables. Beyond improving model transparency, these capabilities provide an evidence-based mechanism for identifying, evaluating, and mitigating algorithmic bias before predictive outputs influence operational policing decisions.

Another consistent finding is that explainability alone is insufficient without corresponding governance and human oversight. The reviewed studies indicate that explanations become operationally meaningful only when they support legal accountability, institutional validation, and transparent decision-making processes. In Nigeria and many other developing-country contexts, comprehensive AI governance frameworks, standardized procedures for algorithmic accountability, and formal notice-and-explanation mechanisms remain underdeveloped [5], [6]. As a result, human-in-the-loop review, transparent operational workflows, and interpretable predictive outputs become essential safeguards for maintaining accountability, public trust, and responsible AI deployment. These findings highlight that explainability should be viewed as both a technical capability and an institutional governance mechanism. Overall, the reviewed evidence demonstrates that the contribution of XAI extends well beyond improving model interpretability. Its greatest value lies in enabling predictive policing systems to satisfy the accountability, governance, and institutional requirements necessary for operational deployment. Without appropriate explainability mechanisms, even highly accurate predictive models may remain unsuitable for real-world policing because their outputs cannot be adequately validated, contested, or trusted by decision-makers.

The synthesis therefore suggests that XAI should be regarded as a core design principle for future predictive policing systems rather than an auxiliary analytical tool. For Nigeria and other developing-country contexts, effective implementation requires interpretable model architectures, post-hoc explanation techniques, fairness evaluation, human-in-the-loop oversight, and governance mechanisms that operate together throughout the AI lifecycle. Such an integrated approach strengthens transparency, accountability, and public trust while supporting the responsible deployment of AI in operational policing environments.

4.6. RQ6: What ethical, governance, and deployment challenges influence the adoption of AI-driven predictive policing in Nigeria and other African contexts?

The reviewed literature consistently indicates that the adoption of AI-driven predictive policing in Nigeria and other developing-country contexts is constrained less by algorithmic capability than by the readiness of the surrounding socio-technical environment. Across the selected studies, successful implementation depends on the interaction among data quality, governance capacity, institutional readiness, explainability, and digital infrastructure rather than predictive accuracy alone. Although comparative evidence from India, Bangladesh, Pakistan, South Africa, the Philippines, and the United Arab Emirates provides valuable insights, Nigeria serves as the primary analytical focus because it reflects many of the operational conditions commonly encountered across developing-country policing environments.

A recurring challenge identified in the literature is ethical readiness, particularly the risk of amplifying historical bias. Predictive policing models trained on historical crime records may unintentionally reproduce patterns of selective enforcement, underreporting, ethnic disparities, and broader socio-economic inequalities when historical data are treated as objective representations of crime [2], [3]. As a consequence, predictive systems may repeatedly classify the same communities as high-risk, reinforcing existing policing practices instead of improving public safety. The reviewed studies further indicate that ethical deployment extends beyond statistical fairness. In many African contexts, predictive policing systems developed using data-rich environments may not adequately reflect community-oriented policing practices, customary justice mechanisms, or locally embedded conflict-resolution processes, resulting in what [76] describes as cultural flattening. These findings emphasize that responsible AI deployment requires sensitivity not only to algorithmic fairness but also to local institutional and cultural contexts.

The literature also highlights governance and institutional readiness as fundamental prerequisites for operational deployment. Nigeria currently lacks comprehensive AI-specific governance frameworks for predictive policing, resulting in fragmented arrangements for accountability, data stewardship, auditability, and legal oversight [5], [6]. Similar governance limitations have been reported across several developing-country settings, where weak database interoperability, inconsistent data standards, and unclear institutional responsibilities hinder transparent decision-making and complicate accountability when AI-assisted decisions contribute to operational errors or violations of individual rights. Collectively, these findings suggest that governance should be regarded as an enabling component of AI deployment rather than a regulatory consideration introduced after system implementation.

A third recurring theme concerns deployment feasibility under resource-constrained conditions. Nigerian policing agencies continue to operate with fragmented crime databases, partial digitization, inconsistent reporting practices, unreliable electricity supply, limited high-speed internet connectivity, and constrained computational resources [2], [13]. These conditions directly influence the operational suitability of different AI architectures. Although deep learning, Transformer-based, and graph-based models frequently achieve superior predictive performance when evaluated using well-curated datasets, their computational demands and dependence on high-quality data limit their routine adoption in many developing-country environments. Evidence from the Maiduguri case study illustrates this trade-off, where Decision Trees produced more reliable operational performance than computationally intensive alternatives when trained on fragmented local datasets [32]. Similar observations reported across other developing-country studies indicate that successful deployment depends on selecting AI models that match local infrastructure and operational capacity rather than simply adopting the most sophisticated available algorithms.

Taken together, these findings demonstrate that the barriers to adopting predictive policing in developing-country contexts are fundamentally socio-technical rather than algorithmic. Ethical safeguards, governance structures, institutional capacity, data ecosystems, and digital infrastructure collectively determine whether AI systems can be deployed responsibly, sustainably, and effectively. This distinction represents one of the central findings of the present review: differences between developed and developing countries arise primarily from deployment conditions rather than from differences in access to AI methodologies. In other words, comparable algorithms may produce substantially different operational outcomes because they are deployed within markedly different institutional, technical, and governance environments.

Overall, the synthesis suggests that responsible predictive policing in Nigeria and other developing-country contexts requires deployment-oriented AI frameworks that extend beyond predictive performance. Future systems should integrate locally validated predictive models, human-in-the-loop decision support, explainability, fairness assessment, robust governance mechanisms, lightweight computational architectures, NLP-enabled transformation of narrative intelligence, and place-based prediction strategies that align with local operational realities. Such an integrated approach provides a practical pathway for translating predictive policing from experimental AI research into trustworthy and sustainable operational practice.

4.7. RQ7: What research gaps and future directions emerge for context-aware, hybrid, and explainable predictive policing systems?

The synthesis of the reviewed literature indicates that the principal limitation of current predictive policing research is no longer the lack of sophisticated AI algorithms, but the limited integration of these technologies into deployment-ready systems capable of operating effectively under real-world policing conditions. Across the 81 reviewed studies, ML, NLP, spatial-temporal learning, and XAI are predominantly investigated as separate research domains, with relatively few studies considering their interaction within a unified operational framework. This fragmented research landscape constrains the translation of predictive policing from experimental models to practical decision-support systems, particularly in Nigeria and other developing-country contexts where deployment is shaped by data quality, institutional readiness, governance, computational resources, and operational feasibility rather than predictive performance alone.

A cross-study synthesis identifies four recurring research gaps. The first is an information integration gap, whereby structured crime records remain weakly connected with narrative intelligence, spatial-temporal reasoning, and explainability mechanisms, resulting in the loss of valuable contextual information. The second is a deployment gap, as predictive models developed using standardized and well-curated datasets frequently experience reduced performance when applied to fragmented records, inconsistent reporting practices, and resource-constrained policing environments. The third is a governance gap, where explainability, fairness assessment, and human oversight are commonly introduced as post-hoc validation mechanisms rather than being embedded throughout the AI deployment lifecycle. Finally, a computational gap persists because increasingly sophisticated deep learning, graph-based, and Transformer architectures often require computational resources and digital infrastructures that exceed the capabilities of many law-enforcement agencies in developing-country settings.

To summarize these findings, Figure 3 presents a conceptual deployment-oriented reference architecture synthesized from the reviewed literature. Rather than introducing a new predictive model, the architecture consolidates the principal methodological and deployment-related considerations identified throughout this review into a unified conceptual workflow. The synthesis illustrates how structured crime records can be complemented by intelligence extracted from police pocket notebooks, crime diaries, investigation case files, emergency call transcripts, and social media. NLP serves as an intelligence extraction layer that transforms unstructured narratives into structured predictive features, which are subsequently integrated with spatial-temporal information through multimodal feature fusion. The architecture further highlights the complementary roles of temporal learning, graph-based reasoning, and XAI in supporting robust, transparent, and operationally feasible predictive policing systems.

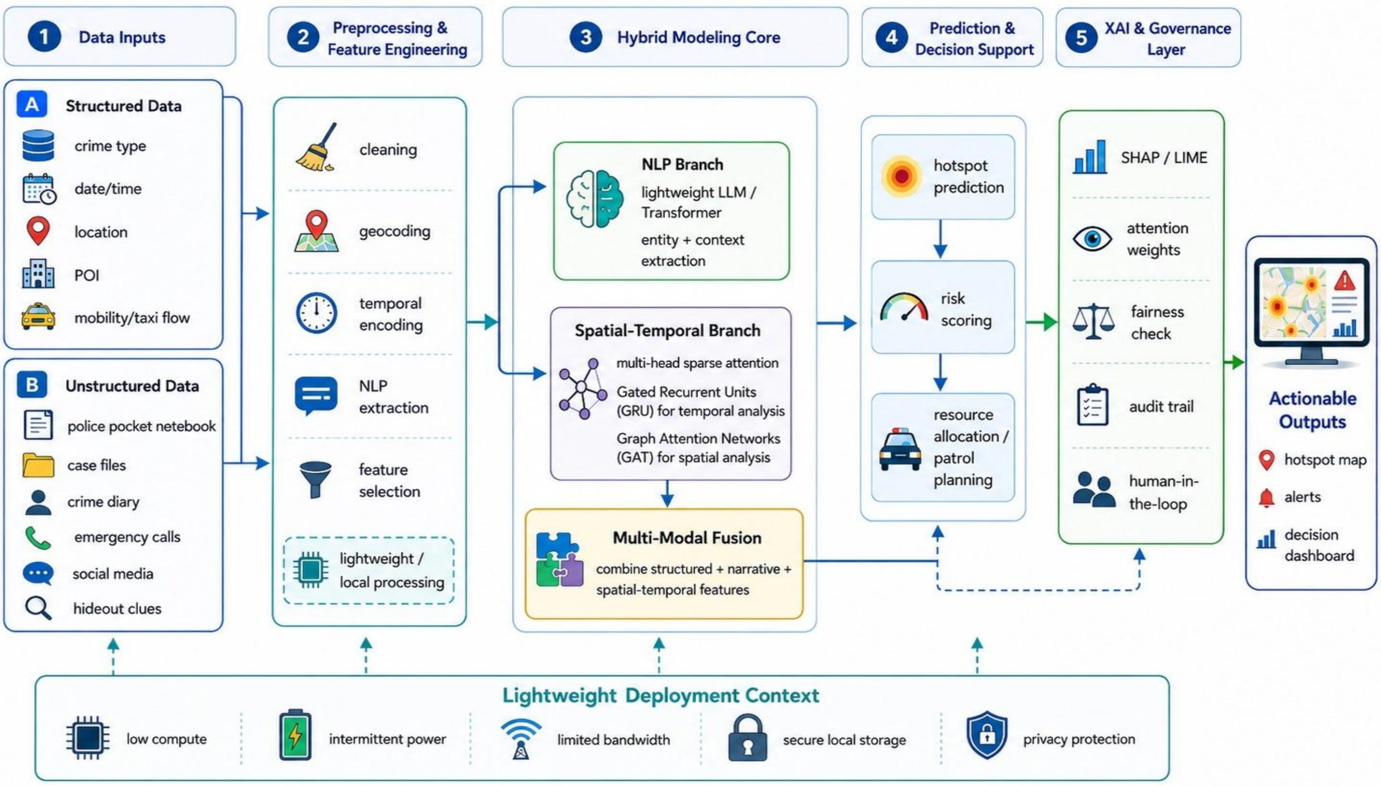

Figure 3. Conceptual deployment-oriented reference architecture synthesized from the reviewed literature.

The conceptual reference architecture summarizes the principal deployment-oriented design considerations emerging from the reviewed literature. It illustrates how structured crime records, narrative intelligence, lightweight NLP, spatial-temporal learning, multimodal feature fusion, and XAI-governance mechanisms can function as complementary components within context-aware predictive policing systems for resource-constrained environments.

As illustrated in Figure 3, the conceptual architecture emphasizes deployment considerations rather than algorithmic novelty. Instead of assuming abundant computational resources, fully digitized databases, or mature governance frameworks, the synthesized architecture reflects the operational realities commonly reported across developing-country policing environments. Accordingly, it highlights several recurring design principles identified throughout the literature, including lightweight computation, local data processing, privacy preservation, robustness to fragmented datasets, human-in-the-loop decision support, explainability, and continuous model adaptation. These principles collectively suggest that successful predictive policing depends on balancing predictive capability with deployment readiness rather than maximizing predictive accuracy alone.

The reviewed literature further suggests several priorities for future research. Beyond reporting predictive performance, future studies should place greater emphasis on cross-jurisdiction model generalization, robustness under noisy and incomplete datasets, multimodal data fusion, explainability fidelity, fairness evaluation, computational efficiency, and long-term operational deployment. Comparative studies involving both developed and developing-country policing environments would also provide stronger evidence regarding how data readiness, governance maturity, institutional capacity, and deployment conditions influence the successful adoption of AI-assisted policing systems.

Overall, the synthesis demonstrates that future advances in predictive policing are likely to depend less on developing increasingly sophisticated standalone algorithms than on designing integrated, deployment-oriented AI ecosystems. The conceptual reference architecture presented in Figure 3 should therefore be interpreted as a synthesis of the reviewed evidence rather than a new predictive framework. More broadly, this review identifies the AI Deployment Gap as the conceptual bridge linking advances in predictive modeling with successful operational deployment, providing a coherent perspective for guiding future research, system design, and responsible AI implementation in Nigeria and other developing-country contexts.

5. Discussion

This section interprets the findings through the AI Deployment Gap framework developed in this review, highlighting how data readiness, institutional capacity, governance, explainability, and deployment feasibility collectively influence the adoption of AI-driven predictive policing in developing-country contexts. The synthesis demonstrates that predictive policing has evolved beyond conventional hotspot mapping toward hybrid, multimodal, and socio-technical systems integrating ML, NLP, spatial-temporal learning, and XAI. However, the review also reveals that advances in predictive modeling do not necessarily translate into successful operational deployment. Instead, the practical value of AI-driven predictive policing depends on its compatibility with local data ecosystems, digital infrastructure, governance mechanisms, and institutional capabilities. This observation is particularly relevant to Nigeria and other developing countries, where fragmented datasets, limited digital infrastructure, evolving governance frameworks, and constrained resources require context-aware and deployment-oriented AI systems rather than the direct transfer of models developed in data-rich environments.

5.1. The AI Deployment Gap as a Design Problem

One of the central findings of this review is that the principal challenge facing predictive policing in developing-country contexts is not the availability of advanced AI techniques but the mismatch between algorithmic capability and deployment readiness. This mismatch, conceptualized as the AI Deployment Gap, arises when predictive models designed for highly digitized and data-rich environments are deployed in policing systems characterized by fragmented datasets, narrative-based records, evolving governance structures, limited computational infrastructure, and constrained institutional capacity. Consequently, the operational success of predictive policing depends as much on deployment conditions as on algorithmic performance.

The comparative evidence reviewed in this study further suggests that the distinction between developed and developing-country contexts should not be interpreted as a difference in algorithmic sophistication. Instead, the primary differences lie in the environments in which AI systems are deployed. Developed-country policing agencies generally operate with standardized digital crime databases, interoperable information systems, mature governance frameworks, and reliable computational infrastructure. By contrast, many developing-country policing environments, including Nigeria, continue to rely on partially digitized records, heterogeneous reporting practices, fragmented information systems, and evolving institutional governance. These differences substantially influence model generalizability, operational reliability, and long-term sustainability, even when similar AI techniques are employed.

These findings have important implications for the design of predictive policing systems. Rather than framing model selection as a search for the highest predictive accuracy, future systems should be designed around deployment readiness. This requires selecting algorithms that match local data characteristics, computational resources, institutional capacity, governance requirements, and operational workflows. Accordingly, lightweight architectures, locally validated models, multimodal data fusion, NLP-enabled extraction of narrative intelligence, and human-in-the-loop XAI represent more appropriate design strategies for resource-constrained policing environments. Viewed from this perspective, the AI Deployment Gap provides a practical design framework that links methodological innovation with operational feasibility, responsible governance, and sustainable deployment.

5.2. Narrative Intelligence and the Information-Loss Problem

Another important finding emerging from this review is the persistent underutilization of narrative intelligence within predictive policing systems. Although police operations routinely generate large volumes of textual information—including crime diaries, investigation reports, witness statements, and intelligence briefings—these data remain largely disconnected from predictive modeling pipelines. As a result, many existing AI systems rely predominantly on structured administrative variables while overlooking contextual information that may be critical for understanding crime emergence, offender behavior, and evolving criminal activities.

The reviewed studies consistently suggest that narrative information should be regarded as a strategic intelligence resource rather than a supplementary data source. When transformed into structured representations through NLP, narrative records provide valuable contextual information regarding offender intent, modus operandi, community tensions, mobility patterns, environmental conditions, and situational risk factors that cannot be captured using conventional structured crime records alone. Integrating these semantic features with spatial-temporal and graph-based models therefore enables a more comprehensive understanding of crime dynamics and supports more informed operational decision-making.

From a deployment perspective, the principal challenge is not merely adopting more advanced NLP models but embedding language-processing capabilities into predictive policing workflows that reflect local reporting practices, institutional procedures, and resource limitations. This finding is particularly relevant in developing-country contexts, where narrative documentation remains the dominant form of operational record keeping. Rather than requiring complete digital transformation before AI adoption, context-adapted NLP provides a practical mechanism for converting existing operational documents into structured predictive intelligence. Such an approach not only reduces information loss but also offers an incremental pathway for strengthening predictive policing capabilities while remaining compatible with existing policing practices and deployment constraints.

5.3. Deployment Robustness as a Socio-Technical Requirement

An important implication of this review is that robust predictive policing cannot be achieved solely through high benchmark performance. Instead, deployment robustness depends on whether AI systems can maintain reliable performance under diverse operational conditions, including variations in data quality, reporting practices, institutional workflows, governance structures, and computational resources. The synthesis demonstrates that changes in these socio-technical conditions may substantially influence model behavior, even when the underlying algorithms remain unchanged. Consequently, deployment robustness should be regarded as a system-level property rather than an algorithmic characteristic.

These findings suggest that local validation should be treated as an integral component of model development rather than a final evaluation step. Predictive models intended for operational deployment should be assessed according to their ability to function effectively with locally available datasets, computational infrastructure, institutional processes, and reporting practices. Such an evaluation perspective moves beyond the conventional objective of maximizing predictive accuracy and instead prioritizes the development of context-aware AI systems capable of balancing predictive performance, interpretability, robustness, and operational feasibility. This perspective reinforces one of the central conclusions of this review: successful predictive policing depends on designing AI systems around deployment conditions rather than assuming that models validated in data-rich environments will generalize across different policing contexts.

5.4. From Predictive Accuracy to Deployment Readiness

One of the most significant insights emerging from this review is that predictive accuracy alone provides an incomplete measure of AI readiness for operational policing. Although benchmark performance remains an important indicator of algorithmic capability, it does not necessarily reflect whether predictive outputs can be interpreted, validated, governed, and incorporated into routine policing practice. As demonstrated throughout this review, the practical value of predictive policing depends on the interaction between predictive capability and the institutional conditions under which AI systems are deployed.

This shift in perspective is particularly relevant for developing-country policing environments, where governance frameworks, institutional capacity, technical infrastructure, and operational oversight continue to evolve. Under these conditions, explainability functions not merely as a desirable model characteristic but as a practical mechanism that enables officers and supervisory authorities to interpret predictions, detect potential bias, justify operational decisions, and maintain public accountability. Accordingly, deployment readiness should be understood as the integration of predictive performance with explainability, governance, transparency, and human judgment rather than as a purely technical property of AI models.

The findings further suggest that future evaluations of predictive policing systems should adopt broader assessment criteria extending beyond conventional predictive metrics. Alongside accuracy, future research should systematically evaluate explainability, robustness, fairness, governance compliance, computational efficiency, and long-term operational usability. Such a multidimensional evaluation framework more accurately reflects the realities of resource-constrained policing environments and provides a stronger basis for assessing whether AI systems are suitable for responsible operational deployment. Ultimately, narrowing the AI Deployment Gap requires aligning advances in algorithm design with institutional readiness, governance capacity, and practical deployment requirements rather than pursuing increasingly complex predictive models alone.

5.5. Ethical and Institutional Implications for Nigeria, Africa, and Developing Countries

The findings of this review indicate that the ethical challenges associated with predictive policing cannot be separated from the institutional environments in which AI systems are deployed. In developing-country contexts, concerns regarding fairness, accountability, transparency, and public trust arise not only from algorithmic design but also from fragmented governance structures, inconsistent data management practices, evolving regulatory frameworks, and limited mechanisms for independent oversight. Consequently, ethical AI deployment should be viewed as a socio-technical challenge that requires institutional preparedness alongside technical innovation.

This interpretation provides an important perspective on the AI Deployment Gap discussed throughout this review. While predictive policing research conducted in developed countries often assumes mature digital infrastructures, interoperable databases, and well-established governance mechanisms, many policing agencies in Nigeria and other African countries continue to operate under conditions characterized by incomplete digitization, heterogeneous reporting practices, fragmented information systems, and limited institutional capacity. Under these circumstances, deployment challenges extend beyond algorithmic bias to include issues related to data provenance, model auditing, procedural accountability, operational transparency, and organizational acceptance.

The review further suggests that successful AI adoption depends on institutional adaptation rather than algorithmic transfer. Importing high-performing predictive models without corresponding governance mechanisms is unlikely to produce sustainable operational benefits. Instead, predictive policing systems should be accompanied by clearly defined policies governing data stewardship, model validation, continuous monitoring, audit procedures, and human decision authority. Such governance arrangements help ensure that AI-generated recommendations remain subject to professional oversight and reduce the risk of automated predictions being adopted uncritically during operational decision-making.

Another important consideration concerns contextual legitimacy. Policing practices across Nigeria and many other African jurisdictions are influenced by local administrative structures, community engagement, customary legal traditions, and diverse conflict-resolution mechanisms. AI systems that overlook these contextual characteristics may demonstrate satisfactory technical performance while remaining difficult to integrate into everyday policing practice. Designing predictive policing systems that incorporate explainability, transparent decision support, human oversight, and operational procedures aligned with local institutional practices is therefore essential for improving organizational acceptance, public trust, and long-term sustainability.

Overall, this review demonstrates that the responsible deployment of predictive policing in Nigeria and other developing-country contexts should be approached primarily as a governance and institutional design challenge rather than solely an algorithmic one. Bridging the AI Deployment Gap therefore requires coordinated advances in AI methodology, institutional readiness, governance frameworks, and operational practice so that technological innovation can be translated into accountable, trustworthy, and sustainable policing systems.

5.6. Implications for Future Smart Policing Architecture

The findings of this review suggest that future predictive policing systems should be designed around deployment readiness rather than algorithmic sophistication alone. Across the reviewed studies, many high-performing AI models assume the availability of standardized digital records, interoperable databases, reliable computational infrastructure, and mature governance mechanisms. However, these assumptions rarely hold in Nigeria and many other developing-country policing environments, where fragmented datasets, heterogeneous reporting practices, limited digital infrastructure, and evolving institutional frameworks remain common. Consequently, future smart policing architectures should explicitly incorporate deployment considerations alongside predictive performance to ensure that AI systems remain operationally feasible, explainable, and adaptable under diverse socio-technical conditions.

Based on the evidence synthesized throughout this review, Figure 3 presents a conceptual deployment-oriented reference architecture that summarizes the principal design considerations identified across the reviewed studies. Rather than introducing a new predictive model, the architecture consolidates existing evidence into a unified conceptual workflow that illustrates how ML, NLP, spatial-temporal learning, and XAI can operate as complementary components within an integrated predictive policing ecosystem.

Structured crime records are combined with intelligence extracted from police narratives and other unstructured sources, enabling richer contextual representation than conventional structured datasets alone. Spatial-temporal learning captures evolving crime dynamics and geographic interactions, while multimodal feature fusion integrates heterogeneous information sources to improve predictive capability under fragmented data conditions.

The conceptual architecture further highlights that explainability and governance should be considered integral design principles rather than post-deployment additions. Human-in-the-loop validation, interpretable prediction outputs, fairness assessment, and audit trails collectively support transparent operational review before AI-generated recommendations influence policing decisions. This synthesis suggests that successful deployment depends not only on predictive capability but also on whether model outputs can be interpreted, validated, and incorporated into existing legal, institutional, and operational decision-making processes.

From a practical perspective, the architecture also summarizes several implementation priorities identified across the reviewed literature for Nigeria and similar developing-country contexts. These include lightweight computational models suitable for limited hardware resources, privacy-preserving local processing for sensitive policing data, modular architectures supporting incremental digitization, and continuous model adaptation as data quality and institutional capacity improve. Collectively, these considerations provide a deployment-oriented roadmap that aligns AI development with the operational realities of resource-constrained policing environments.

Overall, the findings suggest that future predictive policing research should move beyond developing increasingly sophisticated standalone algorithms toward designing deployment-oriented AI ecosystems that integrate predictive modeling, contextual intelligence, explainability, governance, and operational feasibility. In this regard, the conceptual architecture in Figure 3 serves as a synthesis of the reviewed evidence rather than a new predictive framework, providing practical guidance for future research and system development.

5.7. Overall Contribution

This review contributes to the predictive policing literature by shifting the focus from comparing AI algorithms to understanding the conditions required for their successful deployment. Rather than evaluating predictive models solely on the basis of benchmark accuracy, the review synthesizes existing evidence through the AI Deployment Gap framework, demonstrating that the practical value of predictive policing depends on the interaction between algorithmic capability and the socio-technical environments in which AI systems are deployed. Across the reviewed studies, data readiness, institutional capacity, governance, explainability, computational feasibility, and operational readiness consistently emerged as critical determinants of successful implementation, particularly within developing-country policing contexts.

A second contribution lies in integrating methodological and contextual perspectives that have typically been examined independently in previous review studies. Existing surveys commonly discuss ML, DL, spatial-temporal learning, NLP, and XAI as separate methodological developments. In contrast, this review demonstrates that these technologies function most effectively as complementary components within an integrated predictive policing ecosystem. By synthesizing evidence across structured data analytics, narrative intelligence extraction, spatial reasoning, explainability, and governance, the review highlights the importance of designing AI systems that simultaneously balance predictive capability, transparency, adaptability, and institutional compatibility.

This review further contributes a deployment-oriented taxonomy that links methodological choices to operational constraints rather than treating predictive algorithms as universally applicable solutions. The synthesis demonstrates that selecting an appropriate predictive model should be guided not only by benchmark performance but also by data quality, computational resources, governance requirements, legal oversight, and organizational capacity. This perspective helps explain why AI models achieving state-of-the-art performance in highly digitized environments frequently require substantial adaptation before they can be deployed effectively within Nigeria and other resource-constrained policing systems.

Finally, this review synthesizes the accumulated evidence into a deployment-oriented conceptual reference architecture (Figure 3) that summarizes the major design principles identified across the reviewed literature. Rather than proposing a new predictive algorithm, the conceptual architecture illustrates how multimodal intelligence integration, explainability, governance mechanisms, and deployment-aware design considerations can be combined to support future AI-assisted policing systems. More broadly, this review positions deployment readiness as a fundamental criterion for evaluating predictive policing technologies and argues that the principal distinction between developed and developing-country contexts lies not in the sophistication of AI algorithms but in the conditions under which they are deployed. By introducing this deployment-oriented perspective, the review extends existing predictive policing surveys and provides a coherent conceptual foundation for future research, system design, and responsible AI implementation in developing-country policing environments.

6. Conclusions

This review examined the evolution of AI-enabled predictive policing through the analytical perspective of the AI Deployment Gap, highlighting how deployment conditions shape the selection, evaluation, and operational suitability of AI-driven predictive policing systems across developed and developing-country contexts. Rather than comparing algorithms solely in terms of predictive performance, the review synthesized evidence demonstrating that data readiness, institutional capacity, governance maturity, explainability, computational resources, and deployment feasibility collectively determine whether AI models can be translated from research prototypes into practical policing tools. The synthesis indicates that recent advances in ML, DL, spatial-temporal learning, GNNs, NLP, and XAI have substantially expanded the technical capabilities of predictive policing. However, these advances do not necessarily translate into successful operational deployment. In many developing-country contexts, including Nigeria, fragmented information systems, narrative-heavy crime records, partial digitization, evolving governance frameworks, and resource constraints fundamentally influence model robustness, explainability, and long-term sustainability. Consequently, predictive accuracy alone is insufficient as the primary criterion for evaluating predictive policing systems.

A central contribution of this review is the introduction of the AI Deployment Gap as a unifying analytical perspective for interpreting predictive policing research. Unlike previous surveys that primarily categorize AI techniques or compare predictive performance, this review demonstrates that deployment readiness emerges from the interaction between methodological capability and the socio-technical conditions under which AI systems operate. This perspective helps explain why models achieving state-of-the-art performance in highly digitized environments often require substantial adaptation before they can be deployed effectively in developing-country policing systems.

Beyond identifying research gaps, this review synthesizes the existing literature into a deployment-oriented conceptual reference architecture that summarizes the principal design considerations consistently reported across previous studies. Rather than representing a new predictive framework, this conceptual architecture integrates multimodal intelligence extraction, spatial-temporal reasoning, explainability, governance, and deployment-aware design principles into a coherent perspective that may guide future research and system development for resource-constrained policing environments.

The findings further suggest that future predictive policing research should move beyond isolated improvements to individual algorithms and instead investigate integrated AI ecosystems capable of supporting real-world operational deployment. Priority research directions include multimodal fusion of structured and unstructured policing data, lightweight and privacy-preserving AI models, adaptive learning strategies for addressing distribution shifts, explainability mechanisms embedded throughout the prediction pipeline, and comprehensive evaluation frameworks that jointly assess predictive performance, robustness, fairness, computational efficiency, governance, and deployment readiness.

The successful adoption of predictive policing will also require advances beyond algorithm development. Strengthening institutional capacity through AI training for law enforcement personnel, improving inter-agency data interoperability, developing standardized and secure crime information infrastructures, and establishing transparent governance frameworks for algorithmic accountability are essential prerequisites for responsible deployment. Achieving these objectives will require sustained collaboration among researchers, law enforcement agencies, policymakers, judicial institutions, and technology developers to ensure that predictive policing systems remain technically effective while satisfying legal, ethical, and societal expectations.

Overall, this review demonstrates that the primary distinction between predictive policing in developed and developing-country contexts lies not in the sophistication of AI algorithms but in the conditions under which they are deployed. By positioning deployment readiness as a central criterion for evaluating predictive policing technologies, the review provides a coherent conceptual foundation for future research and offers a deployment-oriented perspective for advancing responsible, explainable, and context-aware AI-assisted policing in Nigeria and other developing-country environments.

Author Contributions: Conceptualization: M.O.O. and E.O.O.; Methodology: M.O.O. and O.V.A.; Validation: E.O.O. and O.V.A.; Resources: M.O.O. and O.V.A.; Writing—Original Draft Preparation: M.O.O.; Writing—Review & Editing: O.V.A.; Supervision: E.O.O.; Project Administration: E.O.O. and O.V.A.; Funding Acquisition: All authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors express their sincere gratitude to the Department of Computer Engineering, School of Electrical Systems Engineering, The Federal University of Technology, Akure, Ondo State, Nigeria, and the Department of Information and Communication Engineering, Faculty of Engineering, Elizade University, Ilara-Mokin, Ondo State, Nigeria, for providing a conducive research environment and institutional support throughout this study. The authors also acknowledge the use of ChatGPT as an AI-assisted language support tool to improve the clarity, grammar, and readability of the manuscript. All AI-assisted outputs were carefully reviewed, revised, and verified by the authors, who assume full responsibility for the accuracy, integrity, and originality of the final published work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

[1] A. Adebayo, "Youths unemployment and crime in Nigeria: A nexus and implications for national development," *Int. J. Sociol. Anthropol.*, vol. 5, no. 9, pp. 350–357, Dec. 2013, doi: 10.5897/IJSA2013.0452.

[2] D. R. N. Eke, "Adopting Artificial Intelligence-Based Predictive Policing in Nigeria," *African J. Adv. Sci. Technol. Res.*, vol. 18, no. 1, pp. 142–164, Apr. 2025, doi: 10.62154/ajastr.2025.018.010680.

[3] S. P. Asuquo and A. Sinha, "The use of artificial intelligence in predicting criminal behaviour: A comparative study in Nigeria and India," *African J. Soc. Behav. Sci.*, vol. 15, no. 6, pp. 2559–2578, 2025.

[4] I. Yusuf and L. O. Okpara, "The influence of technological advancement on Nigeria's administration of the criminal justice system," *Sokoto J. Geogr. Stud.*, vol. 2, no. 1, pp. 69–93, 2025, [Online]. Available: https://sjgs.org.ng/index.php/data/article/view/43

[5] A. G. Sam, "Ethical and regulatory framework of artificial intelligence (AI) in Nigeria: The dilemma of global adaptation for sustainable growth," *Int. J. Innov. Soc. Sci. \& Humanit. Res.*, vol. 12, no. 3, pp. 55–65, 2024, [Online]. Available: https://www.seahipublications.org/wp-content/uploads/2024/07/IJISSHR-S-7-2024.pdf

[6] E. Halford, "The Transformer Led Policing model: a framework for applying generative artificial intelligence in policing," *Polic. A J. Policy Pract.*, vol. 19, Jan. 2025, doi: 10.1093/police/paaf027.

[7] E. E. Joh, "Policing by numbers: Big data and the Fourth Amendment," *Washingt. Law Rev.*, vol. 89, no. 1, pp. 35–68, 2014, [Online]. Available: https://digitalcommons.law.uw.edu/wlr/vol89/iss1/3

[8] K. Mohammed and A. Shehu, "A Review of Artificial Intelligence (AI) Challenges and Future Prospects of Explainable AI in Major Fields: A Case Study of Nigeria," *Open J. Phys. Sci. (ISSN 2734-2123)*, vol. 4, no. 1, pp. 1–18, Feb. 2023, doi: 10.52417/ojps.v4i1.458.

[9] G. Kim, Y. Cho, Y. Han, and G. Lee, "Crime mapping in urban environments using explainable AI: A case study of Daegu, Korea," *Sustain. Cities Soc.*, vol. 130, p. 106507, Jul. 2025, doi: 10.1016/j.scs.2025.106507.

[10] F. Olaoye and A. Egon, "Predictive Policing and Crime Prevention," Aug. 2024. [Online]. Available: https://easychair.org/publications/preprint/lQ2r

[11] V. Mandalapu, L. Elluri, P. Vyas, and N. Roy, "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions," *IEEE Access*, vol. 11, pp. 60153–60170, 2023, doi: 10.1109/ACCESS.2023.3286344.

[12] P. Sarzaeim, Q. H. Mahmoud, and A. Azim, "A Framework for LLM-Assisted Smart Policing System," *IEEE Access*, vol. 12, pp. 74915–74929, 2024, doi: 10.1109/ACCESS.2024.3404862.

[13] T., Uduo and O. R., "Evaluating the Effectiveness of Predictive Policing in Nigeria: A Qualitative study of law Enforcement Practices in Lagos State," *African J. Soc. Sci. Humanit. Res.*, vol. 7, no. 4, pp. 187–199, Oct. 2024, doi: 10.52589/AJSSHR-POLHVJ1F.

[14] M. A. Ekaabi, K. Khalid, R. Davidson, A. H. Kamarudin, and C. Preece, "Smart policing service quality: conceptualisation, development and validation," *Polic. An Int. J.*, vol. 43, no. 5, pp. 707–721, Aug. 2020, doi: 10.1108/PIJPSM-03-2020-0038.

[15] A. Sandhu and P. Fussey, "The 'uberization of policing'? How police negotiate and operationalise predictive policing technology," *Polic. Soc.*, vol. 31, no. 1, pp. 66–81, Jan. 2021, doi: 10.1080/10439463.2020.1803315.

[16] M. M. Yamin, A. Shalaginov, and B. Katt, "Smart Policing for a Smart World Opportunities, Challenges and Way Forward," in *Advances in Intelligent Systems and Computing*, 2020, pp. 532–549. doi: 10.1007/978-3-030-39445-5_39.

[17] K. Lum and W. Isaac, "To Predict and Serve?," *Significance*, vol. 13, no. 5, pp. 14–19, Oct. 2016, doi: 10.1111/j.1740-9713.2016.00960.x.

[18] P. J. Brantingham, M. Valasik, and G. O. Mohler, "Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial," *Stat. Public Policy*, vol. 5, no. 1, pp. 1–6, Jan. 2018, doi: 10.1080/2330443X.2018.1438940.

[19] V. P. Orekoya, D. Mathias, E. O. Bennett, and V. I. E. Anireh, "CriPaaP: A geospatial crime pattern analysis and prediction framework integrating DBSCAN, enhanced LSTM, and ST-GNN for urban safety in Nigeria," *Int. J. Comput. Tech.*, vol. 12, no. 5, pp. 628–644, Oct. 2025, [Online]. Available: https://ijctjournal.org/wp-content/uploads/2025/10/CriPaaP-A-Geospatial-Crime-Pattern-Analysis-and-Prediction-Framework-Integrating-DBSCAN-Enhanced-LSTM-and-ST-GNN-for-Urban-Safety-in-Nigeria.pdf

[20] M. Chau, J. J. Xu, and H. Chen, "Extracting meaningful entities from police narrative reports," in *Proceedings of the National Conference on Digital Government Research (DG.O 2002)*, Digital Government Society, Jun. 2002. [Online]. Available: https://repository.arizona.edu/items/5cbc2820-2b48-4ca0-aec2-9124c59aa1f3

[21] R. Kitchin, "Big Data, new epistemologies and paradigm shifts," *Big Data Soc.*, vol. 1, no. 1, Apr. 2014, doi: 10.1177/2053951714528481.

[22] F. Dakalbab, M. Abu Talib, O. Abu Waraga, A. Bou Nassif, S. Abbas, and Q. Nasir, "Artificial intelligence & crime prediction: A systematic literature review," *Soc. Sci. Humanit. Open*, vol. 6, no. 1, p. 100342, 2022, doi: 10.1016/j.ssaho.2022.100342.

[23] Y. Liu, Q. Guo, C. Yang, and Y. Liao, "TIPS: Tailored Information Extraction in Public Security Using Domain-Enhanced Large Language Model," *Comput. Mater. Contin.*, vol. 83, no. 2, pp. 2555–2572, 2025, doi: 10.32604/cmc.2025.060318.

[24] W. Lukmanjaya, C. Halmich, T. Butler, D. Cook, and G. Karystianis, "Computational text analysis on unstructured police data: a scoping review," *Crime Sci.*, vol. 15, no. 1, p. 6, Feb. 2026, doi: 10.1186/s40163-026-00272-2.

[25] X. Xing and P. Chen, "Entity Extraction of Key Elements in 110 Police Reports Based on Large Language Models," *Appl. Sci.*, vol. 14, no. 17, p. 7819, Sep. 2024, doi: 10.3390/app14177819.

[26] M. S. Gerber, "Predicting crime using Twitter and kernel density estimation," *Decis. Support Syst.*, vol. 61, pp. 115–125, May 2014, doi: 10.1016/j.dss.2014.02.003.

[27] P. Hosseini, S. Khoshsirat, M. Jalayer, S. Das, and H. Zhou, "Application of text mining techniques to identify actual wrong-way driving (WWD) crashes in police reports," *Int. J. Transp. Sci. Technol.*, vol. 12, no. 4, pp. 1038–1051, Dec. 2023, doi: 10.1016/j.ijtst.2022.12.002.

[28] T. Wang, C. Rudin, D. Wagner, and R. Sevieri, "Learning to Detect Patterns of Crime," in *Lecture Notes in Computer Science*, 2013, pp. 515–530. doi: 10.1007/978-3-642-40994-3_33.

[29] R. B. Utsha, M. N. Alif, Y. Rayhan, T. Hashem, and M. E. Ali, "Deep Learning Based Crime Prediction Models: Experiments and Analysis," Jul. 27, 2024. [Online]. Available: http://arxiv.org/abs/2407.19324

[30] M. Afzal and P. Panagiotopoulos, "Smart Policing: A Critical Review of the Literature," in *Lecture Notes in Computer Science*, 2020, pp. 59–70. doi: 10.1007/978-3-030-57599-1_5.

[31] A. Meijer and M. Wessels, "Predictive Policing: Review of Benefits and Drawbacks," *Int. J. Public Adm.*, vol. 42, no. 12, pp. 1031–1039, Sep. 2019, doi: 10.1080/01900692.2019.1575664.

[32] B. J. Yusuf and K. A. Samuel, "Crime prevention and detection using predictive algorithms: A case study of Maiduguri Metropolis," *Int. J. Comput. Tech.*, vol. 12, no. 5, pp. 321–335, Sep. 2025, doi: 10.5281/zenodo.17252927.

[33] P. Sarzaeim, Q. H. Mahmoud, A. Azim, G. Bauer, and I. Bowles, "A Systematic Review of Using Machine Learning and Natural Language Processing in Smart Policing," *Computers*, vol. 12, no. 12, p. 255, Dec. 2023, doi: 10.3390/computers12120255.

[34] T. S. Barros, C. E. S. Pires, and D. C. Nascimento, "Leveraging BERT for extractive text summarization on federal police documents," *Knowl. Inf. Syst.*, vol. 65, no. 11, pp. 4873–4903, Nov. 2023, doi: 10.1007/s10115-023-01912-8.

[35] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?' : Explaining the Predictions of Any Classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, NY, USA: ACM, Aug. 2016, pp. 1135–1144. doi: 10.1145/2939672.2939778.

[36] S. S, S. N, V. S, and S. P, "Crime Prediction using Machine Learning," *EAI Endorsed Trans. Internet Things*, vol. 10, Feb. 2024, doi: 10.4108/eetiot.5123.

[37] A. Joseph Albino, J. Ernest Camello, and C. F. Peña, "Spatio-Temporal Crime Prediction Using Dynamic Mode Decomposition and CNN-LSTM: A Comparative Study," in *Matrix Theory and Principal Component Analysis (PCA) Applications in Industry*, IntechOpen, 2026. doi: 10.5772/intechopen.1012879.

[38] K.-J. Kim, C.-H. Lee, S.-E. Bae, J.-H. Choi, and W. Kang, "Digital forensics in law enforcement: A case study of LLM-driven evidence analysis," *Forensic Sci. Int. Digit. Investig.*, vol. 54, p. 301939, Sep. 2025, doi: 10.1016/j.fsidi.2025.301939.

[39] A. O. Adedamola and T. P. Ogundunmade, "Predictive Modelling of Crime Data using Machine Learning Models: A Case Study of Oyo State, Nigeria," *Int. J. Innov. Sci. Res. Technol.*, pp. 1669–1677, Apr. 2025, doi: 10.38124/ijisrt/25apr851.

[40] M. Hou, X. Hu, J. Cai, X. Han, and S. Yuan, "An Integrated Graph Model for Spatial–Temporal Urban Crime Prediction Based on Attention Mechanism," *ISPRS Int. J. Geo-Information*, vol. 11, no. 5, p. 294, Apr. 2022, doi: 10.3390/ijgi11050294.

[41] E. G. İlgün and M. Dener, "Exploratory data analysis, time series analysis, crime type prediction, and trend forecasting in crime data using machine learning, deep learning, and statistical methods," *Neural Comput. Appl.*, vol. 37, no. 18, pp. 11773–11798, Jun. 2025, doi: 10.1007/s00521-025-11094-9.

[42] A. Vaswani *et al.*, "Attention Is All You Need," in *31st Conference on Neural Information Processing Systems (NIPS 2017)*, 2017. [Online]. Available: http://arxiv.org/abs/1706.03762

[43] Z. Qin, B. Wei, and C. Gao, "Innovative LSGTime Model for Crime Spatiotemporal Prediction Based on MindSpore Framework," Apr. 01, 2025. [Online]. Available: http://arxiv.org/abs/2503.20136

[44] W. Safat, S. Asghar, and S. A. Gillani, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques," *IEEE Access*, vol. 9, pp. 70080–70094, 2021, doi: 10.1109/ACCESS.2021.3078117.

[45] S. Brayne and A. Christin, "Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts," *Soc. Probl.*, vol. 68, no. 3, pp. 608–624, Aug. 2021, doi: 10.1093/socpro/spaa004.

[46] C. Catlett, E. Cesario, D. Talia, and A. Vinci, "A Data-Driven Approach for Spatio-Temporal Crime Predictions in Smart Cities," in *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*, IEEE, Jun. 2018, pp. 17–24. doi: 10.1109/SMARTCOMP.2018.00069.

[47] D. Weisburd, E. R. Groff, and S.-M. Yang, *The Criminology of Place*. Oxford University Press, 2012. doi: 10.1093/acprof:oso/9780195369083.001.0001.

[48] J. D. Madia, "Review of Predictive Policing and Artificial Intelligence," *Int. J. Police Sci.*, vol. 1, no. 1, 2022, doi: 10.56331/487529/ipsa5.

[49] S. Chainey, "Crime Mapping," in *Encyclopedia of Criminology and Criminal Justice*, New York, NY: Springer New York, 2014, pp. 699–709. doi: 10.1007/978-1-4614-5690-2_317.

[50] O. I. Odufisan, O. V. Abhulimen, and E. O. Ogunti, "Harnessing artificial intelligence and machine learning for fraud detection and prevention in Nigeria," *J. Econ. Criminol.*, vol. 7, p. 100127, Mar. 2025, doi: 10.1016/j.jeconc.2025.100127.

[51] M. A. Awodire, A. O. Momoh, N. Monteiro, and T. Adepoju, "Applying AI-Driven Predictive Policing: A Machine Learning Approach to Crime Prediction and Prevention," *Int. J. Eng. Comput. Sci.*, vol. 14, no. 06, pp. 27317–27339, Jun. 2025, doi: 10.18535/ijecs.v14i06.5160.

[52] M.-S. Baek, W. Park, J. Park, K.-H. Jang, and Y.-T. Lee, "Smart Policing Technique With Crime Type and Risk Score Prediction Based on Machine Learning for Early Awareness of Risk Situation," *IEEE Access*, vol. 9, pp. 131906–131915, 2021, doi: 10.1109/ACCESS.2021.3112682.

[53] A. Alsubayhin, M. Ramzan, and B. Alzahrani, "Crime Prediction Using Machine Learning: A Comparative Analysis," *J. Comput. Sci.*, vol. 19, no. 9, pp. 1170–1179, Sep. 2023, doi: 10.3844/jcssp.2023.1170.1179.

[54] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang, "Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks," in *ArXiv*, New York, NY, USA: ACM, Aug. 2020, pp. 753–763. doi: 10.1145/3394486.3403118.

[55] A. Malik, R. Maciejewski, S. Towers, S. McCullough, and D. S. Ebert, "Proactive Spatiotemporal Resource Allocation and Predictive Visual Analytics for Community Policing and Law Enforcement," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, pp. 1863–1872, Dec. 2014, doi: 10.1109/TVCG.2014.2346926.

[56] J. V. Devi and K. S. Kavitha, "Automating Time Series Forecasting on Crime Data using RNN-LSTM," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 10, 2021, doi: 10.14569/IJACSA.2021.0121051.

[57] J. H. Friedman, "Stochastic gradient boosting," *Comput. Stat. Data Anal.*, vol. 38, no. 4, pp. 367–378, Feb. 2002, doi: 10.1016/S0167-9473(01)00065-2.

[58] Z. Qin, B. Wei, C. Gao, F. Zhu, W. Qin, and Q. Zhang, "ACSAformer: A crime forecasting model based on sparse attention and adaptive graph convolution," *Front. Phys.*, vol. 13, Jun. 2025, doi: 10.3389/fphy.2025.1596987.

[59] W. Smith, "Policing Words with Machines: A Proof-of-Concept for LLM-Assisted Qualitative Analysis in Law Enforcement Research," *CrimRxiv*, Nov. 2025, doi: 10.21428/cb6ab371.0552f1de.

[60] S. Adhikary, P. Sen, D. Roy, and K. Ghosh, "A case study for automated attribute extraction from legal documents using large language models," *Artif. Intell. Law*, vol. 34, no. 1, pp. 245–266, Mar. 2026, doi: 10.1007/s10506-024-09425-7.

[61] U. Mande, Y. Srinivas, and J. V. R. Murthy, "Criminal mapping based on forensic evidences using generalized Gaussian mixture model," *Int. J. Comput. Sci. & Appl.*, vol. 1, no. 4, pp. 22–31, Jun. 2012.

[62] E. M. A. El-Aziz, S. Mesbah, and K. Mahar, "GIS-based decision support system for criminal tracking," in *2012 22nd International Conference on Computer Theory and Applications (ICCTA)*, IEEE, Oct. 2012, pp. 30–34. doi: 10.1109/ICCTA.2012.6523543.

[63] I. K. E., A. O. Daisi, and O. C. Akpomuode, "Spatial analysis of crime patterns in Nigeria (2021--2023): A GIS-based approach," *Int. J. Res.*, vol. 12, no. 6, pp. 426–450, Jun. 2025, doi: 10.5281/zenodo.15761682.

[64] J. M. Leigh, S. J. Dunnett, and L. M. Jackson, "Predictive policing using hotspot analysis," in *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2016), Volume II*, Hong Kong: International Association of Engineers (IAENG), Mar. 2016, pp. 856–861. [Online]. Available: https://www.iaeng.org/publication/IMECS2016/IMECS2016_pp856-861.pdf

[65] C. Huang, C. Zhang, J. Zhao, X. Wu, D. Yin, and N. Chawla, "MiST: A Multiview and Multimodal Spatial-Temporal Learning Framework for Citywide Abnormal Event Forecasting," in *The World Wide Web Conference*, New York, NY, USA: ACM, May 2019, pp. 717–728. doi: 10.1145/3308558.3313730.

[66] F. Caffaro, L. Bongiovanni, and C. Rossi, "Geo-temporal Crime Forecasting Using a Deep Learning Attention-Based Model," in *Security Informatics and Law Enforcement*, 2025, pp. 323–329. doi: 10.1007/978-3-031-62083-6_26.

[67] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification And Regression Trees*. Routledge, 2017. doi: 10.1201/9781315139470.

[68] Y. Park, R. S. Park, and H. Kim, "Key Information Extraction for Crime Investigation by Hybrid Classification Model," *Electronics*, vol. 13, no. 8, p. 1525, Apr. 2024, doi: 10.3390/electronics13081525.

[69] C. Huang, J. Zhang, Y. Zheng, and N. V. Chawla, "DeepCrime," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, New York, NY, USA: ACM, Oct. 2018, pp. 1423–1432. doi: 10.1145/3269206.3271793.

[70] T. Alshammari, "Extracting data from unstructured crime text to represent in structured occurrence nets using natural language processing," in *Proceedings of the International Workshop on Petri Nets and Software Engineering (PNSE 2024)*, in CEUR Workshop Proceedings, vol. 3730. CEUR-WS.org, 2024, pp. 270–282. [Online]. Available: https://ceur-ws.org/Vol-3730/paper16.pdf

[71] N. Kumar, J. De Beer, J. Vanthienen, and M.-F. Moens, "Intelligent Information Retrieval Tools for Police," in *Lecture Notes in Computer Science*, 2006, pp. 664–665. doi: 10.1007/11760146_80.

[72] C. H. Ku, A. Iriberri, and G. Leroy, "Crime Information Extraction from Police and Witness Narrative Reports," in *2008 IEEE Conference on Technologies for Homeland Security*, IEEE, May 2008, pp. 193–198. doi: 10.1109/THS.2008.4534448.

[73] S. Relins, D. Birks, and C. Lloyd, "Using Instruction-Tuned Large Language Models to Identify Indicators of Vulnerability in Police Incident Narratives," *J. Quant. Criminol.*, vol. 41, no. 4, pp. 647–684, Dec. 2025, doi: 10.1007/s10940-025-09611-z.

[74] R. Juez-Hernandez, L. Quijano-Sánchez, F. Liberatore, and J. Gómez, "AGORA: An intelligent system for the anonymization, information extraction and automatic mapping of sensitive documents," *Appl. Soft Comput.*, vol. 145, p. 110540, Sep. 2023, doi: 10.1016/j.asoc.2023.110540.

[75] S. O. Onazi, R. F. Olanrewaju, and G. Aimufua, "Big Data and National Security Threats in Nigeria: Challenges, Opportunities, and Strategies," Feb. 24, 2025. doi: 10.20944/preprints202502.1883.v1.

[76] J. Ubanyionwu, "Ethical and cultural considerations in artificial intelligence (AI) application to customary law," *African J. Law, Ethics Educ.*, vol. 7, no. 2, pp. 142–159, 2024, [Online]. Available: https://ajleejournal.com/index.php/ajlee/article/view/197/193

[77] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, Nov. 2017, pp. 4768–4777. [Online]. Available: https://dl.acm.org/doi/10.5555/3295222.3295230

[78] A. de Santana Correia and E. L. Colombini, "Attention, please! A survey of neural attention models in deep learning," *Artif. Intell. Rev.*, vol. 55, no. 8, pp. 6037–6124, Dec. 2022, doi: 10.1007/s10462-022-10148-x.

[79] Y. Lu *et al.*, "Unified Structure Generation for Universal Information Extraction," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2022, pp. 5755–5772. doi: 10.18653/v1/2022.acl-long.395.

[80] M. Roy, "Cathy O'Neil. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown Publishers, 2016. 272p. Hardcover, $26 (ISBN 978-0553418811).," *Coll. Res. Libr.*, vol. 78, no. 3, p. 403, Apr. 2017, doi: 10.5860/crl.78.3.403.

[81] Y. Zhang, M.-P. Kwan, and L. Fang, "An LLM driven dataset on the spatiotemporal distributions of street and neighborhood crime in China," *Sci. Data*, vol. 12, no. 1, p. 467, Mar. 2025, doi: 10.1038/s41597-025-04757-8.