ISSN 2007-9737

Artificial Intelligence Applied to Crime Prediction: A Bibliometric and Systematic Review of Predictive Models, Spatio-Temporal Data, and Emerging Frontiers

Rodrigo Italo Cayo-Mamani[1], Javier Gamboa-Cruzado[2], Jorge Nolasco-Valenzuela[3], Carlos Chávez-Herrera[2], Alex Salazar-Marzal[2], Martín Gamboa-Cruzado[4], Erick Quesquén Alarcón[2], Cesar Jesús Núñez-Prado[5,*]

[1] Universidad Nacional Federico Villarreal, Peru

[2] Universidad Nacional Mayor de San Marcos, Peru

[3] Universidad San Ignacio de Loyola, Peru

[4] Universidad Nacional de Trujillo, Peru

[5] Instituto Politécnico Nacional, ESIMEZ, Mexico

2021018461@unfv.edu.pe,
{jgamboac, cchavezh, alexmelecio.salazar, equesquena} @unmsm.edu.pe ,
jorge.nolasco@usil.pe, {martingamboac, cesar.jnprado}@gmail.com

Abstract. Crime prediction using Artificial Intelligence has increased in technical complexity and public relevance; however, the evidence remains fragmented regarding implementation practices, publication quality, and scientific collaboration. This paper synthesizes and characterizes the literature on Artificial Intelligence applied to Crime Prediction, with emphasis on development languages, quartile levels, co-authorship networks, and thematic categories. A systematic review was conducted following Kitchenham and PRISMA 2020 guidelines, with searches in ProQuest, IEEE Xplore, Scopus, ScienceDirect, and SpringerLink through July 22, 2025. After applying inclusion and exclusion criteria and conducting quality assessment with a threshold of at least 24, 57 studies out of 70 (81.4%) were selected for technical and bibliometric analysis. The results show the dominance of Python as the implementation language (60%), with secondary presence of R (20%) and limited adoption of Java, C++, and Julia at approximately 6.7% each. Production is concentrated in Q1 journals (33/57), followed by Q2 (16/57) and Q3 (8/57), suggesting high visibility, although with heterogeneous impact across sources. Co-authorship analysis reveals core and bridge authors connecting subcommunities, while the thematic map highlights specialized lines such as AI Crime Analytics and AI Policing, as well as marginal topics that are still consolidating. In conclusion, the field is progressing toward more integrated approaches but requires comparable standards, reproducible traceability, and governance frameworks for responsible deployment.

Keywords. Artificial intelligence, crime prediction, computational intelligence, crime forecasting, systematic literature review.

1 Introduction

In recent years, artificial intelligence, AI, has consolidated itself as a disruptive technology with strong potential to transform public security systems through data driven predictive analytics. In particular, crime prediction has evolved from traditional statistical approaches toward more sophisticated models supported by machine learning, data mining, and spatio temporal analytics. This shift responds to the growing availability of large volumes of crime data and to the need to anticipate criminal patterns with greater precision and timeliness. Nevertheless, the scientific evidence regarding the real impact of AI on crime prediction remains fragmented and heterogeneous, highlighting the need for a systematic synthesis that clarifies its contributions, limitations, and emerging trends. First, authors [1,58] show that Crime Prediction is strengthened when articulated with data informed simulations, for example agent based approaches, to support decision making and explore urban scenarios in a reproducible manner. Likewise, deep learning models oriented toward spatio temporal dynamics, such as LSTM and attention mechanisms, have achieved consistent improvements over traditional baselines, reinforcing their role as a dominant technical axis in Crime Prediction [4,64]. Complementarily, the incorporation of explicit spatial structures through hybrid architectures, for example GCN LSTM, converges with hotspot detection approaches based on clustering, suggesting an evolution toward models that are more sensitive to spatial dependence and urban heterogeneity [7,37].

Within the context of intelligent analysis of social data, several studies have shown that the use of Machine Learning enables the automation and optimization of processing large volumes of information from social networks. In particular, ML based sentiment analysis has been shown to improve analytical efficiency and reduce operational costs in business environments, demonstrating the potential of these techniques to support data driven decision making [72]. Furthermore, in cybersecurity domains, empirical evidence confirms the consolidation of Machine Learning algorithms, such as Support Vector Machine, Random Forest, and Naive Bayes, for automated threat detection tasks, reinforcing the maturity and transferability of these approaches to predictive analytics scenarios in data driven security contexts [75].

On the other hand, the exploitation of social and textual signals has gained traction. Sentiment analysis on social networks combined with reinforcement learning aligns with reviews that systematize crimes in social media and provide taxonomies and datasets, expanding the spectrum of evidence for predictive systems [5,6]. In contrast, approaches based on LLM, such as GPT, are emerging as an additional layer for classification and crime analysis in intelligent policing, consistent with reviews that highlight the use of ML and NLP in law enforcement applications, although still facing methodological and governance challenges [2,63]. According to evidence centered on video surveillance, models that leverage visual sequences and specific mechanisms, for example ASGCM, are integrated within broader intelligent surveillance frameworks, where the combination of computer vision and Artificial Intelligence structures predictive surveillance and operational analytics [3,59].

In the domain of identification and biometrics, it is relevant that comparative studies of facial recognition versus eyewitness testimony converge with advances in deep networks, such as siamese architectures, to enhance robustness under pose variation, reinforcing the trend toward improving performance and stability under non ideal conditions [10,34]. Despite these advances, the literature emphasizes substantial risks. Bias mitigation, for example age related bias, and the requirement for explainable models are connected with analyses of democratic tensions and discriminatory practices. Therefore, technical quality must be evaluated together with criteria of fairness, transparency, and institutional legitimacy [8,13,14]. Consequently, public security based on video surveillance also benefits from innovations designed for constrained environments, such as weapon detection with efficient architectures. Although high performance is reported in intelligent IoT and deep learning surveillance, challenges related to generalization and responsible deployment persist [24,66].

At the decision support level, the use of multicriteria models to select situational crime prevention interventions, CPTED, complements discussions on aligning Artificial Intelligence with sustainable security and social values, suggesting that operational effectiveness should be integrated with criteria of public acceptance and sustainability [27,11]. In parallel, geospatial approaches present two main strands. One involves spatial and temporal models with statistical inference, and the other includes terrain modeling schemes for socio environmental factors. Together, both contribute to prioritizing resources under uncertainty and operationalizing territorial explanations of crime risk [29,53]. Although classical methods remain relevant, evidence from ensemble approaches, such as stacking with SVM, and comparisons with robust models, for example RF and XGBoost, confirm that the best results strongly depend on context and dataset characteristics. Therefore, cross city portability requires more rigorous validation designs [9,32].

Additionally, the combination of clustering with time series models for hotspots is reinforced by critical reviews on high crime area prediction and detection, explaining the persistence of hybrid pipelines to anticipate trends using longitudinal urban data [48,67]. In contrast, studies on the adoption of Artificial Intelligence based smart city systems dialogue with thematic classifications of the field, including security, mobility, and health, suggesting that productivity peaks respond both to technical maturity and to implementation and regulatory incentives [52,62]. Finally, reviews on intelligent policing and ML techniques for law enforcement agencies connect with specific applications such as crime linkage using Random Forest, evidencing an expansion from prediction toward investigative and operational management tasks [60,61,55]. Moreover, recidivism models based on clustering approaches and explanatory components, together with pedestrian vulnerability prediction, indicate an expansion toward ethical decision making and risk prevention in urban security beyond conventional crime [54,56]. Similarly, anomalous emotion recognition appears as an adjacent line that may provide early warning signals, although its integration must be handled carefully due to potential sociotechnical harms and validation requirements [65,11].

Based on the reviewed state of the art, although there is a consolidated body of reviews documenting the use of AI and machine learning in crime prediction from methodological perspectives, such as types of analysis, predictive models, datasets, and performance metrics, technological perspectives, including video surveillance with computer vision, IoT, and deep learning, and operational perspectives, such as intelligent policing, forensic applications, and ethical considerations related to bias, the available evidence remains fragmented regarding critical dimensions for field maturity and comparability. These dimensions include standardized criteria for evaluating system effectiveness, programming languages used in implementation, the positioning of scientific production according to quartile levels, collaboration and coauthorship dynamics among researchers, and the systematic thematic characterization of the corpus [58-67]. This gap limits both the objective evaluation of performance and technical reproducibility of solutions and the bibliometric understanding of the evolution, quality, and interconnection of research, justifying a systematic review aimed at integrating, organizing, and analyzing the literature from a combined technical and bibliometric approach. Consequently, the distinctive contribution of this review lies in complementing previous syntheses focused primarily on models and applications by explicitly incorporating an analytical framework that consolidates effectiveness criteria, development languages, publication quartiles, coauthorship networks, and thematic categories. This provides a comparative and structured view that enables scientific traceability, quality assessment, and the identification of future research opportunities. The main objective is to analyze the impact of artificial intelligence on crime prediction by evaluating the effectiveness of current predictive models and exploring the social, ethical, and legal implications of their implementation. Within this framework, the paper is organized as follows. Section 2 presents the background, Section 3 describes the methodology employed, Section 4 presents and analyzes the results obtained, and Section 5 offers the conclusions together with recommendations for future research.

2 Background

Before examining in depth the use of artificial intelligence in crime prediction, it is essential to understand the fundamental concepts that support its implementation. The relationship between artificial intelligence and crime prediction involves a series of technological advances that enable pattern analysis and automated decision making. Understanding these elements is crucial to evaluate current trends and the potential impacts of artificial intelligence on the prevention and control of complex criminal activities.

2.1 Artificial Intelligence

Artificial intelligence constitutes a complex concept [10], and the need to explain how a system operates and why it proposes a specific decision has become increasingly relevant [13]. These systems are primarily grounded in artificial intelligence [52]; this involves the design of computational machines equipped with AI generated tools so that they can act in ways comparable to humans [68]. AI may provide significant contributions [12]; however, it represents the simulation by computers of intelligence inherent to humans [59]. These are information processing technologies that integrate models and algorithms capable of learning and executing advanced cognitive tasks [65].

2.2 Crime Prediction

Crime prediction and forecasting approaches have undergone substantial evolution in recent years since the introduction of commercial software packages [25]. The increase in criminal incidents, together with the growing availability of crime related data, has transformed crime prediction into a highly complex challenge [19]. Moreover, advances in object detection and tracking technologies have significantly strengthened the capacity to analyze human trajectories across different environments [40]. In this context, the use of agent based simulations to anticipate criminal patterns in urban environments [1], as well as the application of reinforcement learning and specialized loss functions [5], have contributed to improving the accuracy of predictive models. Likewise, crime prediction is also addressed through information aggregation models based on video surveillance to anticipate criminal activities [3].

2.3 Artificial Intelligence Driven Crime Prediction

Artificial intelligence, AI, has consolidated itself as a transformative factor in current crime prediction models by overcoming the limitations of traditional statistical approaches through machine learning capabilities, large scale data analysis, and automated decision making. Within this framework, AI based technologies enable the identification of spatial and temporal crime patterns with greater precision by integrating information from multiple sources such as video surveillance, police records, and geospatial data, thereby strengthening predictive capacity. From a methodological perspective, the articulation between AI and crime prediction is supported by mechanisms such as deep neural networks, reinforcement learning, agent based models, and data mining techniques, which facilitate the detection of complex regularities and the anticipation of criminal events.

Theoretically, this relationship is grounded in frameworks such as environmental criminology and routine activity theory, which help explain the spatio temporal dynamics of crime. However, AI based crime prediction faces challenges including data quality and bias, model explainability, ethical implications, and validation in real world contexts. Overall, the convergence between artificial intelligence and crime prediction promotes more proactive and evidence based predictive paradigms.

3 Review Method

A systematic literature review, SLR, approach was employed based on the guidelines proposed by Kitchenham [58] and the Preferred Reporting Items for Systematic Reviews and Meta Analysis, PRISMA 2020 guidelines [69] (see Figure 1). Through this approach, the study aims to conduct a comprehensive review of the use of artificial intelligence in crime prediction in order to generate precise responses to the proposed research questions and obtain an integrated and comparative view of trends in the use of Artificial Intelligence in Crime Prediction. The purpose of applying an SLR in this study is to examine in a structured manner the application of artificial intelligence to crime prediction in urban environments. Through this process, the study seeks to synthesize prior contributions, identify gaps between current and previous research, and construct a clear conceptual framework to guide future investigations. The results obtained from the

ISSN 2007-9737

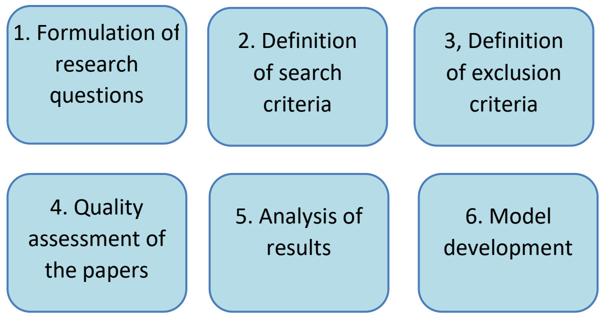

Fig. 1. Stages of the SLR

Table 1. Research Questions and Objectives

Question	Objective
RQ1: Which programming languages are most frequently used to develop Artificial Intelligence solutions?	Evaluate the programming languages employed to develop artificial intelligence systems.
RQ2: Which indicators or metrics are predominantly used to evaluate the performance of Artificial Intelligence based systems applied to Crime Prediction?	Determine the predominant indicators used to evaluate the performance of Artificial Intelligence based systems applied to Crime Prediction.
RQ3: At which quartile levels are scientific publications concentrated regarding the use of artificial intelligence in Crime Prediction?	Analyze the quartile levels of scientific journals where studies on the use of artificial intelligence in crime prediction are published.
RQ4: Which authors present higher levels of coauthorship in research related to Artificial Intelligence and its influence on Crime Prediction?	Identify researchers who frequently collaborate in studies on artificial intelligence and its influence on crime prediction.
RQ5: Which Thematic Categories characterize research addressing Artificial Intelligence and its impact on Crime Prediction?	Analyze the types of thematic categories under which research on the use of artificial intelligence in crime prediction has been developed

rigorous review and evaluation of information sources demonstrate consistency and coherence, reinforcing the reliability of the process. This comprehensive assessment, based on strict quality criteria, ensures that only sources meeting the highest standards are considered. In this way, it is guaranteed that the selected information is relevant, accurate, and appropriate to support the objectives of the present review.

3.1 Problems and Objectives

The following table presents the fundamental questions addressing the use of artificial intelligence in crime prediction in urban contexts. This study has been structured into two essential components, which are clearly detailed in Table 1.

3.2 Information Sources and Search Strategies

For the development of this research, five high quality academic information sources were used: ProQuest, IEEE Xplore, Scopus, ScienceDirect, and SpringerLink. In addition, a table of descriptors and their grouped synonyms was developed in order to conduct systematic searches in each information source, as shown in Table 2. It should be noted that the database search process was conducted through July 22, 2025, thereby establishing the temporal cutoff of the study.

Table 2. Search descriptors and their synonyms by conceptual group

Conceptual Group	Descriptor
Artificial Intelligence	artificial intelligence/ computational intelligence/ intelligent systems/ smart machines/ machine intelligence/ intelligent processing/ artificial cognitive systems/ intelligent algorithms/ artificial neural networks
Crime Prediction	crime forecasting/ crime detection / crime prediction/ crime anticipation/ crime prevention/ criminal forecasting/ crime analysis/ criminal behavior modeling/ predictive crime analysis

Table 3. Information Sources and Search Equations

Information Source	Search Equation
ProQuest	TI ("artificial intelligence" OR "computational intelligence" OR "intelligent systems" OR "smart machines" OR "machine intelligence" OR "intelligent processing" OR "artificial cognitive systems" OR "intelligent algorithms" OR "artificial neural networks") AND AB ("crime forecasting" OR "crime detection" OR "crime prediction" OR "crime anticipation" OR "crime prevention" OR "criminal forecasting" OR "crime analysis" OR "criminal behavior modeling" OR "predictive crime analysis")
IEEE Xplore	("artificial intelligence" OR "computational intelligence" OR "intelligent systems" OR "smart machines" OR "machine intelligence" OR "intelligent processing" OR "artificial cognitive systems" OR "intelligent algorithms" OR "artificial neural networks") AND ("crime forecasting" OR "crime detection" OR "crime prediction" OR "crime anticipation" OR "crime prevention" OR "criminal forecasting" OR "crime analysis" OR "criminal behavior modeling" OR "predictive crime analysis")
Scopus	TITLE-ABS-KEY("artificial intelligence" OR "computational intelligence" OR "intelligent systems" OR "smart machines" OR "machine intelligence" OR "intelligent processing" OR "artificial cognitive systems" OR "intelligent algorithms" OR "artificial neural networks") AND TITLE-ABS-KEY("crime forecasting" OR "crime detection" OR "crime prediction" OR "crime anticipation" OR "crime prevention" OR "criminal forecasting" OR "crime analysis" OR "criminal behavior modeling" OR "predictive crime analysis")
ScienceDirect	TITLE-ABSTR-KEY("artificial intelligence" OR "computational intelligence" OR "intelligent systems" OR "smart machines" OR "machine intelligence" OR "intelligent processing" OR "artificial cognitive systems" OR "intelligent algorithms" OR "artificial neural networks") AND TITLE-ABSTR-KEY("crime forecasting" OR "crime detection" OR "crime prediction" OR "crime anticipation" OR "crime prevention" OR "criminal forecasting" OR "crime analysis" OR "criminal behavior modeling" OR "predictive crime analysis")
SpringerLink	("artificial intelligence" OR "computational intelligence" OR "intelligent systems" OR "smart machines" OR "machine intelligence" OR "intelligent processing" OR "artificial cognitive systems" OR "intelligent algorithms" OR "artificial neural networks") AND ("crime forecasting" OR "crime detection" OR "crime prediction" OR "crime anticipation" OR "crime prevention" OR "criminal forecasting" OR "crime analysis" OR "criminal behavior modeling" OR "predictive crime analysis")

Furthermore, specific search equations were developed for each information source, using the English descriptors and their respective synonyms as a basis in order to broaden the scope of retrieval and identify relevant information more

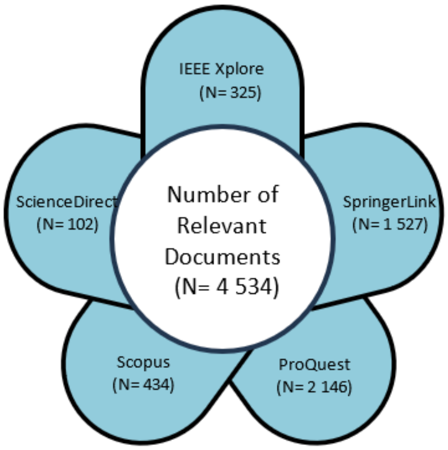

Fig. 2. Consolidated Chart by Number of Source

Table 4. Exclusion Criteria applied in the study selection process

Criterion	Description
EC_1	Documents classified as books, book chapters, theses, systematic review papers, or bibliometric reviews.
EC_2	The papers are more than 7 years old.
EC_3	They are not written in English.
EC_4	They were not published in peer reviewed conferences or journals.
EC_5	The full text of the paper is not accessible.
EC_6	The titles and keywords of the papers are not appropriate.
EC_7	The papers are not original

comprehensively. These equations are presented in Table 3.

3.3 Identified Studies

After executing the search equations in each information source, a general count of the records initially retrieved was conducted without applying restrictions on language, publication period, or document type, in order to estimate the gross volume of available literature. The quantity obtained from each source is presented in Figure 2.

3.4 Selection Criteria

To ensure the methodological rigor, quality, and relevance of the papers included in this review, a set of well defined Exclusion Criteria, EC, was established in accordance with systematic review standards. In parallel, complementary inclusion considerations were applied to guarantee the suitability of the selected studies for the objectives of this paper. The criteria applied during the screening and eligibility phases are presented in Table 4.

3.5 Selection of Studies

The results obtained through the systematic application of the inclusion and exclusion criteria were visually represented in order to remove non relevant records and ensure transparency in the selection process. This procedure makes it possible to clearly illustrate the stages of identification, screening, eligibility, and inclusion of the analyzed studies. The complete workflow of the process is presented in Figure 3.

3.6 Quality Assessment

To ensure the objectivity and consistency of the process, seven quality assessment criteria, QA1–QA7, were defined in alignment with recent standards in systematic reviews on Machine Learning based prediction [71]. These criteria verify aspects ranging from the clear identification of objectives to the transparent reporting of experimental results. Each paper was independently evaluated by two reviewers, ensuring that only sources meeting the highest standards were considered. Table 5 presents the quality assessment applied to the analyzed studies.

The results indicate that most studies meet the established quality criteria. An inclusion threshold of a total score of at least 24 was defined; therefore, studies scoring below this value were removed from the analysis.

After the refinement process, 57 of the 70 evaluated sources (81.4%) were selected, forming

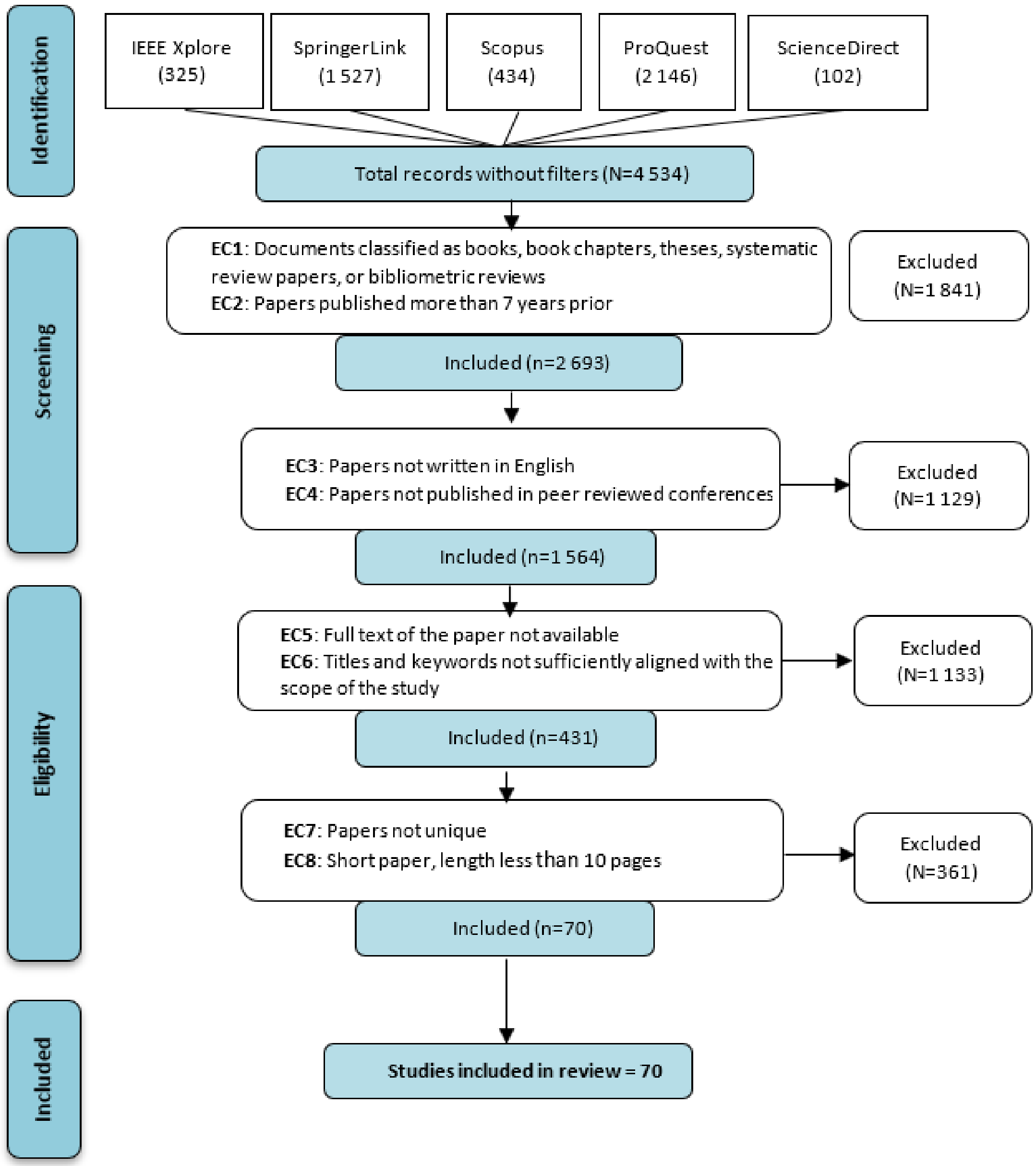

Fig. 3. PRISMA flow diagram

a methodologically robust final corpus for the synthesis of findings.

Scores were assigned independently by two reviewers, and in cases of discrepancy, disagreements were resolved through the intervention of a third reviewer, thereby ensuring the reliability of the evaluation process.

3.7 Data Extraction Strategy

After conducting a thorough and detailed reading of the information sources, the collected records were systematically organized using a specialized bibliographic management tool in order to establish

ISSN 2007-9737

Table 5. Quality Assessment Method

Ref.	Type	QA1	QA2	QA3	QA4	QA5	QA6	QA7	Score
[1]	Journal	4	4	4	4	5	5	4	30
[2]	Journal	3	2	3	3	1	3	3	18
[3]	Journal	5	5	5	5	5	5	5	35
[4]	Journal	4	2	4	4	4	1	4	23
[5]	Journal	5	5	5	5	4	4	5	33
[6]	Journal	3	5	3	3	4	4	3	25
[7]	Journal	1	5	1	5	5	1	5	23
[8]	Journal	3	3	3	3	5	5	3	25
[9]	Journal	1	2	5	2	3	3	5	21
[10]	Journal	4	4	4	4	5	5	4	30
[11]	Journal	1	4	3	1	4	3	5	21
[12]	Journal	3	3	4	3	4	4	3	24
[13]	Journal	5	5	4	5	3	5	5	32
[14]	Journal	5	4	5	5	5	4	5	33
[15]	Journal	3	5	3	3	1	4	3	22
[16]	Journal	5	3	5	5	5	3	5	31
[17]	Journal	1	5	4	1	3	5	1	20
[18]	Journal	4	3	4	4	5	5	4	29
[19]	Journal	3	5	3	3	4	3	3	24
[20]	Journal	5	4	5	5	4	5	5	33
[21]	Journal	4	1	4	4	2	4	4	23
[22]	Journal	5	4	5	5	4	5	5	33
[23]	Journal	3	5	3	3	5	3	3	25
[24]	Journal	5	3	5	5	3	5	5	31
[25]	Journal	3	5	3	3	5	3	3	25
[26]	Journal	5	3	5	5	3	5	5	31
[27]	Journal	4	2	4	4	1	4	4	23
[28]	Journal	4	4	5	4	5	4	4	30
[29]	Journal	3	5	4	5	4	5	5	31
[30]	Journal	5	3	5	3	5	3	3	27
[31]	Journal	4	4	5	5	3	5	5	31
[32]	Journal	5	5	4	3	5	3	3	28
[33]	Journal	3	3	5	5	3	5	5	29
[34]	Journal	5	5	3	1	3	2	4	23
[35]	Dataset	3	3	5	4	5	4	4	28

a structured and hierarchical approach to file handling. For this purpose, Mendeley Desktop was used, facilitating the efficient and organized management of the academic papers included in the review.

Additionally, within the framework of this study, analytical graphs were incorporated into the results and discussion sections, generated through the use of artificial intelligence, thereby optimizing the visualization and comprehension of the data. Furthermore, the research assistant RAj, developed by Dr. Javier Gamboa Cruzado, was employed to support the collection and processing of relevant information, strengthening the quality and depth of the analysis conducted.

[36]	Journal	5	5	3	5	4	5	5	32
[37]	Journal	4	4	5	3	5	3	3	27
[38]	Journal	4	5	4	5	3	5	5	31
[39]	Journal	3	4	3	3	5	3	3	24
[40]	Journal	5	3	2	5	3	1	1	20
[41]	Journal	4	5	4	4	5	4	4	30
[42]	Journal	5	4	5	5	4	5	5	33
[43]	Journal	3	5	3	3	5	3	3	25
[44]	Journal	5	3	5	5	3	5	5	31
[45]	Conference	3	5	3	3	1	3	3	21
[46]	Journal	5	3	1	5	3	1	4	22
[47]	Journal	4	5	4	4	5	4	4	30
[48]	Journal	5	4	5	5	4	5	5	33
[49]	Journal	3	4	5	3	5	3	3	26
[50]	Journal	5	5	4	5	5	4	5	33
[51]	Journal	4	3	5	3	3	5	4	27
[52]	Journal	4	5	3	5	5	3	5	30
[53]	Journal	3	3	5	3	3	5	3	25
[54]	Journal	5	5	3	5	5	3	5	31
[55]	Journal	4	4	5	4	4	4	5	30
[56]	Journal	3	5	4	5	5	5	4	31
[57]	Journal	5	4	5	4	3	3	5	29
[58]	Journal	4	5	4	5	5	5	3	31
[59]	Journal	5	3	5	3	3	5	4	28
[60]	Journal	3	5	3	5	5	3	3	27
[61]	Journal	5	3	5	3	4	5	5	30
[62]	Journal	3	5	3	5	5	4	4	29
[63]	Journal	5	4	5	4	4	5	5	32
[64]	Journal	4	5	4	5	5	4	3	30
[65]	Journal	4	3	5	3	3	5	5	28
[66]	Journal	4	5	3	5	5	3	3	28
[67]	Journal	3	3	5	3	3	5	5	27
[68]	Journal	5	5	3	5	5	3	4	30
[69]	Journal	4	4	5	4	4	5	5	31
[70]	Journal	4	5	4	5	5	4	4	31

ISSN 2007-9737

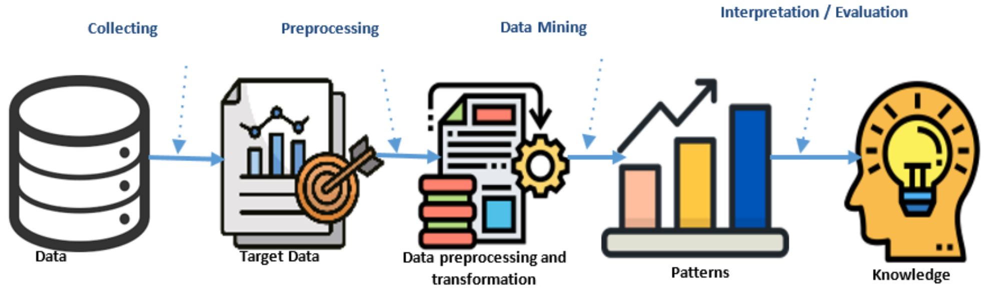

Fig. 4. Processing of Selected Documents

Table 6. Synthesis of Findings by Methodological Category

Method Category	Performance	Key Contributions	Limitations	Emerging Trends	Refs.	Qty. (%)
Agent-Based and Simulation	Accuracy: high predictive power (street-segment crime counts)	Large-scale data-layer integration; Agent decision-making with ML; High-resolution crime pattern simulation	Does not capture individual offender behavioral variability	Agent-based simulation increasingly combined with GIS and ML for fine-grained urban crime pattern modeling	[1]	1 (1.8)
Computer Vision and Surveillance	Accuracy: 82.71%–98.85%; F1-score: 0.607–94.3%; Recall: 90.5%; mAP: 92.3%–98.04%; Latency: 101 ms/image; Average Precision: +10.2% (vs YOLOv8-small)	Real-time surveillance pipelines for crime/violence/weapon detection; Enhanced recognition under challenging conditions (e.g., super-resolution; pose invariance); Edge-deployable models; First AI approach for Robbery Behavior Potential prediction via video cues	Sensitivity to video quality, angles, lighting and unseen environments; Dataset limitations and annotation bias; Limited generalizability beyond evaluated contexts; NR metrics in some studies	Growth of edge-ready CV (Jetson-class) and super-resolution to boost detection; Increasing fusion of tracking (DeepSORT) with detection for behavior risk estimation; Broader use of prompt-based recognition and robustness techniques for adverse surveillance conditions	[3] [10] [17] [22] [24] [33] [34] [40] [43] [46] [49] [56]	12 (21.1)
Deep Learning for Spatio-temporal Prediction	MAE: 0.008–0.02; R²: 0.94–0.95; SMAPE: 0.6%–1.03%; Accuracy: 75%; Recall: strong; False Positives: 4,038	Multi-module feature/decision fusion architectures for cross-city crime prediction; Multimodal fusion of social signals and historical crime data; Deep learning with interpretability/clustering for offender profiling	City-specific training may limit transferability; Dataset/domain dependence; Ethical and error-cost concerns (false positives) and need for context-aware validation; NR in some setups	Increasing multimodal fusion (social + historical) for improved prediction; Wider use of attention-based sequence models for spatio-temporal crime dynamics; Early movement toward interpretable deep models in criminal-justice settings	[4] [38] [54]	3 (5.3)

Ensemble and Classical Supervised Learning	Accuracy: 34.0%–99.5%; F1-score: NR; Log Loss: 1.74; Accuracy: 0.4262; Precision: 90.2%; Recall: 80.1%	Demonstrated competitive baselines and comparative evaluations across classifiers; Identified best-performing models per dataset (e.g., LightGBM vs RF/LR); High-accuracy ensemble stacking for specific crime settings; ML-enabled crime linkage using predictive features	Performance highly dataset- and city-dependent; Lower accuracies on multi-class crime-type prediction; Limited generalization across crime categories/regions; Some metrics not reported	Continued reliance on strong tabular baselines (GBMs and RF) for large historical crime logs; Increased emphasis on comparative benchmarking across classical ML families; Use of supervised ML for linkage tasks beyond forecasting	[9] [19] [21] [32] [55]	
Geospatial and Spatial Statistical Modeling	Accuracy: 92.5%–95.2%; mAP: 92.5%; MAPE: 29.3%; R^2: 0.8; RMSE: 15.3; Hotspots identified (NR metrics)	High-resolution crime probability/risk mapping; Quantified uncertainty and addressed data scarcity via Bayesian spatial modeling; Automated hotspot detection using clustering; Practical geospatial workflows for spatio-temporal crime pattern analysis	Limited to specific cities/contexts; Data quality issues (e.g., handwritten records; reporting bias); Generalizability constraints to non-urban/rural areas; NR metrics in some studies	Stronger coupling of explainability (e.g., SHAP) with spatial risk maps; Greater use of Bayesian/spatial models to represent uncertainty under data scarcity; Expansion of clustering + spatial regression pipelines for hotspot discovery	[18] [29] [37] [47] [53]	
LLMs and NLP for Crime Intelligence	Accuracy: 94.5%; F1 Score: 92.3%; F1 score: 86%	Integrated LLMs to enhance crime prediction capabilities; Demonstrated cross-lingual text-based monitoring aligned with official statistics	Limited exploration of ethical implications for LLM policing; Language/domain constraints (Thai-only) and transfer risks; NR for some implementation details	Early shift toward LLM-enabled classification/prediction with prompt/fine-tune pipelines; Growing use of cross-lingual NLP for near–real-time crime intelligence from open text sources	[2] [16]	2 (3.5)
Multi-Criteria Decision-Making	Statistically validated via case study and sensitivity analysis; NR accuracy metrics	Decision-support frameworks for selection under uncertainty; Integrated MCDM pipelines (CoCoSo + AHP + SVNSs) for complex choices; CPTED element selection model for municipal decision-making	Limited applicability to crime prediction outcomes; Limited external validity beyond case contexts; Often lacks domain-specific social influences; NR on reproducibility	Increased use of uncertainty-aware MCDM to support security-related operational decisions; Growing adoption of fuzzy/neutrosophic extensions to handle ambiguity in evaluation criteria	[7] [8] [11] [20] [26] [27] [28] [36] [45] [57]	10 (17.5)

ISSN 2007-9737

Reviews, Conceptual, Ethics and Regulation	Felony violence reduction: 15%; Homicide reduction: 10%; Fairness Score: 4.41/7; Accuracy: NR (several reviews); Duration: NR (no significant effect reported)	Synthesized challenges/opportunities and proposed taxonomies; Advanced conceptual framing on democratic norms, bias and governance; Empirical evidence on operational impacts (e.g., facial recognition outcomes; report-writing trials); Identified drivers/barriers for AI-enabled system adoption	Limited empirical validation in several reviews/conceptual works; Sparse or NR performance reporting; Context-specific findings; Data scarcity and bias concerns persist	Growing emphasis on explainability, trustworthiness and governance alongside technical modeling; Increased empirical policy-evaluation designs (DiD/RCT) to assess real-world impacts; Expansion of taxonomies and horizon scanning for emerging AI-enabled crime threats	[6] [12] [13] [14] [15] [23] [31] [35] [39] [41] [44] [50] [51] [52]	14 (24.6)
Time Series and Forecasting Models	MAE: 11.47 (vs DBSCAN MAE: 27.03); RMSE: NR (adequate); MAE: NR (adequate); Accuracy: 72.5%; Correctness: 81.7%	Forecasting trends and hotspots using statistical/sequence models; MI-HMM-MAP model outperforming classical baselines; Improved clustering-informed forecasting with lower MAE	Limited to specific cities and crime types; Some metrics reported qualitatively (adequate/NR); Transferability constraints across regions and contexts	Continued combination of clustering with classical forecasting (SARIMA) for hotspot-aware prediction; Persistence of HMM-style models for density/hotspot estimation; Ongoing use of hybrid time-series toolkits (LSTM + ARIMA) for crime trend forecasting	[25] [42] [48]	3 (5.3)

3.1 Synthesis of Findings

The review of the literature on the use of artificial intelligence in crime prediction reveals relevant findings that contribute to understanding the impact of AI based systems on strengthening public security. Among the most notable results is a growing trend in the adoption of Machine Learning algorithms, particularly neural network and deep learning models, aimed at predicting crime patterns and optimizing preventive strategies. Likewise, a positive correlation is identified between the incorporation of advanced technologies and increased prediction accuracy, reinforcing the effectiveness of intelligent solutions in the field of urban security.

4 Results and Discussion

This section presents the results obtained during the study, followed by a detailed analysis thereof. These results are contextualized through a comparison with the existing literature, enabling the identification of both convergences and discrepancies. In addition, they are discussed in relation to the objectives established at the beginning of the research, assessing their fulfillment and relevance within the proposed theoretical framework. For the processing of unstructured text, which constitutes part of this study, the stages illustrated in Figure 4 were followed.

4.1 General Description of the Studies

Table 6 synthesizes the reviewed literature by methodological category, enabling a structured comparison of performance, contributions, limitations, and emerging trends in AI-based crime prediction. This taxonomy-driven aggregation supports a clearer understanding of how the field is evolving across analytical paradigms. Interpreting the distribution of studies (Qty. %)

provides evidence of methodological concentration and maturity patterns within the domain.

Method Category: The distribution reveals a strong concentration in Reviews, Conceptual, Ethics and Regulation (24.6%) and Computer Vision and Surveillance (21.1%), indicating that the field is simultaneously expanding in practical surveillance deployments and in critical governance reflection. The prominence of vision-based approaches is causally linked to the rapid proliferation of CCTV infrastructure and the maturity of deep learning for visual analytics. Conversely, the relatively small share of Agent-Based and Simulation (1.8%) and Reinforcement Learning and Metaheuristics (3.5%) suggests that adaptive and simulation-driven paradigms remain underexplored, likely due to higher modeling complexity and data requirements. Overall, the taxonomy evidences a field still anchored in detection-centric pipelines rather than fully predictive or prescriptive intelligence systems.

Performance: Reported metrics show high peak accuracies in several categories (often >90%), particularly in Computer Vision and Deep Learning pipelines, yet these results are highly heterogeneous and context-dependent. Performance figures typically originate from single-city datasets or controlled environments, which causally explains the wide metric variability and the absence of pooled evaluation. Categories such as Ensemble Learning display broad accuracy ranges (34.0%–99.5%), evidencing sensitivity to class imbalance and dataset structure. The persistence of NR metrics in multiple studies further limits comparability. This fragmentation indicates that the field prioritizes model optimization within narrow contexts rather than standardized benchmarking across jurisdictions.

Key Contributions: Across categories, contributions converge on three dominant axes: (i) improved predictive accuracy through hybrid and multimodal models, (ii) operational support for policing via hotspot mapping and surveillance automation, and (iii) growing attention to explainability and governance. The strong presence of real-time surveillance contributions is causally associated with the operational demands of smart-city ecosystems and public safety investments. Meanwhile, LLM-based studies, although still limited (3.5%), signal a shift toward text-driven intelligence and cross-lingual monitoring. However, relatively few works advance fully integrated decision-support ecosystems, indicating that most research remains component-level rather than system-level.

Limitations: Recurring constraints include limited generalizability, dataset bias, city-specific validation, and incomplete reporting of metrics. These limitations are structurally linked to the heavy reliance on localized crime datasets and proprietary surveillance data. Vision-based systems additionally suffer from environmental sensitivity (lighting, angle, occlusion), while conceptual studies lack empirical grounding. The persistence of NR fields suggests reporting inconsistency across the literature. Collectively, these weaknesses indicate a maturity gap between high-performing prototypes and scalable, transferable crime prediction systems suitable for heterogeneous urban contexts.

Emerging Trends: Several convergent trajectories are evident: (i) fusion of multimodal data (video, text, geospatial), (ii) increasing integration of explainable AI in risk mapping, (iii) edge-ready and real-time surveillance pipelines, and (iv) early adoption of LLMs for crime intelligence. These trends are driven by three causal forces: the expansion of urban sensing infrastructure, advances in deep learning efficiency, and growing regulatory pressure for transparency. Notably, the modest presence of RL and simulation suggests that fully adaptive predictive policing systems remain an emerging frontier rather than a consolidated paradigm.

Refs: The reference distribution demonstrates broad thematic coverage, with the largest clusters associated with surveillance and governance-oriented studies. High-density reference groups in Computer Vision (12 studies) and Reviews/Conceptual work (14 studies) reinforce the dual trajectory of technical intensification and ethical scrutiny. Smaller clusters (e.g., RL and LLMs) confirm that some advanced paradigms are still in early diffusion stages. The traceability across references supports the internal consistency of taxonomy while highlighting uneven research maturity across methodological families.

Qty. (%): The proportional distribution confirms a field dominated by descriptive/analytical reviews (24.6%) and surveillance-driven implementations

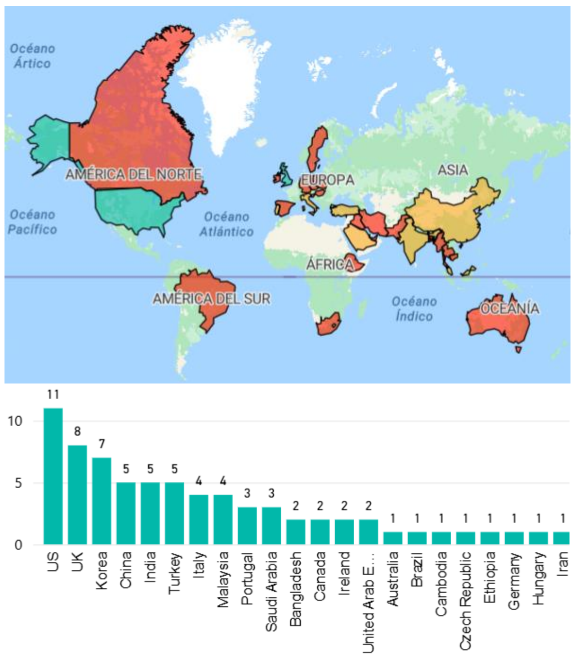

Fig. 5. Distribution of Publications by Country

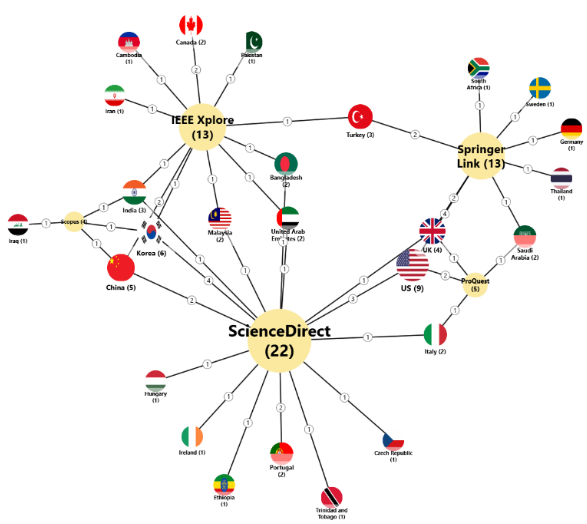

Fig. 6. Distribution of Publications by Source

(21.1%), followed by MCDM decision frameworks (17.5%). Predictive deep learning, classical ML, and geospatial modeling each occupy mid-tier shares (~5–9%), while RL, LLM-based intelligence, and simulation remain marginal (≤3.5%). This asymmetry suggests that technological deployment, especially vision-based detection, has outpaced the development of adaptive, learning-centric, and theoretically grounded predictive ecosystems. The pattern is consistent with funding and infrastructure realities in smart-city security, where immediate detection capabilities are prioritized over long-horizon predictive intelligence.

The evidence indicates that AI for crime prediction is transitioning from detection-centric systems toward more integrated, multimodal intelligence frameworks, but the shift remains incomplete. The dominance of surveillance pipelines suggests strong operational demand but also raises scalability and governance challenges. Methodological fragmentation and inconsistent reporting hinder meta-analytic synthesis and cross-city generalization. Future research should prioritize standardized benchmarks, multimodal fusion with explainability, and adaptive learning paradigms (e.g., RL and simulation). Strengthening these areas is essential for moving from high-performing prototypes to robust, policy-aligned predictive policing ecosystems suitable for real-world deployment.

Figures 5, 6 and Table 7 make it possible to characterize the geographical footprint of the evidence, identifying the countries with the highest publication output, and, in parallel, the dependence on publication sources, indicating where studies are indexed or published and how they are distributed by country. This provides contextual insight into coverage bias and scientific visibility. Taken together, they support a critical interpretation of regional concentration, dissemination channels, and the maturity of the field.

As shown in the results, production is concentrated in a limited number of countries, with the United States (11), the United Kingdom (8), and South Korea (7) leading, suggesting combined effects of installed capacity, including laboratories, data availability, and funding, together with broader access to digitized urban infrastructures such as CCTV systems, police records, and open data platforms.

According to the findings presented, the Asia–Europe block increases its relative weight, with China and India (5 each), Türkiye (5), and Italy and Malaysia (4 each). This pattern is consistent with

ISSN 2007-9737

the expansion of smart city initiatives and the accelerated adoption of advanced analytics in public security, although institutional heterogeneity across countries remains evident.

Based on the observed distribution, country level counts appear to reflect non-exclusive participation, likely due to coauthorship and multiple institutional affiliations. This is consistent with a domain in which datasets and methodological frameworks are shared across international research groups, yet with persistent leadership from economies exhibiting higher scientific output.

Regarding publication sources, ScienceDirect concentrates the largest number of papers (22/57), whereas IEEE Xplore, with fewer papers (13/57), shows higher relative impact, with 334 citations, 26 citations per paper, and 73 citations per year. This suggests that highly cited contributions are often disseminated through technically specialized channels with denser scholarly communities.

Finally, the source–country network reveals the centrality of ScienceDirect as a multinational hub and a segmentation by channel, with IEEE and Springer exhibiting more specific nodes. This pattern points to editorial and indexing mechanisms that condition visibility, citation rates, and replicability, for example through access to datasets, evaluation standards, and subfield maturity.

According to Simisterra-Batallas, Pico-Valencia, Sayago-Heredia, and Quiñónez-Ku [60], India positions itself as the leading country by concentrating 18 sources used in their study, indicating a particularly robust documentary base in that context. In contrast, Tomaz, Fernandes, and Sciutti [62] emphasize an increase in publications in Central and South America, which, although not displacing Asian leadership, suggests a geographical expansion of scientific interest and, consequently, a diversification of research agendas related to machine learning and crime prediction. Despite these differences in emphasis, Dakalbab, Nasir, Bou Nassif, and Abbas [63] concur that the United States and India constitute the principal poles of concentration of studies. Therefore, the development of the field appears driven by ecosystems with high research capacity and data availability. Overall, the comparative findings indicate a dominant core, United States and India, and, although still emerging, a regional growth dynamic in Latin America that deserves attention due to its potential to expand the evidence base across diverse socioterritorial contexts.

Geographical and source concentration suggests that the generalization of results toward Latin America, Africa, and intermediate cities requires multi jurisdiction validation designs and transfer protocols, including comparable metrics, bias assessment, and fairness criteria. These considerations are also applicable to sectors such as transportation, retail, banking and insurance, and logistics, where incident prediction, fraud detection, and operational risk assessment share similar methodological challenges.

To expand impact across other geographical areas and time periods, the incorporation of longitudinal evidence, for example before and after regulatory or technological changes, is recommended, along with comparisons between scenarios with different levels of infrastructure, such as dense CCTV coverage versus limited data environments, thereby reducing smart city centric bias.

The difference between volume, concentrated in ScienceDirect, and relative impact, more visible in IEEE, suggests dual strategies: consolidating broad syntheses in repositories with wider coverage while deepening technically reproducible contributions, including code, datasets, and benchmarks, to accelerate scientific maturity and cross sectoral transfer over time.

Figure 7 and Table 8 present the temporal evolution of scientific production in artificial intelligence applied to crime prediction, integrating publication volume and impact metrics. Their analysis enables the identification of the field’s degree of maturity, as well as recent growth dynamics and academic influence.

As shown in the results, there is a marked acceleration in publications beginning in 2023, with a peak in 2024 (19; 33.3%) and a sustained high level in 2025 (15; 26.3%). This pattern is consistent with the recent mainstreaming of advanced analytics, including deep learning, computer vision, and more recently LLMs, as well as with the growing availability of urban data and computational platforms.

ISSN 2007-9737

Table 7. Distribution of Publications by Source and Country

Source	N° Papers	N° Citations	H-Index	Citations / Year	Citations / Paper
ScienceDirect	22	88	2341	27	4
IEEE Xplore	13	334	3770	73	26
Springer Link	13	162	696	39	12
ProQuest	5	12	378	5	2
Scopus	4	6	353	3	2
Total	**57**	**602**	**7538**	**147**	**11**

Table 8. Annual Distribution of Scientific Production and Bibliometric Impact

Year	N° Papers	% Papers	N° Citations	% Citations	H-Index	% H-Index	Citations/ Paper
2024	19	33,3%	30	5,0%	1886	25,0%	1.6
2025	15	26,3%	9	1,5%	1969	26,1%	0.6
2023	9	15,8%	82	13,6%	1360	18,0%	9.1
2021	7	12,3%	362	60,1%	1326	17,6%	51.7
2022	5	8,8%	56	9,3%	417	5,5%	11.2
2019	2	3,5%	63	10,5%	580	7,7%	31.5
Total	57	100,0%	602	100,0%	7538	100,0%	10.6

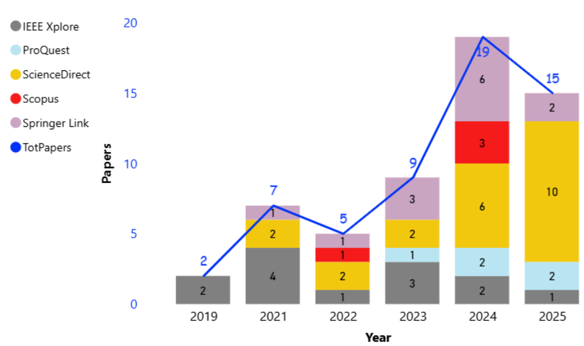

Fig. 7. Distribution of Papers by Year

According to the findings presented, the increase in papers does not translate proportionally into citations. The year 2024 accounts for 33.3% of papers but only 5.0% of citations, with 1.6 citations per paper, while 2025 represents 26.3% of papers with 1.5% of citations, or 0.6 citations per paper. This pattern can be explained by the citation window effect, referring to the maturation time required for citations to accumulate, and by the high turnover of topics and benchmarks within the field.

Based on the observed distribution, 2021 dominates in citations (60.1%) and presents the highest citations per paper ratio (51.7), suggesting that a limited set of early papers functioned as methodological or dataset foundations widely reused, amplifying their centrality through citation dependency.

Furthermore, the annual H index trajectory, from 580 in 2019 to 1886 in 2024 and 1969 in 2025, indicates consolidation of the field. However, the recent concentration of contributions to the H index in 2024 and 2025, approximately 25 to 26 percent, points to rapid expansion with impact that is still in the process of stabilization.

Finally, the pattern by source suggests diversification of dissemination channels in 2024 and 2025, with greater presence in ScienceDirect and contributions from IEEE and Springer. This may reflect convergence between applied research, including implementations and systems, and engineering contributions related to models, evaluation, and deployment, although reporting standards remain non homogeneous.

According to Mandalapu, Elluri, Vyas, and Roy [59], research on machine learning based crime prediction shows accelerated growth beginning in 2018, reaching a peak in 2021. This behavior suggests a phase of rapid technological adoption driven by the maturation of machine learning techniques and greater data availability. In contrast, Simisterra-Batallas, Pico-Valencia, Sayago-Heredia, and Quiñónez-Ku [60], within their systematic review, report the use of 54 sources and identify 2024 as the year with the highest production, 15 papers, which is relevant because it indicates continuity in research dynamism beyond the previously identified peak. Nevertheless, Butt, Letchmunan, Hassan, Ali, Baqir, and Sherazi [61] attribute part of this growth to the expansion of public and private databases, with predominance of public datasets between 2018 and 2019. Therefore, increased data accessibility appears to have acted as a catalyst

ISSN 2007-9737

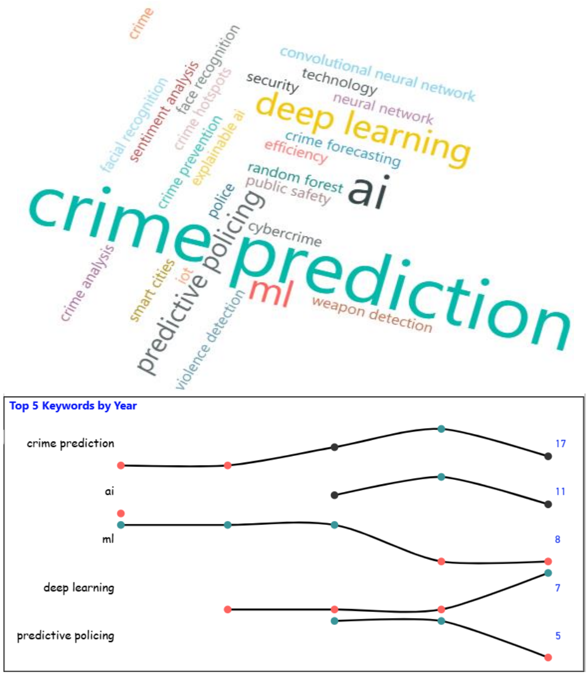

Fig. 8. Top 5 Keywords

Table 9. Distribution of Keywords by Journal Quartile

Keyword	Q1	Q2	Q3	Total
crime prediction	12	2	3	17
ai	6	3	2	11
ml	3	2	3	8
deep learning	4	2	1	7
predictive policing	1	4	0	5
convolutional neural network	2	0	0	2
crime	1	0	1	2
crime analysis	1	1	0	2
crime forecasting	1	0	1	2
crime hotspots	1	1	0	2
crime prevention	1	1	0	2
cybercrime	1	1	0	2
efficiency	0	1	1	2
explainable ai	1	1	0	2
face recognition	1	1	0	2
facial recognition	1	1	0	2
Total	174	88	52	314

for new research. Overall, Tomaz, Fernandes, and Sciutti [62] reinforce this trend by identifying another peak in 2023 for studies on Artificial Intelligence applied to crime prediction. In summary, the comparative evidence suggests sustained but non linear growth, conditioned by data availability, methodological maturity, and institutional interest in predictive solutions.

Recent growth makes it necessary to prioritize comparability criteria, including datasets, metrics, and validation protocols, so that advances are transferable not only to public security but also to financial fraud detection, cybersecurity, operational continuity, retail, and logistics, where risk is likewise spatio temporal and multimodal.

To extend utility to other geographical regions, such as Latin America and Africa, and to different time periods, including before and after regulatory or technological changes, longitudinal and multi city studies are required to measure robustness under distribution shifts, bias, and unequal infrastructure availability, such as CCTV density and open data access.

Given the lag between publication volume and citation accumulation, it is reasonable to expect that 2024 and 2025 will increase their impact in the coming years if they produce reusable artifacts, including code, datasets, and benchmarks, and align with governance frameworks and ethical evaluation standards, thereby reducing the risk of non generalizable or non adoptable results.

Figure 8 and Table 9 synthesize the thematic agenda of the field through the frequency and annual evolution of keywords, enabling the identification of dominant focal areas and conceptual shifts. Their joint interpretation provides evidence regarding which approaches are consolidating and which are emerging, while also signaling potential indexing biases and issues related to terminological standardization.

As shown in the results, crime prediction clearly dominates (Total = 17), followed by AI (11) and ML (8), reflecting a pragmatic orientation toward predictive models and a still generic taxonomy of terms. This pattern is typical of expanding fields in which emphasis is placed on what is done, predicting, rather than on how it is validated or in which context it is deployed.

According to the findings presented, the time series suggests increasing consolidation of crime

Table 10. Programming Languages

Programming language	Reference	Qty. (%)
Python	[1] [13] [19] [21] [25] [33] [35] [38] [48]	9 (60)
R	[12] [16] [31]	3 (20)
Java	[17]	1 (6.67)
C++	[26]	1 (6.67)
Julia	[9]	1 (6.67)

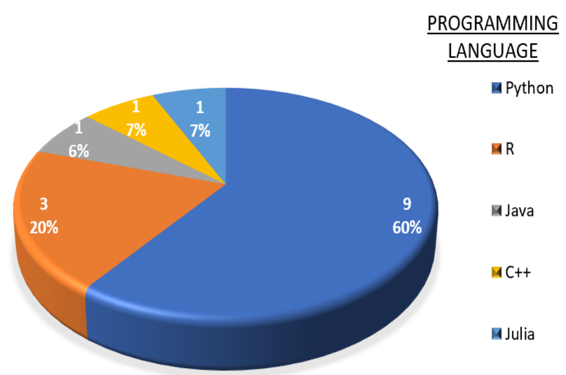

Fig. 9. Programming Languages

prediction and AI, while ML maintains moderate presence. This pattern is consistent with the transition from classical machine learning approaches toward broader Artificial Intelligence frameworks that integrate deep learning, computer vision, and more recently LLMs under a unified conceptual umbrella.

Based on the observed distribution, deep learning (7) has gained relevance in recent years. This may be explained by the availability of larger data volumes, including video, geospatial information, and historical records, as well as improvements in computational capacity, enabling deep architectures with enhanced performance in spatio temporal and multimodal scenarios.

Likewise, predictive policing (5) appears as a persistent yet more circumscribed topic, suggesting that part of the literature avoids this label due to its normative and ethical connotations. Instead, studies may prefer technically neutral terms such as crime forecasting, hotspots, or public safety in order to reduce regulatory and reputational friction.

Finally, the table shows concentration in Q1 to Q3 journals (Total = 314; Q1 = 174), but with semantic dispersion across closely related terms, such as crime forecasting, crime hotspots, crime prevention, and cybercrime, each with relatively low totals. This indicates a lack of terminological standardization and may induce undercoverage in database searches if search equations do not adequately incorporate synonyms and variant expressions.

According to Yaghoubi, Yaghoubi, Khamees, and Vakili [65], the most recurrent keywords in the literature on Artificial Intelligence applied to crime prediction, particularly artificial neural network and artificial intelligence, demonstrate strong thematic concentration on AI based predictive models. It is important to note that this recurrence reflects the maturity and centrality of these approaches within the field.

In contrast, Huamantingo, Cano Lengua, and Rodríguez [64] emphasize the prominence of terms such as crime and machine learning, which, while maintaining thematic coherence, expand the focus toward operational applications of predictive analytics. Collectively, these terminological patterns suggest that research is structured around the convergence between machine learning techniques and security related challenges, thereby consolidating a conceptual core oriented toward data driven crime prediction.

The dominance of general terms such as AI and ML suggests that future reviews and developments should strengthen methodological specificity, including metrics, bias assessment, explainability, and external validation. This recommendation is transferable to other sectors such as banking fraud detection, insurance risk modeling, cybersecurity, retail analytics, and logistics, where prediction likewise depends on heterogeneous signals.

To expand validity across other geographical regions and time periods, the promotion of controlled vocabularies and comparable taxonomies, for example spatio temporal forecasting, hotspot prediction, and surveillance based detection, is advisable, facilitating replication across cities with different infrastructure levels and socioeconomic contexts.

The moderate weight of predictive policing and the rise of deep learning imply that the future agenda must balance performance with governance. More powerful models will require robust auditing frameworks, explainability mechanisms, and impact evaluation procedures, particularly when transferred to sensitive domains and regulatory environments that evolve over time.

4.2 Answers to the Research Problems

Within the framework of this systematic literature review, this section synthesizes and critically analyzes the empirical evidence in order to address the proposed research questions. Based on the cross comparative examination of the selected studies, methodological patterns, technological trends, and persistent gaps in the use of Artificial Intelligence for Crime Prediction are identified. In addition, the predominant approaches, performance metrics, and publication dynamics are contrasted to assess the degree of maturity of the field. Finally, grounded in these findings, theoretical, methodological, and applied implications are derived to guide future research lines and potential implementation scenarios in real world contexts.

4.2.1 RQ1: Which programming languages are most frequently used to develop Artificial Intelligence solutions?

Table 10 and Figure 9 synthesize the programming languages reported in the analyzed studies, serving as an indirect indicator of implementation ecosystems and the reproducibility of Artificial Intelligence based systems.

As shown in the results, Python accounts for the largest proportion of reported implementations (60%), which can be explained by its dominant ecosystem for machine learning and deep learning, the wide availability of libraries, and the low friction associated with prototyping and experimenting with multiple models.

According to the findings presented, R (20%) appears as the second most used option, consistent with its strength in statistical modeling, exploratory analysis, and script based reproducibility, suggesting a relevant presence of quantitative approaches and analytical validations.

Based on the observed distribution, the low frequency of Java, C++, and Julia (approximately 6.7% each) suggests that the field prioritizes rapid development and experimentation over system level performance optimization, industrial deployment, or high performance computing.

Likewise, the limited heterogeneity, reflected in the predominance of two languages, may induce dependence on specific toolkits and pipeline biases, potentially affecting comparability across studies if environments, versions, and reporting practices are not standardized.

Finally, given that only a fraction of studies explicitly declare the programming language used, the distribution may also reflect underreporting of implementation details, which impacts replicability and makes it more difficult to audit the actual robustness of the reported results.

According to Simisterra-Batallas, Pico-Valencia, Sayago-Heredia, and Quiñónez-Ku [60], Python has consolidated itself as the most widely used language for developing Artificial Intelligence solutions, followed by Java. This is particularly relevant because it evidences a preference for an ecosystem with high availability of machine learning and data analysis libraries; therefore, its adoption favors more agile and reproducible development cycles. In contrast, Dakalbab, Abu Talib, Abu Waraga, Bou Nassif, Abbas, and Nasi [63] also emphasize the centrality of Python in their investigations, particularly associated with data processing and analysis, although they identify Weka as the predominant tool, suggesting a methodological emphasis oriented toward integrated data mining environments. Similar findings are reported in other applied domains: [71] identified that Python (32%) leads chatbot development for e commerce, followed by Java (25%), indicating that the choice of Python responds to structural factors, including library ecosystems, agile prototyping, and an active community, rather than to domain specific characteristics. Without prejudice to these differences, collectively these works agree that Python occupies a transversal position in Artificial Intelligence implementation; consequently, the observed divergences are more closely related to

Table 11. Effectiveness Criteria

Criterion	Reference	Qty. (%)
Prediction Accuracy	[3] [7] [8] [13] [16] [17] [23] [25] [26] [28] [33] [35] [48] [54]	14 (45.16)
Model performance	[1] [2] [5] [7] [16] [18] [19] [21] [26] [32] [36] [38] [53] [55]	14 (45.16)
Model Explainability	[17]	1 (3.23)
F1 Score	[8]	1 (3.23)
Area Under Curve	[20]	1 (3.23)

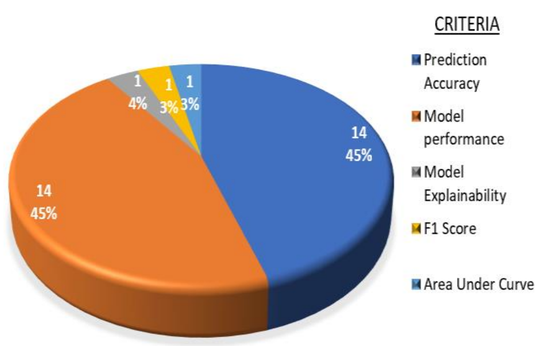

Fig. 10. Effectiveness Criterion

tooling and workflow preferences than to a displacement of Python's role, ultimately reaffirming its status as a de facto standard in the field. Similar evidence is reported in other applied domains: [73] identified that classical algorithms such as Decision Tree (21.25%), K means (20%), and Logistic Regression (17.5%) lead ML implementation in API deployment, suggesting that tool selection responds to factors such as library maturity, ease of prototyping, and documentation availability, rather than to the specificities of the application domain.

The dominance of Python facilitates the replication and transfer of solutions to other sectors, including finance, retail, healthcare, logistics, and cybersecurity, where Artificial Intelligence requires rapid experimentation cycles. However, it also demands the strengthening of engineering practices, including testing, traceability, and MLOps, to ensure production readiness.

To expand applicability across other geographical regions and time periods, it is advisable to promote reusable artifacts, such as code, dependency specifications, and containerized environments, along with deployment guidelines that reduce infrastructure gaps and enable longitudinal comparison of results under changing data and contextual conditions.

Given the minority use of C++, Java, and Julia, future research should explicitly justify language selection when objectives involve low latency, edge computing, or institutional integration, preventing technological decisions from limiting scalability, interoperability, or the long term sustainability of the solution.

4.2.2 RQ2: Which indicators or metrics are predominantly used to evaluate the performance of Artificial Intelligence based systems applied to Crime Prediction?

Table 11 and Figure 10 synthesize the evaluation metrics used to quantify the performance of Artificial Intelligence based systems in Crime Prediction, allowing identification of the criteria that dominate the reported evidence.

This analysis is essential to assess comparability across studies and the degree of methodological maturity within the field.

As shown in the results, Prediction Accuracy and Model Performance account for virtually the entire reported evaluation (45.16% each; 14/31), suggesting a dominant orientation toward global effectiveness metrics and aggregated results rather than toward fine grained diagnostics of model behavior.

According to the findings presented, the preference for these metrics may be explained by their interpretative simplicity and convenience for rapid comparison across approaches, particularly when studies employ different datasets, time horizons, and operational definitions of crime.

Based on the observed distribution, Explainability, F1 Score, and AUC appear marginally (3.23% each; 1/31), a pattern consistent with a field that prioritizes “how well it predicts” over "why it predicts," even though public safety decisions require traceability and bias control.

Likewise, the limited use of F1 and AUC suggests that evaluation does not always align with typical class imbalance scenarios and asymmetric cost structures, such as false positives versus false negatives, in which accuracy alone may be misleading.

Overall, the observed metric fragmentation reflects experimental heterogeneity and possible underreporting, limiting quantitative synthesis and reinforcing the need for standardized evaluation frameworks for Artificial Intelligence based Crime Prediction.

According to Simisterra-Batallas, Pico-Valencia, Sayago-Heredia, and Quiñónez-Ku [60], accuracy is positioned as the central indicator for evaluating predictive models in Crime Prediction, which is relevant because it privileges the system's overall correctness. In a convergent line, Tomaz, Fernandes, and Sciutti [62] also emphasize accuracy, although they incorporate F1 Score as a complementary criterion, responding to the need to balance precision and recall when performance cannot be adequately described by a single metric.

In contrast, Dakalbab, Abu Talib, Abu Waraga, Bou Nassif, Abbas, and Nasi [63] highlight the utility of mean error, particularly under class imbalance scenarios; therefore, they orient evaluation toward metrics sensitive to the magnitude and distribution of error. The predominance of Prediction Accuracy and Model Performance as evaluation criteria in crime prediction reflects a transversal trend in Artificial Intelligence applications. Similar findings are reported in other domains: [74] identified that accuracy, latency, and energy efficiency are the most frequently used metrics in AI systems for precision agriculture, suggesting that the choice of indicators responds to structural factors, including technical effectiveness, resource consumption, and interpretability, rather than to domain specific characteristics.

Despite these differences, collectively the studies suggest that a robust performance assessment requires combining correctness measures with metrics that capture stability and sensitivity to non uniform distributions; consequently, the predominance of accuracy coexists with a gradual shift toward more comprehensive evaluation schemes, ultimately better aligned with real operational conditions.

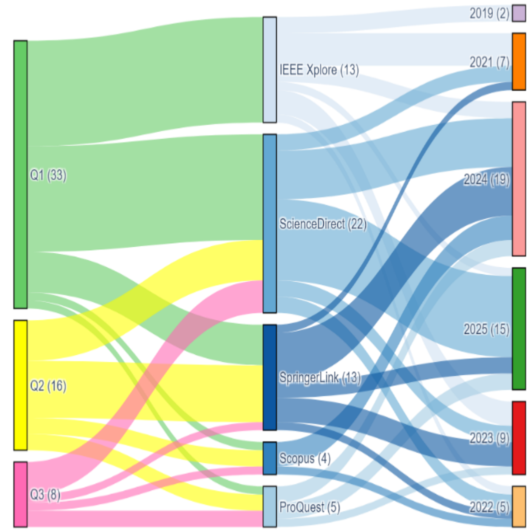

Fig. 11. Distribution of Papers by Quartile, Source, and Year

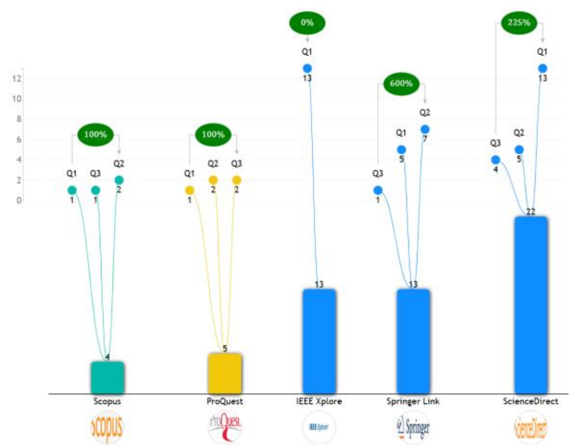

Fig. 12. Papers by Quartile and Source

To improve comparability and transferability to other sectors, including healthcare, financial fraud detection, cybersecurity, and predictive maintenance, it is recommended to complement Accuracy with imbalance robust metrics such as F1, AUC, precision, and recall, along with cost oriented error analysis to strengthen external validity. In other geographical regions and time periods, where crime patterns and data quality vary, it is critical to report stratified evaluation and temporal and spatial validation, avoiding inflated conclusions derived from unrealistic data partitions or sampling biases.

Given the limited presence of explainability, future studies should incorporate XAI techniques and bias auditing as standard evaluation criteria, since operational adoption requires justifiability and governance, not merely aggregated performance.

4.2.3 RQ3: In which quartile levels are scientific publications related to the use of Artificial Intelligence in Crime Prediction concentrated?

Figures 11 and 12, together with Tables 12 and 13, integrate three key dimensions of the evidence: (i) concentration by quartile level (Q1–Q3), (ii) dissemination channel by editorial source, and (iii) annual evolution. This integration makes it possible to simultaneously position perceived quality, visibility, and bibliometric impact of research on Artificial Intelligence applied to Crime Prediction. Such triangulation is essential to interpret not only where research is published, but also why it is published there and with what citation performance.

As shown in the results, production is concentrated in Q1 (33/57; 57.9%) and Q2 (16/57; 28.1%), accounting for 86.0% of the total. This suggests that the topic has achieved sufficient technical maturity and social relevance to surpass high editorial thresholds, including methodological rigor, reproducibility, and ethical discussion.

According to the findings presented, impact is not distributed proportionally to volume. Q1 gathers the largest citation mass (475) and the highest H Index (5933), whereas Q3, although minority (14.0%), exhibits Citations per Paper = 15, which is compatible with a small number of studies that are potentially highly referenced, such as focused reviews or widely reusable methodological contributions.

Based on the observed distribution, source structure shapes impact. IEEE Xplore contributes fewer papers than ScienceDirect (13 versus 22) but concentrates more citations (334) and higher Citations per Paper (26), a pattern consistent with citation dynamics typical of engineering communities, including metric driven comparisons and reuse of frameworks and benchmarks.

Table 12. Distribution of Papers and Bibliometric Metrics by Quartile

Quartile	No. of Papers	No. of Citations	Citations/ Paper	H-Index
Q1	33	475	14	5933
Q2	16	8	1	953
Q3	8	119	15	652
Total	57	602	11	7538

Table 13. Distribution of Papers and Bibliometric Metrics by Source

Source	No. of Papers	No. of Citations	Citations/ Paper	H-Index
ScienceDirect	22	88	4	2341
IEEE Xplore	13	334	26	3770
Springer Link	13	162	12	696
ProQuest	5	12	2	378
Scopus	4	6	2	353
Total	57	602	11	7538

Likewise, the yearly flow shows a recent surge (2024: 19; 2025: 15), consistent with two causal mechanisms: (i) expansion of urban data, including sensors, open records, and video streams, along with increased computational capacity, and (ii) acceleration of advanced approaches, such as deep learning and NLP based LLMs, which increase publishable value in higher quartile venues.

Overall, the evidence suggests a field that predominantly publishes in high quartile journals, yet whose comparability remains conditioned by heterogeneity in datasets, crime definitions, and evaluation metrics. Therefore, concentration in Q1 and Q2 should be interpreted as a signal of visibility and editorial rigor, not as an automatic guarantee of evaluative standardization.

According to Tomaz, Fernandes, and Sciutti [62], the reviewed evidence was deliberately grounded in Q1 indexed journals, identifying 23 publications at that level. It is important to highlight that this methodological decision aims to maximize the robustness and reliability of findings and, consequently, to reduce uncertainty associated with literature of lower visibility or editorial rigor. In

ISSN 2007-9737

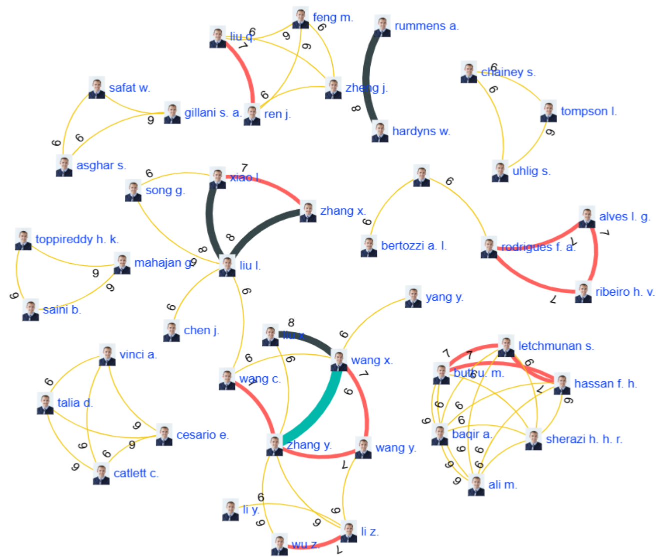

Fig. 13. Collaboration Network Among Authors

Table 14. Centrality and Cohesion Metrics in the Coauthorship Network of the Studies

Citation	Degree	Strength	Clustering Coefficient	Betweenness
liu l.	0.39	790	0.07	0.03
zhang y.	0.41	788	0.05	0.10
zhang x.	0.33	662	0.09	0.02
brantingham	0.32	646	0.07	0.04
wang y.	0.33	642	0.04	0.09
xiao l.	0.32	635	0.09	0.01
wang x.	0.29	581	0.05	0.05
rodrigues f. a.	0.29	579	0.11	0.01
hussain a.	0.3	566	0.1	0.01
hardyns w.	0.28	556	0.1	0.01
rummens a.	0.28	556	0.1	0.01
butt u. m.	0.28	555	0.12	0.00
hassan f. h.	0.28	555	0.12	0.00

a consistent line, Huamantingo, Cano Lengua, and Rodríguez [64] also prioritize first quartile sources as an essential component of their systematic review, although their emphasis lies more on the selection criterion than on reporting a specific volume of studies per quartile. Despite these reporting differences, both works converge in suggesting that research on Artificial Intelligence applied to Crime Prediction tends to concentrate, or to be preferentially evaluated, in Q1 channels; therefore, the comparative synthesis indicates a field seeking legitimacy through high quality evidence and demanding editorial standards. In sum, there is a clear orientation toward higher impact publications as the basis for building robust conclusions.

In other sectors, including financial fraud detection, cybersecurity, public health, and logistics, this concentration in Q1 and Q2 suggests that methodological designs and evaluation frameworks used in Artificial Intelligence based Crime Prediction may be transferable, provided they are adapted to asymmetric error costs, privacy requirements, and bias auditing needs.

Across other geographical regions and time periods, the pattern indicates that high impact publication will require spatio temporal validation, addressing data drift, socioeconomic changes, and seasonality, along with comparable reporting across cities and countries, avoiding conclusions dependent on a single contextual setting.

Given the weight of sources with high citations per paper, such as IEEE, future agendas should prioritize benchmark standards, data traceability, and governance frameworks in order to sustain scientific impact and operational applicability without degrading fairness, transparency, and institutional legitimacy.

4.2.4 RQ4: Which authors present the highest levels of coauthorship in research related to Artificial Intelligence and its influence on Crime Prediction?

The bibliometric network (Figure 13) and Table 14, which reports network metrics including Degree, Strength, Clustering Coefficient, and Betweenness, make it possible to identify core authors and bridge authors who structure scientific collaboration in Artificial Intelligence applied to Crime Prediction. This approach is relevant because coauthorship reflects not only productivity, but also the capacity to articulate teams, datasets, methodological frameworks, and research lines.

As shown in the results, the highest levels of coauthorship are concentrated in Zhang Y. (Degree = 0.41; Strength = 788) and Liu L. (Degree = 0.39; Strength = 790), suggesting leadership roles in large teams and or sustained participation in multiple technical collaborations.

According to the findings presented, Zhang Y. also exhibits the highest Betweenness (0.10), indicating that his coauthorship is not only extensive but also strategically interconnects subgroups, a pattern typical of researchers who integrate methodologies or datasets across communities.

Based on the observed distribution, Wang Y. also acts as an articulating node (Strength = 642; Betweenness = 0.09), consistent with dynamics in which certain authors facilitate the transfer of approaches, for example from spatial analytics to learning based models, and sustain repeated collaborations.

Likewise, authors with high Strength but low to moderate Betweenness, such as Brantingham P. J. (646; 0.04) or Xiao L. (635; 0.01), suggest intense collaboration within relatively consolidated research lines, with a more limited bridging function across clusters.

Overall, the coexistence of clusters with relatively low to moderate Clustering Coefficient values (approximately 0.04 to 0.12) indicates collaborative networks that combine strong cores with selective connections, likely driven by the need for urban datasets, computational infrastructure, and contextual validation, which typically require interdisciplinary cooperation.

According to Kaur and Saini [66], the coauthorship pattern in Artificial Intelligence research applied to Crime Prediction presents a clearly dominant actor, as Zhang J. appears as the most frequent coauthor and is positioned at the core of the network. It is important to note that his larger node size and connectivity with multiple authors from different groups suggest an articulating role in collaboration and knowledge circulation. In contrast, Yaghoubi, Yaghoubi, Khamees, and Vakili [64] shift the focus from the individual level to the macro dimension of collaboration, showing that India concentrates the largest number of international links, with prominent connections to the United States, the United Kingdom, and several countries across Asia, Europe, and Africa. Consequently, the country operates as a structural hub facilitating transnational cooperation. Although both analyses rely on different scales, author versus country, collectively they converge on the same interpretation: coauthorship in this field tends to be organized around central nodes with high intermediation; therefore, scientific production is driven by collaborative leadership and dense international networks. In sum, individual centrality and country to country connectivity emerge as complementary mechanisms explaining the highest levels of coauthorship observed.

In other sectors, including financial fraud detection, cybersecurity, public health, and retail

ISSN 2007-9737

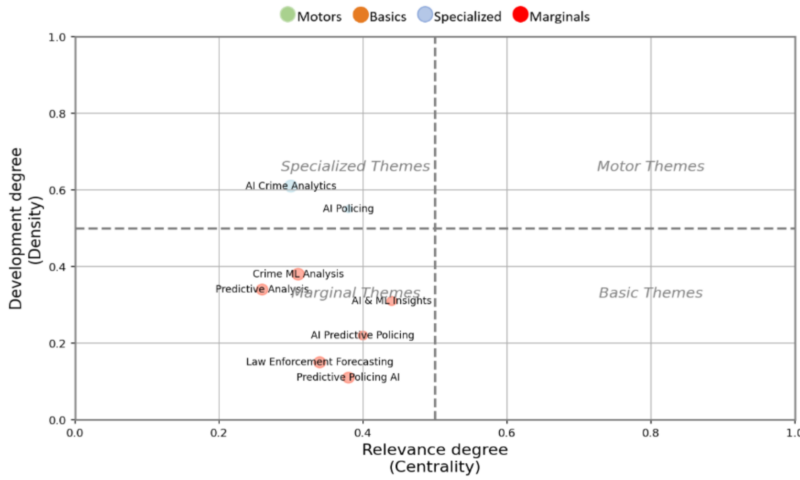

Fig. 14. Thematic map

Table 15. Thematic Clusters and Strategic Positioning based on Density and Centrality Indicator

Topic	Density	Centrality	Total Citations	Total Documents	Category
AI Crime Analytics	0.61	0.30	62	12	Specialized
AI Policing	0.55	0.38	21	7	Specialized
Crime ML Analysis	0.38	0.31	60	10	Marginal
Predictive Analysis	0.34	0.26	52	8	Marginal
AI & ML Insights	0.31	0.44	36	9	Marginal
AI Predictive Policing	0.22	0.40	34	8	Marginal
Law Enforcement Forecasting	0.15	0.34	54	9	Marginal
Predictive Policing AI	0.11	0.38	52	8	Marginal

analytics, identifying bridge authors with high betweenness is useful for transferring methodological frameworks and accelerating university industry partnerships oriented toward responsible and auditable predictive models. Across other geographical regions and time periods, the network suggests that research scalability will depend on multi city consortia and data sharing agreements; therefore, longitudinal replications and cross regional comparisons should prioritize teams capable of interconnecting clusters.

To strengthen impact and adoption, future agendas should encourage collaborations that integrate explainability metrics, bias assessment, and operational validation, avoiding domain or dataset closed networks and promoting standardization across technical and public policy communities.

4.2.5 RQ5: Which Thematic Categories characterize research addressing Artificial Intelligence and its impact on Crime Prediction?

Figure 14 and Table 15 synthesize, based on keyword analysis, the conceptual structure of the field of study on Artificial Intelligence and its impact on Crime Prediction. They integrate a strategic interpretation through the thematic map with quantitative evidence derived from the table. Taken together, these visualizations make it possible to identify which research lines exhibit greater relevance, reflected in centrality, and which demonstrate a higher degree of maturity, reflected in density, thereby avoiding fragmented interpretations based solely on term frequency.

As shown in the results, AI Crime Analytics appears as a specialized theme, characterized by high density and moderate centrality, suggesting a technically consolidated subfield that is not yet fully articulated as a transversal axis of the domain. The presence of AI Policing also within the specialized quadrant indicates methodological maturity driven by operational deployments, for example analytics for patrol management and resource allocation. However, its intermediate centrality suggests dependence on institutional and regulatory frameworks to scale toward the core of the field.

Based on the observed distribution, most topics are positioned as marginal, including Predictive Analysis, Crime ML Analysis, AI & ML Insights, AI Predictive Policing, and Predictive Policing AI. This reflects semantic fragmentation and diversity of approaches that still lack a unified vocabulary capable of increasing their centrality.

In particular, the moderate centrality of several marginal topics suggests that they function as connecting lines, or bridge keywords, yet their low density indicates that they still lack stable research agendas, comparable datasets, and consistent evaluation protocols.

Finally, the coexistence of semantically close themes, such as “predictive policing” versus “predictive policing AI,” points to a field in consolidation, in which terminological standardization and methodological convergence could shift these topics toward more central and developed quadrants.

According to Kaur and Saini [66], the thematic map suggests that data mining, crime prediction, and big data analytics constitute fundamental axes that are still in the process of consolidation, since their relevance is not accompanied by an equivalent level of development; consequently, they operate as basic lines that sustain the field, although they still require greater methodological maturity. In the same interpretation, machine learning and cybersecurity maintain relevance, although they exhibit lower density, indicating only partial articulation with dominant conceptual cores. In contrast, Garzón L., Ruiz O., Castro D., and Pérez Pertuzd highlight that artificial intelligence, law enforcement, and decision support systems function as motor themes, combining high centrality and high density; therefore, they structure both the research agenda and operational applications. Despite these differences, both analyses converge in identifying deep learning, and associated NLP approaches, as expanding lines that are not yet fully stabilized, while IoT tends to occupy more peripheral or specialized positions. In sum, the thematic categories of the field are organized around an applied core oriented toward decision support and policing practices, complemented by layers of data analytics, machine learning, and emerging technologies whose integration still shows heterogeneity and varying degrees of maturity.

These patterns suggest that, in order to mature the field, it is advisable to promote common standards for keywords, metrics, and experimental design, thereby facilitating comparison and transfer to other sectors, including financial fraud detection, cybersecurity, industrial safety, and logistics. The centrality density reading enables prioritization of lines with high applicability across other geographical contexts, such as cities with different data infrastructure levels, and across different time periods, including pre and post AI adoption, promoting longitudinal replicability.

Moreover, the transition of marginal topics toward basic or motor themes will depend on mechanisms such as data harmonization, multi city evaluation, and governance frameworks, preventing growth from being merely technical and ensuring that it becomes generalizable and institutionally sustainable.

5 Conclusions and Future Research

In summary, the systematic review confirms that research on Artificial Intelligence applied to Crime Prediction is advancing with a strong bias toward operational solutions and empirical validations, yet still with methodological heterogeneity that restricts inter study comparability.

With respect to RQ1, the marked concentration in Python (60%) suggests practical standardization driven by its library ecosystem, reproducible pipelines, and availability of skilled professionals, thereby reducing implementation friction compared to less widespread alternatives. At the same time, the secondary presence of R (20%) indicates that part of the field maintains a statistical and analytical orientation, consistent with scenarios in which modeling, inference, and data exploration precede industrial scale deployment. In contrast, the low frequency of Java, C++, and Julia (approximately 6.67% each) suggests that their adoption is confined to niches, such as embedded systems, extreme performance environments, or scientific prototyping, rather than serving as the development standard in the domain.

Regarding RQ3, the predominance of publications in Q1 (33/57) demonstrates that the topic is consolidating in high demand venues, likely due to its interdisciplinary character and increasing pressure for transparency, robustness, and responsible evaluation. Nevertheless, the relevant presence of Q2 (16/57) and Q3 (8/57) suggests a funnel effect, in which emerging proposals and proof of concept studies first mature in mid tier quartiles before stabilizing and scaling to leading venues. Moreover, distribution by source and bibliometric impact indicates that quality does not depend solely on quartile level, but also on the type of community, such as engineering versus applied social sciences, and on the reporting standards encouraged by each database. Consequently, the quartile pattern should be interpreted as a proxy for field maturity, but it does not replace critical evaluation of data bias, error costs, including false positives, and external validity, especially in multi city contexts.

Concerning RQ4, the coauthorship network reveals core authors with high connectivity and bridge authors with high betweenness, indicating that advancement in the domain depends on leadership capable of integrating heterogeneous datasets, methods, and teams. This structure is consistent with a data intensive field, in which cooperation becomes a causal mechanism for accessing urban data, infrastructure, and contextual validation; therefore, networks tend to concentrate around nodes with coordination capacity. According to the evidence, high coauthorship reflects not only productivity but also methodological transfer capacity across subcommunities, such as spatial analytics, deep learning, and intelligent surveillance, thereby accelerating technical convergence. At the same time, the international dimension of collaboration suggests that certain countries operate as hubs due to data availability, critical mass of research groups, and funding, which may bias research agendas toward specific urban realities.

With regard to RQ5, the thematic map reveals a conceptual structure still in consolidation. Few topics appear specialized or mature, while several remain marginal due to the absence of unified vocabulary and comparable protocols. This pattern suggests a mechanism of semantic fragmentation, in which closely related terms coexist as separate lines, diluting centrality and hindering cumulative evidence, despite growing production volume. Therefore, the transition of marginal topics toward basic or motor themes will depend more on keyword standardization, multi context evaluation, and availability of shared benchmarks than on isolated performance improvements within a single dataset. It is important to emphasize that the applicability of the findings is conditioned by declared limitations, including a restricted number of sources, a bounded temporal window, and possible indexing bias, which may underrepresent regional or technical literature not indexed in major databases. Consequently, critical reading of the field requires balancing technical advancement with governance: without common reporting criteria, including metrics, fairness, explainability, and error costs, adoption may grow faster than legitimacy and generalizability. In sum, the systematic review provides a coherent closure: the domain progresses through practical standardization, reflected in programming languages, editorial consolidation, reflected in quartiles, collaborative leadership, reflected in coauthorship, and conceptual maturation, reflected in thematic structures, yet urgently requires common frameworks to produce transferable evidence.

Future research should design multi city and longitudinal evaluations, applying consistent criteria and reporting standards, in order to distinguish real improvements from dataset artifacts and to validate transferability across regions and time periods. It is also necessary to systematically integrate explainability, bias assessment, error costs, and governance as primary outcomes, ensuring that technical effectiveness is not reported in isolation from institutional legitimacy and social risk. Finally, analysis should be extended to other sectors and business areas, including financial fraud detection, cybersecurity, retail, logistics, and public health, as well as to underrepresented geographical regions, verifying whether the same motor themes and collaboration patterns replicate or diverge.

References

1. **Rosés, R., Kadar, C., Malleson, N. (2021**). A Data-Driven Agent-Based Simulation to Predict Crime Patterns in an Urban Environment. Computers, Environment and Urban Systems, Vol. 101, pp. 101660.doi: 10.1016/j.compenvurbsys.2021.101660**.**
2. **Sarzaeim, P., Mahmoud, Q.H., Azim, A. (2024).** A Framework for LLM-Assisted Smart Policing System. IEEE Access, pp. 3404862. doi: 10.1109/ACCESS.2024.3404862.
3. **Li, Z., Zhang, X., Xu, F., Jing, X., Zhang, T. (2023).** A Multi-Scale Video Surveillance-Based Information Aggregation Model for Crime Prediction. Alexandria Engineering Journal. doi: 10.1016/j.aej.2023.04.045.
4. **Tasnim, N., Toufique Imam, I., Hashem, M.M.A. (2022).** A Novel Multi-Module Approach to Predict Crime Based on Multivariate Spatio-Temporal Data Using Attention and Sequential Fusion Model. IEEE Access, pp. 3171843. doi: 10.1109/ACCESS.2022.3171843.
5. **Yang, L., Guofan, J., Yixin, Z., Qianze, W., Jian, Z., Alizadehsani, R., Pławiak, P. (2024).** A Reinforcement Learning Approach Combined with Scope Loss Function for Crime Prediction on Twitter (X). IEEE Access, pp. 3473296. doi: 10.1109/ACCESS.2024. 3473296.
6. **Drury, B., Drury, S.M., Rahman, M.A., Ullah, I. (2022).** A Social Network of Crime: A Review of the Use of Social Networks for Crime and the Detection of Crime. Online Social Networks and Media, pp. 100211. doi: 10.1016/j.osnem.2022.100211.
7. **Shan, M., Ye, C., Peng, S., Peng, S. (2025).** Ada-GCNLSTM: An Adaptive Urban Crime Spatiotemporal Prediction Model. Journal of Neural and Spatial Systems Research, Vol. 11. doi: 10.1016/j.jnlssr.2024.11.003.
8. **Almasoud, A.S., Idowu, J.A. (2024).** Algorithmic Fairness in Predictive Policing. Artificial Intelligence and Law. doi: 10.1007/s43681-024-00541-3.
9. **Kshatri, S.S., Singh, D., Narain, B., Bhatia, S., Quasim, M.T., Sinha, G.R. (2021).** An Empirical Analysis of Machine Learning Algorithms for Crime Prediction Using Stacked Generalization: An Ensemble Approach. IEEE Access, pp. 3075140. doi: 10.1109/ACCESS.2021.3075140.
10. **Kleider-Offutt, H., Stevens, B., Mickes, L., Boogert, S. (2024).** Application of Artificial Intelligence to Eyewitness Identification. Cognitive Research: Principles and Implications, Vol. 9, pp. 19. doi: 10.1186/s41235-024-00542-0.
11. **Malmio, I. (2024).** Artificial Intelligence and the Social Dimension of Sustainable Development: Through a Security Perspective. Journal of Sustainable Development. doi: 10.1007/s43621-024-00677-6.
12. **Dubravova, H., Cap, J., Holubova, K., Hribnak, L. (2024).** Artificial Intelligence as an Innovative Element of Support in Policing. Procedia Computer Science, Vol. 5. doi: 10.1016/j.procs.2024.05.101.
13. **Ersöz, F., Ersöz, T., Marcelloni, F., Ruffini, F. (2025).** Artificial Intelligence in Crime Prediction: A Survey with a Focus on Explainability. IEEE Access, pp. 3553934. doi: 10.1109/ACCESS.2025.3553934.
14. **Shad, K.B. (2023).** Artificial Intelligence-Related Anomalies and Predictive Policing: Normative (Dis)Orders in Liberal Democracies. AI & Society. doi: 10.1007/s00146-023-01751-9.
15. **Adams, I.T. (2025).** Automation and Artificial Intelligence in Police Body-Worn Cameras: Experimental Evidence of Impact on Perceptions of Fairness among Officers. Journal of Criminal Justice, Vol. 97, pp. 102373. doi: 10.1016/j.jcrimjus.2025.102373.
16. **Tuarob, S., Tatiyamaneekul, P., Pongpaichet, S., Tawichsri, T., Noraset, T. (2025).** Beyond Administrative Reports: A Deep Learning Framework for Classifying and Monitoring Crime and Accidents Leveraging

Large-Scale Online News. Neural Computing and Applications. doi: 10.1007/s00521-024-10833-8.

17. **Shah, N., Bhagat, N., Shah, M. (2021).** Crime Forecasting: A Machine Learning and Computer Vision Approach to Crime Prediction and Prevention. Crime Science. doi: 10.1186/s42492-021-00075-z.
18. **Kim, G., Cho, Y., Han, Y., Lee, G. (2025).** Crime Mapping in Urban Environments Using Explainable AI: A Case Study of Daegu, Korea. Sustainable Cities and Society, pp. 106507. doi: 10.1016/j.scs.2025.106507.
19. **Hussain, F.S., Aljuboori, A.F. (2025).** Crime Prediction Based on Classification Approaches. Procedia Computer Science. doi: 10.1016/j.procs.2025.04.095.
20. **Yin, J. (2023).** Crime Prediction Methods Based on Machine Learning: A Survey. Computers, Materials & Continua. doi: 10.32604/cmc.2023.034190.
21. **Alsubayhin, A., Ramzan, M.S., Alzahrani, B. (2024).** Crime Prediction Model Using Three Classification Techniques: Random Forest, Logistic Regression, and LightGBM. International Journal of Advanced Computer Science and Applications. doi: 10.14569/IJACSA.2024.0150123.
22. **Mukto, M.M., Al Mahmud, M., Haque, I., Ahmed, M.A., Jabid, T., Ali, M.S., Rashid, M.A., Islam, M.M., Islam, M. (2024).** Design of a Real-Time Crime Monitoring System Using Deep Learning Techniques. Intelligent Systems with Applications, pp. 200311. doi: 10.1016/j.iswa.2023.200311.
23. **Hou, F. (2022).** Echoing Mechanism of Juvenile Delinquency Prevention and Occupational Therapy Education Guidance Based on Artificial Intelligence. Computational Intelligence and Neuroscience, pp. 9115547. doi: 10.1155/2022/9115547.
24. **Berardini, D., Migliorelli, L., Galdelli, A., Marín-Jiménez, M.J. (2025).** Edge Artificial Intelligence and Super-Resolution for Enhanced Weapon Detection in Video Surveillance. Engineering Applications of Artificial Intelligence, pp. 109684. doi: 10.1016/j.engappai.2024.109684.
25. **Safat, W., Asghar, S., Gillani, S.A., Martínez, V.F. (2021).** Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques. IEEE Access, pp. 3078117. doi: 10.1109/ACCESS.2021.3078117.
26. **Wu, J., Abrar, S.M., Awasthi, N., Frias-Martinez, E. (2022).** Enhancing Short-Term Crime Prediction with Human Mobility Flows and Deep Learning Architectures. EPJ Data Science. doi: 10.1140/epjds/s13688-022-00366-2.
27. **Choi, W., Na, J., Lee, S. (2024).** Evaluating Intelligent CPTED Systems to Support Crime Prevention Decision-Making in Municipal Control Centers. Applied Sciences. doi: 10.3390/app14156581.
28. **İlgün, E.G., Dener, M. (2025).** Exploratory Data Analysis, Time Series Analysis, Crime Type Prediction, and Trend Forecasting in Crime Data Using Machine Learning, Deep Learning, and Statistical Methods. Neural Computing and Applications. doi: 10.1007/s00521-025-11094-9.
29. **Coleman, T., Mokilane, P., Rangata, M. et al. (2024).** Exploring the Usefulness of the INLA Model in Predicting Levels of Crime in the City of Johannesburg, South Africa. Crime Science, Vol. 13, pp. 25. doi: 10.1186/s40163-024-00219-5.
30. **Kim, D., Kan, Y., Aum, Y., Lee, W., Yi, G. (2023).** Hotspots-Based Patrol Route Optimization Algorithm for Smart Policing. Heliyon, pp. e20931. doi: 10.1016/j.heliyon.2023.e20931.
31. **Johnson, S.D. (2024).** Identifying and Preventing Future Forms of Crimes Using Situational Crime Prevention. Security Journal. doi: 10.1057/s41284-024-00441-5.
32. **Bandpey, Z., Piri, S., Shokouhian, M. (2025).** Integrating Machine Learning Techniques for Enhanced Safety and Crime Analysis in Maryland. Applied Sciences. doi: 10.3390/app15094642.
33. **Sultana, T., Wahid, K.A., Awasthi, N. (2019).** IoT-Guard: Event-Driven Fog-Based Video Surveillance System for Real-Time Security

Management. IEEE Access, pp. 2941978. doi: 10.1109/ACCESS.2019.2941978.

34. **Gao, Y., Xiong, N., Yu, W., Lee, H.J. (2019).** Learning Identity-Aware Face Features Across Poses Based on Deep Siamese Networks. IEEE Access, pp. 2932760. doi: 10.1109/ACCESS.2019.2932760.
35. **Karn, R.R., Kudva, P., Elfadel, I.M. (2021).** Learning without Forgetting: A New Framework for Network Cyber Security Threat Detection. IEEE Access, pp. 3115946. doi: 10.1109/ACCESS.2021.3115946.
36. **Ippolito, A., Lozano, A. (2023).** Machine Learning in Crime Prediction. Journal of Ambient Intelligence and Humanized Computing. doi: 10.1007/s12652-023-04530-y.
37. **Cesario, E., Lindia, P., Vinci, A. (2024).** Multi-Density Crime Predictor: An Approach to Forecast Criminal Activities in Multi-Density Crime Hotspots. Journal of Big Data. doi: 10.1186/s40537-024-00935-4.
38. **Tam, S., Tanriöver, Ö.Ö. (2023).** Multimodal Deep Learning Crime Prediction Using Tweets. IEEE Access, pp. 3308967. doi: 10.1109/ACCESS.2023.3308967.
39. **Adams, I.T., Barter, M., McLean, K., Boehme, H., Geary, I.A. (2024).** No Man's Hand: Artificial Intelligence Does Not Improve Police Report Writing Speed. AI & Society. doi: 10.1007/s00146-024-09644-7.
40. **Jeon, H., Kim, H., Kim, D., Kim, J. (2024).** PASS-CCTV: Proactive Anomaly Surveillance System for CCTV Footage Analysis in Adverse Environmental Conditions. Expert Systems with Applications, pp. 124391. doi: 10.1016/j.eswa.2024.124391.
41. **Johnson, T.L., Johnson, N.N., Topalli, V., McCurdy, D., Wallace, A. (2024).** Police Facial Recognition Applications and Violent Crime Control in US Cities. Cities, pp. 105472. doi: 10.1016/j.cities.2024.105472.
42. **Robertson, D.L., Goodridge, W.S. (2022).** Predicting Density of Serious Crime Incidents Using a Multiple-Input Hidden Markov Maximization a Posteriori Model. Machine Learning with Applications, pp. 100231. doi: 10.1016/j.mlwa.2021.100231.
43. **Pouyan, S., Charmi, M., Azarpeyvand, A., Hassanpoor, H. (2023).** Propounding First Artificial Intelligence Approach for Predicting Robbery Behavior Potential in an Indoor Security Camera. IEEE Access, pp. 3284472. doi: 10.1109/ACCESS.2023.3284472.
44. **Csatlós, E. (2024).** Prospective Implementation of AI for Enhancing European (In)Security: Challenges in Reasoning of Automated Travel Authorization Decisions. Computer Law & Security Review, pp. 105995. doi: 10.1016/j.clsr.2024.105995.
45. **Kaur, M., Saini, M. (2024).** Role of Artificial Intelligence in the Crime Prediction and Pattern Analysis Studies Published over the Last Decade: A Scientometric Analysis. Artificial Intelligence Review, Vol. 57, pp. 202. doi: 10.1007/s10462-024-10823-1.
46. **Saxena, N., Varshney, D. (2021).** Smart Home Security Solutions Using Facial Authentication and Speaker Recognition through Artificial Neural Networks. International Journal of Cognitive Computing in Engineering. doi: 10.1016/j.ijcce.2021.10.001.
47. **Mekonnen, Z.A., Engda, E.M., Deraro, K.F., Agegnehu, N., Reda, T.M., Jothimani, M. (2025).** Spatio-Temporal Crime Pattern Analysis in Addis Ketema, Addis Ababa, Ethiopia: GIS and R-Based Approaches. Scientific African, pp. e02656. doi: 10.1016/j.sciaf.2025.e02656.
48. **Butt, U.M., Letchmunan, S., Hassan, F.H., Ali, M., Baqir, A., Sherazi, H.H.R., Koh, T.W. (2021).** Spatio-Temporal Crime Predictions by Leveraging Artificial Intelligence for Citizens Security in Smart Cities. IEEE Access, pp. 3068306. doi: 10.1109/ACCESS.2021. 3068306.
49. **Kim, H.-B., Choi, N., Kwon, H.-J., Kim, H. (2023).** Surveillance System for Real-Time High-Precision Recognition of Criminal Faces from Wild Videos. IEEE Access, pp. 3282451. doi: 10.1109/ACCESS.2023.3282451.
50. **Cabrera, B.M., Luiz, L.E., Teixeira, J.P. (2025).** The Artificial Intelligence Act: Insights Regarding Its Application and Implications. Procedia Computer Science. doi: 10.1016/j.procs.2025.02.116.

51. **Haley, P. (2025).** The Impact of Biometric Surveillance on Reducing Violent Crime: Strategies for Apprehending Criminals While Protecting the Innocent. Sensors. doi: 10.3390/s25103160.
52. **Ferreira dos Santos, J.P., de Matos, C.A., Groznik, A. (2025).** The Role of Artificial Intelligence in Smart City Systems Usage: Drivers, Barriers, and Behavioural Outcomes. Technological Forecasting and Social Change, pp. 102867. doi: 10.1016/j.techfore.2025.102867.
53. **Abdulkareem, F.A.A. (2024).** The Role of Predictive Policing in Predicting Spatio-Temporal Crime Mapping. Sociology and Criminology Studies, pp. 112. doi: 10.21638/spbu14.2024.112.
54. **Cavus, M., Benli, M.N., Altuntas, U., Sari, M., Ayan, H., Ugurluoglu, Y.F. (2025).** Transparent and Bias-Resilient AI Framework for Recidivism Prediction Using Deep Learning and Clustering Techniques in Criminal Justice. Applied Soft Computing, pp. 113160. doi: 10.1016/j.asoc.2025.113160.
55. **Halford, E., Gibson, I. (2025).** Using Machine Learning to Conduct Crime Linking of Residential Burglary. International Journal of Law, Crime and Justice, pp. 100716. doi: 10.1016/j.ijlcj.2024.100716.
56. **Ham, J.-S., Kim, D.H., Moon, J. (2025).** WatchoutPed: A Dataset and Model for Vulnerable Pedestrian Anticipation in Surveillance Videos. Knowledge-Based Systems, pp. 113352. doi: 10.1016/j.knosys.2025.113352.
57. **Podoletz, L. (2023).** We Have to Talk About Emotional AI and Crime. AI & Society. doi: 10.1007/s00146-022-01435-w.
58. **Kitchenham, B. (2004).** Procedures for Performing Systematic Reviews. Joint Technical Report, Keele University and NICTA.
59. **Mandalapu, V., Elluri, L., Vyas, P., Roy, N. (2023).** Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions. IEEE Access, Vol. 11, pp. 60153–60170. doi: 10.1109/ACCESS.2023.3286344.
60. **Tomaz, M.S.L.S., Fernandes, B.J.T., Sciutti, A. (2025).** Identification of Anomalous Behavior through the Observation of an Individual's Emotional Variation: A Systematic Review. IEEE Access, Vol. 13, pp. 32927–32949. doi: 10.1109/ACCESS.2025.3540034.
61. **Simisterra-Batallas, C., Pico-Valencia, P., Sayago-Heredia, J., Quiñónez-Ku, X. (2025).** Internet of Things and Deep Learning for Citizen Security: A Systematic Literature Review on Violence and Crime. Future Internet, Vol. 17, No. 4, pp. 159. doi: 10.3390/fi17040159.
62. **Butt, U.M., Letchmunan, S., Hassan, F.H., Ali, M., Baqir, A., Sherazi, H.H.R. (2020).** Spatio-Temporal Crime Hotspot Detection and Prediction: A Systematic Literature Review. IEEE Access, Vol. 8, pp. 166553–166574. doi: 10.1109/ACCESS.2020.3022808.
63. **Fatima, D., Qassim, N., Manar, A.T., Omnia, A.W., Ali, B.N., Sohail, A. (2022).** Artificial Intelligence and Crime Prediction: A Systematic Literature Review. Social Sciences & Humanities Open. doi: 10.1016/j.ssaho.2022.100342.
64. **Huamantingo, R., Cano-Lengua, M., Rodriguez, C. (2025).** Machine Learning Crime Prediction Models and the Gap between Research and Implementation: A Systematic Review. Karbala International Journal of Modern Science, Vol. 11, No. 3, Article 11. doi: 10.33640/2405-609X.3419.
65. **Yaghoubi, E., Yaghoubi, E., Khamees, A. et al. (2024).** A Systematic Review and Meta-Analysis of Artificial Neural Network, Machine Learning, Deep Learning, and Ensemble Learning Approaches in the Field of Geotechnical Engineering. Neural Computing and Applications, Vol. 36, pp. 12655–12699. doi: 10.1007/s00521-024-09893-7.
66. **Kaur, M., Saini, M. (2024).** Role of Artificial Intelligence in the Crime Prediction and Pattern Analysis Studies Published over the Last Decade: A Scientometric Analysis. Artificial Intelligence Review, Vol. 57, pp. 202. doi: 10.1007/s10462-024-10823-1.
67. **Dahmane, W.M., Ouchani, S., Bouarfa, H. (2025).** Smart Cities Services and Solutions: A Systematic Review. Data and Information

Management, Vol. 9, No. 2, pp. 100087. doi: 10.1016/j.dim.2024.100087.

68. **Garzón Santos, J.L., Ruiz Otálora, J.H., Castro Duarte, J.G., Pérez Pertuz, J.J. (2024).** Importancia de la inteligencia artificial para los cuerpos de policía: un análisis bibliométrico. Revista Logos Ciencia & Tecnología, Vol. 17, No. 1, pp. 119–134. doi: 10.22335/rlct.v17i1.2009.

69. **PRISMA (2025).** Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

70. **Gamboa-Cruzado, J.A., Cristomo-Castro, R., Vila-Buleje, J., López-Goycochea, J., Valenzuela, J. (2024).** Heart Attack Prediction Using Machine Learning: A Comprehensive Systematic Review and Bibliometric Analysis. Journal of Theoretical and Applied Information Technology, Vol. 102, No. 5, pp. 1930–1944.

71. **Gamboa-Cruzado, J.A., Méndez-Morales, C., Franco del Carpio, J., López-Goycochea, J., Alva, A. (2023).** Use of Chatbots in E-Commerce: A Comprehensive Systematic Review. Journal of Theoretical and Applied Information Technology, Vol. 101, No. 4, pp. 1172–1183.

72. **Alegre-Veliz, R., Gaspar-Ortiz, P., Gamboa-Cruzado, J.A., Rodríguez Baca, L., Grande Pizarro, W., Méndez Mueras, R. (2022).** Machine Learning for Feeling Analysis in Twitter Communications: A Case Study in HEYDRU!, Perú. International Journal of Interactive Mobile Technologies, Vol. 16, No. 24, pp. 126–142. doi: 10.3991/ijim.v16i24.35493.

73. **Rojas Valdivia, J., Gamboa-Cruzado, J., de la Cruz Vélez, P. (2023).** Systematic Literature Review on Machine Learning and Its Impact on APIs Deployment. Computación y Sistemas, Vol. 27, No. 4, pp. 1107–1124. doi: 10.13053/CyS-27-4-4371.

74. **Andrade-Mogollon, T., Gamboa-Cruzado, J., Amayo-Gamboa, F. (2025).** Systematic Literature Review of Generative AI and IoT as Key Technologies for Precision Agriculture. Computación y Sistemas, Vol. 29, No. 2, pp. 857–882. doi: 10.13053/CyS-29-2-5738.

75. **Cárdenas-Quispe, A., Vergaray-Mezarina, R., Gamboa-Cruzado, J. (2021).** Machine learning para la detección de malware en Android: revisión sistemática de la literatura. RISTI – Revista Ibérica de Sistemas e Tecnologias de Informação, No. 45, pp. 318–331.

Article received on 28/07/2025; accepted on 13/11/2025.
**Corresponding author is Cesar Jesús Núñez-Prado.*